All correspondence regarding *Voting matters* should be addressed to:

B A Wichmann, The Editor, *Voting matters*, The Electoral Reform Society, 6 Chancel Street, London, SE1 0UU.

Voting matters

for the technical issues of STV

The Electoral Reform Society

Volume 1, Issues 1-10 March 1994 - March 1999

Editorial

In recent years *Representation* has tended to shy away from articles of a technical nature and restrict itself to the non-technical. While there may be some advantages in this course of action, it has left those with technical things to say on voting systems without a suitable outlet for their ideas and arguments. The members of the Electoral Reform Society's Technical Committee, and others, have been unhappy about this. Hence this new venture, which it is intended to circulate to those Society members who request it.

In this first issue, we reprint some earlier articles that deserve a wider circulation. Those by B L Meek, originally published over 20 years ago in French, have been available in English only as a typed and duplicated version containing many errors. These are classic papers which have led to much discussion in recent years. Whether one agrees with Meek's conclusions or not, it cannot be denied that those who argue about his method need to know what he did actually say.

The article by D R Woodall was also printed with an error originally and this reprint includes the necessary correction. Although Woodall's method is basically the same as Meek's, it was entirely independently derived and it is interesting to see his different approach.

The article by C H E Warren has not been published before. It is a slightly rewritten version of a paper first submitted in 1983, but not then accepted. Warren's method is similar in spirit to the other two, but differs in the way it performs. Each of the two counting methods has an advantage over the other in some circumstances so, although a majority of the ERS Technical Committee prefer the Meek/ Woodall formulation, the Warren alternative is worth bearing in mind. The final paper discusses the differences.

I D Hill

Chairman, ERS Technical Committee

A New Approach to the Single Transferable Vote

Paper I: Equality of Treatment of voters and a feedback mechanism for vote counting.

B L Meek, Computer Unit, Queen Elizabeth College, London.

Brian Meek is now at King's College London.

With some differences in presentation, the paper was originally published in French in Mathématiques et Sciences Humaines, No 25, pp13-23, 1969.

Abstract

It is shown that none of the counting methods so far used in single transferable vote elections satisfies the criterion that all votes should, as far as possible, be taken equally into account. A feedback method of counting is described which does satisfy this criterion within the general limitations imposed by the STV system. This counting method, though very laborious for manual counting, would be feasible in automated elections.

1. Introduction

While the preferential voting system known as the Single Transferable Vote (STV)[1] has been criticised on various grounds, the following advantages claimed for it do not seem to have been seriously challenged:

(A) The number of ‘wasted’ votes in an election (i.e, which do not contribute to the election of any candidate) is kept to a minimum.

(B) As far as possible the opinions of each voter are taken equally into account.

(C) There is no incentive for a voter to vote in any way other than according to his actual preference.

It is the purpose of this and a subsequent paper to consider (A), (B) and (C) from a decision-theoretic viewpoint, within a single constituency; it will be shown that (A), (B) and (C) in fact do not hold in present STV procedures, but may be made to hold, within certain overall limitations, by appropriate modification of the counting method.

2. The wasted vote

An essential feature of an STV election is the ‘quota’. If there are s vacancies to be filled, the quota q is the *smallest* number such that, if s candidates have q votes each, it is not possible for an $(s+1)$th candidate to have as many as q votes. Thus if the total votes are T, then $T-sq < q$, but $T-s(q-1) \geq q-1$, whence $q = [1+T/(s+1)]$, where the square brackets denote ‘integer part of’.

Candidates with more than q votes are elected, and have their surplus votes transferred according to the next preferences marked; if there are no such candidates, the bottom candidate is eliminated and all his votes so transferred. Repeated application of these rules ensures that at the end of the count s candidates have at least q votes each and so the total wasted vote w satisfies $w < T/(s+1)$.

Given s and T, it is clear from the definition of q that condition (A) is satisfied provided the next preference at each transfer is always given. It is possible for the above inequality, and hence condition (A), to be violated, if w is increased by the addition of votes which are non-transferable because no next preference has been indicated. In this paper we shall assume that this does not occur; it will be shown in a second paper that it is possible still to satisfy (A) in such cases by modifying the definition of q.

3. Equality of treatment

The discussion of condition (A) shows that, in general, there will be some wasted votes, except in the trivial cases when $s \geq T$. It is therefore not possible under STV to guarantee that *all* votes will be taken equally into account (e.g. votes with first preferences for runner-up candidates), although all are taken indirectly into account when calculating the quota.[2]

Within this obvious limitation, attempts have been made to eliminate possible sources of inequity of treatment by various modifications of the counting rules. Such sources include:

(i) the choice of which votes to transfer from the total for a candidate who has exceeded the quota

(ii) errors introduced by taking whole-number approximations to fractions of totals for transfer − particularly in elections with small total vote

(iii) calculation of the proportion for transfer from an elected candidate on the basis of the last batch of votes transferred to him, and not on his total vote.

The common way of overcoming difficulties (i) and (ii) is to use the variant of STV known as the Senate Rules. Each vote is divided into K parts (usually K = 100 or 1000) and each part treated as a separate vote (of value $1/K$) with identical preference listings.

Difficulty (i) is overcome by transferring the appropriate proportion of each divided vote, while the method clearly reduces the errors involved in (ii) by the factor $1/K$. If $K=10^n$ this is simply working to n decimal places. The value of K has only to be increased until the errors are too small to affect the result of the election.[3] The method is equivalent to transferring the *whole* vote at an appropriately reduced value, and it is this interpretation we shall use from now on.

Difficulty (iii) is slightly more technical, and warrants further explanation. Suppose at some stage a candidate has obtained x ($<q$) votes. By transfer from another (elected or eliminated) candidate he now acquires a further y votes, where $x+y \geq q$. His surplus is now $z=x+y-q$. It would appear that his $x+y$ votes should now be transferred, with value reduced by the factor $z/(x+y)$.

It is, however, common practice for only the y votes to be transferred, with value reduced by the factor z/y. The reason for adopting this procedure is simply the practical one, in a manual count, of reducing as much as possible the rescrutiny of ballots for later preferences. However, neither this nor the argument that 'the difference is unlikely to affect the result' are particularly relevant to a decision-theoretic discussion, though we shall return to the practicability problem later.

Of more importance here is the argument 'in STV a vote only counts for one candidate at a time, and should count for the first preference where possible'. If accepted, this would of course also render difficulties (i) and (ii) irrelevant, and the Senate Rules unnecessary; the first part of it is in fact sometimes used as a 'proof' that STV satisfies condition (B). But even without the Senate Rules the statement is false; however the surplus votes are chosen for transfer, *it is the existence of the untransferred votes which makes the transferred votes surplus.* A vote not only counts directly for one candidate; it can indirectly affect the progress of the count, the pattern of transfers, and ultimately the *election or non-election of other candidates.*[4]

It is this fact which is at the root of the failure of STV to satisfy condition (B).

In the specific situation described above, the candidate achieves election not only because of the accession of the *y* new votes, but because of the existence of the *x* previous votes; hence for condition (B) to be satisfied, all *x+y* votes should be transferred at the appropriate reduced value.

However, there is yet a fourth difficulty, one which does not seem to have been recognised hitherto.

(iv) In determining the next preference to which a vote is to be transferred, *elected as well as eliminated candidates are ignored.*

Let us suppose that of *y* votes to be transferred, *y/2* are marked next to go to candidate A, and *y/2* to candidate B. Let us further suppose that A has already been elected; under STV the *y/2* votes which would otherwise go to him are transferred to the next candidate marked (assumed C in every case) provided that that candidate is not also already elected. Thus *y/2* go to B, and *y/2* to C. The inequities are plain; the votes for A which enabled the *y/2* to go to C rather than A had no say in their destination, while C obtains these votes at the same value as B receives his. Suppose these *y* votes were originally first-preference votes for a candidate D, now eliminated; those who voted for A next and then C at least have had their second choice elected, while those who voted next for B have not − yet these votes go, under STV, to both B and C at full value.

In section 6 we shall describe a counting mechanism which overcomes all these difficulties.

4. Making the most of one's vote

Any system which contains wasted votes contains at least some element of incentive to vote in other than his preferred way; the case for (C) in STV is that it is difficult for a voter to be sure (rightly or wrongly) that his vote will be wasted, both because the number of wasted votes is relatively small, and because the wasted votes are those for the non-elected but non-eliminated candidates − i.e. of the stronger, not the weaker, runners-up. However, it is also possible for voters to take advantage of the features of STV described in section 3, provided they are sufficiently well informed, by voting in a sophisticated manner. This is most easily shown by an example:

Let *T*=3599, *s*=3, *q*=900, and the unsophisticated first-preference votes for the six candidates A, B, ... F be as follows:

A	B	C	D	E	F
1020	890	880	589	200	20

In this case the 120 surplus votes of A divide 60 to B, 20 to C, 40 to D and the elected candidates are A, B and C.

Suppose there are 170 voters who above voted A, D, C ... It is known that the second-preference votes of F will go to C, and of E to D. Then the sophisticated way for these 170 to vote is F, A, D, C,... *in order to prevent A from being elected on the first count.*

A	B	C	D	E	F
850	890	880	589	200	190

On the elimination of F, his original 20 votes go to C, and the 170 sophisticated votes return to A. However, the 120 surplus is now taken entirely from this batch (see (iii) in section 3) and goes to D. C having no surplus, E must be eliminated and D is elected.

A different type of sophisticated voting is given below:

T=239, *s*=2, *q*=80.

Unsophisticated case: C and A elected:

C,A,B...	C,B,A...	B,A....	A,B.....
120	80	31	8

Sophisticated case: C and B elected:

C,A,B...	C,B,A...	E,B,A...	B,A....	A,B.....
120	50	30	31	8

It seems to be a new result that sophisticated voting is possible in STV, though it is well-known that it can occur in other voting systems and considerable work has been done on decision processes using a games-theoretic approach. Black[5] in his discussion of STV does mention the possibility of ‘an organised minority (perverting) the use of the system’ but only in connection with a candidate with just the quota on first preferences who is rated last by the rest of the electorate. STV supporters would claim that if a candidate can obtain a quota this *ipso facto* entitles him to be elected, particularly if he gets the quota on first preferences, and it is certainly difficult to understand what Black means by ‘pervert’ in this context.

5. Other considerations

At this point we shall mention some other aspects of STV, mainly in order to define the limitations of the present discussion. Proper treatment of the points raised in this section are well outside the scope of the present work, and is the subject of a projected further, more general paper.

The conditions (A), (B), (C) discussed so far were chosen simply because they seem to be specific to STV among constituency-type systems in parliamentary elections. However, other conditions could be applied, notably those specified by Arrow in his General Possibility Theorem.[6]

As STV elections are multi-vacancy, the preferences between candidates listed by the voters do not as they stand represent an ordering of *independent* alternatives, and so Arrow's analysis is not directly applicable. The deduction from the voter’s ordering of candidates of his ordering of the actual independent alternatives (the possible subsets of the set of all candidates who might actually be elected) is by no means straightforward. Nevertheless, at some stage of the count the process reduces to electing one candidate to one remaining vacancy, and so the consequences of the theorem, and the Condorcet paradox, cannot be escaped. Using the alternatives as they stand, even though they are not independent, STV clearly satisfies Arrow's conditions 1, 4, and 5. The condition 3 of independence of irrelevant alternatives is not satisfied, nor is condition 2 (the positive association of social and individual values). This can be seen from the above analysis.

A related point, and probably the strongest decision-theoretic argument against STV, is the fact that a candidate may be everyone's second choice but not be elected. This difficulty is not overcome by the feedback method, and it does not seem to the author to be possible to do so while retaining a system which would be recognisably a ‘single’ transferable vote.

Virtually all other discussion of STV, both for and against, seem to have been about political and not decision-theoretic considerations.

For example, Black[5] does discuss STV from what he terms the ‘statical’ point of view, but although he does express some disquiet about the ‘heterogeneity’ involved in STV (basically, that some votes count for first preferences, others for second or later preferences), he does not go into the problem in detail and concludes ‘in spite of those drawbacks (STV) has merits ... it is not difficult to see why many people, *regarding it purely as a statical system*, (Black's italics) should hold (it) in esteem’. The italicised phrase is to introduce other, ‘dynamical’ arguments against STV.[7] Black does not discuss the conditions mentioned here; though the germ of the idea of inequity is contained in the word ‘heterogeneity’; in fact as section 3 shows, the heterogeneity which worries him is more apparent than real, and the feedback method described in section 6 eliminates what there is. Nor − oddly − does the ‘everyone's second choice’ problem, even though this is closely connected with the doubts mentioned at the end of the last section.

6. The feedback process

One of the criticisms of STV which is often made is that its rules are too complicated, and are not derived from principles which can be simply stated. The above discussion shows that this is not surprising; the rules are in many cases little more than rules of thumb, designed for practical convenience rather than theoretic merit. The feedback process, however, is derived from simply-stated principles:

Principle 1. If a candidate is eliminated, all ballots are treated *as if that candidate had never stood.* [8]

Principle 2. If a candidate has achieved the quota, he retains a fixed proportion of every vote received, and transfers the remainder to the next *non-eliminated* candidate, the retained total equalling the quota.

Principle 1 is the one which leads to the feedback mechanism. For, suppose a voter marks his ballot A, B, C,.. and A is eliminated, the ballot, by Principle 1, is henceforward treated as if it read B, C,.. on the assumption that if A had not stood at all, the voter would have ordered the other candidates as before and B would have been first preference[9]. But suppose that B has at an earlier count reached the quota. Then this ballot must now be treated as an original first preference for B; that is, according to Principle 2, the same proportion of this vote must be retained by B as for the others, passing the rest to C (instead of the *whole* vote going to C as in previous methods). However, this will mean that the total retained by B is now greater than the quota. Thus the proportion of B's votes to be retained must be recalculated, and will in fact drop − in other words we must go back to the beginning, with A now eliminated. This is the feedback process.

Note that the proportion of each of B's votes to be transferred is increased by this accession of support; B's supporters have a say in the transfer of the extra surplus, since it is their existence which has made it surplus. All support for B is now treated equally, being divided proportionately to leave him with exactly the quota.

Consider now the effect of Principle 2. The transfer of B's vote may lead to another candidate, D, being elected. All votes, new and old, for D, have now to be divided, leaving D with the quota and distributing the rest to the next *non-eliminated* candidate. Some ballots may have B, another elected candidate, as next candidate. Under previous rules, only continuing (i.e. non-eliminated and *non-elected*) candidates can receive transfers. Now these votes are regarded as extra support for B: he takes the proportion allotted him by D, *retains the proportion that he keeps of all he receives*, and transfers the rest − now the *third* marked candidate. Formerly the third candidate would get all of the proportion transferred by D (see (iv) section 3).

It can be seen that B will once more have more than the quota if he does not again reduce the proportion which he retains. However, the increased proportion transferred may in part go to D who will therefore have to reduce the proportion *he* retains. This will react back on B, and it is clear that we have an infinite regression. However, it is also clear that the proportions for transfer do not increase without limit, there being only a finite total surplus available from B and D, who must each retain a quota. The problem is in fact a mathematical one of determining the proportions to be retained by each which will leave them both with a quota, taking into account the extent of mutual support. If p_B is the proportion B transfers, and p_D that which D transfers, supporters of both B and D have their votes transferred to third preferences at value $p_B p_D$. Those putting B first have $1-p_B$ retained by him and $p_B(1-p_D)$ retained by D; those putting D first have $1-p_D$ retained by him and $p_D(1-p_B)$ retained by B.

We now, as examples, give the formulae for the proportions for transfer in the cases of 1, 2, 3 and 4 elected candidates:

One candidate

$$t_1(1-p_1)=q$$

This is the same formula as before, except that t_1 now contains all effective first-preference votes for the candidate, including those obtained from eliminated candidates, who by Principle 1 are now ignored. The proportion p_1 is recalculated every time t_1 is increased by the elimination of a candidate.

Two candidates

The first elected candidate has t_1 first preference votes, of which t_{12} have the second elected candidate as second preference. Hence $p_1 t_{12}$ are passed on to that candidate. Similarly $p_2 t_{21}$ are received from the second candidate. Thus

$$(t_1+p_2t_{21})(1-p_1)=q$$

$$(t_2+p_1t_{12})(1-p_2)=q$$

Three candidates

The votes received by candidate 1 are now his first-preference t_1, second-preference p_2t_{21} from candidate 2 and p_3t_{31} from candidate 3, and third-preference $p_2(p_3t_{321})$ from candidate 3 (1st), 2 (2nd) and $p_3(p_2t_{231})$ from candidate 2 (1st), 3(2nd).

Thus:

$$[t_1+p_2t_{21}+p_3t_{31}+p_2p_3(t_{321}+t_{231})](1-p_1)=q$$

Two similar formulae hold, obtained by cyclic permutation of the suffices.

Four candidates

The formula now is:

$$[t_1+\sum_{i=2}^{4}p_it_{i1}+\sum_{i=2(i\neq j)}^{4}\sum_{j=2}^{4}p_ip_jt_{ij1}+p_2p_3p_4\sum{}'t_{(234)1}](1-p_1)=q$$

where Σ' indicates summation over all permutations of (234); there are three similar formulae.

The extension to any number of candidates is straightforward. It should be noted:

(i) The formulae for n candidates may be reduced to those for $n-1$ candidates by eliminating the nth equation and putting $p_n=0$ in the others;

(ii) Full recursion is not necessary on the elimination of a candidate if none of the totals or subtotals in the formulae in use at that stage are changed as a result.

7. Calculating the proportions

It can be seen that one of the difficulties involved in the feedback process arises from the need to calculate the proportions for transfer. However, a simple iterative procedure enables this to be done to any required accuracy. We shall take as the simplest example the position with two elected candidates, where the equations to be solved are, as above:

$$(t_1+p_2t_{21})(1-p_1)=q \qquad (1)$$

$$(t_2+p_1t_{12})(1-p_2)=q \qquad (2)$$

In these equations only the p_i are unknown. Suppose we guess a value of p_2 which is too low; then $(1-p_1)$ will be too large in equation (1), that is p_1 will also be too small. If we substitute this in equation (2) it will similarly give a value of p_2 which is too low.

The total vote for the two candidates is t_1+t_2; for them both to be elected $t_1+t_2 \geq 2q$. Suppose the strict inequality holds; in a non-trivial case t_{12}, t_{21} are both non-zero. Further, at least one of t_1, t_2 is greater than q; assume it is t_1. If we put p_2=0 in (1) we can solve for p_1, giving a value p_1>0. This p_1 is the proportion to be transferred if candidate 1 were the only elected candidate; thus $t_2+p_1t_{12} \geq q$ or candidate 2 would not be elected. If the equality holds, candidate 2 only just gets the quota and so p_2=0 from equation (2); thus the equations are solved.

If the strict inequality holds, we get a value of $p_2 > 0$ which is too small. Substituting in (1) increases the coefficient of $(1-p_1)$ and hence increases p_1; the new value of p_1 is increased (but is still too low). Substitution in (2) gives similarly an increased, but too low, value of p_2. Thus the iterative process gives monotonically increasing sequences of values p_1, p_2 bounded above, which hence tend to limits which are the solutions of the equations. A cycle of iterations which leads to two successive sets of values the same to the given accuracy is taken as the approximate solution required. Note that the approximate values may be slightly smaller than the exact ones, but this is exactly what we want; otherwise too much of the support for the candidate concerned would be transferred and he would be left with less than the quota. The process can also be easily shown to work in the limiting case, t_1+t_2=2q.

It is clear that the success of this iterative procedure depends on the fact that all the quantities in the totals (the coefficients of $(1-p_i)$ in each equation) are non-negative, and that therefore it will work for any number of equations provided they are solved cyclically in order of election − this condition being necessary to avoid getting negative values of p_i. Since the counting process can only increase the totals of support for elected candidates, it is also clear that the p_i for those candidates can only increase as the count progresses;[10] thus it is safe to take as starting values of the p_i the ones obtained at a previous stage, putting p_i=0 initially for newly-elected candidates only (in which case, as mentioned above, the equations reduce to the ones at the previous stage and hence will yield, at the beginning of the iteration, the same answers).

It can be shown fairly simply that the convergence rate of the iterative process is likely to be unsatisfactory only when both of the following conditions hold; that all the p_i are small, and the cross-totals t_{ij} etc, are as large as possible. This would not cause difficulty even on the rare occasions on which all these conditions were satisfied, since the occurrence of slow convergence can be detected in advance and allowed for, while at a later stage in the count some at least of the p_i are likely to rise sufficiently to accelerate to the true convergence satisfactorily.

8. Conclusions

It is obvious even from the above example that the feedback process is a much more laborious method of arriving at a result than any at present in use; in a full-scale election with thousands of ballots to scrutinise, it would be very lengthy indeed. However, even the present methods are sufficiently lengthy to make it worthwhile using computers to help in the counting,[11] and if this is done, then complex counting methods are no problem.

It may be argued that the actual results of any election would be different so infrequently that the additional complication is unnecessary. This is a matter for conjecture, or preferably, for further investigation. However, the method has been tried out in two cases, once using figures obtained by a quasi-random process, and once in an actual STV election. In both, there were differences in the candidates elected.[12] Particularly since STV supporters lay such emphasis on the criterion of equality of treatment (condition (B)), it would seem worthwhile in automated counting to adopt the feedback method.

To sum up, the feedback method does satisfy the criterion, subject to the limitations imposed by the basic STV system − i.e. the theoretical minimum of wasted votes, and the elimination of candidates. There is one further limitation not so far discussed, imposed by the voters themselves if they take advantage of the possibility allowed by STV of listing only some of the candidates in preference order. The extension of the feedback method to cover this is dealt with in Paper II; it turns out that the extension also, as a bonus, allows voters to express their views much more accurately than under previous STV methods.[13]

References and Notes

1. For a complete description of STV see E Lakeman and J Lambert: Voting in Democracies (Faber and Faber 1955). *(The current edition in 1994 is E Lakeman: How Democracies Vote (4th edition, Faber and Faber 1974).)*

2. This is nevertheless more than can be said for some common voting systems, such as the simple majority system.

3. This cannot, of course, cope with the case of exact equality, where some other method has to be used, if only drawing of lots.

4. To argue, in connection with a *transferable* system, that a vote should where possible not be transferable, seems inconsistent, particularly in view of the strong arguments put forward by STV supporters against the single non-transferable vote system, where an elector may choose only one from a list of candidates even though more than one are to be elected. See Lakeman

and Lambert, *op, cit.*[1]

5. Duncan Black: Theory of Committees and Elections (2nd edition, Cambridge, 1963, pp 80-83).

6. K Arrow: Social Choice and Individual Values (2nd edition, Wiley 1962).

7. The case for the other side may be found in Lakeman and Lambert, *op cit* [1].

8. The similarity of this principle to Arrow's condition of independence of irrelevant alternatives is obvious. However, the interdependence of the alternatives here means that the condition is not in fact satisfied.

9. This innocent-looking assumption is open to major criticism. Full discussion is outside the scope of this paper; it is hoped to include this in the projected more general paper mentioned in section 5.

10. Clearly Arrow's condition 2, the positive association of individual and social values, is now satisfied by the non-independent alternatives.

11. For a feasibility study in general terms, see P Dean and B L Meek: the Automation of Voting Systems; Paper I; Analysis (Data and Control Systems, January 1967, p16); Paper II; Implementation (Data and Control Systems, February 1967, p22), and B L Meek: Electronic Voting by 1975? (Data Systems, July 1967, p12) − the date in the last source referring to the UK. For a description of the actual use of computers in STV elections in the United States, see Walter L Pragnell: Computers and Conventions (The Living Church, 20th August 1967, p12).

12. For obvious reasons the work on the actual election cannot be made public!

13. These papers are the result of a problem posed by Miss Enid Lakeman, Director of the Electoral Reform Society, London; the author wishes to thank her for her encouragement in the progress of the work. Thanks are due also to Professor W B Bonnor, Mr Robert Cassen, Mr Peter Dean, Mr Michael Steed and Professor Gordon Tullock for valuable discussions, correspondence and advice.

A New Approach to the Single Transferable Vote

Paper II: The problem of non-transferable votes

B L Meek, Computer Unit, Queen Elizabeth College, London.

Brian Meek is now at King's College London.

The original version of this paper was dated 21 March 1968 and was published in French in Mathématiques et Sciences Humaines No 29, pp 33-39, 1970. This note, and note 8 in its present form, have been added in this reprint.

Abstract

The feedback counting method used for Single Transferable Vote elections, developed in an earlier paper, is extended to cover situations in which there are non-transferable votes. It is shown that present counting methods, on the other hand, may not satisfy the condition that the number of wasted votes be kept to a minimum in such situations. The extension of the method to permit voters to give equal preferences to candidates is also described.

1. Introduction

In an earlier paper[1] (hereafter referred to as Paper I) the Single Transferable Vote (STV) system of voting was considered from the point of view of certain conditions, the main one being that as far as possible the opinions of all voters are taken equally into account; it was shown that present STV counting methods do not satisfy this condition. A 'feedback' counting mechanism was suggested which would overcome this problem. In Paper I, however, we confined ourselves only to the cases where, whenever a vote is rescrutinised for transfer, a next preference is always given. In this paper we shall show how the feedback method can be extended to cope with situations where no such preference is available. We shall here adopt the reverse procedure to Paper I; we shall consider the application of the feedback mechanism to these cases first, and only then discuss present counting methods in the light of the conditions.

2. Rules for vote-casting

Even within the same voting system major differences can be made simply by changing the rules governing what constitutes a valid ballot. For example, in a multiple-vacancy election by simple majority where each voter has one independent vote for each vacancy, the result can be totally different if the voter is forced to use all of his votes (in effect to vote against his favourite candidates) instead of using only some.[2] In STV the equivalent requirement would be that all candidates should be listed in preference order. However, in the simple-majority case distortions can arise in that some votes may not be genuine, having only been added in order to make up the correct number; in STV a voter may only wish to express his preferences for a few candidates, being indifferent to the remainder. Normal STV practice is in fact to accept as valid any ballot showing a unique first preference; thereafter the voter may, optionally, give further preferences for as many or as few of the remaining candidates as he wishes. In STV the feedback mechanism could be applied as it stands simply by declaring as invalid any votes which do not give preferences for all candidates or (relaxing this somewhat) declaring invalid during the progress of the count any vote encountered for which a next preference is required but not available, and then restarting the count. However, it would clearly be more satisfactory not to impose additional restrictions on the voter if this can be avoided.

3. Extension of the feedback method

We recall here the two principles of the feedback mechanism stated in Paper I:

Principle 1. If a candidate is eliminated, all ballots are treated *as if that candidate had never stood*.

Principle 2. If a candidate has achieved the quota, he retains a fixed proportion of every vote received, and transfers the remainder to the next *non-eliminated* candidate, the retained total equalling the quota.

Since transfers are only made from eliminated or elected candidates, non-transferability only arises when all the marked candidates are eliminated or elected. The simplest case to consider is that when all the marked candidates are eliminated. By Principle 1, such a ballot has to be treated as if those candidates had never stood; and hence as if the ballot is invalid. This implies that the total *T* of valid ballots is reduced; this in turn implies that, on the elimination of any candidate, if non-transferable ballots occur the feedback should include the recalculation of the quota, using the reduced value of *T*.

The case of a ballot with marked candidates who are elected is less straightforward. Suppose an elected candidate C receives a total *x* of votes with no further preferences marked on them (any marked eliminated candidates can, by Principle 1, be ignored). By Principle 2, C must pass on a fixed proportion *p* of these, as all other, votes and retain the rest as part of his quota. The difficulty arises because it is not clear to whom these votes should be transferred.

If the difficulty were to be avoided by increasing the proportion transferred of votes for which a next preference is marked, to enable all *x* votes to be retained by C, this would clearly reintroduce inequities of the kind Principle 2 was designed to eliminate. Not to transfer the proportion at all would mean leaving C with more than the quota (see also section 4). The two possible ways of strictly obeying Principle 2 are

(a) to divide the otherwise non-transferable proportion equally between the remaining (i.e. unmarked and uneliminated) candidates; or

(b) to subtract this quantity from the total *T* of votes cast, and recalculate the quota with the new value.

Method (a) is based on the view that the voter regards the unmarked candidates as of equal merit, which is why he has not given preferences. The second method is based on the view that the voter's action is a partial abstention; he has not sufficient knowledge of these candidates to judge between them, and prefers to leave the choice to the other voters. It should be noted that the two methods are not equivalent; in the first the totals of the unmarked candidates, in particular the non-eliminated ones, are raised equally, whereas in the second the quota increases the proportions transferred from the elected candidates, and the increase in the votes of non-elected candidates will vary according to these values.

For the moment we shall resolve the (apparent) dilemma by making the (apparently) arbitrary decision to adopt the second method. The *prima facie* case for this is that in general some unmarked candidates will be elected candidates, and hence the adoption of the first method will in any case involve the recalculation of the quota. However, the real justification will appear in section 6, when it will be shown that the dilemma need not, in fact, exist at all.

4. Current STV practice

Current STV procedure in dealing with non-transferable votes involves different rules in different circumstances. The main rules are

(i) If a vote is not transferable from an eliminated candidate, it is set aside; such votes play no further part in the count.

(ii) If the number of votes non-transferable from an elected candidate is not greater than the quota, those votes are included in the quota and only the transferable votes determine the distribution of the surplus. If the number is greater than the quota, then the transferable votes are transferred (at unreduced value), the difference between the non-transferable votes and the quota increasing the non-transferable total.

In Paper I we considered STV from the point of view of three conditions. Condition (C) we shall discuss later; the others were

(A) The number of wasted votes in an election (i.e. which do not contribute to the election of any candidate) is kept to a minimum.

(B) As far as possible the opinions of each voter are taken equally into account.

It is clear at once that, when there are non-transferable votes, condition (B) cannot be satisfied even by the feedback counting method unless recalculation of the quota is included, for otherwise candidates at a later stage of the count, when a number of non-transferable votes have accumulated, need less that the original quota to be elected. Indeed, if as many as q votes become non-transferable, it is impossible for the last elected candidate to achieve a full quota.

We saw in Paper I that condition (A) is satisfied when there are no non-transferable votes. When votes do become non-transferable these have to be added to the 'wasted' total W, and the formula in Paper I becomes

$$W \leq T/(S+1) + T_0$$

where T_0 is the non-transferable total. However, this is derived from a quota calculated on the total T and not on the total available vote $T' = T - T_0$. Thus with recalculation of the quota we have

$$W' < T'/(S+1) + T_0 = W - T_0/(S+1) < W$$

i.e. condition (A) is violated unless the quota is recalculated[3].

It is clear that rule (ii) above is an attempt to satisfy condition (A), but it only does so at the cost of violating condition (B); for example, if a candidate E is elected with $q + x$ votes, q of which are non-transferable, the x remaining votes will be transferred at unreduced value to the next preference even though their earlier preference for E has been satisfied. Further, the present rule that votes cannot be transferred to an elected candidate (see Paper I) means that both by rule (i) and by rule (ii) many whole votes may be declared completely non-transferable, thus swelling T_0 and W above, whereas the feedback method allows each vote to count partly for the elected candidates marked and only a fraction becomes non-transferable.

Thus, on two grounds, current STV counting methods violate condition (A). It could perhaps be argued that the feedback method cannot satisfy condition (A) unless method (a) rather than method (b) of section 3 is used when dealing with unmarked candidates. We shall discuss this point in section 6.

5. Recalculating the quota

It can be seen that in recalculating the quota and having to apply it in retrospect to candidates already elected, the same difficulties occur as in the simple feedback situation, without non-transferable votes, described in Paper I. We consider first the case of an elected candidate. If some of his votes are non-transferable, the appropriate proportion is subtracted from the total vote, and the quota recalculated. The reduction in the quota makes more of the elected candidate's votes surplus, which increases the proportion for transfer; this increases the non-transferable proportion to be subtracted from the total, which further reduces the quota, and so on. The equations to be solved are

$$q = [(T - p_1 t_{10})/(S+1) + 1] \qquad (1)$$

$$t_1(1 - p_1) = q \qquad (2)$$

where, as in Paper I, S is the number of vacancies, T is the total votes (now ignoring any which mark only eliminated candidates), t_1 the total for the elected candidate, p_1 the proportion he transfers, t_{10} the total vote for the candidate not transferable to others, and q is the quota.

These two equations can be solved easily for p_1 and q by equating the expressions for q; however, if there is more than one elected candidate the iterative method of finding the p_i, described in Paper I, will be needed, and it is convenient to discuss the extension of the iterative process to include the recalculation of the quota in terms of the simplest case, above. Equation (1) with $p_1 = 0$ gives the original value of q. Equation (2) then gives a first value of $p_1 > 0$. Substitution of this value in (1) gives a new value of q smaller than before; use of the new q in (2) gives a larger p_1, and so on. Thus we have a monotone increasing sequence of values for p_1, bounded above by 1, and a monotone decreasing sequence of values of q bounded below by 0; these sequences must therefore tend to limits which are the solutions to the equations. The convergence rate is satisfactory; simple analysis shows that the errors are multiplied in each cycle by a factor which is at most $1/(S + 1)$.

The process is extended to the case of n elected candidates by adding to the equations in Paper I the equation

$$q = [T_n/(S+1) + 1]$$

which must be evaluated for q first in each iterative cycle. $T_n = T_n(p_1, p_2, \ldots, p_n)$ is the total available for transfer in each case; for $n = 1, 2, 3$ it is given by

$$T_1 = T - p_1 t_{10}$$

$$T_2 = T - \{p_1 t_{10} + p_2 t_{20} + p_1 p_2 (t_{120} + t_{210})\}$$

$$T_3 = T - \{\Sigma_1 p_i t_{i0} + \Sigma_2\, p_i p_j t_{ij0} + \Sigma_3\, p_1 p_2 p_3\, t_{(123)0}\}$$

In these formulae $t_{ij\cdots k0}$ is the total transferable from candidate i to candidate j, to ..., to candidate k but not further; $\Sigma_1$denotes summing over i; Σ_2 denotes summing over all i, j, $i \neq j$; Σ_3 denotes summing over all permutations of (123).

The reader will easily derive equivalent formulae for higher values of n; putting $p_n = 0$ in the expression for T_n gives the expression for T_{n-1}.

6. Equal preferences

In section 2 we discussed briefly the effect of different validity rules on otherwise identical voting systems. The usual STV counting procedures depend on the existence at each stage of a unique next preference, the only deviation allowed being, as we have seen, that the absence of further preferences does not make the vote as a whole invalid. It is standard practice to accept as valid a vote with a unique first preference, and to accept further preferences provided one and only one is marked at each stage; if no, or more than one, next preference is given at any point, all markings at and past this point are ignored.

For the simplest form of STV counting, involving the physical transfer of ballot papers from pile to pile, the need for a unique next preference is obvious. However, with the feedback method such a restriction is no longer necessary, and indeed it is not necessary even with Senate Rules counting. A vote can be marked A1, B1, C2, ... with A and B as equal first preferences and credited at 0.5 each to A and B. If A is elected or eliminated the 0.5 is transferred at reduced or full value to the next preference − which of course is B and not C. In effect, such a vote is equivalent to two normal STV votes, of value 0.5 each, marked A,B,C... and B,A,C... respectively. Similarly, if A, B, C are all marked equal first, this is equivalent to 6 (= 3!) votes of value 1/6 each, marked A,B,C...; A,C,B...; B,A,C...; B,C,A...; C,A,B...; and C,B,A... . It is easy to see that this can be extended to equal preferences at any stage, and that K equal preferences correspond to $K!$ possible orderings of the candidates concerned, each sharing $1/K!$ of the value at that stage.

Such an extension of the validity rules enables us to resolve the dilemma between the methods (a) and (b) in section 3 of dealing with non-transferable votes. A voter who, at a certain stage, wishes his vote, if transferred, to be shared equally between the remaining candidates, can simply mark those candidates as equal (i.e. last) preferences. Thus the dilemma does not after all exist; both of the methods can be used, and the voter himself can determine which is to be used for his own ballot by the way that he marks it; failure to rank a candidate indicates a genuine (partial) abstention.

This extension of the validity rules also enables condition (C) of Paper I to be satisfied more closely. The condition was:

(C) There is no incentive for a voter to vote in any way other than according to his actual preference.

Here we are interpreting this condition in a particular way not discussed in Paper I: the STV voting rules not merely encourage but force a voter to vote other than according to his preference in the restricted sense that, e.g. if he rates two candidates as equal first he is not allowed to vote accordingly, but must assign a preference order between them which may well be arbitrary. In view of the importance of first preferences in STV, this is undesirable. A voter is similarly forced to make an unreal ordering of candidates to which he is indifferent if, for example, he has listed his real preferences but wishes to give the lowest ranking to a candidate he particularly dislikes. This kind of voting is very common.

Permitting equal preferences thus gives much greater flexibility to the voter to express his ordering of the candidates, and is thus a desirable reform whether the feedback method is used for counting or the Senate Rules retained.[4]

7. Concluding remarks

Two distinct problems arise in the development of a voting system; the information with regard to the choices which is required from each voter, and the way in which this information is to be processed to arrive at "the social choice".

The first problem is mainly outside the scope of these papers, but has been touched on in the last section. It is a basic assumption of STV that the individual preference orderings of each voter is sufficient information[5] to obtain the social ordering, and the voting rule extensions described above follow naturally from this principle, and indeed bring STV more closely into line, in a certain sense, with the work of Arrow.[6]

The possible development of (preferential, transferable) voting systems which use further relevant information is the subject of continuing work.[8]

The second problem is the classical problem of decision theory. Assuming the basic STV structure, these papers have shown that the feedback method of counting is needed to satisfy the declared aims of STV as a decision-making procedure more consistently.

This improvement can be made without causing any more difficulty to the voter, and allows the counting procedure to be described by two simple principles instead of by a collection of rules, some of which are rules of thumb.

The disadvantage of the method is the need for many repetitive calculations, which for reasons of sheer practicality rules it out for manual counting except when the numbers of vacancies, candidates and votes are small. However, as pointed out in Paper I, an STV count is already a sufficiently tedious process for it to be worthwhile to use a computer, and the additions to the feedback method described in this paper would be simple to add to the computer program.

As E G Cluff has pointed out,[7] one advantage of election automation is that one is not restricted in the choice of voting system to what is practically feasible in a manual count. The feedback method can lead to different results from the Senate Rules in non-trivial cases, and is therefore a choice to be considered when the automation of STV elections is being implemented.

References and notes

1 B L Meek, A new approach to the Single Transferable Vote I: Equality of treatment of voters and a feedback mechanism for vote counting, Mathématiques et Sciences Humaines No 25, pp 13-23, 1969.

2 It can in fact lead to the defeat of a candidate who is first choice of a majority of the electorate. It is depressing to note that a public election in England was held under precisely these rules as recently as 1964.

3 This kind of inequity can be found most often in elections with large numbers of candidates and vacancies − e.g. for society committees − and can lead to disillusion with STV as a voting system which has little relation to its merits or demerits.

4 The possibility of a voter sharing his first preference other than equally between a number of candidates would take us too far afield, into the realm of multiple transferable voting systems − the subject of continuing work on more general preferential voting systems. In STV the task of the voter is in comparison a straightforward one, in some ways made easier by allowing equal preferences.

5 And, indeed, necessary information!

6 K Arrow, Social choice and individual values, 2nd edn, Wiley 1962. For what is meant by "in a certain sense" see Paper I.

7 See B L Meek, Electronic voting by 1975?, Data Systems July 1967, p 12.

8 See also note 4. This further work was later published (in English) as: B L Meek, A transferable voting system including intensity of preference, Mathématiques et Sciences Humaines No 50, pp 23-29, 1975.

Computer counting in STV elections

D R Woodall, Department of Mathematics, University of Nottingham.

Douglas Woodall is Reader in Pure Mathematics at Nottingham University.

Reprinted [with minor corrections] from Representation 23 (1983), 4-6.

The Single Transferable Vote is by far and away the fairest form of electoral system. Nevertheless, when the counting in STV elections is carried out by hand, rather arbitrary decisions have to be made in order to simplify the count, and these introduce anomalies. Although small in comparison with anomalies present in other electoral systems, these anomalies may affect the result, and are certainly annoying to the purist.

The biggest anomaly is caused by the decision, always made, not to transfer votes to candidates who have already reached the quota of votes necessary for election. This means that the way in which a given voter's vote will be assigned may depend on the order in which candidates are declared elected or eliminated during the counting, and it can lead to the following form of tactical voting by those who understand the system. If it is possible to identify a candidate W who is sure to be eliminated early (say, the Cambridge University Raving Loony Party candidate), then a voter can increase the effect of his genuine second choice by putting W first. For example, if two voters both want A as first choice and B as second, and A happens to be declared elected on the first count, then the voter who lists his choices as 'A B ...' will have (say) one third of his vote transferred to B, whereas the one who lists his choices as 'W A B ...' will have all of his vote transferred to B, since A will already have been declared elected by the time W is eliminated. Since one aim of an electoral system should be to discourage tactical voting, this seems to me to be a serious drawback.

If, on the other hand, one agrees that surpluses will be transferred to candidates who have already reached the quota, then one has to do something to avoid the never-ending transfer of progressively smaller surpluses between two candidates. Whatever strategy one adopts, it is bound to introduce other anomalies, albeit smaller than the one already described.

If the counting is carried out by computer, however, no such arbitrary decisions are necessary, as the never-ending transfer can be carried out to completion, or at least until the surpluses remaining to be transferred are less than (say) a millionth of a vote. The resulting procedure is described in the next paragraph in a different way. It is comparatively simple in concept, and the undoubtedly long calculations are all safely hidden inside the computer.

The counting is divided into rounds, in each of which one candidate is eliminated. In each round of the elimination, a scaling factor is assigned to each candidate, representing the proportion that will actually be credited to him out of the votes potentially available to him, in such a way that:

1) a candidate who has already been eliminated in a previous round is assigned scaling factor 0, so that no votes will be credited to him in the current round;

2) a candidate whose fate is undecided at the end of the current round is assigned scaling factor 1, so that all the votes potentially available to him are credited to him; and

3) a candidate who by the end of the current round has at least the quota of votes necessary for election (and so is certain to be elected) is assigned a scaling factor less than or equal to 1 so that the number of votes credited to him is brought down exactly to the quota.

The candidate with the smallest number of votes is then eliminated, and the process is repeated until the number of candidates remaining is equal to the number of places to be filled.

For example, suppose that, in a given round of the counting, candidates A and B are certain of election and have scaling factors of two thirds and three quarters respectively, and candidates C, D and E have already been eliminated in previous rounds, whereas the fates of the remaining candidates remain undecided. Then a voter who lists the candidates in the order C, A, D, B, E, F will, in the current round, have none of his vote assigned to C. The whole of his vote will be passed down to A, who will retain two thirds of it. The remaining third of his vote will be passed over D and down to B, who will retain three quarters of it (that is, one quarter of a vote). The twelfth of a vote that is still unassigned will be passed over E and down to F, who will retain all of it.

The calculation of the scaling factors, which would be prohibitively long to do by hand, could be carried out quite easily by computer. However, once the computer had done the work, it would be possible to check by hand that the computer was correct; certainly this would take no longer than carrying out the whole count by hand as at present.

(This situation is not unusual in mathematics. Suppose, for example, that you were asked to find a number x between 1 and 2, accurate to seven places of decimals, such that (say)

$$x^5 + x^4 - 4x^3 - 3x^2 + 3x + 1 = 0.$$

You would find it very tedious to do so by hand, even with the aid of a pocket calculator. Suppose, however, that a computer were to do the work and tell you that the answer is 1.6825071; then it would take you only a few minutes to check that the computer was correct.)

The size of computer required would depend on the size of the electorate, on the number of places to be filled and, to a lesser extent, on the number of candidates. In the case of an election with both a very large electorate and a large number of places, it might even be impossible to carry out the calculations in a reasonable time with the present generation of computers.

However, for parliamentary elections, there would be no problem: the calculations could be done quite easily even on a mini-computer.

Since proposing the above method, I have learnt that it is not new; a differently worded but exactly equivalent method was proposed by Brian Meek in 1969.[1,2] I hope it will be possible to agree that, whenever computer counting is used in STV elections, this method should be used.

References

1 B L Meek, Une nouvelle approche du scrutin transférable, Mathématics et Sciences Humaines 25 (1969), 13-23.

2 B L Meek, Une nouvelle approche du scrutin transférable (fin), Mathématics et Sciences Humaines 29 (1970), 33-39.

Counting in STV elections

C H E Warren

Hugh Warren is a retired aeronautical engineer, past member of the ERS Council and a present member of its Technical Committee.

Introduction

Whatever criticisms may be levelled against First-Past-The-Post as a system of voting, at least the system has the merit that, although the count may be conducted in many ways, all ways give the same result. The Single Transferable Vote is demonstrably a better system of *voting*, but the system has the disadvantage that the result depends upon how the *counting* is conducted.

Counts have been done in many ways, and in some peculiar ways by some well-meaning, but unversed, enthusiasts for STV. One of the commonest methods of conducting the count, and indeed the method that the Electoral Reform Society uses, is that given by Newland and Britton.[1] Their paper tells one *how* to conduct a count by their method, but not *why* they make many of the arbitrary decisions that they do. Woodall[2] has suggested that they are made for expediency − to simplify the count − and he goes on to propose another method, which he advocates whenever computer counting is used. As Woodall points out, his method would be prohibitively long with human counting. As Woodall also states, a differently worded but an exactly equivalent method to his had been proposed by Meek in 1969.[3,4]

The object of this paper is, first, to consider some of the principles that are felt to be important in deciding upon a method for conducting the count, and then to go on and propose a method that meets these principles.

Principles

The first principle of the STV system is that election is by quota. A candidate is deemed elected when the vote assigned to him attains a given quota. The quota is chosen as the minimum vote which will not allow more than the required number of candidates to be elected. This is the Droop quota, and is the total valid vote divided by one more than the number of candidates to be elected.

The second principle concerns the transference of a voter's vote to the preferences later than his first preference. The voter needs to be assured that his later preferences will in no way upset the voter's earlier preferences. Equally a voter's later preferences should not be considered unless, in regard to each earlier preference candidate, either the voter has borne an equal share with other voters who have voted for that candidate in giving him the necessary quota, or that earlier preference candidate has been eliminated. The way in which Newland and Britton conduct a count does not meet this principle.

The third principle concerns the elimination of candidates. Unfortunately no-one appears to have proposed a principle in this regard. So what is usually done is that, when no candidate has a surplus above the quota, in order to allow the count to continue, the candidate whose vote is least is eliminated.

Method

If, after counting the first preference votes, the votes for one or more candidates exceed the quota, then the essential feature of the method proposed here is that these candidates are allowed to retain only part of the vote that had been expressed for them such as will give each candidate just the necessary quota. The part of the vote that the candidate retains is called the 'amount retained'. The voters who have voted for one of these candidates, for whom the amount retained is x_1, say, then have an amount remaining of $(1–x_1)$, which is then transferred to the voters' expressed second preferences. If an expressed second preference has an amount retained of x_2, say, and if x_1+x_2 is less than unity, then the voter still has an amount remaining of $(1–x_1–x_2)$, which is then transferred to the expressed third preference, and so on. Proceeding in this way, the end of the first stage of the count is reached when some candidates have just the quota, whereas the remainder have varying amounts of vote less than the quota.

The candidate whose vote at the end of the first stage is least is eliminated. This means that, wherever his name appears on a ballot paper, it is 'passed over', and, in effect, all the later preferences are 'moved up one'. Elimination of a candidate will usually cause the votes for some other candidates to exceed the quota. The amount to be retained by each candidate is then reduced to such lower value as will give each candidate just the necessary quota. Voters who have voted for these candidates with reduced amount retained will then find that they have more vote remaining for transference to later preferences. Proceeding in this way, at the end of each stage of the count, some candidates will have just the quota, whereas the remainder will have varying amounts of vote less than the quota.

Eventually the number of non-eliminated candidates will be reduced to one more than the number to be elected. When the amounts to be retained are now recalculated so as to reduce each candidate's vote to the necessary quota, all candidates will have just the quota, so the one candidate who has an amount retained of just 1 is the one eliminated. The remaining candidates are deemed elected.

If at any stage a ballot paper does not contain sufficient preferences for transference to be made, then the balance of vote is ascribed 'non-transferable', and the quota is recalculated excluding the non-transferable vote.

The main question that the proposed method of conducting the count poses is: how does one decide upon the amount to be retained by each candidate at each stage? From what has been said, the amounts retained have to be such that, when the count is made, each candidate to whom an amount to be retained of less than 1 has been assigned achieves just a quota. The problem of finding the amounts retained, and the associated quota, is a mathematical one which is relatively straightforward, even if protracted, but which a computer can help to solve. Here we are concerned only with the principle, not with precisely how the task be done. However, it is not necessary for everyone to know *how* to assign the amounts retained. As Woodall[2] has exemplarily pointed out, it is only necessary for anyone to be able to *check* that the assigned amounts retained do in fact achieve the desired result.

References

1 R A Newland and F S Britton, How to conduct an election by the Single Transferable Vote, second edition, Electoral Reform Society of Great Britain and Ireland (1976).

2 D R Woodall, Computer counting in STV elections, Representation, Vol.23, No.90 (1982), 4-6.

3 B L Meek, Une nouvelle approche du scrutin transférable, Mathématics et Sciences Humaines 25 (1969), 13-23.

4 B L Meek, Une nouvelle approche du scrutin transférable (fin), Mathématics et Sciences Humaines 29 (1970), 33-39.

Meek or Warren counting

I D Hill.

David Hill is a retired statistician, member of ERS Council and Chairman of its Technical Committee.

The Meek system and the Warren system for counting an STV election are very similar, but whereas the fractions of a vote retained by successive elected candidates are multiplicative under Meek, they are additive under Warren. For example, if candidate A is keeping 1/2 of everything received and candidate B is keeping 1/3, a vote reading AB... will, under Meek, give 1/2 of a vote to A and 1/6 of a vote to B (i.e. 1/3 of the remaining 1/2), leaving 1/3 of a vote to be passed on further. With those same fractions under Warren, a similar vote will give 1/2 of a vote to A and 1/3 of a vote to B, leaving 1/6 of a vote to be passed on further. (It should be noted, though, that in any actual case the fractions will not usually be the same under the two systems). The Warren system will often lead to the situation where not enough vote remains for the fraction required; in such a case all that remains is taken and nothing remains to go any further.

There is no difference in the ease of writing a computer program to satisfy the one system or the other; the choice can be made solely on which is regarded as better in principle. It should also be reported that in real examples of STV elections, as distinct from artificially constructed examples, no case has yet been found where the two elect a different set of candidates, so the difference for real life seems to be slight.

There has been much argument over which system is to be preferred. In the end, we have settled on a particular example which demonstrates that each system can be said to suffer from a difficulty that the other one solves. It must therefore be a matter of judgement which difficulty is regarded as the more serious, rather than a firm decision of one always being better than the other.

The Meek rationale is that all transfers from a surplus should be in proportion to the 'votes-worth' put into that surplus. Thus 5 identical votes, each of current value 1/5, should have the identical effect to that of 1 complete vote for the same preferences. The Warren rationale is that no voter should be allowed to influence the election of an additional candidate until having contributed as much as any other voter to the election of each candidate who has already been elected and is named earlier in the voter's preferences. Thus the 5, each of value 1/5, are to be treated as 5, not as the equivalent of 1.

The example that shows the differences has 5 candidates for 3 seats and 32 votes, leading to a quota of 8.0. The votes are:

12 ABC, 12 BE, 7 C, 1 D.

Meek supporters can point out the Warren anomaly that A and B each had a substantial surplus on the first count, yet the 12 ABC votes are given by the Warren system entirely to A and B and, in consequence, C fails to get the 1 extra vote needed for election and E takes the third seat. Under Meek, C easily beats E.

Warren supporters can point out the Meek anomaly that if the 12 ABC voters had voted BAC instead, the Meek system would have behaved exactly like the Warren system, and E would have beaten C. It seems illogical that the choice of C or E should depend upon the ordering by those 12 voters as ABC or BAC when A and B were both elected anyway.

Deciding between the two systems must therefore remain a matter of personal preference.

It may be of interest to see exactly how each of the two systems would treat this example. Each would note that A and B are both elected on the first count, each having 12 first preferences for a quota of 8.

The Meek system would calculate that A needs to keep 2/3 of everything received whereas B needs to keep 1/2, these fractions being derived so that each of A and B keeps exactly a quota. The 12 ABC votes would be allocated as 2/3 of 12 = 8 to A, 1/2 of the remaining 4 = 2 to B, the remaining 2 to C. The 12 BE votes would be allocated as 1/2 of 12 = 6 to B, the remaining 6 to E. At the next count the current votes would therefore be A 8, B 8, C 9, D 1, E 6. The third seat is thus assigned to C and no more needs to be done.

The Warren system would calculate that A's amount retained needs to be 2/3 and B's 1/3, again derived such that (under the different counting method) each of A and B keeps exactly a quota. The 12 ABC votes would be allocated as 2/3 of 12 = 8 to A, 1/3 of 12 = 4 to B. The 12 BE votes would be allocated as 1/3 of 12 = 4 to B, the remaining 8 to E. At the next count the current votes would therefore be A 8, B 8, C 7, D 1, E 8. The third seat is thus assigned to E and no more needs to be done.

Issue 2, September 1994

Editorial

Voting matters is concerned with the implementation of the Single Transferable Vote. However, STV is merely one method of analysing ballot papers in which preferential voting is used. In consequence, other methods of analysis could provide some insight into STV. In this issue, one particular problem of STV is highlighted, namely that of the elimination of a popular candidate with few first-preference votes. David Hill and Simon Gazeley provide algorithms to 'overcome' this problem and discuss the consequences. Due to the impossibility of satisfying apparently simple requirements, Douglas Woodall has shown that overcoming the above problem is bound to introduce other anomalies.

Brian Wichmann.

STV with successive selection — An alternative to excluding the lowest

S Gazeley

Simon Gazeley is a civil servant working in Bath. He is a member of Council and of the Technical Committee of the ERS.

The problem with current STV systems

A feature of STV which is not shared by other preferential voting systems is election on attaining a certain number of votes (the 'quota'). If the number of candidates who have a quota of first preference votes is insufficient to fill all the seats being contested, those which are left are filled by candidates whose quotas contain votes which have been transferred from other candidates. These transfers take two forms: of surpluses above the quota for election from candidates who are already elected, and of all the votes previously standing to the credit of candidates who have been excluded in accordance with the rules.

When it is necessary to withdraw a candidate from contention, all versions of STV currently in use exclude the one who has fewest votes at that time. It is contended that the consequences of this rule in conventional STV formulations can be haphazard and therefore unjust in their effect. Consider the following count:

```
AD  35
BD  33
CD  32
```

There are here 3 separate and substantial majorities: against A, against B and against C. The only thing that all the voters agree on is that D is preferable to two out of the other three candidates; yet STV excludes D first, however many seats are being contested. Unfairness and anomalies such as this arise because candidates are excluded before the full extent of the support available to them has been investigated. Even though every ballot-paper may have the same candidate marked as the next available preference, that candidate will not survive if they do not have enough votes now.

An even more serious consequence of the 'exclude the lowest' rule is that it is possible for voters to assist their favoured candidates by withholding support rather than giving it. Consider the following election for one seat:

```
AC  13
BC   8
CA   9
```

Having been excluded, B's votes go to C, who now has an absolute majority and gets the seat. But suppose that two of A's supporters had voted BC instead:

```
AC 11
BC 10
CA  9
```

Now C is excluded first and A gets the seat.

Is it possible, then, to remove this anomaly without introducing another? The answer, unfortunately, is 'no'. Woodall[1] proposed that every count under any reasonable electoral system should have the following four properties:

1. Increased support, for a candidate who would otherwise have been elected, should not prevent their election;

2. a. Later preferences should not count against earlier preferences;

   b. Later preferences should not count towards earlier preferences;

3. If no second preferences are expressed, and there is a candidate who has more first-preference votes than any other candidate, that candidate should be elected;

4. If the number of ballots marked X first, Y second plus the number marked Y first, X second is more than half the total number of ballots, then at least one of X and Y should be elected.

He then proved that no such system can be devised.

We have already noted that current STV systems can (but usually do not) fail on Woodall's first property; this is the failure that in Dummett's[2] eyes precludes consideration of STV as a possible option for public elections in the UK. As no system can have all four properties, a price for having one has always to be paid in terms of lacking at least one other. Under the system proposed below, some counts (but by no means all) may fail to have Woodall's first or second property, but all will have the other two. Whether the price is worth paying is a question to which no definitive answer can be given: it is ultimately a matter of personal preference.

STV by successive selection (SS)

The object of exclusion in current STV formulas is to release votes from one candidate to be transferred to others so that one or more of them will get a quota. STV(SS) retains the transfer of votes from candidates who are not yet elected, but differs from present STV systems in that no candidate is permanently withdrawn from contention. When it becomes necessary to release a candidate's votes, that candidate is ‘suspended’ (withdrawn temporarily) after being identified as the one whose election to the next vacant seat would be least appropriate.

Manual STV systems need to keep within reasonable bounds both the time taken to count an election and the scope for human error and this need can give rise to anomalies. Meek[3] and Warren[4] have devised schemes without these anomalies for distributing votes which would be impracticable using manual methods. STV(SS) is designed (but not yet programmed) to be run on a computer using either of these schemes, but only one should be used in any one election.

In addition to Woodall's four properties, every count under a reasonable system would have the property that of a set of *d* or more candidates to which *d* Droop quotas of voters are solidly committed, more than (*d*-1) should be elected; if the set contains fewer than *d* candidates, all of them should be elected. According to Dummett, a group of voters are ‘solidly committed’ to a set of candidates if every voter in the group prefers all candidates within the set to any candidate outside it. STV(SS) and other STV formulas achieve proper representation of sets of candidates by withdrawing from contention candidates who have less than a quota of votes and by transferring surplus votes from those candidates who have more than a quota.

The principle underlying STV(SS)

STV(SS) is predicated on the proposition that when no surpluses remain to be transferred, there is only one candidate (barring ties) who is the most appropriate occupant of the next seat. Appropriateness depends among other things on who has been elected already: if Candidate X is the ‘most appropriate’ and Candidate Y is the ‘next most appropriate’ at any given point, it does not follow when X is elected that Y is now the ‘most appropriate’. The next candidate to be elected is the one who can command a quota and for whose election the other non-elected candidates need to sacrifice the smallest proportion of their votes.

Under STV(SS), each non-elected candidate in turn is tested to see what proportion of the votes of the other non-elected candidates have to be passed on in addition to the surpluses of the elected candidates to give them the quota. Of those who can command a quota, the candidate who requires the smallest proportion of the others' votes is the ‘most appropriate’ to be elected next. The process is best illustrated by an example. Consider the following votes for one seat:

```
A  49
BC 26
CB 25
```

No candidate has a quota, but instead of excluding the lowest we test each candidate in turn to see which is the ‘best buy’. Let us test A first. The quota is 50 and B and C have 51 votes between them; we therefore change their Keep Values (KVs: see the Annex for further details) from 1.0 to 50/51 (0.9804). At the second distribution the votes look like this:

```
A   49.0000
B   25.9708
C   25.0096
```

The new total of votes is 99.9804, making the quota 49.9902. A still has not got the quota, so the count proceeds. The final distribution looks like this:

```
A   49.0000
B   24.8216
C   24.1784
```

At this point, we record the fact that the common KV of B and C is 0.8020. If we now test B, we find that the final common KV of A and C is 0.5152; when we test C the common KV of A and B is 0.5050.

At first sight, A seems the obvious choice to get the seat: however, if A were to be successful, Woodall's fourth property would be lacking. No candidate should be elected who cannot command a Droop quota of the votes which are active at the time of their election. If we remove C from contention (C is 'least appropriate' as the other candidates had to give up the greatest proportion of their own votes to secure C's quota) and redistribute C's votes, B now secures a Droop quota and is elected.

But why make the selection on the basis of the other candidates' final KVs? The reason is that these represent the degree of support that exists for the proposition that a given candidate should be added to the set of elected candidates. Suppose that some of the votes in an election were cast as follows:

```
AC   54
BC   45
```

(there may be other candidates and other votes, but these need not concern us) and that it is necessary for 33 of these votes to be passed from A and B to C. This is achieved by setting the common KV of A and B at 0.6667 − A and B have to pass on 0.3333 of the current value of each incoming vote to secure C's quota. But suppose the votes had been

```
ABC 54
BAC 45
```

the other votes and candidates being the same. This time, to give 33 votes to C, the common KV of A and B has to be 0.4226 i.e. 0.5774 of the current value of each incoming vote has to be passed on, over 1.7 times as much. The lower a candidate is in the order of preference of the average vote being considered at any point, the lower the common KV of the other non-elected candidates has to be in order to give that candidate a quota.

How STV(SS) works

STV(SS) has two parts: detailed instructions to the computer are given in the Annex. What follows is a general description and explanation of their functions.

The first part

In the first part, the non-elected candidates are ranked in 'order of electability', which forms the basis on which candidates are elected or suspended. All the non-elected candidates are sub-classified at the start as 'contending'. There are two further sub-classifications, namely 'under test' and 'tested'; only one candidate at a time is under test. The object is to ascertain for the candidate under test what proportion of the votes of the contending and tested candidates it is necessary to pass on to give them the current quota. Each non-elected candidate in turn is classified as under test. If a candidate under test is classified as elected, the first part is repeated.

When the candidate under test and the elected candidate have Q or more votes each, the candidate under test has recorded against their name the common KV of the contending and tested candidates: this is that candidate's 'electability score'. When all the non-elected candidates have been tested, they are ranked in descending order of electability score: this ranking is for use in the second part. An electability score of 1.0 indicates that the candidate needs to take no votes from other unelected candidates to get the quota, so there is no reason not to classify that candidate as elected at once.

The second part

In the second part, the next candidate to be elected is identified on the basis of their ranking from the first part and their ability to command a Droop quota of votes. The highest candidate in the ranking is elected as soon as it is shown that they can command a Droop quota of currently active votes. If the highest candidate cannot, the second highest non-suspended candidate gets the seat instead. In this part, non-elected candidates are sub-classified as 'contending', 'protected' (contending candidates become protected when they get a quota) and 'suspended'; they are all classified as contending at the start. Suspended candidates have a KV of 0.0. At the end of the procedure, all the candidates' KVs are reset at 1.0.

Contending candidates are suspended in reverse order of ranking: protected candidates cannot be suspended before the next candidate is classified elected. The fact that a candidate has a Droop quota of currently active votes now does not necessarily indicate that they will achieve one at a subsequent stage and vice-versa. The rankings obtained in each pass through the first part are crucially dependent on which of the previously contending candidates was elected in the preceding second part.

An example

Let us see how STV(SS) works on the examples on page 1:

Count 1		Count 2	
AC	13	AC	11
BC	8	BC	10
CA	9	CA	9

In Count 1, the ranking is A (the common KV of the other two candidates would be 0.7962), C (0.7143) and B (0.2023), so B is suspended first and C gets the seat. The Count 2 ranking is C (0.7143), A (0.6311) and B (0.2929); B is once more the first to be suspended so C again gets the seat.

Conclusion

As specified above, the system appears to be long-winded: there are possible short-cuts, but these would obscure essentials and have been excluded.

STV(SS) is a logical system which is submitted as a contribution to the continuing debate on what the characteristics of the best possible system might be. Refinements are necessary (for instance, a way of breaking ties has to be devised), but there is here the basis for a debate.

References

1 D R Woodall, *An Impossibility Theorem for Electoral Systems*, Discrete Mathematics 66 (1987) pp 209-211

2 Michael Dummett, *Towards a More Representative Voting System: The Plant Report*, New Left Review (1992) pp 98-113

3 i. B L Meek, *Une nouvelle approche du scrutin transférable*, Mathématiques et Sciences Humaines, No. 25, pp 13-23 (1969)

ii. B L Meek, *Une nouvelle approche du scrutin transférable (fin)*, Mathématiques et Sciences Humaines, No. 29, pp 33-39 (1970).

4 C H E Warren, *Counting in STV Elections*, Voting matters, No. 1, pp 12-13 (March 1994)

Annex

STV(SS) — Detailed Instructions

The first part

1. If there is any candidate for whom no voter has expressed any preference at all, treat every such candidate as having withdrawn. If fewer than (N+1) candidates remain, end the count; otherwise, set the ranking of every remaining candidate to equal first.

2. Classify every non-elected candidate as contending and repeat the following procedure until there are no contending candidates left:

a. Set every candidate's KV at 1.0 and select a contending candidate to be the candidate under test.

b. Examine each ballot-paper in turn and distribute the value of the vote in accordance with the voter's preferences and the KVs of the candidates as follows:

Either

i. *The Meek Formulation*. Offer the vote to each candidate for whom the voter has voted in order of preference expressed. Multiply the fraction of the vote which has not yet been allocated by the KV of the candidate to whom it is being offered, and allocate that proportion of the vote to that candidate. Any part of the vote left over after all the candidates for whom the voter has voted have received their share is non-transferable.

or

ii. *The Warren Formulation*. Offer the vote to each candidate for whom the voter has voted in order of preference expressed. Award to each candidate in turn a fraction of the vote equal to that candidate's KV; if the fraction of the vote remaining is less than the KV of the current candidate, award all that is left to that candidate. Any part of the vote left over after all the candidates for whom the voter has voted have received their share is non-transferable.

c. Calculate the quota according to the formula $Q=V/(N+1)$, where V is the total number of votes credited to all the candidates and N is the number of seats being contested.

d. If the elected candidates and the candidate under test have at least Q votes each, go to Step e. Otherwise, calculate new KVs for all the candidates as follows:

i. For all the elected candidates and the candidate under test, multiply the current KV by Q and divide the result by that candidate's current total of votes.

ii. Multiply the common KV of the contending candidates and the tested candidates by $(V–(E+1)Q)/T$, where E is the number of candidates elected so far and T is the total of the votes credited to the contending and tested candidates.

If any new KV exceeds 1.0, reset it at 1.0. Go to Step b.

e. Record the common KV of the contending and tested candidates against the name of the current candidate under test; let this be that candidate's ‘electability score’. Classify that candidate as tested.

3. If no tested candidate has an electability score of 1.0, rank the tested candidates in their existing order within descending order of electability score and go to Step 5. Otherwise, classify as elected every tested candidate whose

electability score is 1.0.

4. If there are *N* elected candidates, end the count. Otherwise, go to Step 2.

The second part

5. Classify every non-elected candidate as contending and set every candidate's KV to 1.0. Repeat the following procedure until either the highest-ranked contending or protected candidate and the elected candidates have *Q* or more votes each, or there are only *N* non-suspended candidates.

a. Examine each ballot-paper in turn and distribute the value of the vote in accordance with the voter's preferences and the KVs of the candidates as follows:

Either

i. *The Meek Formulation*. Offer the vote to each candidate for whom the voter has voted in order of preference expressed. Multiply the fraction of the vote which has not yet been allocated by the KV of the candidate to whom it is being offered, and allocate that proportion of the vote to that candidate. Any part of the vote left over after all the candidates for whom the voter has voted have received their share is non-transferable.

or

ii. *The Warren Formulation*. Offer the vote to each candidate for whom the voter has voted in order of preference expressed. Award to each candidate in turn a fraction of the vote equal to that candidate's KV; if the fraction of the vote remaining is less than the KV of the current candidate, award all that is left to that candidate. Any part of the vote left over after all the candidates for whom the voter has voted have received their share is non-transferable.

b. Calculate the quota according to the formula $Q=V/(N+1)$, where *V* is the total number of votes credited to all the candidates and *N* is the number of seats being contested. Classify any contending candidate with *Q* or more votes as ‘protected’.

c. If any candidate has more than *Q* votes, calculate a new KV for each such candidate by multiplying their present KV by *Q* and dividing the result by their present total of votes. Otherwise, suspend the contending candidate who is ranked lowest.

6. Classify as elected the highest-ranked contending or protected candidate.

7. If *N* candidates are elected, end the count: otherwise, go to Step 2.

Sequential STV

I D Hill

David Hill is a retired statistician. He is a member of Council and Chairman of the Technical Committee of the ERS.

The Meek system for counting an STV election overcomes most of the troubles encountered in using older systems designed for counting by hand, but the problem of premature exclusion remains. Premature exclusion of a candidate occurs when someone is the lowest because hidden behind another who, in the end, is also not going to succeed. If A, who would otherwise have been elected, fails because B stood and was elected instead, it is bad luck for A but there is nothing disturbing about it in principle. If, however, A fails because B stood, but then B does not get in either, that is disturbing.

Exclusion of the lowest candidate, when an exclusion is necessary, is the trouble. After all, if the so-called first past the post is not necessarily the right person to elect, then neither is the last past the post necessarily the right one to exclude. Is there some other way of handling things that would do better? What is needed is a mechanism to discover initially which candidates have some hope of election and which have virtually none, and to get rid of the ‘no-hopers’ at the start of the count. Others cannot then suffer from their presence.

Let the election be to fill *k* seats from *n* candidates, and let $m = n - k$. Sequential STV then consists of a number of main-phases and sub-phases, each being an STV election for *k* seats but with varying selections of candidates. The STV elections are preferably conducted using Meek-style counting but other rules could be used.

Main-phase 1. All *n* candidates, but instead of dividing into elected and excluded, divide them into probables and others respectively. Set all *n* candidates to unmarked.

Sub-phase 1.1. The *k* probables plus any other one candidate. Set the winners to marked.

Sub-phase 1.2. The same *k* probables plus any other one candidate not yet tested. Set any unmarked winners to marked.

etc.

Sub-phase 1.*m*. The same *k* probables plus the last candidate not yet tested. Set any unmarked winners to marked.

If at any sub-phase there is a tie that has to be settled using random selection, then all *k* + 1 of the candidates involved are set to marked.

Main-phase 2. All marked candidates, dividing into probables and others. If the resulting set of probables is the same as a previous set, those candidates are elected and the process finishes. Otherwise reset all *n* candidates to unmarked and continue.

Sub-phases 2.1 - 2.m. As 1.1 - 1.m but using the new probables.

Main-phase 3. As main-phase 2.

etc. etc.

It may be noted that anyone getting a quota of first preferences on the original count is, in fact, certain to be elected in the end, but to be classified for the time being as probable does no harm.

The process must terminate because there is only a finite number of sets of *k* that can be formed from *n*. Usually it will terminate with two successive main-phases showing the same set of *k* probables. In that case the result is firmly established. If, however, the two showing the same set are not successive it will mean that the system is cycling in Condorcet-paradox style. In that case it may be that a better rule could be devised than taking the first set to occur twice but it has to be recognised that a totally satisfactory answer is impossible.

Each candidate is given a fair chance by being tested against each new set of probables and since each sub-phase consists of only *k* + 1 candidates for *k* seats, exclusion is never necessary during the sub-phases so the 'exclude the lowest' rule is not operative there.

Example

With 5 candidates for 2 seats, suppose the votes

```
104 AEBCD
103 BECDA
102 CEDBA
101 DEBCA
  3 EABCD
  3 EBCDA
  3 ECDBA
  3 EDBCA
```

It is evident that E is a strong candidate, in that if any one of A, B, C or D were to withdraw, E would be the first elected. Yet under simple STV the first action is to exclude E, and B and C are elected. Under sequential STV we find

Phase	Candidates	Winners	Probables	Marked
1	ABCDE	BC	BC	
1.1	BCA	BC		BC
1.2	BCD	BC		
1.3	BCE	BE		E
2	BCE	BE	BE	
2.1	BEA	BE		BE
2.2	BEC	BE		
2.3	BED	BE		
3	BE	BE	BE	

B and E are consequently elected. It will be noted that some elections may be repeats of ones already done (main-phase 2 and sub-phase 2.2 in the above example are both repeats of sub-phase 1.3). The result may of course merely be copied down without actually repeating any calculations.

Should it be used?

If any scheme is to be adopted to get rid of (or at least to ease) the problem of premature exclusion, I believe that this is about as good as can be devised. Yet, after much consideration, I do not recommend it for general use, because it breaks the rule, which simple STV always obeys, that a voter's later preferences ought not to interfere with that voter's earlier preferences.

The following example to demonstrate this trouble is derived from those that Douglas Woodall devised to prove his 'impossibility' theorem. Let there be 3 candidates for 1 seat and votes

```
25 A
17 BC
16 C
```

Phase	Candidates	Winner	Probable	Marked
1	ABC	A	A	
1.1	AB	A		A
1.2	AC	C		C
2	AC	C	C	
2.1	CA	C		C
2.2	CB	B		B
3	BC	B	B	
3.1	BA	A		A
3.2	BC	B		B
4	AB	A	A	

So A is elected. But if the A voters had put in C as a second preference, we get

```
25 AC
17 BC
16 C
```

Phase	Candidates	Winner	Probable	Marked
1	ABC	A	A	
1.1	AB	A		A
1.2	AC	C		C
2	AC	C	C	
2.1	CA	C		C
2.2	CB	C		
3	C	C	C	

and C is elected. So the A voters have failed to elect A because they gave C as a second preference.

Even if this is a rare event, it still means that we cannot assure voters that their later preferences cannot upset their earlier preferences. I believe that this is too high a price to pay. There is not much point in reducing the frequency of one type of fault if, in doing so, you introduce another fault as bad.

Only one seat

The system is really intended, as is STV in general, for situations where there are several seats to be filled, but it can also be used in place of Alternative Vote for a single seat. Trying it out on many examples suggests that, for realistic voting patterns, it is almost certain to elect the Condorcet winner if there is one, but artificial examples can be devised to demonstrate that there is no guarantee that it will do so.

For example, let there be 4 candidates for 1 seat and votes

```
98 ADCB
98 CDBA
99 BDAC
 3 ACBD
 2 CBAD
```

Phase	Candidates	Winner	Probable	Marked
1	ABCD	A	A	
1.1	AB	B		B
1.2	AC	A		A
1.3	AD	D		D
2	ABD	B	B	
2.1	BA	B		B
2.2	BC	C		C
2.3	BD	D		D
3	BCD	C	C	
3.1	CA	A		A
3.2	CB	C		C
3.3	CD	D		D
4	ACD	A	A	

So A is elected, even though D would be the Condorcet winner (for the results of AD, BD and CD are all D). It should be emphasised, though, that this is not likely in practice but only with carefully devised artificial examples.

Acknowledgement

I acknowledge that, since I first produced this scheme, Dr David Chapman has produced an almost identical scheme entirely independently.

Two STV Elections

B A Wichmann

Brian Wichmann is the Editor of *Voting matters*, and is a Visiting Professor in Computing at The Open University.

I believe two STV elections may be of interest to the readers of *Voting matters*, due to the implications of the results on the properties that an ideal STV algorithm should (perhaps) have.

The first election is the Eurovision Song contest for 1992 which is an interesting election to analyse since the votes are publicly available, in spite of the voters not knowing of the other votes. Each country votes for the songs of other countries by awarding 12, 10, 8, 7, 6, 5, 4, 3, 2 and 1 points, which can be transcribed into STV preferences.

The points system gave for those over 100: Ireland (155), UK (139), Malta (123), and Italy (111). Since the points total is given after each country has voted, the commentator (Terry Wogan) reported that Ireland was unbeatable by the UK before the last few countries voted. An analysis of the votes by other means is quite different.

The ERS hand counting rules declare the UK as the winner, as does the Meek STV algorithm. However, more countries preferred Ireland to the UK than the contrary (by 12 to 11, rather close). Indeed, by the Condorcet rules, Ireland would be the winner, since Ireland is preferred to any other country by a majority. The reason that the ERS rules elect the UK is that Ireland is eliminated earlier, leaving the last contest between Malta and the UK, which the UK wins. The Meek algorithm is similar, but with Italy being the last to be eliminated.

One STV algorithm due to Tideman considers all possible pairs of results. In the case of a single seat, Tideman will elect a Condorcet winner (assuming there is one) and hence chooses Ireland in this case. One is therefore left to wonder if an ‘ideal’ STV algorithm should always elect a Condorcet winner, assuming there is one.

The second election is one for which I acted as returning officer for a rather unusual ‘election’ at my place of work.

The research institute at which I work has had a library for a group of about 60 scientists for at least 30 years. As the research has changed over the years, new journals have been ordered. However, except in obvious cases, it has not been clear which journals should be cancelled — especially since a complete ‘run’ of a journal will be lost. I therefore proposed that an STV election be run to determine which journals should be cancelled and which new ones to order.

The management agreed to this proposal and hence I ran the election as follows: A list was obtained of the (about) 200 journals, which were assigned a code. The scientists were asked to place up to 40 journals in preferential order, being given about a month to place their ballot.

Quite a bit of effort is necessary to fill in the ballot paper. Nobody attempted more than the 40 preferences, the average being about 20. About half of those eligible voted, which I thought was quite reasonable, since quite a few would have no direct use for the library.

The ballot revealed that 4 journals were in the library but not on the list provided. Eight journals were written in by electors which were not in the library.

The analysis of the results proved very interesting. With 31 people voting for a total of 198 journals, the quota is a lot less than 1. This implies that about the first six preferences would be selected for any reasonable number of journals. However, there was not a fixed number of ‘seats’, and hence I had to decide what threshold to set. Due to the difficulty for the electors, I did not interpret the ballot papers according the usual ERS rules. In one case in which one preference was unclear, I omitted that preference but did not ignore subsequent preferences. In two other cases in which a journal was selected twice, I merely ignored the second choice.

An initial analysis showed that 27 Journals did not appear in any position on the ballot papers. This gave an instant selection of journals to cancel. I ran the ballot with the option to cancel 10 and 20 further journals.

I have several STV algorithms available on my home computer which I used to compute the result. I had decided in advance that I would use the Meek algorithm for the election, but the other versions could be used to see what difference it made.

The first problem was that the programs I had, required a trivial modification to handle as many as 200 ‘candidates’. After having made that modification, it was found that the programs would not work on my PC because the full results over-filled my floppy discs! A further modification was needed to output only the final table and a summary of the eliminations and elections.

The three versions of STV were:

1) The Meek algorithm, as published in the Computer Journal (1987, Vol 30, p277)

2) The ERS hand-counting rules (as programmed by David Hill)

3) The Tideman algorithm, as approximated by my program.

The ERS results were quite unacceptable which shows that the hand-counting rules do not seem to have been used upon such an election. The problem is that if the election is run with the same number of seats as those selected in any preference, the algorithm does not select just those selected by the electors! This problem can be expected of any algorithm that does not see subsequent preferences.

The other two algorithms produced virtually identical results. With the reduction to 20 fewer journals than those selected, one difference was found between Meek and Tideman. A manual inspection of the results with the two journals in question, showed no clear distinction.

After producing the result, I computed for each of the 31 ballots, the way in which the final stage of the ballot had divided up the vote. This information was given to each elector. It created further interest in the STV algorithm. Those who had given more preferences had, in general, a lower non-transferable loss. However, the variations were very large. For instance, a person would gave the largest number of preferences (36) had a small loss, while a person would gave 15 preferences had no non-transferable loss.

I conclude from this election that STV can be used for such selections, but that the ERS hand-counting rules are not appropriate. Also, any STV algorithm approved by ERS in future should not suffer from this noted defect. Namely, if only N candidates are represented in the preferences and N is the number of seats, then the algorithm should elect those N. This requirement does not seem to lead to additional problems. It appears that the STV algorithms which recompute the quota can satisfy this requirement, since in the particular circumstances the entire ballot papers are then processed.

An STV Database

B A Wichmann

Since we know that no single algorithm for STV can have all the properties one might like, it appears that some statistical analysis may be needed to select an optimal algorithm. People do not vote at random and therefore any effective analysis must take into account voting patterns. For instance, if voters always voted strictly along party lines, proportional representation among such parties would be an important factor.

Collections of ballot papers from real elections would be useful for any practical analysis. There is a *de facto* standard for the representation of ballot papers in a computer, being the form used by the Meek algorithm. Hence collection of such data is practical and useful. Both David Hill, Nicholas Tideman and myself had such collections, accumulated informally over several years. I have now put this collection into a consistent framework so that the material can be provided to anybody who would like it — merely post a floppy disc to me, and I can return the disc with this data.

The data available has been classified in a number of ways as follows:

Real: Data here is that from real elections, with the possible exception that a statistical sample of the total ballot papers would be acceptable. The reason for this is that it presents a means of providing 'real' data without providing the total information. There are potential dangers in analysis of real data, since an alternative algorithm could elect a different person, giving rise to concerns about the election itself, rather than the principles involved. Another reason for accepting a subset of all the votes is that this is all that may be feasible for a large election. Obviously, this data is provided in a form which precludes the identification of the election involved. There are currently 46 data sets in this class.

Mock: This is data from genuine elections, except that no position or office is at stake. Mock elections are often used to educate people into the principle of STV. There are currently 2 sets in this class.

Semi: Elections in this class are not genuine elections, but are clearly related to real elections. Examples in this class are 'ballot' papers derived from published STV elections (from Northern Ireland), elections from the Eurovision Song Contest and elections in which there was no fixed number of 'seats'. There are currently 21 data sets in this class.

Test: Data in this class are not derived from any election but have been constructed to demonstrate the difference between some algorithms, show a bug in a computer algorithm, or some similar purpose. There are currently 129 in this class.

I would very much welcome additional data, especially from real elections in which some 'party' aspect is involved. The data can be provided in a form in which the origin cannot be traced. I have analysed an Irish election to produce a single data set in the **Semi** class, but this is very time consuming and has to make a number of assumptions to produce anything like the actual ballot papers. Hence real data is much superior.

Is a feedback method of calculating the quota really necessary?

R J C Fennell

Robin Fennell is a retired radar technician with the RAF and more recently, a Customs Officer. He has been a member of ERS since an abortive attempt to introduce STV into his union elections. He is currently active in transport and defence campaigning as well as electoral reform.

The March issue of *Voting matters* reprinted papers by B L Meek, [1,2] D R Woodall [3] and C H E Warren.[4] In this paper I will propose that their feedback method of calculating the quota is not necessary. To do this I will consider some of the basic principles of the Single Transferable Vote (STV) system.

One problem identified[5] is that if a candidate is elected any further preferences for that candidate are passed over. The question to be considered is 'are elected candidates continuing in the election or should they be considered as no longer available to receive votes'?

In other words is the purpose of a vote in the Single Transferable Vote system to try to elect candidates in the order the voter wishes or to place candidates in popularity order and have this order respected whatever the outcome of the rounds of the count? I suggest that it is the former. Once a candidate has been elected he has achieved the aim of participation in the election and, henceforth should take no further part in the election. Under these circumstances the manual counting method is satisfactory.

We will take a voting paper that shows preferences A,B,C,D and assume that B was elected on the first round. The transfer of B's surplus elects A on the second round. The question now arises on our paper, should the transfer of A's surplus go to B or C. Let us assume that our voter had future vision when deciding the preferences and knew that B would be elected in the first round; would our voter put B as the second preference? I suggest that anyone so gifted would select the preferences A,C,D thus maximising the transfers to the candidates they wished to see elected. Of course this foresight is not available to voters so to cover all possibilities the voter will elect to keep to the original selection knowing that the counting system will not waste any part of a vote by transferring it to a candidate already elected.

In practice few voters would take the risk of excluding a candidate on the grounds that they are certain to be elected. If too many did then B would not be elected. Voters can be expected to behave in a rational fashion and vote for the candidates of their choice in the order they wish. When a candidate has been elected they have achieved the aim of both the candidate and the voter. The voter will now wish any surplus votes to be concentrated on the unelected choices.

Another problem identified by Meek[6] is how to treat unmarked candidates. He suggests that they should be considered either as being of equal merit, or that the voter wishes to leave the ordering of these candidates to others. Meek ignores the third possibility that the voter does not wish these candidates to have any part of the vote. The omission of the third alternative in Meek's paper is possibly due to the voting instructions that take a form similar to ‘place the candidates in order until you can no longer differentiate between them’. If the instructions were changed to a form similar to ‘place the candidates in order until you no longer wish the remaining candidates to have your vote’ it would be clear how the voter required unmarked candidates to be treated. Under these circumstances the manual counting method is satisfactory.

If STV is to be used in local or parliamentary elections many voters will only want to vote for their particular party. They will not wish any proportion of their vote to go to candidates of a party with an opposing view to theirs. If votes are apportioned to all non-selected candidates, voters will have no way of ensuring that they do not vote for candidates of a party whose policies they cannot agree with.

The other problem foreseen by Meek[7] that I will consider is the possibility of voters indicating the same preference for two or more candidates. He suggests that this should be allowed and the counting system modified to accommodate it. The Electoral Reform Society (ERS) supports the Single Transferable Vote system, not the Transferable Multi-vote of Unity Value System. This second system may exist but it is not that supported by the Society and therefore should not be considered. The Single Transferable Vote system requires voters to cast a single vote, all or part of which may be transferred. That a multiple vote may have unitary value is irrelevant, it is a single vote which must be utilised.

D R Woodall[8] raises a different problem, that of the tactical voter. He postulates a situation where there are several Sensible Party candidates, say A,B,C and one Silly Party candidate, W. The tactical voter decides that W will be excluded and in order to maximise the transfer of votes after the first round he will vote W,A,B,C rather than A,B,C which is the real preference. The problem for the tactical voter comes when several voters take the same line. Assume in this election that the quota is 200. If 201 voters vote tactically and put W first then W will be elected reducing the vacancies available for Sensible Party candidates. In these circumstances the tactical voters will be as silly as W's party. The only way to avoid this is to place preferences in the order the voter wishes the candidates to be elected and not to attempt to vote tactically.

One of the main advantages of STV is that attempts to vote tactically are likely to end in a result that will not suit the tactical voter. The situation above could happen irrespective of the number of candidates or the size of the quota. The only safe way for voters to use their vote successfully is to vote according to preference.

The three works printed in the March issue of *Voting matters* may be mathematically rigorous but are they required? My contention is that if the basic principles of the Single Transferable Vote system are carefully considered then the feedback method of counting is unnecessary. The manual method used to date is satisfactory to ensure the correct result.

There is one further matter to be considered. If the feedback method is to be used, the constant recalculations necessary will require computers to be used. It is recognised in the papers supporting the method that it is too laborious to use hand counting. While the ERS has voted to use both computer and manual counting for its internal elections, I doubt if a system which cannot reasonably be counted by hand will be accepted by the general public. Computers are quick but they rely on the integrity of their programming. Computer technology is not yet at a state where incorrect programming, whether by accident or intent, will always be exposed. While it is not possible to say that the currently accepted Newland/Britton hand counting rules will always produce the correct result, they will produce a satisfactory result. I can see no reason to change the current system of counting.

References

1. B L Meek, A New Approach to the Single Transferable Vote, Paper I, *Voting matters*, March 1994.

2. B L Meek, A New Approach to the Single Transferable Vote, Paper II, *Voting matters*, March 1994.

3. D R Woodall, Computer Counting in STV Elections, *Voting matters*, March 1994.

4. C H E Warren, Counting in STV Elections, *Voting matters*, March 1994.

5. As 1, section 3, item (iv).

6. As 2, section 3, penultimate paragraph.

7. As 2, section 6.8. As 3, second paragraph.

8. As 3, second paragraph.

Issue 3, December 1994

Editorial

In this issue we have a mixture of papers. There is a continuing debate about revisions to the ERS rules, which arose from Fennell's paper in the last issue.

Hill and I, in separate papers, consider the effect of small changes — steadiness or stability. Global properties and local properties are the topic of Woodall's paper which I hope could be used as a basis for terminology and analysis in further issues of *Voting matters*.

It would be nice to automate all suggested algorithms for STV and compare them against a library of test cases. Unfortunately, the effort involved often precludes this which means that choices are being made on less than perfect information (not unlike elections themselves).

Brian Wichmann.

Comparing the stability of two STV algorithms

B A Wichmann

Brian Wichmann is the Editor of *Voting matters* and a visiting professor of The Open University

The problem of stability

This note does not consider the usual properties of STV algorithms that have been the subject of Woodall's analysis, but that of *stability.* For a mechanical system modelled by continuous variables, the analysis of stability is an application of differential calculus. We cannot use such an approach with STV algorithms since the system is discrete, and we know that some small changes are bound to produce a discrete change in those elected.

For an STV algorithm, we could have too much stability in that part of the ballot papers are simply ignored — for instance, by only using the first preference. On the other hand, we could have an algorithm which lacks stability in vital respects by changing the result for inconsequential changes to a ballot paper.

One change made to a ballot paper can be regarded as *small*, due to the nature of the preferential system. Since the usual means of balloting does not provide for the voter to give equal preference, when the ballot paper records ABC, this might be because A and B were regarded as equal, but the voter specified A first arbitrarily. Hence the voter could equally have written BAC instead. Hence given the ballot ABC, the voter's true intentions could perhaps have been expressed as BAC or ACB. In general, given n preferences, $n-1$ ballot papers constructed by interchanging neighbouring preferences could be regarded as *small* differences.

Now consider two algorithms for STV which have broadly similar properties (as do all serious contenders). Figures 1 and 2 represent graphically these two algorithms.

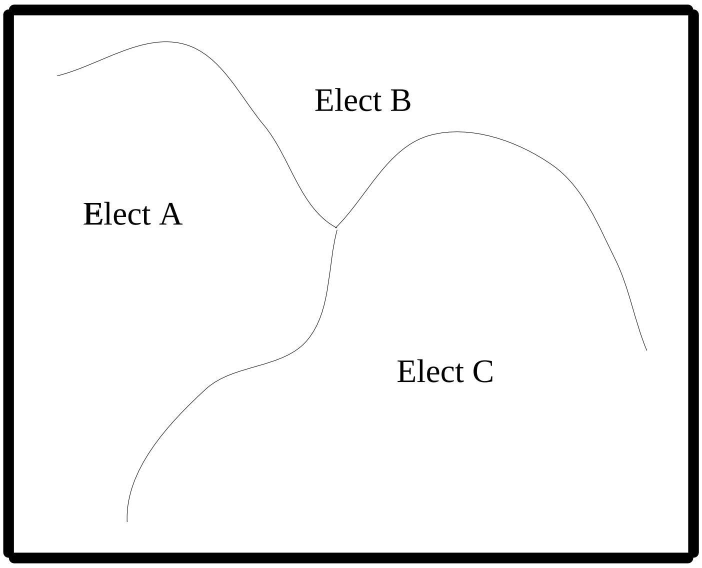


Figure 1

Figure 1 represents a stable algorithm since small changes are unlikely to change the result of an election, while Figure 2 represents an unstable algorithm. If we were operating in two dimensions, then the property of stability could be measured rather like the game of shove-halfpenny: one would measure the probability that a small circle placed at random on the figure crossed one of the dividing lines.

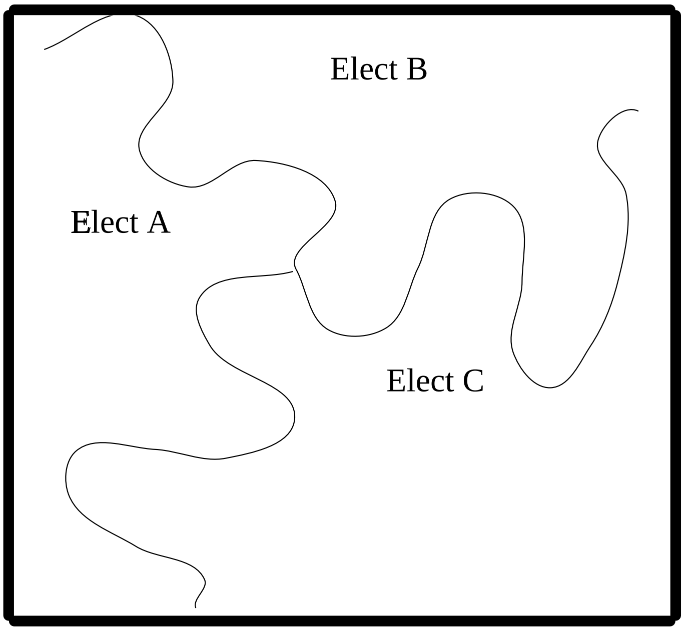


Figure 2

In the case of STV algorithms, we do not have a simple two-dimensional system, and hence the figures are a crude diagrammatic representation. To measure the probabilities we must conduct a suitably controlled experiment. Fortunately, we can use a computer to aid this process so that we can perform the equivalent of shove-halfpenny sufficiently often to obtain results which are likely to be meaningful.

The experimental method

We now specify the experimental method to compare the ERS hand counting rules versus the Meek algorithm. (Any two algorithms could be chosen, but this seems the most interesting pair.)

We select an actual election for which the ballot papers are available. We also choose a number, about 20, which is the number of ballot papers from the full set that is to be selected at random. (We return to the choice of this number *n* later.)

From each real election, we derive 100 mini-elections by randomly selecting *n* ballot papers. The experimental method is to analyse the effect of making small changes to these mini-elections. The analysis is as follows. Firstly, we compute the result of the two algorithms from the mini-election. (The result need not be the same for the two algorithms, nor the same as for the full election.) We now consider all the possible similar mini-elections derived by making one small change to one of the ballot papers. (This is potentially hundreds of elections — hence the computer.) This particular mini-election is on the edge if a specific criterion is met, say at least one of the small changes produces a different result.

The choice of *n* is important. If *n* is very small (say 1), then it is clear that the mini-election will not be representative of the real election. On the other hand, if *n* is large (say the full election), then the computation of the ‘edge’ becomes too large, and also the number of possible mini-elections becomes too small (in this case only 1). Care must be taken over the specific criterion for being on the edge. If one takes something like the ERS council elections (i.e., several posts to fill with no parties, so that small changes are likely to make a difference to the outcome), with the criterion that *any* small change resulting in a difference implies being on the edge, then there is a danger that *all* mini-elections are on the edge!

For the 100 random mini-elections we perform a different analysis in each of the three experiments given here. If one could assume statistical independence, then it would be a simple matter to undertake a χ^2 test to see if the result is significant. Unfortunately, we do not have elections with a large enough number of ballot papers to ensure the independence, and therefore we must be content with a non-statistical treatment.

The programs and test data

Two programs have been written, one for using the ERS algorithm and the other the Meek version. Apart from the STV algorithm in use, both work in an identical fashion. They read 100 mini-elections in the conventional format. Firstly, the result is computed for this election, then every possible small change is made, and for each such change, the number of changes to those elected is recorded.

The number of changes to those elected for one small change is usually 0 (no change), but is sometimes 1, rarely 2 and very rarely 3. Hence for each ballot paper in the mini-election, $n-1$ integers are output, representing the number of changes arising from each of the $n-1$ possible interchanges of adjacent preferences, where *n* is the number of preferences marked on the ballot paper. This implies that the output is of similar length to the input — an important consideration, since if complete results were printed for each election result computed, hundreds of pages of material would be produced.

The analysis is most easily seen by considering an example. A mini-election from election R038 is as follows:

```
 17 5
1 11 9 10 0
1 10 17 5 9 11 16 0
1 6 16 2 1 14 17 10 9 11 5 8 4 12 13 15 0
1 4 8 12 15 13 0
1 17 5 11 1 16 10 2 0
1 5 9 10 11 17 0
1 3 7 9 14 17 0
...
```

```
1 8 10 13 11 17 0
1 9 5 6 17 0
1 11 10 5 17 9 0
1 6 4 15 14 16 8 1 0
1 6 14 16 1 2 4 13 12 8 15 0
1 13 4 15 12 8 0
0
"A.1 " "B.2 " "C.3 " "D.4 "
"E.5 " "F.6 " "G.7 " "H.8 "
"I.9 " "J.10 " "K.11 " "L.12 "
"M.13 " "N.14 " "O.15 " "P.16 "
"Q.17 "
"1R038: H3H "
```

The above data is for an election with 17 candidates for 5 seats, in which the first ballot paper selects candidate 11 (K) as the first preference, then 9 and lastly 10. The names of the candidates are the letters A-Q, a convention used throughout.

The program computes the effect of making all possible interchanges of adjacent preferences, which for Meek gives:

```
v1 +F-L-B-O-P-H-G-N-E-M-C+I+Q-K+J+A-D 68
0 1 1 1
2 1 1 1 1 1 1 0 0
2 1 1 0 0 1
0 0 1 1
...
2 1 1 1
0 1 0 1
1 0 0 1
1 0 0 0 0 0
2 1 0 1
1 1 0 2 0 0 1 1 0 1 0 0 1 1
1 1 0 1 1
0 1 m
```

The first line gives the result (with Meek) for this mini-election, where `+F-L` means F is elected and L is excluded, etc. (The v1 and 68 are not relevant.) Then, starting with the last ballot paper and working back towards the beginning, the number of differences to the result is printed for each possible interchange. Hence the last ballot paper has four possible interchanges, the first one giving no difference, but the last three each making a single difference. So in this case, interchanging the first two preferences makes no difference, but interchanging the 2nd and 3rd preferences does change the result by one candidate. The 'm' relates to the third experiment and is explained later.

One other program is needed which selects *n* ballot papers at random from a real election, and repeats this 100 times. This program is fast and straightforward.

For the main election data, six real elections have been chosen from the data already available (see *Voting matters*, Issue 2). The statistics from these elections are as follows:

Identifier	Papers	Candidates	Seats	*n*
R006	239	9	2	20
R008	261	10	3	25
R010	270	9	5	27
R017	479	8	1	15
R033	196	14	7	25
R038	177	17	5	20

Unfortunately, none of the elections in the data base are from elections involving parties, and so such elections could not be selected for this study.

We can now summarise the results obtained by example. For election R017, 100 mini-elections are computed by selecting 15 ballot papers from the actual 479. For each of these mini-elections, we compute what difference (if any) would be made by a single transposition of a preference. This is repeated for each possible transposition, which in this case, involves the analysis of 4585 elections!

Experiment 1

We now consider the issue raised initially — that of the ‘size’ of the edge dividing the line between different election results. We therefore need to devise a criterion for being on the ‘edge’, and compare the results for the six elections with the two algorithms.

Criterion: Some change for any transposition

Election	ERS edge	Meek edge
R006	74	65
R008	80	74
R010	95	87
R017	69	74
R033	99	95
R038	100	100

This table means, for instance, that for the 100 mini-elections derived from R006, 74 are on the ‘edge’ for ERS and 65 for Meek — which implies that there were 26 or 35 elections for which no change was made by any transpositions. Hence a very high proportion of the mini-elections are on the ‘edge’, over three quarters in almost all cases. However, even the most optimistic assumption shows that there is not much difference between the two algorithms.

We now change the criterion for being on the edge so that a lower proportion are on the edge.

Criterion: More than three transpositions make a change

Election	ERS edge	Meek edge
R006	41	38
R008	55	46
R010	76	57
R017	32	49
R033	91	75
R038	94	91

We again conclude that there is not much difference between the two algorithms.

We need to look at aspects other than the actual size of the edge to see significant differences.

Experiment 2

In this experiment, conducted with the programs and data as before, we look at properties of the edges rather than their actual magnitude.

Given a mini-election which is on the edge, then we know at least some transpositions of the preferences will change the result. It is therefore natural to ask which specific transpositions can change the result. Clearly, it is more likely that transposing the first two preferences will alter the result, but what about the subsequent transpositions? We therefore analyse the number of times a transposition makes a change, against the position of the transposition (p_i).

```
Combined results
      p1  p2  p3  p4  p5  p6  p7  p8  p9 p10 p11 p12
R006
 ERS 283  31  14   6   4   0...
Meek 310 147  61  39  23  16  14   0...
R008
 ERS 452  56  11   8   5   4   0   1   3   0...
Meek 393 161  70  42  25  13  12  11   0...
R010
 ERS 668 173  36  21   9   4   2   0...
Meek 423 174  82  34  21  18  13   0...
R017
 ERS 214  27   4   5   2   1   0...
Meek 279 210 123 119 104  94   0...
R033
 ERS 979 227  78  31  17   8   3   2   1   0   0   1
Meek1876 392 225 144 117  91  69  61  57  41  41  34  37
R038
 ERS 734 203  44  31  17   6   3   2   0   1   1   0
Meek 723 502 376 346 157 138 107  97  91  44  36  33
```

In the table above, for each of the six elections, the number of times a transposition makes (at least) one change to the result is tabulated against the preference position for all the 100 mini-elections. The difference between ERS and Meek is now obvious. The number of changes for the first preference between the two algorithms is similar and is surely not significant. However, in all subsequent preferences, many more changes arise from Meek than from ERS.

In examining the subsequent preferences, there is no natural scale to work to, since a change in preference *n* is more significant if there are *n* candidates than 2*n* candidates. The number of seats is also relevant to this scale. Hence in analysing the table above, both the number of candidates and seats must be considered.

We can add up the results from each election for those positions beyond the number of seats (*s*) for each election, giving the following results:

	Position			
	s +1	*s*+2	*s*+3	>*s*+3
ERS	61	21	15	11
Meek	530	364	293	596
ratio	8.7	17	19	54

Hence we conclude that transposing preferences beyond the number of seats has virtually no effect with ERS as compared with Meek.

Experiment 3

In a paper in this issue of *Voting matters*, Woodall defines the property mono-raise. For the elections analysed by the experiments undertaken here, we can determine the extent to which a weaker property than mono-raise is violated. Since our analysis determines the effect of a single interchange in the preferences, given a preference pair A,B which is replaced by B,A, the raising of the order of B should not disadvantage B. This implies that if the election with A,B elects B, then that with B,A should also elect B. If this condition is not satisfied, then mono-raise is violated, and is marked by 'm' in the output files.

We can now compare the violation rate for ERS and Meek, which is as follows:

Election	ERS violations	Meek violations
R006	0	32
R008	2	29
R010	5	8
R017	5	78
R033	5	70
R038	0	141

Hence there is no question that Meek violates mono-raise much more than ERS. This is likely to be due to the increased sensitivity of Meek to the effects of late preferences.

Conclusions

The analysis undertaken in this paper has led to the following conclusions:

1. There is no evidence that the ERS and Meek algorithms are any different with respect to the size of the boundary between the election of different candidates.

2. Making small changes by transposing preferences later than the number of seats makes virtually no difference with ERS but a substantial difference with the Meek algorithm.

3. Meek violates mono-raise much more than ERS.

Point 2 indicates that the Meek algorithm is much more sensitive to the voter's wishes than ERS, and moreover this sensitivity is not at the expense of making the algorithm less stable. However, the fact that Meek violates mono-raise so much more than ERS might question the extra sensitivity of Meek. It would appear that an ideal algorithm would have the sensitivity of Meek, but would only violate mono-raise with the same frequency as ERS. I suspect that it is actually the extra sensitivity of Meek that gives rise to the mono-raise violations, so that the best of Meek and ERS is not possible.

It appears that the results presented here have some limitations. Firstly, the mini-elections necessarily have a small number of ballot papers and so the results need not apply to larger elections. Secondly, a consequence of the small number of ballot papers is that in many cases, random choices are made by both the ERS and Meek algorithms.

The comparative steadiness test of electoral methods

I D Hill

David Hill is a regular contributor to *Voting matters.*

In comparing one electoral method with another it is useful to examine their comparative steadiness. It should be noted that it is only a comparative test and does not give a "goodness" score for any individual method on its own but only for one method relative to another. Nor does the fact that any method comes out as the better of the two by this test indicate that it is necessarily better in any other way.

To use it, first run each method for the same number of seats and the same given set of votes and see whether they both elect the same candidates. If they do, this test is not applicable. Otherwise, see whether there is one or more candidate whom neither method elects. If there is no such candidate, again the test is not applicable. In particular, the test can never be applicable if the number of candidates is only 1 greater than the number of seats, but the fact that it is often not applicable does not destroy its value in those cases where it does apply.

If the test is applicable, then treat all candidates who failed to be elected by either method as withdrawn, and re-run each method. If each method continues to elect the same candidates as before, then there is nothing to choose between them on this test for this particular set of votes. If, however, one method makes no change in whom it elects while the other makes a change, then the no-change method gains a point in comparison with the other.

For example, if there are 5 candidates for 3 seats, and the votes are:

```
51 ABC
44 ABD
 5 EABD
```

the current ERS rules will elect A, B and D whereas the Meek rules will elect A, B and C. They agree that E is not elected, so the comparative steadiness test treats E as having withdrawn and re-runs the election. Now the Meek rules still elect A, B and C, but the ERS rules switch to electing A, B and C too. Meek therefore shows greater steadiness for this particular set of votes.

While such artificial elections are important as illustrations, what most matters is which rules are steadier for real elections. Taking the 57 real elections that I have available, I find the test to be applicable for only 10 of them. In 4 of those, these two systems are both steady, neither changing its result when the relevant candidates are withdrawn. In the other 6, however, the Meek system remains steady but the ERS system changes. By this test, the Meek system seems to be superior, so far as the evidence goes, though a few more results in the same direction would help to make more certain that the difference is not just a chance effect.

It should be noted, of course, that discovering a lack of steadiness must not be used to change the result of a real election, which must always be in accordance with the rules as laid down for that election. The test is only for research purposes, not to interfere with a result.

Editorial Note: It is possible to apply the steadiness test even when an election gives the same result. This can be done by selecting random ballot papers from the election in the manner of the mini-elections in the previous paper. With the 100 mini-elections from the real election R006, 17 of these elect different candidates so that the steadiness test can be applied. Of these 17, none were steady for the ERS rules, while 13 were steady according to Meek. One mini-election could not be considered since a random choice was made. For the remaining 3 mini-elections, neither were steady, and in one case, the removal of the no-hope candidates causes the two algorithms to interchange the results!

Response to the paper by R J C Fennell

P Dean

Peter Dean is a Trustee of the ERS Ballot Services.

I was surprised to see the existing manual system defended by R J C Fennel as being beyond reproach.

The basic flaw in the manual method is that it allows for the election of candidates receiving less than the quota. This has led to Tasmania requiring at least 7 preferences, combined with a rotated ballot paper since 1973. Even in our own elections some are elected with 4 fewer votes when there are 8 non-transferable votes.

There is a refinement which could easily be introduced in manual elections. This is that when the stage is reached when some candidate(s) fail to reach the quota, a recount takes place with only those remaining taking part. This means that votes previously wasted on candidates with no chance can then influence the result by being allotted to a lower preference for a candidate previously elected. The result will then be demonstrably fair. Taking an actual mock election in the Solent area in 1989 as an example, in which there were 20 candidates for 5 seats. The manual result gave a lead of 4.88 to the last elected — although short of the quota. The Meek system elected the runner-up instead by a margin of 2.01 votes. If a further 5 counts has been added the manual system would have come to a similar result, but by an even larger margin of 7.42. The new result is demonstrably fair — with the last candidate having 2.53 over the quota.

Sometimes the unfair result is even obvious to the public. Such a case occurred in Cork East in 1954. The two Fine Gael candidates received 153 more votes than the two Fianna Fail candidates (1162 non-transferable) yet Fianna Fail won 2 of the 3 seats. Such results discredit the whole system.

The current mechanised system is quite unsuitable for small elections. For instance, with 9 votes and 18 candidates for 3 seats it proceeds to eliminate 7 candidates with 1 vote completely at random. It is quite clear that a different order of exclusion would give a different result. Personally I favour a points method based upon the preferences expressed which would give some form of ranking order to be used instead of the random method.

Are better STV rules worthwhile?

A reply to R J C Fennell

I D Hill

David Hill is the Chairman of the ERS Technical Committee.

R J C Fennell's article, in *Voting matters* issue 2, raises a number of matters that deserve reply.

Taking a voting paper naming ABCD in order of preference, where B has been elected on the first count and A is elected on the second count as a result of transfers from B, he asks whether that paper's surplus should go to B or to C. He appears to have failed to notice that, in the current ERS rules, it is totally immaterial whether the vote is taken as if it were ACD instead of ABCD, because that paper's surplus does not go anywhere. The voter's second and subsequent preferences are completely ignored, the whole paper remains with A, while only the new votes that A has received are redistributed.

Let us look instead at the point that he was trying to make. Suppose, in that same election, that C was also elected on the first count, and we have a paper naming BACD. That paper will pass from B to A and will be further redistributed, at a suitable value. Should it go to the next choice C, or jump over C straight to D as currently happens? He suggests that such a voter with future vision would not have put C into the list, so it is right to jump to D. But all voters ought to be treated alike, and therefore, if we are to treat one as if future vision existed, we must do so to all others too, and most voters would wish to change their votes if they knew what was going to happen; nobody would vote for the runner-up, of course. But such a change would make sense only if nobody else changes; if we treat everybody as though allowing them to change, the assumed future vision would collapse, no individual could then know how to change and the whole system would become wildly unstable. There is only one satisfactory way out, and that is to treat each vote in strict accordance with what it says, and not by what we assume that it might have said if only the voter had known what would happen.

Transferring to a candidate who has already been elected, as in Meek-style STV, does not waste votes, as is suggested, because the same size surplus is passed on in any case. The change is only to whether the surplus is taken fairly, from all relevant groups in proportion to their current totals of votes, or unfairly in some other way. To change the example, suppose that there are 100 AC votes and 10 BAD votes in a situation where the quota is 77. A is elected on the first count giving 77 to stay with A for quota, 23 to be transferred to C.

If, later on, B is excluded then in current ERS rules the 10 all pass to D. The Meek alternative is that only 3 pass to D, while 7 more of the original ACs pass to C. The two methods are

```
ERS    100 votes    77 stay with A    23 go to C
further 10 votes                      10 go to D

Meek   100 votes    70 stay with A    30 go to C
further 10 votes     7 stay with A     3 go to D
```

In either case 77 have been kept and 33 redistributed, but I do not see how anyone could claim that the first method is satisfactory if we are able to operate the second. The article suggests that the voters ‘will wish any surplus votes to be concentrated on the unelected choices’. That is to say that the BAD voters would like the first alternative. Of course they would; that is not in dispute. But it is not fair to the AC voters to allow it.

The next point addressed by Fennell is the treatment of ‘short lists’, that is to say votes that would be transferred if they had a next choice, but do not show one. He mentions the two possible treatments discussed in detail by Meek, but says that Meek's papers ignore a third possibility, and it is evident that he is thinking of something like the current ERS rule. He is wrong to say that Meek ignored this; his paper said ‘If the difficulty were to be avoided by increasing the proportion transferred of votes for which a next preference is marked, to enable all x votes to be retained by C, this would clearly reintroduce inequities of the kind Principle 2 was designed to eliminate’. I agree with Meek that this possibility does not deserve any more discussion than that, but many people have failed to see that this method is wrong in principle, and a far greater quantity of writing has gone into it in the last few years than can be reproduced here. I can well see that people might take the wrong decision on this at a first quick glance, but the number who continue to do so even after thought and discussion is quite extraordinary.

I disagree with Meek that the voters should be given the choice between the two methods he discusses in detail. This would have to mean explaining to them the different effects of each, a task that I would not wish on anyone. Meek points out that the two can give different results; usually they do not but, in the few cases we know of where they do, to give the relevant surplus to ‘non-transferable’ and reduce the quota to compensate is always the preferable option.

Fennell suggests that these voters may not wish other candidates to have any part of the vote. I agree with that — indeed I insist that, whatever those voters wish, we have no right to assume what their wishes are, but only to obey what their ballot papers say, namely that if they become entitled to a further choice they wish to abstain from making one. It is true that, in the current ERS (and most other) rules, the ballot papers are not physically transferred to any other candidate, but what matters is not what is done with pieces of paper, but the effect of the rule. Consider the simple case, with 4 candidates for 2 seats, and votes

```
40 AB
17 CD
 3 DC
```

The quota, in current ERS rules, is 20. So 20 votes go to A and A's surplus of 20 goes to B, and A and B are elected, but the situation is ‘on a knife-edge’ for, if D were to withdraw before the count, A and C would be elected. Now with a knife-edge situation any relevant change in one direction must settle the matter, and it is certainly a relevant change if half B's support is lost, to give

```
20 AB
20 A
17 CD
 3 DC
```

Yet the current ERS rules take no notice whatever, but still give 20 to A and 20 to B.

It is sometimes argued that if the AB voters had had pre-vision, they would have gone straight to B but, as argued above, we cannot allow that without allowing pre-vision to other voters too. Given pre-vision, the DC voters would have voted for C. Given pre-vision, the A plumpers need not have bothered to vote at all. The only fair thing to do is to take what the ballot papers actually say, and everything in proportion to the numbers involved. That means that half A's surplus must go to B, and half to non-transferable, which gives C the second seat.

Provided that the quota can be changed to allow for the non-transferables, as in Meek's method, it can be shown that this does not waste any extra votes at all. What one method wastes in non-transferable, the other wastes on leaving more votes with elected candidates than they now need to be sure of election. With hand-counting methods, where true quota-reduction is not practicable, it could be the case that the present rule does more good than harm, but I know of no evidence to support such a view.

Turning to the discussion of whether voters should be allowed to express equality of preference if they wish, rather than a strict ordering, I cannot agree that it ought not to be considered. It is undoubtedly the case that the absence of this feature is regarded by many as a major disadvantage of STV. There are some difficulties of implementing it, and it would complicate the instructions to voters. I believe that it is something that we ought to consider introducing one day, but that there are more important things to be done first.

Fennell then discusses the ‘Silly Party candidate’ method of tactical voting discussed in Woodall's paper. He is right, of course, that trying to utilise it may be to the voter's disadvantage if a wrong guess is made, and would certainly be troublesome if too many voters tried to do it. It is a bad feature though that it should be possible at all. Furthermore it is not necessary for there to be a silly party candidate or tactical voters for the effect to occur. It always happens to those voters who put as first preference the first candidate to be excluded, no matter how sincere their choice, causing a distortion that should be avoided if possible.

Fennell is correct that it is not practicable, save for very small elections, to use methods such as those proposed without doing the count by computer but, in this electronic age, can that really be an adequate reason for putting up with second-best results? He queries whether computer-generated results would be trusted and this is certainly something to which attention has to be given. There are two distinct ways in which things might go wrong. The first is in the input of the data from the ballot papers, but this could be subject to repetition if a recount is requested.

The second possibility of error is in the program to calculate the results but, in a public election it could be arranged that, once the data input has been agreed as correct, each candidate would be given a copy of the data on floppy disc. Each party would have its own program, each independently written from the rules specified in the Act of Parliament, and its own computer near at hand. Within a few minutes each could have checked that the official result is agreed. Such a system would lead to much greater protection against errors than anything that could be done with hand-counted STV.

The article concluded that ‘the currently accepted Newland/Britton hand counting rules ... will produce a satisfactory result. I can see no reason to change the current system of counting’. What is meant by a satisfactory result? In comparing the results of real elections by hand-counted rules and by Meek rules the result is different more often than not, and all the indications are that the result is not merely different but better, in more accurately representing the voters' wishes. Now that the ability exists to do something better than can be done by hand, it would be absurd to try to exist in the past. Does it not matter to the Electoral Reform Society (or others) whether we get the best result or not?

Properties of Preferential Election Rules

D R Woodall

Douglas Woodall is Reader in Pure Mathematics at Nottingham University. In this article, he argues that more attention should be paid to properties of electoral systems, and less to procedures. He lists many properties that a preferential election rule may or may not have, and discusses them with reference to STV.

1. Introduction

I have often been struck—and never more than in the last year—by how much the types of argument used by the supporters of the Single Transferable Vote (STV) differ from those used by its opponents. When it comes to the details of the count, the supporters of STV almost invariably try to defend its procedures directly, on the grounds that they follow certain principles, or that they do with each vote exactly what the voter would want done with it, if the voter were able to be present at the count and to express an opinion. Unfortunately, there is no guarantee that adopting sensible procedures, at each stage of the count, will lead to a system with sensible properties, and the opponents of STV often emphasize its less desirable properties. In particular, it is now well known that STV is not monotonic: that is, that increased support, for a candidate who would otherwise have been elected, can prevent that candidate from being elected. It was ostensibly because of this and related anomalies that the Plant Report rejected STV.

Properties of electoral systems can be thought of as "performance indicators", and like any other performance indicators they need to be used with care. If one chooses a set of performance indicators in advance, it may well be possible to manufacture a high score on those indicators in an artificial way, which does not represent good performance in any real sense. Nevertheless, it seems to me that the Electoral Reform Society needs to pay more attention to properties if it is not to be sidelined in the electoral debate. In particular, since different desirable properties often turn out to be mutually incompatible, it is important to discover which sets of properties can hold simultaneously in an electoral system. Only then will it be possible to decide whether there are electoral systems that retain what is essential in STV while avoiding some of the pitfalls.

The purpose of this article is to introduce a long list of technical properties that an election rule may or may not have, to invent snappy descriptive names for them all, and to discuss them with special reference to STV. Except where otherwise indicated, statements made about STV apply equally well to the Newland-Britton and Meek versions of

STV. In a later article I hope to address the question of monotonicity in more detail.

2. Notation and terminology

As is usual in the Social Choice literature, I shall use lower-case letters *a, b, c,...* to denote candidates (or choices). Each voter casts a *ballot* containing a *preference listing* of the candidates, which is written as (for example) *abc*, to denote that the voter places *a* first, *b* second and *c* third, with no fourth choice being expressed. A preference listing is *complete* if all candidates are included in it and *truncated* if some are left out. (A preference listing that leaves out just one candidate will be treated by most election rules, including STV, as if it were complete; but one should not call it complete, since some election rules may not treat it as such.) A *profile* is a set of preference listings, such as might represent the ballots cast in an election. Profiles may be represented in either of the forms shown for Elections 1 and 2 below, indicating either the proportion, or the absolute number, of ballots of each type cast.

The term *outcome* will be used in the sense of "possible outcome" (assuming there are no ties). Thus in an election to fill two seats from four candidates *a, b, c, d,* there are six outcomes, corresponding to the six possible ways of choosing the two candidates to be elected: {*a*, *b*}, {*a*, *c*}, {*a*, *d*}, {*b*, *c*}, {*b*, *d*} and {*c*, *d*}.

```
  Election 1          Election 2
   (1 seat)            (2 seats)
ab   0.17          a   9    ea   4
ac   0.16          b   9    eb   4
bac  0.33          c  10    fc   1
cb   0.34          d  10    fd   1
                            fe   6
```

An *election rule* is usually thought of as a method that, given a profile, chooses a corresponding outcome—or, in the event of a tie, chooses two or more outcomes, one of which must then be selected in some other way (such as by tossing a coin). However, this description is not quite adequate to deal with the complexities of ties. Consider Election 1 above, with the votes counted by STV (or, rather, by the Alternative Vote (AV), which is the rule to which STV reduces in a single-seat election). No candidate reaches the quota of 0.5, and there is an initial tie for exclusion between *a* and *b*. If *b* is excluded then *a* is immediately elected, whereas if *a* is excluded then *b* and *c* tie for election. Thus *a* is elected with probability ½, and *b* and *c* are elected with probability ¼ each.

A similar situation arises in Election 2, again under STV. There are 54 votes cast, so the quota is 18, and there is an initial tie for exclusion between *e* and *f*. If *e* is excluded then *f*, *c* and *d* must also be excluded, and *a* and *b* are elected; whereas, if *f* is excluded, then *a* and *b* must also be excluded, and then *e* is elected and *c* and *d* tie for second place. Thus the outcome {*a*, *b*} is chosen with probability ½, and the outcomes {*c*, *e*} and {*d*, *e*} are chosen with probability ¼ each.

Because of examples like these, I define a (preferential) *election rule* to be a procedure that, given a profile, associates a corresponding non-negative probability with each outcome, in such a way that the probabilities associated with all possible outcomes add up to 1. The "normal" situation is that all the outcomes are given probability 0 except for one, which has probability 1 (meaning that that outcome is chosen unequivocally). If anything else happens, then we say that the result is a *tie* between all the outcomes that have non-zero probability.

3. Axioms

There are so many properties that an election rule may have, that it is useful to categorize them in some way. Four in particular seem sufficiently basic to deserve to be called *axioms*. The first is more or less implicit in the above definition of an election rule; but it has a name, and so for completeness I include it here.

Anonymity. The result should depend only on the number of ballots of each possible type in the profile (and not, for example, on the order in which they are cast, or on extraneous information such as the heights of the candidates).

Neutrality. If some permutation is applied to the names of all the candidates on all the ballots in the profile, then the same permutation should be applied to the result. For example, since STV is neutral, if *a* is replaced by *c* and *c* by *a* on every ballot in Election 2 above, then STV would choose {*b*, *c*} with probability ½ and {*a*, *e*} and {*d*, *e*} with probability ¼ each. One consequence of neutrality is that a tie in a single-seat election cannot be resolved simply by electing the first in alphabetical order among the tied candidates.

A rule that is both anonymous and neutral is called *symmetric*.

Homogeneity. The result should depend only on the *proportion* of ballots of each possible type. In particular, if every ballot is replicated the same number of times, then the result should not change. It is this property that enables us to describe profiles as in Election 1 above, showing the proportion, rather than the absolute number, of ballots of each type cast.

Discrimination. If a particular profile P_0 gives rise to a tie, then it should be possible to find a profile P that does not give rise to a tie and in which the proportion of ballots of each type differs from its value in P_0 by an arbitrarily small amount. This rules out, for example, the following method of electing one candidate from three: elect the candidate who beats both of the others in pairwise comparisons, if there is such a candidate, and otherwise declare the result a three-way tie. For in that case, not only would the profile in Election 3 below give rise to a tie, but anything at all close to it would also give a tie, contrary to the axiom of discrimination.

```
                abc  1/3
Election 3:     bca  1/3
(1 seat)        cab  1/3
```

A *proper* election rule is one that satisfies the above four axioms; that is, one that is anonymous, neutral, homogeneous and discriminating. The term "axiom" is used rather freely in the literature as a synonym for "property", but I shall restrict its use to these four, which I regard as genuinely axiomatic, in the sense that I am not interested in any rule that does not satisfy them.

A word of warning is needed about homogeneity. In any practical election where the count is carried out by computer, there will be a limit to the number of decimal places that the computer can hold accurately. Thus there are bound to be situations in which two numbers that are not really equal are regarded as equal by the computer program, because they become equal when rounded to the appropriate number of decimal places. In this case, if every ballot were replicated the same, sufficiently large, number of times, then the difference between the two numbers of votes would become significant, and the computer might give a different result. However, this is a minor problem, introduced by the practical need to round numbers; the axiom of homogeneity should be applied to the underlying theoretical rule, with no rounding.

With this interpretation, STV is a proper election rule.

4. Global or absolute properties

It is convenient to divide properties into *global* or *absolute* properties on the one hand, and *local* or *relative* properties on the other. The former say something about the result of applying an election rule to a single profile, whereas the latter say something about how the result should (or should not) change when certain changes are made to the profile. Not all properties fall unambiguously into one of these two classes, but sufficiently many do for the distinction to be useful.

The most important single property of STV is what I call the *Droop proportionality criterion* or *DPC*. Recall that if v votes are cast in an election to fill s seats, then the quantity $v/(s + 1)$ is called the *Droop quota*.

> DPC. If, for some whole numbers k and m satisfying $0 < k \leq m$, more than k Droop quotas of voters put the same m candidates (not necessarily in the same order) as the top m candidates in their preference listings, then at least k of those m candidates should be elected. (In the event of a tie, this should be interpreted as saying that every outcome that is chosen with non-zero probability should include at least k of these m candidates.)

In statements of properties, the word "should" indicates that the property says that something should happen, not necessarily that I personally agree. However, in this case I certainly do: DPC seems to me to be a *sine qua non* for a fair election rule. I suggest that any system that satisfies DPC deserves to be called a *quota-preferential* system and to be regarded as a system of proportional representation (within each constituency)—an STV-lookalike. Conversely, I assume that no member of the Electoral Reform Society will be satisfied with anything that does not satisfy DPC.

The property to which DPC reduces in a single-seat election should hold (as a consequence of DPC) even in a multi-seat election, and it deserves a special name.

> Majority. If more than half the voters put the same set of candidates (not necessarily in the same order) at the top of their preference listings, then at least one of those candidates should be elected.

The following rather weak property was formulated with single-seat elections in mind, but it makes sense also for multi-seat elections and, again, it clearly holds for STV.

> Plurality. If some candidate a has strictly fewer votes in total than some other candidate b has first-preference votes, then a should not have greater probability than b of being elected.

The next property has been suggested to me by Brian Wichmann in the light of his experiences reported in the last issue of *Voting matters*[6].

> No-support. A candidate who receives no support at all (that is, who is not listed by any voters in their preference listings) should not be elected unless every candidate who receives some support is also elected.

This is not satisfied by STV with the Newland-Britton rules. For example, if x receives no support at all, and the only support that y receives is on ballots marked *ay*, where a reaches the quota as a result of transfers from other candidates, then x and y will both be recorded throughout as having no votes (since the *ay* ballots are not re-examined when a reaches the quota), and so y is as likely to be excluded as x. It seems that no-support is satisfied by Meek's version of STV, although I do not have a formal proof of this.

The remaining three global properties consist of *Condorcet's principle*, which was proposed by M. J. A. N. Caritat, Marquis de Condorcet (1743-1794), and two modern strengthenings of it. We say that a voter, ballot or preference listing *prefers a* to b if he, she or it lists a above (before) b, or lists a but not b. Let $p(a, b)$ denote the number of voters who prefer a to b. We say that *a beats b* (in pairwise comparisons) if $p(a, b) > p(b, a)$; that is, if the number of voters who prefer *a* to *b* is greater than the number who prefer *b* to *a*. We say that *a ties with b* (in pairwise comparisons) if $p(a, b) = p(b, a)$. A *Condorcet winner* is a candidate who beats every other candidate in pairwise comparisons. A *Condorcet non-loser* is a candidate who beats or ties with every other candidate in pairwise comparisons; note that if there is more than one Condorcet non-loser then all the Condorcet non-losers must tie with each other.

Note that there need not be a Condorcet winner, or even a Condorcet non-loser. In the profile shown in Election 3 above, *a* beats *b*, *b* beats *c* and *c* beats *a*, all by the same margin of 2/3 to 1/3. This is the so-called *Condorcet paradox* or *paradox of voting*: even though each voter provides a linear ordering of the candidates, the result when the votes are totalled can be a cyclical ordering. The *Condorcet top tier* is the smallest nonempty set of candidates such that every candidate in that set beats every candidate (if any) outside that set. In Election 3, the Condorcet top tier consists of all three candidates. If there is a Condorcet winner, then the Condorcet top tier consists just of the Condorcet winner. If there is a Condorcet non-loser, then the Condorcet top tier contains all the Condorcet non-losers, but it may possibly contain other candidates as well.

Condorcet's principle and the two strengthenings of it given below were formulated originally for single-seat elections in which every voter provides a complete preference listing; but I have reworded them here so that they make sense (although they are not necessarily sensible) for all preferential elections.

Condorcet[1]. If there is a Condorcet winner, then the Condorcet winner should be elected.

Smith-Condorcet[4]. At least one candidate from the Condorcet top tier should be elected.

Exclusive-Condorcet (see Fishburn[2]). If there is a Condorcet non-loser, then at least one Condorcet non-loser should be elected.

Note that Smith-Condorcet and exclusive-Condorcet both imply Condorcet, and Smith-Condorcet also implies majority. Smith-Condorcet seems a very natural extension of Condorcet. Exclusive-Condorcet is also very natural, but it is of much less importance since it differs from Condorcet only when there is a "tie" for first place under pairwise comparisons, and that will not happen very often.

Election 4 (1 seat)		Election 5 (2 seats)	
abc	0.30	ad	0.36
bac	0.25	bd	0.34
cab	0.15	cd	0.30
cba	0.30		

STV does not satisfy Condorcet, and so it certainly does not satisfy either of the above two extensions of it. This can be seen in Election 4 above. Under STV (AV), *b* is excluded and *a* is elected. However, *b* is the Condorcet winner, beating both *a* and *c* by the same margin of 0.55 to 0.45. This example highlights a fundamental difference in philosophy between STV and Condorcet-based rules. Loosely speaking, STV tries to keep votes near the tops of the ballots. Thus the preferences of the *cba* voters for *b* over *a* will not even be considered under STV until *c* is excluded, which means that in this example they are not considered at all, since *b* is excluded before *c*. In contrast, Condorcet's principle requires that, right from the outset, the preferences of the *cba* voters for *b* over *a* should be given equal weight with the similar preferences of the *bac* voters. However, despite this difference in philosophy, Condorcet and majority are not actually incompatible in single-seat elections: if one wishes, one can use AV (or any other system of one's choice) to select a candidate from the Condorcet top tier. Any such rule clearly satisfies Smith-Condorcet, and hence satisfies both majority and Condorcet, although it is a moot point whether it is really any better than AV on its own. In multi-seat elections, Condorcet is undesirable, in my opinion, because it is incompatible with DPC, as shown by Election 5 above. Here the quota is $0.3\dot{3}$, and so DPC requires that *a* and *b* should be elected, whereas *d* is the Condorcet winner.

5. Local or relative properties: monotonicity

Local or relative properties are concerned with what happens when a profile is changed in some way. We shall say that a candidate is *helped* or *harmed* by a change in the profile if the result is, respectively, to increase or to decrease the probability of that candidate being elected.

As we saw in Election 4, under STV the later preferences on a ballot are not even considered until the fates of all candidates of earlier preference have been decided. Thus a voter can be certain that adding extra preferences to his or her preference listing can neither help nor harm any candidate already listed. Supporters of STV usually regard this as a very important property, although it has to be said that not everyone agrees; the property has been described (by Michael Dummett, in a letter to Robert Newland) as "quite unreasonable", and (by an anonymous referee) as "unpalatable". There are really two properties here, which we can state as follows.

Later-no-help. Adding a later preference to a ballot should not help any candidate already listed.

Later-no-harm. Adding a later preference to a ballot should not harm any candidate already listed.

We come now to the different versions of monotonicity. The basic theme is that a candidate *x* should not be harmed by a change in the profile that appears to give more support to *x*; but one gets different flavours of monotonicity if one specifies different ways in which the profile might be changed.

Monotonicity. A candidate *x* should not be harmed if:

(mono-raise) *x* is raised on some ballots without changing the orders of the other candidates;

(mono-raise-delete) *x* is raised on some ballots and all candidates now below *x* on those ballots are deleted from them;

(mono-raise-random) *x* is raised on some ballots and the positions now below *x* on those ballots are filled (or left vacant) in any way that results in a valid ballot;

(mono-append) *x* is added at the end of some ballots that did not previously contain *x*;

(mono-sub-plump) some ballots that do not have *x* top are replaced by ballots that have *x* top with no second choice;

(mono-sub-top) some ballots that do not have *x* top are replaced by ballots that have *x* top (and are otherwise arbitrary);

(mono-add-plump) further ballots are added that have *x* top with no second choice;

(mono-add-top) further ballots are added that have *x* top (and are otherwise arbitrary);

(mono-remove-bottom) some ballots are removed, all of which have *x* bottom, below all other candidates.

There is also the following property, which is not strictly a form of monotonicity but is very close to it. It is an extension to multi-seat elections of a property proposed by Moulin[3] for single-seat elections.

Participation. The addition of a further ballot should not, for any positive whole number *k*, reduce the probability that at least one candidate is elected out of the first *k* candidates listed on that ballot.

These properties are not all independent. For example,

mono-raise-random implies both mono-raise and mono-raise-delete;

mono-raise and later-no-help together imply mono-raise-delete;

mono-raise-delete and later-no-harm together imply mono-raise-random;

mono-sub-top implies mono-sub-plump;

mono-sub-plump and later-no-harm together imply mono-sub-top;

mono-append and mono-raise-delete together imply mono-sub-plump;

mono-append and mono-raise-random together imply mono-sub-top;

mono-add-top implies mono-add-plump;

mono-add-plump and later-no-harm together imply mono-add-top;

participation implies mono-add-top.

Moreover, in single-seat elections,

participation implies mono-remove-bottom.

Also, if truncated preference listings are not allowed, then mono-raise-random implies mono-sub-top.

```
                    ab     10
Election 6:         bca     8
  (1 seat)          ca      7
```

STV satisfies mono-append but none of the other properties, although in single-seat elections AV satisfies mono-add-plump and mono-add-top. To see that AV does not satisfy mono-raise, mono-raise-delete, mono-raise-random, mono-sub-plump, mono-sub-top or mono-remove-bottom, consider its effect in Election 6 above. As it stands, *c* is excluded and *a* is elected. But if two of the *bca* ballots are removed, or replaced by *a* or by *abc* or by anything else starting with *a*, then *b* is excluded and *c* is elected instead of *a*.

```
Election 7         Election 8
(2 seats)          (2 seats)
ab       30        ac       207
ac       90        bd       198
bd       59        bdac      12
cb       51        cd       105
d        70        dc       105
```

To see that STV does not satisfy mono-add-plump or mono-add-top, consider Election 7. The quota is 300/3 = 100, so that *a* is elected with a surplus of 20. This is divided 5 to *b*, 15 to *c*, and so *b* has 64 votes to *c*'s 66, *b* is excluded, and *d* is elected. Suppose now that we add a further 24 ballots with *d* top. The quota is now 324/3 = 108, so that *a*'s surplus is now only 12. This is divided 3 to *b*, 9 to *c*, and so *b* has 62 votes to *c*'s 60, *c* is excluded, and *b* is elected instead of *d*.

Although all the monotonicity properties look attractive, I do not think that mono-remove-bottom is desirable in multi-seat elections. Consider Election 8. The quota is 627/3 = 209, and so DPC requires that we elect *b* and either *c* or *d*. It seems to me that {*b*, *c*} is clearly the better result (although STV gives {*b*, *d*}). But if we now remove the 12 *bdac* ballots, then the quota drops to 205, so that we must elect *a* and either *c* or *d*. It seems to me that now {*a*, *d*} is the better result (although STV gives {*a*, *c*}). Thus the removal of the 12 ballots that have *c* bottom *should*, in my opinion, harm *c*.

All the monotonicity properties seem desirable in single-seat elections. However, I proved[7] that no rule simultaneously satisfies mono-sub-plump, later-no-help, later-no-harm, majority and plurality. Since I do not think anyone would seriously consider a rule that did not satisfy both majority and plurality, this shows that in order to have mono-sub-plump one must sacrifice either later-no-help or later-no-harm (or both). Whether or not this is desirable may depend on what other properties one can gain at the same time.

Mono-raise-random, mono-sub-top and participation are very strong properties, and it is possible that they are incompatible with DPC. If one could find a reasonable-looking "STV-lookalike" rule that satisfied all the other monotonicity properties (except for mono-remove-bottom when there is more than one seat), then I personally might well prefer it to STV itself. But we are a long way from finding such a rule at the moment.

While on the subject of monotonicity, I should mention one other monotonicity property, if only to dismiss it immediately.

House-monotonicity. No candidate should be harmed by an increase in the number of seats to be filled, with no change to the profile.

This seems to me to be plain wrong. Consider the profile in Election 5, for example, which is a very slight modification of one suggested to me by David Hill. If one were using this profile to fill a single seat, then clearly *d* should be elected (although that is not the result achieved by AV). But if this same profile were used to fill three seats, then clearly *a*, *b* and *c* should be elected; thus *d* is harmed by the increase in the number of seats.

Another property that is related to monotonicity is known in the literature as *consistency*[8] or *reinforcement*[3], but I prefer to call it by its mathematical name:

Convexity. If the voters are divided into two districts and the ballots from each district are processed separately and the results in the two districts are the same, then processing the ballots of all voters together should give the same result.

		(a)	(b)	(a)+(b)
	ab	6	3	9
Election 9:	bc	4	4	8
(1 seat)	cb	3	6	9

STV does not satisfy convexity. Again, I cannot do better than to quote an example of David Hill's (Election 9). In district (a), *c* is excluded and *b* is elected. In district (b), *a* is excluded and *b* is elected. But when the ballots from the two districts are processed together, *b* is excluded and *c* is elected.

Convexity is one of the best-understood of all properties. Young[8] proved that a symmetric preferential election rule for single-seat elections satisfies convexity if and only if it is equivalent to a point scoring rule (in which one gives each candidate so many points for every voter who puts them first, so many for every voter who puts them second, and so on, and elects the candidate with the largest number of points). Since no point scoring rule can possibly satisfy DPC, it follows that convexity and DPC are mutually incompatible. This is a pity, because convexity implies several of the monotonicity properties; but, sadly, it is of no use to us.

Of course, the absence of convexity will hardly ever be noticed in practice, since elections are not counted both in separate districts and together as a whole. But it is worrying inasmuch as it may suggest that something odd is going on.

6. Further properties

A question that is sometimes asked about STV is, is a truncated preference listing treated as if all the remaining candidates were placed equal last? Since STV (in its usual formulation) does not allow for equality of preference, the question does not really make sense. But one can make sense of it as follows. The *symmetric completion* of a truncated preference listing is obtained by taking all possible completions of it with equal weight, chosen so that the total weight is 1. For example, suppose that there are five candidates, *a, b, c, d, e*. Then

the symmetric completion of a ballot marked *abcd* is a single ballot marked *abcde*, with weight 1;

the symmetric completion of a ballot marked *abc* consists of two ballots, each with weight ½, one marked *abcde* and the other marked *abced*;

the symmetric completion of a ballot marked *ab* consists of six ballots, each with weight 1/6, completed in the six different possible ways: that is, *abcde, abced, abdce, abdec, abecd* and *abedc*;

the symmetric completion of a ballot marked *a* consists of 24 ballots, each with weight 1/24, completed in the 24 different possible ways; and so on.

Symmetric-completion. A truncated preference listing should be treated in the same way as its symmetric completion.

It is not difficult to see that AV satisfies symmetric-completion. Although AV is usually described in terms of a quota, it can alternatively be described as follows: repeatedly exclude the candidate with the smallest number of votes, until there is only one candidate left. The effect of replacing truncated preference listings by their symmetric completions is simply that, at each stage in the count, the votes of all non-excluded candidates are increased by the same amount. It follows that the order of exclusions is not affected, nor therefore is the eventual winner.

```
                a       60
Election 10:    ab      60
(2 seats)       b       14
                c       46
```

To see that STV does not satisfy symmetric-completion in general, consider Election 10. The quota is 180/3 = 60, so that *a* is elected with a surplus of 60. Under the Newland-Britton rules, the whole of *a*'s surplus goes to *b*, who is elected. Under Meek's method, the transfer of *a*'s surplus ends with the quota reduced to (180 − 36)/3 = 48, with 36 non-transferable votes going to ‘excess’, and 36 votes transferred to *b*. Either way, *a* and *b* are elected. However, if each ballot is replaced by its symmetric completion, then, of *a*'s surplus of 60 votes, 45 go to *b* and 15 to *c*, and *c* is elected instead of *b*.

```
Election 11     Election 12
(2 seat)        (3 seats)
ab      40      ab      40
ba       2      ba       2
cd      12      cd      12
dc       6      dc       6
                e      180
```

David Hill has sent me an example, which I have modified slightly above, to show that quota reduction is preferable to symmetric completion in STV. In Election 11 the quota is 60/3 = 20, and so *a* and *b* are elected. In Election 12 the quota is 240/4 = 60, so that *e* is elected with a surplus of 120. Under symmetric completion, this would be used to increase the votes of the remaining candidates by 30 each, so that *a* would be elected first, after which *d* would be excluded and *c* would be elected. However, if the quota is reduced to 20 after the election of *e* then *a* and *b* are elected as in Election 11. To paraphrase David's comments slightly, "Election 12 has one extra candidate, one extra seat, and a large number of extra voters whose sole wish (apparently) is to put that extra candidate into that extra seat. It is nonsense that the original 60 voters should get *a* and *c* elected in Election 12 instead of the *a* and *b* they would have got from Election 11."

The remaining properties are all concerned with the avoidance of "wrecking candidates". A "wrecking candidate" is a candidate who is not elected but who, by standing for election and so "splitting the vote", prevents someone else from being elected. One might naïvely hope to avoid wrecking candidates altogether, which would result in the Independence of Irrelevant Alternatives, or IIA:

IIA. If a candidate *x* is not elected, then the result of the election should be as if *x* had not stood for election.

However, it is easy to see that no discriminating election rule can satisfy both IIA and majority. For, consider Election 3 above. By the axiom of discrimination, there must be a profile arbitrarily close to this one that does not give rise to a tie. If this profile results in the election of *a*, then *b* is a wrecking candidate: for, if *b* had not stood for election, then *c* would have been elected (by majority, since roughly two thirds of the voters prefer *c* to *a*); thus the result of the election is not as if *b* had not stood. In a similar way, if *b* is elected then *c* is a wrecking candidate, and if *c* is elected then *a* is a wrecking candidate.

In an attempt to find a property weaker than IIA but expressing a similar idea, I came up with the following.

Weak-IIA. If *x* is elected, and one adds a new candidate *y* ahead of *x* on some of the ballots on which *x* was first preference (and nowhere else), then either *x* or *y* should be elected.

Unfortunately I do not know of any sensible election rule that satisfies even this. Certainly STV does not. For example, if there are 15 ballots marked *x* and 14 marked *z*, then AV (and any sensible rule) will elect *x*; but if 10 of the 15 *x* ballots are now changed to read *yx*, then AV will exclude *x* and elect *z* instead.

An alternative weakening of IIA has been proposed by Tideman[5]. In his terminology, a number of candidates form a *set of clones* if every preference listing that contains one of them contains all of them, in consecutive positions (but not necessarily always in the same order). He says that a single-seat election rule is *independent of clones* if it satisfies the following properties, which I have reformulated here so that they make sense for multi-seat elections as well.

Clone-in. The expected number of candidates elected from any given set of clones should not increase if one member of the set is deleted from every ballot containing it.

Clone-no-help. Replacing a candidate *x* by a set of clones should not help any other candidate *y*.

Clone-no-harm. Replacing a candidate *x* by a set of clones should not harm any other candidate *y*.

```
                xx'a    13
                x'xa    11
Election 13:    abc     10
(2 seats)       bc      12
                c       14
```

It is not difficult to see that AV satisfies all the clone properties. I am fairly sure that STV also satisfies clone-in in multi-seat elections, although I do not have a formal proof of this. To see that STV does not satisfy the other two clone properties, consider Election 13. The quota is 60/3 = 20. Nobody having reached the quota, *a* is excluded and *b* is elected; then *x*' is excluded and *x* is elected. However, if the clones *x* and *x*' are replaced by a single candidate *x*, then *x* has 24 votes initially and so is elected, and the surplus of 4 votes goes to *a*; therefore *b* is excluded, and *c* is elected instead of *b*. So replacing *x* by a pair of clones helps *b* and harms *c*.

Clone-no-harm is actually incompatible with DPC. To see this, note that if only two candidates stand in a 2-seat election, where the voting is (say) *x* 70, *y* 30, then both must be elected. But if *x* is replaced by a pair of clones and the voting is now *xx*' 35, *x'x* 35, *y* 30, then DPC requires that *x* and *x*' should both be elected. This suggests that clone-no-harm is not a desirable property for multi-seat elections—and Tideman never suggested that it was. But clone-in and clone-no-help both look sensible to me, even for multi-seat elections.

References

1. Marquis de Condorcet, *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*, Paris, 1785.

2. P. C. Fishburn, Condorcet social choice functions, *SIAM Journal on Applied Mathematics* 33 (1977), 469-489.

3. H. Moulin, Condorcet's principle implies the no show paradox, *Journal of Economic Theory* 45 (1988), 53-64.

4. J. H. Smith, Aggregation of preferences with variable electorate, *Econometrica* 41 (1973), 1027-1041.

5. T. N. Tideman, Independence of clones as a criterion for voting rules, *Social Choice and Welfare* 4 (1987), 185-206.

6. B. A. Wichmann, Two STV Elections, *Voting matters* (The Electoral Reform Society) 2 (1994), 7-9.

7. D. R. Woodall, An impossibility theorem for electoral systems, *Discrete Mathematics* 66 (1987), 209-211.

8. H. P. Young, Social choice scoring functions, *SIAM Journal on Applied Mathematics* 28 (1975), 824-838.

Issue 4, August 1995

Editorial

Readers will have noticed that there has been a significant delay in the appearance of this issue. The reason is very simple — a lack of sufficient material. Also, in reading this issue, you will see many familiar names amongst the authors. The conclusion is that we need a wider base for the authorship than we have currently. Hence could I ask all readers to ensure that friends with similar interests subscribe to *Voting matters*?

In the last paper in this issue, Douglas Woodall uses barycentric coordinates to present the analysis of election results with three candidates. Unfortunately, this elegant method of presentation is regarded by the media as too complex for general use. In consequence, in the recent three-way by-election, the comparison between the previous general election and the by-election was hard to understand. Perhaps this is an advantage to the three party managers who could all claim a ‘victory’.

Brian Wichmann.

Progressive Elimination

P Dean

Peter Dean is a Trustee of ERS Ballot Services

In my previous article [Issue 3, page 6] I took the Solent mock election of 1989 to show that electing 5 candidates from a field of 20 gave a different result to choosing 5 from the last 6.

It occurred to me that a computer used in a progressive elimination (19 from 20, then 18 from 19 and so on) could give a different result. Dr Hill proved this to be the case, though he did not favour this method.

Whereas all systems elected candidate Nos 1, 7, 9 and 18; the normal manual method elected No 2, but electing 5 from the last 6 preferred No 20. The progressive elimination finally elected No 19 with No 14 as the runner-up. An examination of the first 5 preferences on each ballot paper revealed that No 19 came 2nd (60), No 20 - 6th (45), No 14 - 7th (37), and No 2 - 8th (34).

This demonstrates that a candidate with considerable secondary support can easily lose out in such an election. No 19 was originally 9th to be eliminated, and No 14 was 13th to go out, being less than a vote behind his running mate - No 2.

Taking only the top 8 based on the first 5 preferences produced that same result as the progressive elimination process.

Meek and monotonicity

I D Hill

David Hill is Chairman of the ERS Technical Committee

In *Voting matters* issue 3, B A Wichmann reported that, using data sampled from real voting patterns, ‘Meek violates mono-raise much more than ERS’. Is this something that Meek supporters should worry about?

We know: (1) that all electoral systems have to suffer from some anomaly or other; (2) that STV's anomaly is that it can fail on monotonicity i.e. a change of vote in a candidate's favour can cause that candidate's defeat; (3) that traditional rules do not even look at a voter's second or subsequent preferences if the first preference is elected later than the first count. So the way to make Meek run into an anomaly where traditional rules do not is to find a case where monotonicity trouble occurs among the preferences that such rules ignore.

Although the numbers of such violations reported are indeed considerably greater for Meek, it should be remembered that these arise from examining many thousands of pseudo-elections, and the proportions of occasions are small. For example, the greatest number of Meek violations found was 141 from a data set called R038, but that number comes from 12421 comparisons of one pseudo-election with another. Furthermore each of these pseudo-elections has only 20 voters, which is very few for electing to 5 places from 17 candidates. So the degree of trouble should not be exaggerated, but nevertheless 141 Meek violations were found and no ERS violations in comparisons derived from that particular data set.

It should be borne in mind that the method used to form these pseudo-elections from any given data set involved sampling each time from the same set of votes and thus there are many repetitions, of particular votes being used more than once. This makes it difficult to judge what the results would be from truly independent samples.

I have examined one case in detail to see what it shows and, to avoid all bias in choosing which case to examine, I decided to take the first one found in the data sets available to me. This involved 14 candidates (A − N) for 7 seats, and contained among its votes one for EJICDNG in that order of preference. Those elected are EFGHIJN by Meek rules, but if that one vote is changed to read EIJCDNG (all other votes being unchanged) which should be to I's advantage, those elected become CEFGHJN, and I has lost the seat to C.

The current ERS rules elect EFGHJCN with 25% probability, EFGHJAI with 58% probability and EFGHJCI with 17% probability, depending upon how two random choices come out. That they reach the same result, given the same random choices, irrespective of whether the one vote is as in the original data set or changed, is inevitable because the only vote changed is from EJICDNG to EIJCDNG. At the first count E has 3 votes where the quota is 3.13 and so is not yet elected. At the second count 2 votes starting GE are transferred to E each at value 0.55, to give E a total of 4.10 and a surplus of 0.97, but that surplus is redistributed solely from the 2 newly-received votes. Whether J or I comes next in the vote that is changed is never even looked at.

Using Meek rules with either set of votes GEFHJ are elected and BDKLM are excluded. At that point with the original votes A has 2.145 while C has 2.100, C is excluded, N and I elected and A left as runner-up. With the modified votes, A has 2.053 while C has 2.060, so A is excluded and nearly all A's votes pass to C. This results in C and N elected, I as runner-up. Either way it is a very close-run thing, but who is ahead, of A and C, happens to reverse and the result unfortunately causes the observed lack of monotonicity.

Should all this worry Meek supporters? I think no more than the fact that lack of monotonicity is an upsetting feature of all STV. We could get rid of that feature by abandoning STV altogether and refusing ever to look at preferences beyond the first, but we know that what is lost by so doing far exceeds what is gained. Similarly if we do not look at later preferences some of the time (even when they are relevant) then we can get rid of the feature some of the time, but again, what is lost by doing that far exceeds what is gained. In general, looking at voters' later preferences whenever they are relevant helps to meet those voters' wishes; that it is occasionally troublesome is a pity but cannot be helped. It remains true that the voter concerned could not possibly anticipate such an effect, so it cannot lead to tactical voting, and also that even if such votes were to arise in reality, the lack of monotonicity would never be noticed except by detailed research of the ballot papers such as is hardly ever performed.

In case anyone wishes to examine this data set further, here are the original votes in Wichmann-Hill format. For those not used to this:

14 7 means 14 candidates for 7 seats;

1 5 10 9 3 4 14 7 0 means a vote for candidates 5 10 9 3 4 14 7 in that order, the initial 1 meaning 1 vote and the 0 terminating it, and so on;

Following all the votes there is an extra 0 to terminate them all and then the names of candidates in the order of their reference numbers, and a title for the election.

To get the modified votes, change the first one to start 1 5 9 10 instead of 1 5 10 9, and change the title on the last line.

```
14 7
1 5 10 9 3 4 14 7 0
1 3 5 13 12 7 1 4 8 0
1 8 7 10 12 13 3 6 4 14 11 9 1 2 5 0
1 5 11 14 7 9 0
1 6 7 10 11 12 3 0
1 8 7 5 13 12 14 6 3 1 2 0
1 6 7 10 12 0
1 7 9 5 8 10 14 3 4 1 2 6 11 12 13 0
1 10 7 12 5 8 3 6 9 14 0
1 7 5 11 6 0
1 1 12 3 14 8 6 13 5 0
1 7 5 12 10 14 4 3 9 6 0
1 9 0
1 7 6 10 12 9 14 0
1 1 12 3 8 14 6 5 13 0
1 10 1 12 8 6 3 9 0
1 8 5 12 3 9 1 7 13 10 11 4 6 0
1 3 4 7 10 0
1 7 10 8 12 3 4 9 14 1 13 2 6 11 5 0
1 14 11 5 10 0
1 14 13 2 1 3 9 12 4 5 8 0
1 7 8 9 5 6 0
1 7 12 4 9 8 14 3 11 0
1 5 14 7 0
1 6 7 10 12 0
0
"A""B""C""D""E""F""G"
"H""I""J""K""L""M""N"
"Original"
```

Trying to find a winning set of candidates

I D Hill

In *Voting matters* issue 2, I introduced the idea of Sequential STV and came to the conclusion that it should not be recommended for general use. But there remains something very attractive in trying to find a set of candidates, of the right size for the number to be elected, such that if an STV election were conducted with that set plus any other one candidate, all other candidates being treated as withdrawn, that set would always be the winners.

We know from Condorcet's paradox that in the one-seat case, where the set is of size 1, there may not be any winner who fulfils the criterion, but at least if we can find such a winner, the result is unique.

In the multi-seat case, we can still get results where no set satisfies the criterion. For an example, consider 4 candidates for 2 seats and votes 1 AB, 1 BC, 1 CD, 1 DA. If we choose AB to test we find that ABD leads to AD as winners; testing AD we find that ACD leads to CD; testing CD we find BCD leads to BC; testing BC we find that ABC leads to AB. So round in circles we go.

But now things are far worse for, even where a set to satisfy the criterion is found, it may not be unique. Again consider 4 candidates for 2 seats and votes 6 A, 6 B, 5 C, 5 D, 4 DA, 4 DB, 4 CA, 4 CB, 4 BC, 4 BD, 4 AC, 4 AD. If we choose AB as potential winners, we find that ABC elects AB and ABD elects AB, which would seem to confirm the choice; but if we choose CD we find that ACD elects CD and BCD elects CD, so that choice is also confirmed. Looking at the votes we can see that AB is, in fact, the better choice, but merely to find any set that fulfils the criterion is not adequate.

Can we then say that, having found a potential winning set, we need only look at disjoint sets to see if there are any others? Again things are not as easy as that. Consider 6 candidates for 4 seats, with the same votes as in the last example, with the addition of 20 E, 20 F. Then if we choose ABEF as potential winners, we find that ABCEF elects ABEF and ABDEF elects ABEF, seeming to confirm the choice; but if we choose CDEF we find that ACDEF elects CDEF and BCDEF elects CDEF, so that choice is also confirmed, and the sets are not disjoint as they both contain E and F.

It is clear therefore that there cannot be a universally best algorithm. For everyday practical use, I believe that simple STV by Meek's method should remain the algorithm of choice.

A simple model of voter behaviour

B A Wichmann

Brian Wichmann is the Editor of *Voting matters* and a visiting professor of The Open University

Voting patterns

The additional information provided by preferential voting means that it is difficult to characterise voter behaviour. For instance, one cannot state that a voter supports party A merely because his first preference is for party A. The total information provided in a preferential ballot is very much larger than in X voting, although the result sheet only provides a small fraction of this information.

An obvious question to raise is if the information provided in a ballot can somehow be simplified to provide the essential content. In this paper, a simple model is proposed which appears to provide the essential information from a preferential ballot.

An example

The principle behind the model is most easily understood by means of an example. The model does not depend upon the number of seats to be filled (indeed, should this value alter the voting patterns?).

Hence we consider the case with four candidates: Albert, Bernard, Clare and Diana, with the votes cast as follows:

```
20  AB
15  CDA
 4  ADC
 1  B
```

From this data, we compute the number of each pair of preferences, adding both the starting position and a terminating position. For instance, the number of times the preference for A is followed by B is 20, and the number of times the starting position is ‘followed by’ A is 20+4=24. The complete table is therefore:

	A	B	C	D	e
s	24	1	15	0	-
A	-	20	0	4	15
B	0	-	0	0	21
C	0	0	-	15	4
D	15	0	4	-	0

Obviously, a preference for X cannot be followed by X, resulting in the diagonal of dashes. The entry under `s-e` could represent the invalid votes.

Having now computed this table, we can use it to characterise voting behaviour. For instance, 24 out of 40, or 60% of voters gave A as their first preference. More than this, we can use the table to compute ballot papers having the same statistical properties. For example, if the first preference was A, then the second row of the table shows that the subsequent preference should be B, D or e in the proportions of 20:4:15. Due to the fortunately large number of zeros in the table, we can easily compute the distribution of all the possible ballot papers which can be constructed this way. Putting these in reducing frequency of occurrence we have:

AB	30.8%	(50.0%)
A	23.1%	
CDAB	16.9%	
CDA	12.7%	(37.5%)
C	7.9%	
ADC	6.1%	(10.0%)
B	2.5%	(2.5%)

The figures in brackets are the frequencies from the original data — which can be seen to be quite different.

A number of points arise from this example:

1. The computation of the frequencies needs to take into account the valid preferences. For instance, the frequency of the ballots starting AD is 0.6×4/39 = 6.1%; the next preference can only be C, since the other option in the table is A which is invalid.

2. The large percentage that plump for A is due to the combination of the large percentage having A as the first preference, and the large percentage having A as the last preference, even though plumping for A does not occur in the original ballot. One would not expect this to occur in practice.

3. In this example, the table seems to be larger than the original ballot papers in information content. Exactly the opposite would occur with real elections with hundreds or more ballot papers.

4. Note that the number of occurrences of A in the ballot papers is the sum of the column A and also the sum of row A (which are therefore equal).

5. It is clear that the ballot papers constructed this way do not have the same distribution of the number of preferences as the original data. However, the mean number of preferences is similar, but smaller (2.19 for the computed data, 2.45 for the original). Clearly, when all ballot papers give a complete set of preferences, the computed data will rarely, but sometimes, give plumping.

6. If the voters voted strictly according to sex (A,B or C,D), then this characteristic would be preserved by the model. Similarly, the model does characterise party voting patterns.

The conclusion so far is that the model characterises some aspects of voter behaviour, but does not mirror other aspects. However, from the point of view of preferential voting systems, we need to know if the characterization influences the results obtained by a variety of STV algorithms. The property can be checked by comparing sets of ballot papers constructed by the above process against those produced by random selection of ballot papers from the original data.

We take the ballot papers from a real election which was to select 7 candidates from 14, being election R33 from the STV database. From this data, which consists of 194 ballot papers, we select 100 elections of 25 votes by a) producing random subsets of the actual ballots, or by b) the process described above.

For each of the 200 elections we determine 4 properties as follows:

1. Determine if the Condorcet top tier consists solely of the candidate G. This was a property of the actual election.

2. Determine if the Meek algorithm elects candidate C. This was a property of the actual election.

3. Determine if the ERS hand counting rules elects candidate N. This was a property of the actual election.

4. Determine if Tideman's algorithm elects candidate E. This was not a property of the actual election. Unfortunately, computing the result from this algorithm can be very slow, and hence the result was determined for 50 elections rather than the 100 for the other three cases.

The results can be summarised by the following table:

	Subset	Process	Number
Condorcet (G)	75	67	100
Meek (C)	42	34	100
ERS (N)	56	47	100
Tideman (E)	14	20	50

I believe that the four properties above are sufficiently independent, and the elections themselves independent enough to undertake the χ-squared test to see if the two sets of elections could be regarded as having come from the same population. Passing this test would indicate that the statistical construction process is effective in providing ‘election’ data for research purposes.

The statistical testing is best done as a separate 2×2 table test of each line. The first line, for example, gives the table

	Condorcet (G)	Analysis other	
Subset	75	25	100
Process	67	33	100
	142	58	200

The four tables give P = 0.28, 0.31, 0.26 and 0.29 respectively, using a two-tailed test. So, so far as this test goes, these show no significant differences in the two methods.

Acknowledgement

Dr David Hill provided the statistical analysis above.

Monotonicity — An In-Depth Study of One Example

D R Woodall

Douglas Woodall is Reader in Pure Mathematics at Nottingham University.

Here is a fairly typical example of the way in which monotonicity can fail with STV (or, as in this case, AV). Consider the pair of single-seat elections below. In Election 1, no candidate has reached the quota of 15, and so *c*, the candidate with the smallest number of first-preference votes, is excluded. All *c*'s votes are transferred to *a*, and so *a* is elected. However, just before the result is announced, it is discovered that two of the ballots placed in the pile labelled *bca* are not in fact marked *bca* at all, but *abc*, so that the true situation is as in Election 2. Naturally *a* is delighted with this increased support. But now *b* has the smallest number of first-preference votes, and so, when the count is redone, *b* is excluded instead of *c*. All *b*'s votes go to *c*, and so *c* is elected instead of *a*. So the effect of this increased support for *a* is to cause *a* not to be elected.

	Election 1	Election 2
abc	11	13
bca	10	8
cab	9	9
Excluded	*c*	*b*
Elected	*a*	*c*

This is the sort of anomaly that has caused some people to reject the whole idea of STV. The question I want to discuss here is, how serious is it really? Certainly nobody is going to pretend that it is desirable; but is it really as bad as some people have been making out?

The first thing to notice is that nobody has been wrongly elected. One might object that it cannot possibly be the case that *a* is the right person to elect in Election 1 and that *c* is the right person to elect in Election 2, in which *a* clearly has more support. But it does not really make sense to talk about "the right person to elect" in these elections. In Election 1, for example, there are 19 voters who prefer *c* to *a*, and only 11 who prefer *a* to *c*, so that *c* seems a better candidate to elect than *a*. But then there are 21 voters who prefer *b* to *c*, and only 9 who prefer *c* to *b*, and so *b* seems a better candidate to elect than *c*. But then again, there are 20 voters who prefer *a* to *b*, and only 10 who prefer *b* to *a*, and so *a* seems a better candidate to elect than *b*. Whichever candidate you choose to elect, someone else can claim to be better! (Of course, this is just an example of the famous Condorcet paradox.) In this situation one should not talk about which is *the right* candidate to elect, but, rather, about which candidates it would be *permissible* to elect. It seems to me that in either of these elections it would be perfectly permissible to elect any one of the three candidates. In this situation STV really does no more than make a somewhat arbitrary selection from among the permissible candidates. It is certainly unfortunate that it chooses *a* in Election 1 and *c* in Election 2, where *a* clearly has more support; but it is in the nature of such processes that this sort of thing will happen.

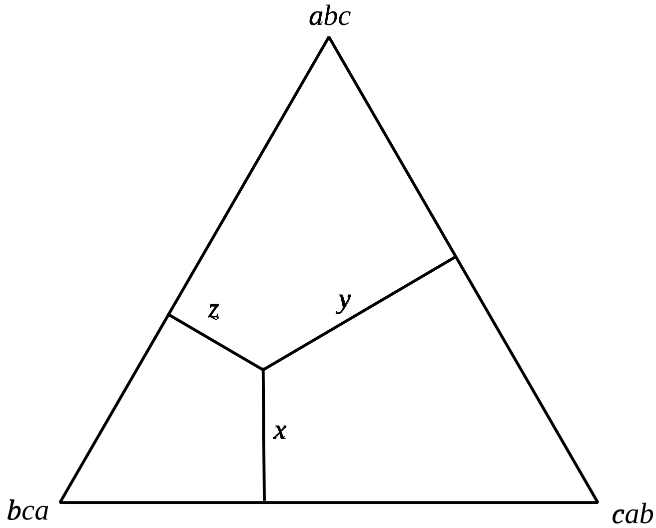


Figure 1

Let us examine more closely what is going on here. Because there are only three different types of ballot present, we can represent the situation diagrammatically, using what are known as *barycentric coordinates* in a triangle. Suppose we draw an equilateral triangle of unit height (Figure 1). If we put a point inside the triangle and drop perpendiculars from it, of lengths x, y and z, to the three sides of the triangle, then it is easy to prove that $x + y + z = 1$, the height of the triangle. So if we label the three corners of the triangle with the three different types of ballot, as in Figure 1, then we can use the point depicted to represent an election in which the proportion of voters voting *abc* is x, the proportion voting *bca* is y, and the proportion voting *cab* is *z*. Thus, for example, the top vertex of the triangle represents an election in which all the voters vote *abc*; the mid-point of the left side represents an election in which half vote *abc* and half vote *bca*; and so on.

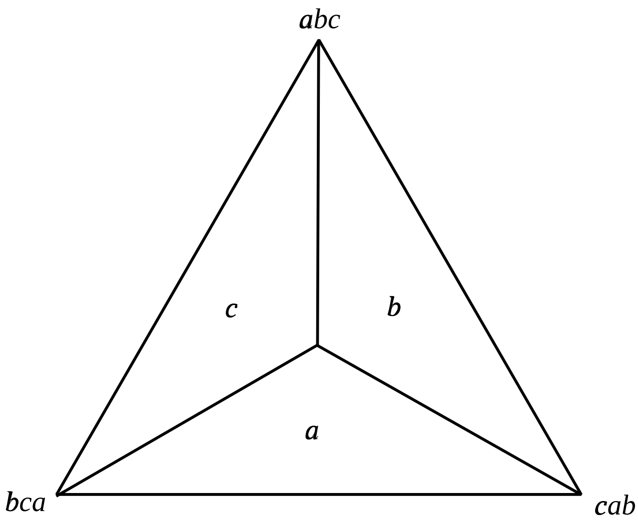


Figure 2: candidate excluded

Suppose now that we exclude the candidate with the smallest number of first-preference votes. Figure 2 shows which candidate is excluded. For example, to the right of the vertical line through *abc* there are more *cab* than *bca* ballots; and above the middle of the three lines through *cab* there are more *abc* than *bca* ballots. So in the region marked *b*, there are fewer *bca* ballots than ballots of either of the other two types, which means that *b* has fewest first-preference votes. Similar remarks apply to the other two regions.

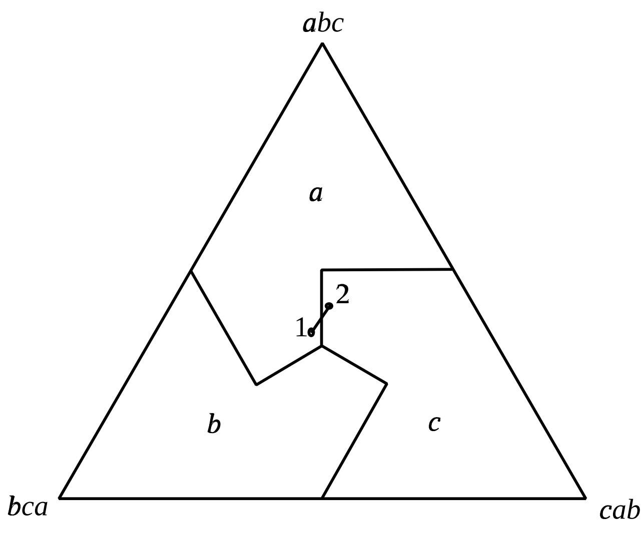


Figure 3: candidate elected

Now consider what happens if *b* is excluded. All of *b*'s votes are transferred to *c*. So the only way that *a* can win is if more than half the ballots are marked *abc*; that is, we are above a horizontal line drawn half way up the triangle. (Of course, in this case *a* will be elected outright — one would not normally exclude *b* first; but it would make no difference to the outcome if one did.) Similar remarks apply to the other two regions, and so the result of the election is as indicated in Figure 3. Figure 3 also shows the points representing Elections 1 and 2. Election 1 is in the region where *a* is elected. Election 2 is obtained from it by converting two *bca* ballots into *abc*, hence by moving parallel to the left edge of the triangle. This takes us into the region in which *c* is elected. Of course, if one continues a bit further in the same direction, then one gets back into the region in which *a* is elected.

The problem is caused, in a sense, by the fact that the regions are not convex. However, one cannot make them convex without violating the spirit of STV. Their convexity is equivalent to the property called Convexity in Woodall[1]; and, as mentioned there, the only election rules that possess this property are point-scoring systems, which do not conform to the spirit of STV.

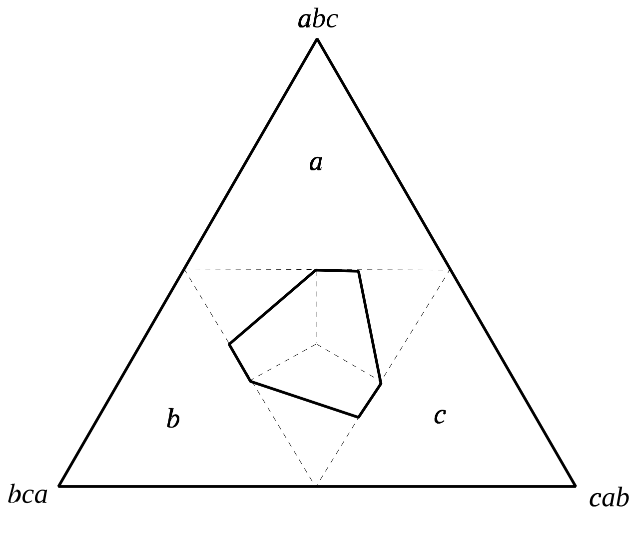


Figure 4: where monotonicity fails

This representation also gives us a way of visualizing where monotonicity fails. If there are two elections (involving only these three types of ballot) that between them show this type of failure of monotonicity, then both elections must lie inside the central region indicated in Figure 4. Note that this region is completely contained within the large dotted triangle, which is where the Condorcet paradox arises. So, in this example, monotonicity does not fail except when there is a Condorcet paradox. However, it is important to stress that, in general, monotonicity can fail even when there is no Condorcet paradox.

Figure 4 suggests the following interpretation. There are certain regions in which it is quite clear who ought to be elected, and in these regions STV elects the candidate that one would expect. But in the middle there is a grey area, where it is not at all clear who ought to be elected, and it is in this grey area that STV behaves in a somewhat haphazard manner; it is really doing no more than making a pseudo-random selection from the appropriate candidates, and it is here that small changes in the profile of ballots can cause perverse changes in the result.

The effect of this is to blur the result of an STV election. Nobody is being wrongly elected, because the problem only arises in the region where one cannot say for certain who ought to be elected anyway. And there is no systematic bias that would, for example, favour one political party rather than another. But the accuracy with which the person or persons elected in an STV election can be said to represent the views of the voters is less precise than it would be if this sort of anomaly did not arise.

The obvious question at this point is whether one can find a system that retains the essential features of STV while avoiding this sort of anomaly. The answer depends on what one regards as the essential features of STV. As we shall see in a later article, it is not possible to avoid this anomaly without sacrificing at least one property that many supporters of STV regard as essential. Nevertheless, I shall describe there a system for single-seat elections that gains so many forms of monotonicity, while sacrificing only one property of STV, that I personally would be willing to recommend it as a better system than the Alternative Vote. Unfortunately, it is not feasible when the votes are to be counted by hand. Also, it is not clear whether it can be extended in any sensible way to multi-seat elections; this is a crucial question, which I have so far been unable to answer.

Reference

1. D R Woodall, Properties of preferential election rules, *Voting matters* Issue 3 (1994), 8-15.

Issue 5, January 1996

Editorial

In this issue, two long and one short article appear which I hope will be of substantial interest to readers. In the first, Crispin Allard produces some estimates of the likely rate of non-monotonicity, based upon a mock election. Secondly, Hugh Warren gives an interesting example of the Condorcet paradox which can only serve to show the inherent complexity of preferential voting. Lastly, I report on a program which attempts to produce plausible election data from STV result sheets.

Estimating the Probability of Monotonicity Failure in a UK General Election

Dr C Allard

Crispin Allard holds a PhD in statistics from the University of Warwick, and is a member of the ERS Council.

1. Summary

Three years ago, the Plant Report rejected STV as a system worth considering for elections to the House of Commons, citing evidence submitted by Michael Dummett (based on an example originating from Reference 2) on the grounds that it could be non-monotonic. In this paper I attempt to estimate the probability of a monotonicity failure which affects the number of seats won by a party. I estimate the probability of this occurring in a multi-member constituency in one election as: 2.5×10^{-4}, equivalent to less than once every century across the whole UK. [*This result was first reported in Reference 1 as 2.8×10^{-4}. I have revised this down as a result of a refinement in the method.*]

2. Representing the problem

Consider an n-member STV constituency, in which n-1 candidates have so far been elected, and the three remaining candidates (denoted A, B and C), one each from the Conservative, Labour and Liberal Democrat Parties are competing for the final place. The conditions for monotonicity failure are as follows:

1. A is ahead of B, and B is ahead of C;
2. When C is eliminated, his transfers put B ahead of A, so that B is elected;
3. If a number of voters switch their relevant preference from A to C, so that both A and C are ahead of B, then when B is eliminated, A is ahead of C, so that A is elected;

for any ordering of A, B and C.

Writing these conditions down in mathematical terms we get:

1. $a > b > c$.
2. $a < b + \alpha c$.
3. There exists x such that: $a - x > b$, $c + x > b$, $a > c + 2x + \beta b$

where

a = the proportion of votes credited to A

b = the proportion of votes credited to B

c = the proportion of votes credited to C

$\alpha = T_{CB} - T_{CA}$

$\beta = T_{BC} - T_{BA}$

T_{ij} = the proportion of i's votes which transfer to j if i is eliminated.

(α and β can be considered as the level of advantage which one party can expect to gain over another as a result of the exclusion of a candidate from a third party).

The following conditions are equivalent to 1-3 above:

M1. $b > c$

M2. $a < b + \alpha c$.

M3. $a > \max\{2b - c,\ (2 + \beta)b - c\}$

Using barycentric coordinates (and denoting each point of the triangle to represent one candidate having all the votes), these conditions are illustrated in figure 1.

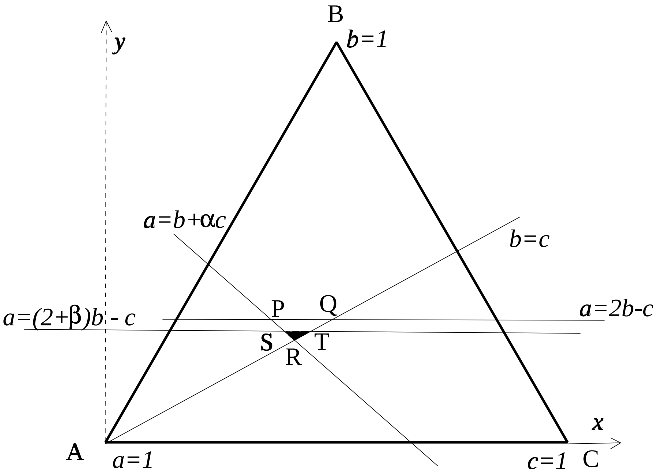


Figure 1

Thus, if we assume a uniform distribution, the probability of this type of monotonicity failure is the ratio of the area of the small triangle (either PQR or STR, whichever is the smaller) to the area of the large one (ABC). To see why we must take the smaller triangle, note that to satisfy condition M3, a point in Figure 1 must be below both the lines:

$a = (2 + \beta)b - c$ and $a = 2b - c$.

Note that if $\beta > \alpha$, conditions M1-M3 cannot simultaneously be satisfied, so in this case we define: Area (STR) = 0.

Switching to Cartesian coordinates,

$x = c + b/2$

$y = \sqrt{3}\ b/2$

the areas of the three triangles are found to be:

$$\text{Area(ABC)} = \sqrt{3}/4$$

$$\text{Area(PQR)} = \frac{\alpha^2}{12\sqrt{3}(3+\alpha)(1+\alpha)}$$

$$\text{Area(STR)} = \frac{\sqrt{3}(\alpha-\beta)^2}{4(3+\beta)^2(3+\alpha)(1+\alpha)} \quad \text{if } \alpha > \beta$$

$= 0$ otherwise

So if we let p be the probability of monotonicity failure, we can find its value as follows:

$$\alpha > 0 \geq \beta \Rightarrow p = \frac{\alpha^2}{9(3+\alpha)(1+\alpha)}$$

$$\alpha > \beta > 0 \Rightarrow p = \frac{(\alpha-\beta)^2}{(3+\beta)^2(3+\alpha)(1+\alpha)}$$

else p=0

Or, by substituting,

$\gamma = \max\{\alpha, \beta, 0\}$

$\delta = \max\{\min\{\alpha, \beta\}, 0\}$

we obtain a single equation for p:

(P1) $$p = \frac{(\gamma-\delta)^2}{(3+\delta)^2(3+\gamma)(1+\gamma)}$$

3. Estimating the transfer patterns

Clearly we need to know the likely pattern of transfers between candidates from different parties, which requires access to the ballot papers of a typical British electorate voting by STV for real political parties. Last year an ERS/ MORI exit poll of 3,983 London voters was conducted during the European Parliament elections, in which they were asked to cast preferential votes in two multi-member constituencies. The results form by far the best available data on the likely behaviour of British voters in an election conducted by STV.

Details of the poll may be found in Reference 3, which includes tables of terminal transfers (transfers of votes from a candidate whose party has no further candidates left who are still eligible to receive votes). Unfortunately, there is no terminal data from Conservative candidates, since none occurred in the count of the mock vote, so this data cannot be used.

Instead I try to consider all the possible transfers of votes which could have taken place. For each of the two constituencies (London North and London South), and for every ordered triple of candidates (Conservative, Labour, Lib Dem), the following data extracted from the poll results is used.

The number of votes which would transfer to the Labour candidate (if the Conservative were to be eliminated leaving only the Labour and Lib Dem candidates); the number which would transfer to the Lib Dem candidate in such circumstances; and the number which would be non-transferable.

This data is repeated for the each of the Labour and Lib Dem candidates being eliminated, providing 840 data sets (sadly not independent!) on which to base the estimate of transfer patterns, and hence estimate p. The number of data sets arises from 216 ordered triples in London North (6-seater), 64 in London South (4-seater), and three data sets for each ordered triple.

4. Method

In outline, I employ the following method (using an Excel spreadsheet):

i) For each data set (representing the potential transfers from one candidate from one party to two candidates from the other parties), the proportions T_{ij} of votes transferred to each of the surviving candidates are calculated.

ii) These proportions are then adjusted using the following approximate shrinkage equation:

$$T'_{ij} = \frac{\bar{\bar{T}}_{ij}/t + nT_{ij}/s}{1/t + n/s}$$

where:

T'_{ij} represents the shrunken estimate of the proportion of i's votes which transfer to j if i is eliminated.

$\bar{\bar{T}}_{ij}$ is the weighted sample mean of T_{ij} based on exclusions of candidates from the same party in a particular constituency.

s is the sample variance of T_{ij}.

n is the size of the data set (the number of first preferences credited to the excluded candidate).

$t = 0.0004$

Note that this is based on a two-stage hierarchical model, in which (for a given constituency and party) there is a party mean value of T_{ij}, with variance 0.0004, about which the candidates' T_{ij} values are distributed.

iii) Based on the values of T'_{ij}, γ, δ and p are calculated, using the above definitions and equation P1.

iv) For each ordered triple of candidates, the three values of p (one for each potential elimination) are summed to allow for all the possible ways in which monotonicity might fail, giving a total probability P.

v) For each constituency, a weighted mean of the probabilities is calculated.

vi) Finally a weighted mean of the probabilities for the two constituencies is taken to produce the result:

$$\varepsilon(P) = 2.5 \times 10^{-4}.$$

So, if the UK is divided into 138 multi-member constituencies, as proposed in Reference 4, and assuming an average of one General Election every four years, we would expect one instance of final-stage monotonicity failure affecting party standing under STV roughly every 115 years.

5. Justifying the approach

The problem of estimating the probability of monotonicity failure under STV is complicated, involving political considerations and statistical judgement as well as pure mathematics. So inevitably I have had to make a number of assumptions and simplifications. I will now attempt to identify all the potential objections to my approach and answer some of the possible criticisms.

5.1 Only monotonicity failure affecting parties is considered.

It is almost certainly true that the probability of affecting individual candidates within a party is much greater. For a start, far more voters are prepared to transfer within a party than between parties. This is supported if we look at ERS Council Elections (which are like elections between candidates of the same party since all support electoral reform), where potential instances have been observed.

Nevertheless, given that STV is the only system which even attempts to represent intra-party opinion, any minor ‘imperfection’ in this respect is irrelevant to the choice of an electoral system. And it is certainly the case that most of the opponents of STV are far more concerned with party representation.

Finally, it is a necessary simplification since intra-party transfer patterns are notoriously unpredictable and difficult to model.

5.2 The model only covers three parties and final-stage transfers.

Of course, earlier stages and a greater number of parties allow more opportunities for monotonicity failure. However, I claim that the probability of this making a difference to the final result is tiny compared to the figure I have calculated above.

To see this, consider the diagram in figure 1. The effect is only possible when there are three candidates with very similar votes (Q is the geometric centre). Thus, if there are four candidates competing for the final place, a candidate who ‘benefits’ in the penultimate stage is still very unlikely to benefit in terms of election (and one who ‘loses’ probably would not have been elected anyway).

If there are four candidates competing for two places, with three in danger of elimination, then the fourth may be discounted (as a certainty), and we are back to the original problem. Only in the case where there are four or more candidates all with similar votes might a relevant situation arise; it is reasonable to ignore such *n*th order terms.

5.3 The method assumes a Uniform (prior) distribution of votes between the three parties.

This assumes that the three parties each have the same marginal distribution. In a one-member constituency this is highly unlikely, but in a multi-member constituency the relevant distribution to consider is the remainder, once *n*-1 seats have been ‘allocated’, and the appropriate number of quotas deducted from each party's vote.

Therefore, in order for the assumption to be reasonable, all we need is to have across the country three parties capable of achieving proportions of votes over a range of at least one quota. This would typically be achieved by a party receiving 10% or more of the national (or regional) vote.

A similar principle is at work behind the Wichmann-Hill pseudo-random generator, where the sum of a number of variables is known to tend to normality, but the fractional part of the sum remains rectangular. There is room here for someone to conduct a proper analysis, which I am confident would uphold my assumption.

5.4 The results are based on an opinion poll conducted only in London.

This represents probably the biggest area of doubt about the result and, since this is the best data available so far, there is no way of avoiding it. The STV ballot paper was constructed by listing (nearly) all candidates in each of the Euro-constituencies represented. Since this was an election for MEPs, recognition of most individual candidates must have been relatively low.

However, we can only speculate on how voters would react in a General election conducted by STV, and it is by no means obvious that voting patterns would be substantially different. The same applies to the London factor. While the relative positions of the parties would vary across the country, there is no reason to suppose that the nature of voting patterns would be any different.

5.5 Why has shrinkage been applied in this way?

Shrinkage is one of the results of Bayesian analysis which has been accepted by non-Bayesian statisticians as representing a true effect which does not appear in more traditional models. I have judged that a hierarchical model is relevant to this situation, so we must take account of shrinkage. A reference will be given in the next issue of *Voting matters* to provide an explanation of shrinkage for non-statisticians.[Not produced?]

If the charge is that I have not defined a full Bayesian hierarchical model, with detailed multivariate prior distributions etc., then I plead guilty. This was done deliberately to avoid specifying prior distributions which might obscure the argument. The value of *t* is arbitrary but, I believe, reasonable. A little sensitivity analysis shows that it does not affect the final result by more than a tenth.

5.6 The weightings used in the final calculation do not allow for some votes having a greater effect.

Rather than try to work out what effect the voting patterns might have had in this particular election, I wanted to gain an estimate of overall voting patterns. This means considering both first and last place candidates, since in different constituencies each party will have somewhere between 0 and 5 ‘safe’ seats, so the candidate involved in a three-way battle could be anyone between the first and sixth most popular in that party.

The best way to cope with such uncertainty is to assign equal weightings to each elector.

6. Conclusions

Using the best data available and using reasonable assumptions I have estimated the probability that monotonicity failure would arise in a UK General Election conducted by STV. That probability turns out to be extremely small. In political terms it may as well be zero. Opponents of STV will need to come up with better reasons if they wish to reject it out of hand.

Acknowledgements

I am grateful to Professor Shaun Bowler of the University of California at Riverside for his help in supplying data from the ERS/MORI poll, and to Richard Wainwright and others for their encouragement and interest in this research.

References

1. Crispin Allard, ‘Lack of Monotonicity - Revisited’, *Representation* 33:2 (1995), pp48-50.

2. G. Doron and R. Kronick, *American Journal of Political Science* 21, pp303-311.

3. Shaun Bowler and David M. Farrell, 'A British PR Election: Testing STV with London's Voters', *Representation* 32:120 (1994/5), pp90-3.

4. Robert A. Newland, *Electing the United Kingdom Parliament*, 3rd Edition (ERS, 1992).

Appendix: Summary Statistics

Below is a table showing the transfer trends in North and South London. The transfers are weighted means, expressed as percentages of the respective first preference votes. The advantages (corresponding to α or β) are given after adjusting for shrinkage. See section 4 for a full explanation. Each party is shown with the number of first preference votes cast in the poll for candidates of that party.

Of the 3,983 voters polled, 3,013 expressed a valid first preference for a candidate from one of the three main parties, of whom 1,778 were from North London and 1,235 from South London. The overall probabilities of monotonicity failure were found to be 0.00013 in North London and 0.00043 in South London, giving a (weighted) mean of 0.00025 and a sample variance of 2×10^{-8}.

	%Transfers				Advantage		
Votes	Con	Lab	LD	NT	Mean	High	Low
North London							
Con:512	-	2.4	5.1	92.6			
Advantage(LD-Lab)					2.7	9.7	-9.8
Lab:1049	1.4	-	8.1	90.5			
Advantage(LD-Con)					6.7	22.5	-5.7
LD:217	4.8	11.3	-	83.9			
Advantage(Lab-Con)					6.5	20.6	-8.6
South London							
Con:400	-	3.0	11.6	85.4			
Advantage(LD-Lab)					8.6	23.2	-8.9
Lab:598	1.9	-	13.9	84.2			
Advantage(LD-Con)					12.0	21.2	3.8
LD:237	7.7	12.5	-	79.9			
Advantage(Lab-Con)					4.8	18.3	-13.0

An example showing that Condorcet infringes a precept of preferential voting systems

C H E Warren

Hugh Warren is a member of the ERS Council.

It is one of the precepts of preferential voting systems that a later preference should neither help nor harm an earlier preference. The purpose of this paper is to show that the Condorcet system of preferential voting infringes this precept.

Consider an election for one seat in which there are 3 candidates:

A is a Catholic Conservative White

B is a Protestant Labour White

C is a Catholic Labour Asian

There are 99 voters:

17 want Labour, they prefer a White to an Asian, and they are indifferent as to sect, so they vote BC.

16 want Labour, they prefer an Asian to a White, and they are indifferent as to sect, so they vote CB.

15 want a Catholic, they prefer Labour to Conservative and they are indifferent as to race, so they vote CA.

17 want a Catholic, they prefer Conservative to Labour, and they are indifferent as to race, so they vote AC.

16 want a White, they prefer Conservative to Labour, and they are indifferent as to sect, so they vote AB.

15 want a White, they prefer Labour to Conservative, and they are indifferent as to sect, so they vote BA.

1, whom we shall call Voter X, wants primarily a Conservative, and wants also an Asian and a Protestant, so is undecided whether to vote AC or AB, but settles for AC.

1, whom we shall call Voter Y, wants primarily a Protestant, and wants also a Conservative and an Asian, so is undecided whether to vote BA or BC, but settles for BA.

1, whom we shall call Voter Z, wants primarily an Asian, and wants also a Protestant and a Conservative, so is undecided whether to vote CB or CA, but settles for CB.

Accordingly the votes are as follows:

```
AB 16
AC 18
BA 16
BC 17
CA 15
CB 17
```

The Condorcet method for the election yields the following results:

C beats B by 50-49

A beats C by 50-49

B beats A by 50-49

Accordingly we see that Condorcet produces a paradox.

(Incidentally, the Single Transferable Vote, which amounts to the commonly called Alternative Vote in this case, would 'exclude the lowest', C, and hence would elect B.)

If the paradox is resolved by electing A, then, if instead of voting AC Voter X had voted AB, Candidate B would have beaten Candidate C, and accordingly by the Condorcet method Candidate B would have been elected. Therefore changing the second preference of Voter X from C to B works to the detriment of his first preference A.

If the paradox is resolved by electing B, then, if instead of voting BA Voter Y had voted BC, Candidate C would have beaten Candidate A, and accordingly by the Condorcet method Candidate C would have been elected. Therefore changing the second preference of Voter Y from A to C works to the detriment of his first preference B.

If the paradox is resolved by electing C, then, if instead of voting CB Voter Z had voted CA, Candidate A would have beaten Candidate B, and accordingly by the Condorcet method Candidate A would have been elected. Therefore changing the second preference of Voter Z from B to A works to the detriment of his first preference C.

Therefore, no matter how the paradox is resolved, the precept that a later preference should not harm an earlier preference is infringed.

Producing plausible party election data

B A Wichmann

The STV database lacks any data from public elections which involves political parties[1]. This is hardly surprising due to the legal constraints on public election data. However, from the point of view of election studies, this omission is very unfortunate. Statistical studies of real election data are important, since we know that desirable logical properties cannot be universally satisfied.

For public elections, the only information available is that of the result sheet. Unfortunately, this information is very much less than that contained in the ballot data itself. Only a few preferences expressed by votes are actually exercised in the counting process and therefore can be reconstructed from the result sheet. It is possible to produce minimal ballot papers which will give the same effect as the result sheet, but such ballot data is very unlike the (unknown) ballot data itself. In contrast, we are here attempting to produce ballot data which appears similar to the actual data, so that our constructed data can be used instead of the real data.

In this study, we are using the Irish election data for the years 1969 and 1973, since this is available in a convenient book format which is easy to process, see Knight and Baxter-Moore[3]. The first election in the book, is that for Carlow-Kilkenny. For this, we have:

	Information content
Result	9 bits
Result sheet	800 bits
Election data	800,000 bits

It might therefore appear that we have a hopeless task since the result sheet contains a thousand times less information than that of the (missing) election data.

However, we established[2] that if we can provide a matrix giving the probabilities of A being followed by B (for all candidates A, B), then election data can be constructed which appears to have the statistical properties one would expect, at least as far as the election results are concerned with the usual STV algorithms. Hence if we can produce an estimate for the A-B probabilities, we can construct plausible data.

Taking the result sheets for all the Irish elections for 1969, we can study just the first transfers made. These transfers are not restricted in the potential choice that can be made by the elector, and therefore can provide a basis for the probabilities we wish to estimate. To compare one constituency with another, we label the candidates FF1, FF2,.. for Fianna Fail in order of the first preferences, and similar for Fine Gael (FG1, etc), Labour (LA1, etc) and others (OT1, etc). (Fortunately, this is exactly the order listed in [3]) We only need to consider the three main parties since they account for around 97% of the first preference votes. However, the 'other' candidates must be taken into account with transfers, and hence appear as a notional party.

Table 1

All first transfers,

Irish elections 1969

Consistuency	FF1	FF2	FF3	FF4	FG1	FG2	FG3	FG4	LA1	LA2	LA3	LA4	OT1	OT2	OT3	NT
Carlow-Kilkenny	-411	0	313		43	7	2	1	21	22						0
Cavan	169	171	117		237	294	72		-1255	-495			458			232
Clare	0	64	23		19	42			346	-533						39
Clare-S Galway	15	51	20		82	10	18		348	-561						17
Cork NW	-4815	3828			114	120			128	75			550			0
Cork SE	-3679	3182			152	122			86	138						0
Mid Cork	-1165	490	391		91	41	15		41	96						0
NE Cork	-1159	450	454		69	3	44	4	127	8						0
SW Cork	141	87			1719	784	-3216		395							90
NE Donegal	-1539	1422			47	50			20							0
Dublin C	-935	662	168		20	6	3	4	14	6	4	8	21	18	1	0
Dublin NC	-3254	1743	676	630	48	64	28		0	42	23					0
Dublin NE	-4268	2054	1710		98	48	23	12	0	55	41	24	168	35		0
Dublin NW	89	46	27		57	99	35	21	-2305	535	719	677				0
Dublin SC	10	23	11		13	14	8	3	11	8	10	8	10	19	-149	0
Dublin SE	46	12	25		-1731	1469			132	21			19	7		0
Dublin SW	4	6	5		11	16	4	5	33	22	10	4	26	-154		6
NC Dublin	0	52	8		214	134	175	-688	31	8	19	36				11
SC Dublin	17	14	16		62	25	17		331	330	-830					18
Dun Laoghaire -	0	24	25		-3317	2030	956		102	53	35		59	33		0
NE Galway	19	26	9		-477	203	168		52							0
W Galway	-780	445	189		27	44	10		28	18			19			0
N Kerry	242	403	69		934	304			351				-2425			122
S Kerry	-1583	1243			57	122	18		143							0
Kildare	197	188	74		305	118	193		178	146			-1496			97
Laois-Offaly	0	55	34		-2075	444	688	487	91	68	65		44			0
E Limerick	12	8	7		19	18	5		112	131	-366		50			4
W Limerick	-3358	1098	1695		175	73	93		144	60			20			0
Longford-W	50	10	6		2	29	25	35	107	108	-420		33			15
Louth	18	89			-1048	614	244		63	20						0
E Mayo	74	58	39		226	145	233		-869							94
W Mayo	36	28	46		122	144	21		-445							51
Meath	99	82	49		107	25	32		981	-1408						33
Monaghan	64	30	22		68	76	33		372	-699						34
Roscommon -	28	36	4		-525	197	224		25	8	3					0
Sliogo-Leitrm	11	18	203		158	8	3		29	51	-506					25
N Tipperary	-1533	628	480		102	71			222	14	16					0
S Tipperary	-1942	1208	462		88	40	13		74	38			19			0
Waterford	1071	679	-2118		35	93			156	24						60
Wexford	51	23	21		39	101	13		343	29	24		112	-813		57
Wicklow	-1010	272	544		36	37			80	41						0

Table 2

Transfers from Fianna Fail

Consistuency	FF1	FF2	FF3	FF4	FG1	FG2	FG3	FG4	LA1	LA2	LA3	LA4	OT1	OT2	OT3	NT
Carlow-K *	-411	313			43	7	2	1	21	22						0
Cork NW	-4815	3828			114	120			128	75			550			0
Cork SE	-3679	3182			152	122			86	138						0
Mid Cork	-1165	490	391		91	41	15		41	96						0
NE Cork	-1159	450	454		69	3	44	4	127	8						0
NE Donegal	-1539	1422			47	50			20							0
Dublin C	-935	662	168		20	6	3	4	14	6	4	8	21	18	1	0
Dublin NC	-3254	1743	676	630	48	64	28		0	42	23					0
Dublin NE	-4268	2054	1710		98	48	23	12	0	55	41	24	168	35		0
W Galway	-780	445	189		27	44	10		28	18			19			0
S Kerry	-1583	1243			57	122	18		143							0
W Limerick	-3358	1098	1695		175	73	93		144	60			20			0
N Tipperary	-1533	628	480		102	71			222	14	16					0
S Tipperary	-1942	1208	462		88	40	13		74	38			19			0
Waterford *	-2118	1071	679		35	93			156	24						60
Wicklow	-1010	272	544		36	37			80	41						0

Table 3

Transfers of 1,000 preferences from Fianna Fail

FF1	FF2	FF3	FF4	FG1	FG2	FG3	FG4	LA1	LA2	LA3	LA4	OT1	OT2	OT3	NT
599	222	18		35	28	8	1	38	19	2	1	24	2	0	2

Table 4

Transfers from Fine Gael

FF1	FF2	FF3	FF4	FG1	FG2	FG3	FG4	LA1	LA2	LA3	LA4	OT1	OT2	OT3	NT
19	29	8	5	527	244	51		68	14	9	3	9	3	0	8

Table 5

Transfers from Labour

FF1	FF2	FF3	FF4	FG1	FG2	FG3	FG4	LA1	LA2	LA3	LA4	OT1	OT2	OT3	NT
59	54	54	0	108	86	46	5	296	125	63		51			53

Table 6

Transfers from other parties

FF1	FF2	FF3	FF4	FG1	FG2	FG3	FG4	LA1	LA2	LA3	LA4	OT1	OT2	OT3	NT
100	128	36		258	110	43	2	182	41	9	1	29	6		56

Table 1 gives the first transfers for all* the 1969 Irish elections. The candidates are labelled as above and NT (for Non-Transferable). A blank in the relevant columns indicates no such candidate. Others are listed in the order given in Knight and Baxter-Moore[3].

Table 2 shows the transfer from Fianna Fail alone. The star against the Waterford entry represents a change from the original. In this case alone, the FF transfer was by elimination; but we wish to put under FF1 the candidate from which transfers were made, which implies permuting the columns as shown. Again, the star against the Carlow-Kilkenny entry represents a change from the original. Here, the candidate FF2 already had the quota, and therefore was not eligible for transfers (or rather any such transfer would have been ignored) and hence the transfer to FF3 is regarded as being for FF2, being the next available FF candidate.

The columns can now be added up to see what the average transfers are. (The total transfers are 33,549, but we express this as votes transferred per thousand.) This result is shown in Table 3, where FF1 here represents the first Fianna Fail candidate to which transfers could be made. As expected, this indicates weak cross-party voting and that the most popular person within a party is that based on first-preference votes.

Tables 4, 5 and 6 give the corresponding transfers of 1,000 votes from Fine Gael, Labour and the other parties respectively.

Hence we now have estimates for our A-B probabilities, although these figures are very crude for the following reasons:

1. The tables show large variations between constituencies.

2. Comparing constituencies with different numbers of candidates for each party is dubious.

3. Grouping all other candidates into a notional party is clearly dubious also.

Nevertheless, we now have some estimates that are probably as good as we can get in the circumstances.

The next process is to use the above estimates for providing default transfer probabilities in those cases in which the result sheet does not provide this information.

For each of the Irish elections for 1969, we compute the transfer probabilities that can be found from the result sheet. For the other values, we use our estimates. This then allows for plausible ballot data to be computed by program.

----

* Donegal-Leitrim is excluded since this has the Speaker of the Dail elected unopposed, so comparisons are difficult.

The computer program does need to reduce the ballot data to manageable proportions. For Carlow-Kilkenny in 1969, there were 46,073 ballot papers. If we constructed this number of ballot papers individually by program, we would have a 750K bytes data file — too big to process rapidly. We can reduce the data file to a more manageable size by having piles of identical papers, which all the computer algorithms can handle rapidly. The program uses piles of 500, 100, 50, 10, 5 and 1 paper(s), adjusted so that the correct number of total ballot papers is produced, and the first preference counts are the same as the result sheet. The data file is now reduced to about 11K bytes.

The program also attempts one further adjustment. The ballot papers match the first preferences and the total votes cast exactly, but the match to subsequent transfers is only similar in terms of the proportion of the occurrence of A-B's in the papers. To obtain a better, but not identical fit, the program computes many examples using different seeds for the random number generator, and selects the best example. Determining the fit between a ballot paper set and the result sheet is not straightforward. To undertake the comparison properly would require a computer version of the Irish STV rules which was not available. Instead, the ERS rules were used, which has a number of differences from the Irish version. The most obvious difference is rounding the votes to whole numbers (single ballot papers are transferred), rather than one hundredths; but this makes little difference in this case with over 10,000 votes cast in each election.

To summarise, the program takes as input:

1. The transfers between parties deduced from a set of elections.

2. The result sheet from a specific election from that set, giving the party affiliation of each candidate.

3. Seeds for the random number generator, and a number of trials from which to select the ballot set with the best fit.

From this, the program outputs a set of ballot papers giving a ‘good’ fit to the specified election. Note that by changing the seeds for the random number generator, slightly different sets of ballot papers will be produced.

This program was then used to construct plausible ballot sets for the 1969 and 1973 Irish elections. The elections in 1973 were regarded as distinct from 1969, so that the same process as illustrated above was used to construct another table of transfers per thousand votes between parties.

A summary of the results from analysing the election data appears below. The meaning of the entries in the table are as follows:

D*n* On my home computer, I have nine different STV-like algorithms. Listed here is the number of algorithms giving a different result from the actual Irish election. A result of D0 is not printed.

C*n* A Condorcet ranking is computed from the election data. From this, the lowest-ranked candidate is found who was elected. C*n* is the number of un-elected candidates ranked at least as high as that candidate.

P*n* From the Condorcet ranking, a Condorcet paradox is evident. P*n* indicates the number of candidates involved in the paradox. The plus sign indicates that the paradox involves both elected and un-elected candidates. (Note that a Condorcet paradox involving the 'top' candidate is undoubtedly a problem when electing a single candidate, but not necessarily in other cases.)

IEM Of the nine STV algorithms that were used to analyse the data, two are of special interest: Meek and the ERS hand-counting rules. Of the three when the Irish result is compared, the odd-one-out is noted (by a single letter). (Note that in the single case of Dublin SW for 1969, all three algorithms gave a different result, so there was not an 'odd-one-out'.)

Consistuency	Result 69				Result 73			
Carlow-Kilkenny		C1	P3+		D2		P5	
Cavan		C1			D2		P3	
Clare	D6	C2		M		C1	P4+	
Clare-S Galway		C1			D2	C1		
Cork NW	D3	C3	P4+		D2	C1		
Cork SE	D2				D2	C1		
Mid Cork	D2					C1		
NE Cork	D6	C1		M	D7	C1		M
SW Cork	D1	C2	P3+		D1	C2	P4+	
NE Donegal	D1							
Dublin C		C1						
Dublin NC	D1	C1				C2	P5+	
Dublin NE		C2			D2	C2	P3	
Dublin NW		C1	P5+			C1		
Dublin SC	D1	C4			D8	C1		M
Dublin SE		C1			D8	C1		M
Dublin SW	D9	C3	P6+	ME		C2	P4+	
NC Dublin						C1	P3+	
SC Dublin	D1				D1	C1		
Dun Laoghaire -					D2			
NE Galway						C1		
W Galway					D3	C1		M
N Kerry					D6	C2	P5+	I
S Kerry		C1	P4+			C1		
Kildare		C1						
Laois-Offaly	D6	C1	P4+	M	D3	C1		E
E Limerick		C1				C1		
W Limerick					D1			
Longford-W	D1	C4			D8	C4	P3,3+	M
Louth					D1		P3	
E Mayo		C1	P4+		D1	C1		
W Mayo		C1						
Meath	D1	C1				C1	P3+	
Monaghan	D2	C1			D8	C2		M
Roscommon -							P3	
Sliogo-Leitrm						C1		
N Tipperary		C1			D3	C1	P3+	
S Tipperary		C2						
Waterford	D7	C1	P3+	I	D2	C2	P5+	
Wexford		C3	P5+					
Wicklow	D9	C2		I				

The method of construction implies that it would be unwise to assume that there was an actual Condorcet paradox for South West Cork, since this property is dependent in part upon the data which has been added by statistical means. However, it would be reasonable to suppose that the fraction of elections in Ireland having a Condorcet paradox is about one third, and about a quarter have a paradox involving elected and unelected candidates.

In many cases, the election result is clearly marginal between two candidates, and hence differences between the STV algorithms is not surprising.

Two elections stand out as being very different. For Dublin South West for 1969, all three main algorithms gave a different result. After the top candidate, the next six were in a Condorcet paradox. It seems clear that this seat is a potential example of non-monotonicity. I have been unable to determine if this is so, since I do not know of any computationally feasible way of determining the property. As an exercise for the readers, I have reproduced the result sheet, together with the fit my program produces, to allow others to determine if non-monotonicity occurs. I have been able to simplify the data by reducing the number of piles substantially, and also reduced the number of votes by a factor of ten, but this still does not provide an easy way of determining this vital property. David Hill has commented on this by noting that perhaps the property is not so important if it is impractical to determine its validity for a specific election.

The other unusual result is that for Longford-Westmeath for 1973. This is the only case in which there were two sets of candidates involved in Condorcet paradoxes in one election.

There is only a weak correlation between those elections having C≠0 and those having D≠0. There is some correlation between the C's and P's, which is hardly surprising due to the underlying dependence upon Condorcet. A Condorcet paradox involving both elected and unelected candidates is no guarantee that any of the STV algorithms will produce a different result as can be seen from Dublin North Central for 1973.

All the computer data produced in this study is available from me on request.

Acknowledgement

This work would not have been possible without the excellent work of J Knight and N Baxter-Moore in tabulating and presenting the results of the 1969 and 1973 Irish elections.

References

1 B A Wichmann. An STV Database. *Voting matters*, issue 2, p9.

2 B A Wichmann. A simple model of voter behaviour. *Voting matters*, issue 4, pp3-5.

3 J Knight and N Baxter-Moore. *Republic of Ireland: The General Elections of 1969 and 1973*. The Arthur McDougall Fund. London. 1973.

Appendix

The table below is the Irish result sheet as from Knight and Baxter-Moore, except that additionally the results computed by the program from the plausible data are shown in italics.

The actual event elected FF1, LA1, LA2 and FF2. The ERS rules with the plausible data elected FF1, LA1, LA2 and FG1, while the Meek algorithm with the plausible data elected FF1, LA1, LA2 and FG2.

There is a single Condorcet winner in LA1, but the set of candidates FF1, FF2, FG1, FG2, LA2 and OT1 are in a Condorcet paradox with the plausible data

Candidate	Stage I	Stage II		Stage III		Stage IV		Stage V		Stage VI		Stage VII		Stage VIII		Stage IX		Stage X		Stage XI		Stage XII	
Dowling	5724	4	5728	12	5740	22	5762	15	5777	651	6428	-589	5839		5839		5839		5839		5839		5839
FF1	*5724*		*5724*		*5725*		*5726*		*5726*		*5726*		*6186*		*5839*		*5839*		*5839*		*5839*		*5839*
Lemass	2512	6	2518	11	2529	13	2542	3	2545	771	3256	520	3776	43	3819	28	3847	15	3862	73	3935	772	4707
FF2	*2512*		*2564*		*2564*		*2564*		*2564*		*3487*		*3757*		*3917*		*3983*		*4020*		*4210*		*5147*
Sherwin	1643	5	1648	12	1660	14	1674	5	1679	-1679													
FF3	*1643*		*1643*		*1643*		*1643*		*1643*														
O'Keeffe	1331	11	1342	341	1683	5	1688	242	1930	22	1952	2	1954	21	1975	88	2063	23	2086	-2086			
FG1	*1331*		*1343*		*1800*		*1800*		*2050*		*2050*		*2050*		*2050*		*2100*		*2174*				
McMahon	1203	16	1219	193	1412	18	1430	579	2009	43	2052	8	2060	26	2086	91	2177	22	2199	1539	3738	767	4505
FG2	*1203*		*1220*		*1320*		*1320*		*1933*		*1983*		*2021*		*2021*		*2021*		*2206*		*3689*		*5594*
Lowe	856	4	860	94	954	10	964	-964															
FG3	*856*		*862*		*963*		*963*																
Redmond	759	5	764	-764																			
FG4	*759*		*760*																				
O'Connell	5273	33	5306	38	5344	169	5513	31	5544	61	5605	10	5615	435	6050		6050	-211	5839		5839		5839
LA1	*5273*		*5298*		*5298*		*5509*		*5509*		*5609*		*5609*		*6359*		*6359*		*5839*		*5839*		*5839*
Dunne	5136	22	5158	23	5181	468	5649	20	5669	129	5798	23	5821	1065	6886	-1047	5839		5839		5839		5839
LA2	*5136*		*5149*		*5150*		*5771*		*5771*		*5781*		*5781*		*6459*		*5839*		*5839*		*5839*		*5839*
Butler	1643	10	1653	10	1663	136	1799	11	1810	10	1820	4	1824	-1824									
LA3	*1643*		*1649*		*1649*		*1659*		*1759*		*1809*		*1809*										
Farrell	893	4	897	1	898	-898																	
LA4	*893*		*894*		*894*																		
Corcoran	2066	28	2094	24	2118	29	2147	45	2192	38	2230	22	2252	186	2438	195	2633	90	2723	196	2919	-2919	
OT1	*2066*		*2086*		*2186*		*2186*		*2186*		*2236*		*2274*		*2444*		*2906*		*3128*		*3568*		
McKeown	154	-154																					
OT2	*154*																						
Non transferable		6	6	5	11	14	25	13	38	14	52		52	48	100	645	745	61	806	278	1084	1380	2464
NT			*1*		*1*		*52*		*52*		*52*		*54*		*105*		*147*		*149*		*210*		*936*
Total	29193		29193		29193		29193		29193		29193		29193		29193		29193		29193		29193		29193

Dublin South West, 1969

Issue 6, May 1996

Editorial

A survey has been conducted of the readership of *Voting matters* which has resulted in a number of changes; these changes are reported on page 9. I have written individually to all those that took the trouble to write to ERS. Please write again if you have further suggestions, and especially if you have material for potential inclusion.

This issue contains five articles. The first is a republication of a further article by Brian Meek. Readers should take note of the preface which points out the very different nature of this article from the other two that *Voting matters* has republished. The second article contains a description of mine of a two-tier form of STV. I am not advocating this, since it appears to be inferior to standard STV.

The third article is a very detailed analysis of the degree of representativity in Irish STV elections by Philip Kestelman. Please note the use of the term *magnitude* to mean the number of seats in a multi-seat election.

Douglas Woodall's article is a very detailed analysis of the rules that could be used for single-seat elections. The importance of this work in my view is that of questioning the desirability of the property that later preferences should not harm or help earlier ones. Whatever your own views are, I hope you will note the consequences of the various impossibility theorems which shows that, even with just one seat, conflicting properties abound. This article does define a large number of terms but I hope readers will find the explanation of those terms adequate.

The last article is by David Hill which analyses the results which have previously been reported in *Replaying the 1992 General Election*. This paper illustrates the difficulties in producing accurate predictions for an STV election when only 9,614 ballot papers are available for all of the UK.

Brian Wichmann

Note in this reprinting

Brian Meek died in 1997: see end of this issue, page 17.

A transferable voting system including intensity of preference

B L Meek

Brian Meek is now at the Computing Centre, King's College London, Strand, London WC2R 2LS. This article was originally published in *Mathématiques et Sciences Humaines*, 13, No 50 1975 pp23-29.

Preface to this republication

After I wrote the two papers describing what has since become known as ‘Meek's method’ — published (in French) in *Mathématiques et Sciences Humaines* in 1969 and 1970, and republished in English in *Voting matters* No.1 — I went on to write a third paper, which the same journal published (in English!) in 1975. Some people have been aware of the existence of this third paper, and this led to a request that it too be republished in *Voting matters*. I have no objection to this being done, but it is important to stress that its status is quite different from the other two.

The first two papers present my analysis of STV counting, and how it can be made as accurate as possible. The method totally accepts the basis of STV as it is, and does not alter or challenge its fundamental assumptions at all. (It does seem to challenge some people's own assumptions about STV, but that's not the same thing at all!) As such, ‘Meek's method’ was always intended as a practical method for conducting an STV count, albeit an expensive one at that time — far less so now, of course. Years later, David Hill, Brian Wichmann and Douglas Woodall demonstrated beyond question that it is a practical method, and earned my eternal gratitude for so doing.

This third paper does not have that status at all. It is in fact no more than an academic exercise, exploring an issue which arises from time to time in the literature on aggregation of individual preferences. It demonstrates that a method of taking account of intensity of preference is possible. This is very far from advancing it as a practical method for implementing an election.

I have never regarded it as a practical method. I do not advocate its adoption, and I shall be very annoyed if anyone attempts to present it as (say) ‘Meek's proposal’ or otherwise imply that I advocate its use. It should not even be linked to ‘Meek's method’ (e.g. by alleging it is an *extension* to my method), at least without very careful qualification. The reason is that ‘Meek's method’ is STV, whereas the process described in this paper is *not* STV. (It is certainly not a ‘single’ transferable vote, for a start.) The way that votes are cast and interpreted is quite different from STV.

To be sure, the vote *counting* shares some similarities, but that is only because the same logic that led to the invention of STV and to the Meek method has been applied to the aggregation process. The individual votes being aggregated are, however, not STV votes. The consequence is that the result can end up very far from STV, as the paper itself clearly shows.

So the paper should be read for what it is, a mathematical demonstration that individual preferences can be fairly aggregated while still taking intensity of preference into account, and not as a suggested practical method for conducting elections. If that is done, there should be no misunderstandings. A voting system, derived from the STV (Single Transferable Vote), is described which includes intensity of preference while avoiding difficulties due to inter-personal comparison of utilities. It is shown that this system allows the voters some control over the method used to aggregate their preferences.

Introduction

This paper describes a voting procedure with a number of interesting properties. Chief among these are the inclusion of intensity of preference in a non-controversial manner — i.e. in a way which avoids the difficulty of inter-personal comparison of utilities — and that in various limiting cases the procedure is equivalent to well-known voting systems such as simple majority, the single transferable vote, the single non-transferable vote, etc. The paper first describes the voting procedure, then looks at the properties mentioned, and finally shows that the procedure offers a partial solution to the problem of determining which voting procedure to use in some decision situation.

The procedure

Any voting procedure consists of two parts — that of vote casting, and that of vote counting. In this case the vote casting procedure for the elector is to assign weights to the different candidates to indicate the order and strength of his preferences between them. It is a basic assumption that strength of preference is transitive, e.g. that if a voter thinks that he prefers A twice as much as B, and B three times as much as C, then he prefers A six times as much as C and can express his preferences by assigning weights to A, B, and C in the ratio 6:3:1.

The vote counting procedure begins by normalising all the weights w_{ij} which the *i*th voter gives to the *j*th candidate, so that

$$\sum_{j=1}^{j=c} w_{ij} = 1, \text{ all } i$$

c being the number of candidates. This is the key, as we shall see later, to the avoidance of troubles due to inter-personal comparison of utilities, since it ensures that as far as possible each voter has an equal say in the voting procedure.

The count proceeds by summing all the weights for all the candidates, i.e. calculating

$$W_j = \sum_{i=1}^{i=v} w_{ij} \quad \text{for all } j,$$

v being the number of voters. Thereafter the count proceeds much in the same way as in the single transferable vote, as modified by the proposals in two earlier papers[1,2]. An STV-type quota is calculated according to the formula

$$\left[q = \frac{W}{s+1} + 1 \right]$$

where *s* is the number of seats to be filled and

$$W = \sum_{j=1}^{j=c} W_j$$

is the total vote, and the brackets indicate that the integral part is to be taken. *q* is the minimum number such that, if *s* candidates have that number, any other candidate must have less than that number.

(In practice it is likely that working will be to fractions of votes — say three decimal places, in which case the "+1" in the formula for *q* is replaced by "+0.001", or equivalently the weights w_{ij} are normalised to sum to 1000 for each voter and the formula for *q* is unaltered.)

The count may proceed by one of two steps. If no W_j exceeds *q*, i.e. no candidate has reached the quota, then the candidate with lowest W_j, say candidate *x*, is eliminated. All the w_{ij} are then renormalised with all w_{ix} made equal to zero. The principle adopted is that if a candidate is eliminated the count proceeds as if that candidate had never stood; the assumption is that the elimination of a candidate does not alter the voter's relative preferences between the remaining candidates. (It is of course quite possible to take issue with this assumption.)

If, however, a candidate, say *y*, has W_y greater than *q*, another renormalisation takes place so that W_y is reduced to *q*. This means that all w_{iy} are reduced by the factor q/W_y, and all w_{ij}, $j \neq y$, are increased by the factor $(1 + w_{iy}q/W_y)/(1 - w_{iy})$. By this means the weights allocated by each voter *i* are adjusted in a quite natural way, so that those supporting *y* give him no more support than is necessary to ensure his election.

Counting continues by the application of one or other of these rules until the requisite number *s* of candidates are elected. Once elected and allocated the quota *q* the weights for that candidate are of course not included in the recalculation. This makes the procedure somewhat simpler than in the modified form of STV described in [1]. However, if all of a voter's choices — i.e. those candidates he has allotted a positive weight — are eliminated, the quota *q* has to be recalculated as in [2] so that this undistributable vote is not included; similarly, when all a voter's choices have been elected and allotted recalculated weights, the residue is non-distributable and also must be subtracted from *W*. Recalculation of the quota does of course imply recalculation of the weights of elected candidates, and an iterative procedure as described in [2] can be used to obtain the new *q* and w_{iy} to any desired accuracy.

Intensity of preference

When expressed crudely in the form “It is of more benefit to me to have A rather than B than it is for you to have B rather than A”, inter-personal comparison of utilities is patently invidious. Nevertheless in actual voting situations intensity of preference is often taken into account. If A and B want to go to a museum when C wants to go to the funfair, the collective choice is frequently the funfair, without any sense of dictatorship or lack of democracy, simply because all know that C's preference is much the most intense.

Lest this be regarded as too trivial an example, it is often the case in committee that the collective choice for chairman is X, even though a majority prefer Y, simply because a substantial minority strongly object to Y. Any theory of voting which does not allow for intensity of preference is certainly incomplete, and any voting system which does not permit its expression cannot be wholly satisfactory.

The present system is based on two principles: that the only person who can gauge the intensity of his preferences is the voter himself; and that as far as possible each voter should contribute equally in the choice of those elected. In a multi-vacancy election ($s > 1$) there is more than just a single choice involved, and so it makes sense to allow a voter to express his preference intensities by contributing all his voting power to the choice of one candidate, or to share this power between the choices of different candidates. Of course, it is possible to regard an *s*-vacancy election as a single choice from the ${}^nC_s$ possible combinations of *s* candidates out of *n* elected, but this view invalidates the assumption that elimination of a candidate does not alter the voter's relative preferences. This is because each combination is independent; a voter may rank candidates individually A, B, C, D in that order, but rate them in pairs AB, BD, CD, BC, ... since he thinks A will only work satisfactorily in combination with B. This kind of multiple election is essentially the election of a ***team*** of *s* people, rather than *s* individuals. STV, and the present system, is concerned with choosing a set of *s* independent individuals from a larger set of *c* candidates. An STV vote is a vote for one individual (the first choice) and only subsidiarily and in special circumstances for lower choices. The present system enables the voter to have a say in all the *s* choices if he wishes, but his share in the whole decision process remains the same, up to the wastage involved in nontransferable votes or those given to unelected candidates who remain when the *s* winners have been chosen.

Equivalence to other voting systems

(a) STV

Let $1 > \varepsilon > 0$. Let the voters order their choices $1-\varepsilon$, $\varepsilon-\varepsilon^2$, $\varepsilon^2-\varepsilon^3$,..... $\varepsilon^{c-2}-\varepsilon^{c-1}$, ε^{c-1}. Then the closer ε is to 0 the closer the actual voting process becomes equivalent to STV. For example, suppose there are 5 candidates and $\varepsilon = 0.01$. A voter's choice will be in the proportions 0.99, 0.0099, 0.000099, 0.00000099, 0.00000001, counting 99% for his first choice. If his first choice is eliminated, the four lower votes remain, and total 0.01. These have to be renormalised to add up to 1, and so are multiplied by 100 to give 0.99, 0.0099, 0.000099, 0.000001. A similar argument applies to votes transferred from elected candidates.

(b) Single non-transferable vote

This, trivially, is when the voter gives 1 to his first choice and 0 to all the others.

(c) Simple majority with multiple vote

Here the voter gives $1/s$ to each of *s* candidates, or perhaps $1/k$ to each of *k* candidates, $k < s$. These are special cases of giving α to *k* candidates and β to $c-k$ candidates, where $\alpha k+\beta(c-k) = 1$ giving a weighting between a more preferred and a less preferred group.

(d) Cumulative vote

In this case the voter gives $\alpha_1, \alpha_2, ... \alpha_k$, to *k* candidates respectively, such that

$$\sum_{i=1}^{i=k} \alpha_i = 1$$

For an exact analogy to the cumulative vote each α_i must be a multiple of $1/s$.

The choice of voting procedure

Such a voting system would require a more than usual sophistication on the part of the voter. This being so, one can consider a further sophistication. The choice of voting procedure is one of immense importance in the democratic process, and no system is wholly stable wherein a substantial minority is dissatisfied with the voting procedure in current use. The required consensus may either be achieved through ignorance or habit, or by general agreement that a system is fair even though another may be advantageous to many, perhaps even a majority. In situations where there is awareness of and controversy about the different properties of voting systems, the present system offers a possible way out of deadlock. For, if most voters use the system in one of the ways described in the last section, then the election will be largely determined according to that voting procedure. Looking at it from the point of view of parties, each party can urge the voters to use the method they favour of filling in the ballot forms. However, it is a weakness in this area that voting systems are so often argued about in terms of fairness to parties or candidates, seldom in terms of fairness to voters. The present system, whose main fault is its complexity, has the virtue of that fault in that each voter can specify as precisely as he wishes the way his vote is to be counted, without this being imposed by others on him or on others by him. Most voting systems allow some such flexibility; the virtue of this system is the much greater precision with which the sophisticated voter can specify his wishes, without his being able by strategic voting to exercise more influence on the final result than is implied by his actual possession of a vote.

Concluding remarks

Despite the scope for manipulation which the system offers, it is clearly derived from and shares the principles of the STV system, particularly with the concept of the quota and the transferability of votes above the quota. One of the chief objections to STV is that it does not guarantee the election of a Condorcet winner, e.g. when one candidate is everyone's *second* choice. While the present system does not guarantee the election of such a candidate (this is obvious, since as shown earlier the system can approach arbitrarily closely to STV), it does render it more likely, and will ensure it provided that the weights given to the candidate are large enough i.e. if the candidate is considered a *good enough* substitute for their first choice by a sufficient number of electors. The price that one has to pay for this improvement to STV is the greater complexity, particularly for the voter.

Bibliography

1. B L Meek, Une nouvelle approche du scrutin transférable I: égalite de traitement des électeurs et technique à rétroaction utilisée pour le depouillement des votes, *Mathématiques et Sciences Humaines*, 25, 1969, pp 13-23. Reproduced in *Voting matters*, Issue 1, pp1-6.

2. B L Meek, Une nouvelle approche du scrutin transférable II: le problème des votes non-transférables *Mathématiques et Sciences Humaines*, 29, 1970, pp 33-39. Reproduced in *Voting matters*, Issue 1, pp7-11.

A form of STV with single-member constituencies

B A Wichmann

Brian Wichmann is the editor of *Voting matters*

One over-riding concern that appears in the Plant report is the desire to retain single-member constituencies in any reform of the electoral system for the House of Commons. A natural question is if STV can be adapted in some way to retain single-member constituencies, but avoiding the non-proportionality of the Alternative Vote (AV). This paper presents such an adaptation.

The basic idea is to use a two-tier system in a similar manner to the German system by having single-member constituencies augmented by members elected in a more proportional manner. The second tier is a group of constituencies which, for convenience, we call a *county*. The electors provide two ‘votes’ by giving the usual preferences to the candidates in their constituency and then also providing a preference vote to *all* the candidates in their county.

The election proceeds in two stages, firstly each constituency is considered individually using STV (which degenerates to AV). However those votes which have *not* been used to elect a candidate here are forwarded to the county vote (or second stage). The county vote first eliminates those candidates already elected at the first stage, and then uses STV to fill the county seats.

The main parameters of this voting system are the number of single-member constituencies used to form a *county*, and the number of seats available at the county level. It appears that about 5 (or more) constituencies should be grouped into a county in order to provide reasonable proportionality and that the number of county seats should be not less than 2 for the same reason.

This system is quite different from conventional STV for a number of reasons:

1. This system, like FPTP has *safe* seats, whereas STV has no such equivalent. For instance, in the Irish elections, almost every constituency has a Fianna Fail or Fine Gael candidate who is *not* elected. My reason for concluding this is that I believe that the main parties, even for safe seats, would not propose more than one candidate since this would appear to present a divided party.

2. The elector's ability to select within a party is restricted. If you are a Conservative party supporter in a safe Conservative constituency with a male candidate, you could not select a woman candidate (given the restriction noted above of a single candidate). On the other hand, if you were in a Labour constituency, your vote would be *wasted*, allowing you to select a woman candidate from the county list as your first preference.

3. Minority interests would be represented at the county level. These interests would be accumulated as wasted votes and hence would have a good chance of representation, depending upon the number of county seats.

Of course, the advantage of this system is that there is no reliance upon the ordering of a party list which is outside direct voter control.

There are some technical details to resolve. I have based my proposal on the use of the Meek algorithm for STV, although this is not strictly essential. However, it is clearly important to compute the fraction of each vote which is wasted (from the first stage) in order to conduct the second stage. This is straightforward since for each voter who contributes to the elected candidate, the percentage wasted is simply the percentage of votes above the quota. This implies that about ½ of the votes would go forward to the second stage. This might imply that about half of the seats should be at county level, but a smaller number is probably satisfactory.

My belief is that this proposal would be quite easy to implement, at least using the Meek algorithm. However, since we have no similar system, it does not seem possible to construct realistic data with which to do any serious study of its suitability.

Is STV a form of PR?

P Kestelman

Philip Kestelman is keen on measuring electoral representativity, and works in the area of family planning

Introduction

In my view, Single Transferable Voting (STV) is the best electoral principle: whether electing one representative by Alternative Voting (AV), or several representatives by multi-member STV. The Collins English dictionary succinctly defines *proportional representation* (PR) as “representation by parties in an elective body in proportion to the votes they win”.

The 1937 Irish Constitution prescribes that both the President and parliamentary deputies (TDs) shall be elected “on the system of proportional representation by means of the single transferable vote”. Of course, PR is *not* an electoral system; but a principle, to which different elections approximate to widely varying degrees.

Accordingly, the basic question is whether STV achieves PR; and if so, how far? To answer this question, we need some overall measure of electoral representativity (‘proportionality’); of which the simplest is the *Rose Index*[12]. For reasons which will become apparent, I have renamed the Rose Index, *Party Total Representativity* (PTR).

Party

Table 1 demonstrates the calculation of PTR, for the 1994 European Parliamentary Election in the Irish Republic. Notice that the total over-representation of all over-represented party votes (+23.7% of first preferences) is equal and opposite to the total under-representation of all under-represented party votes (−23.7%). This overall deviation is the *Loosemore-Hanby Index* (LHI) of party disproportionality[10], — “the most widely used measure of disproportionality”[9].

Thus LHI measures the total under-representation of all under-represented party-voters. Complementing LHI is the Rose Index, PTR = 100.0 − 23.7 = 76.3% of first preference votes. For comparison, in the 1994 European Parliamentary Election in Britain (First-Past-the-Post), PTR = 70.4%. This low British PTR (definitely *not* PR) approximated the Irish PTR (76.3%); and the corresponding STV final count PTR (81.7%) was little higher.

Cole[3] over-estimated final count PTR by excluding non-transferable votes. Moreover, non-transferable votes are under-counted by conventional STV proportionating Droop Quota surplus votes among transferable next preferences (ie. continuing candidates only[11]). Besides, “using later-stage figures overstates the proportionality of STV”[6].

Table 1: Party Representativity analysis of the European Parliamentary Election, Irish Republic 1994.

Party	Votes (V%) first	Votes (V%) final	Seats (S%)	Deviation (S%-V%) first	Deviation (S%-V%) final
Total	100.0	100.0	100.0	0.0	0.0
Fianna Fáil	35.0	37.4	46.7	+11.7	+ 9.3
Fine Gael	24.3	30.8	26.7	+ 2.4	- 4.1
Labour	11.0	4.2	6.7	- 4.3	+ 2.5
Green	7.9	8.9	13.3	+ 5.4	+ 4.4
Cox (Munster)	2.5	4.6	6.7	+ 4.2	+ 2.1
Others/Non-transferable	19.4	14.2	0.0	-19.4	-14.2
Over-represented	69.6		93.3	+23.7	
		55.1	73.3		+18.3
Under-represented	30.4		6.7	-23.7	
		44.9	26.7		-18.3

Source: *Irish Times*, 14 June 1994.

In the first four European Parliamentary elections (1979-94), the Irish PTR ranged from 76.3% to 87.0% of STV first preferences; hardly more representative than the British PTR, ranging from 70.4% to 78.6%. In the 1990 Irish Presidential Election, PTR increased from 38.9% of first preferences to 51.9% of final preferences yet nobody regards AV as PR!

Indeed, none of the foregoing STV elections has achieved anything like PR. However, in the last six Irish general elections (1981-92), PTR has ranged from 90.1% to 96.9% of first preferences, as may be seen in Table 2.

Apparently, multi-member STV is only ‘semi-proportional’. More remarkably, three and five member STV constituencies mediated comparable representativity. This refutes the widespread belief that “political science research establishes conclusively that PR electoral districts must elect at least four MPs before they deliver proportional outcomes”[5]. Indeed, four member STV constituencies proved invariably *less* representative than either three or five member constituencies, although the differences were small.

Table 2: Party Total Representativity by district magnitude in Irish Republic general elections.

	District magnitude			
Date	All	3	4	5
1981	94.2	95.4	89.8	94.7
1982 (Feb)	96.6	97.4	95.6	95.8
1982 (Nov)	95.8	97.4	92.8	95.3
1987	90.1	89.5	89.1	89.9
1989	92.9	94.0	91.1	92.2
1992	91.8	90.2	89.5	91.5

Source: Dáil Éireann[4]

Cumbency

Bogdanor[1] observed that STV advocates prefer to secure “proportional representation of opinion ... which cuts across party lines. But since they do not give a clear operational definition enabling one to measure ‘proportionality of opinion’, it becomes difficult to offer any evaluation of their claim”. Nonetheless, published election results provide some usable, *non*-party data for each candidate, including *Cumbency*: that is, whether incumbent (immediately previous representative) or non-incumbent (‘excumbent’).

Analogously to party, consider the relationship between cumbency first preferences and seats. Instead of PTR, incumbent and excumbent candidates are treated as representing two different parties; and *Cumbency Total Representativity* (CTR) is calculated, as in Table 3.

Table 3: Cumbency Total Representativity by district magnitude for Irish Republic general elections

	District magnitude			
Date	All	3	4	5
1981	86.2	85.0	83.8	88.6
1982 (Feb)	85.0	77.9	92.6	83.0
1982 (Nov)	83.7	72.6	87.0	87.3
1987	81.6	80.2	79.9	83.8
1989	87.9	96.2	82.2	86.9
1992	85.4	76.7	90.4	86.3

Source: Dáil Éireann[4]

Such low CTRs arise from incumbents invariably over-representing their first preferences (high incumbent S%−V%). Notice the distinction between this finding and the unsurprisingly, greater electability of incumbent candidates (high incumbent S%−C%, where C% is the fraction of incumbent candidates).

Of course, incumbents are far more likely than excumbent candidates to be men. Hence the importance of disentangling cumbency from gender.

Gender

At the 1992 Irish General Election, 19% of candidates were women: 8% of incumbents and 24% of excumbent candidates. Among elected candidates (TDs), only 12% were women: 8% of incumbents, and 23% of excumbent TDs. Thus allowing for cumbency, TDs fairly represented candidates by gender.

What of the relationship between votes and seats, by gender (electoral representativity proper)? In 1992, voters cast 13% of their first preferences for women candidates: slightly under-represented by women TDs (12%).

Table 4: Gender Representativity Ratio by district magnitude in Irish Republic general elections 1981-89

	District magnitude			
Cumbency	All	3	4	5
All	0.94	1.20	1.04	0.80
Incumbent	1.10	1.01	1.19	1.02
Excumbent	1.07	2.26	0.98	0.89

Source: Dáil Éireann[4]

As with cumbency, we could calculate a *Gender Total Representativity* (GTR) for each election and district magnitude. However, because there are only two genders (non-transferable!), and so few women candidates (and hence votes for women), it seems more illuminating to aggregate the previous five general elections (1981-89); and to calculate *Gender Representativity Ratios* (GRRs).

GRR is the ratio of female seats per vote to male seats per vote (first preference). Table 4 gives GRR, by district magnitude and cumbency.

In 1981-89 overall, first preferences for women candidates were slightly under-represented (GRR = 0.94). However, allowing for cumbency, women TDs slightly over-represented their first preferences (excumbent GRR = 1.07).

Of particular interest, three member STV constituencies over-represented votes for women by 20% (GRR = 1.20); leaving them under-represented in five member constituencies by 20% (GRR = 0.80) overall. Among excumbent candidates in three member constituencies, first preferences for women were over-represented even more spectacularly; only 5% of votes electing 10% of the TDs (GRR = 2.26). By contrast, in five member constituencies, 15% of the voters for excumbent candidates preferred women, represented by 14% of the TDs (GRR = 0.89).

Alphabetic bias

It is widely believed that candidates appearing high on ballot-forms enjoy some electoral advantage. On Irish general election ballot-forms, candidates' names are listed in surname-alphabetic order. Voters' preferences for (less familiar) excumbent candidates may well be more vulnerable to 'Positional Voting Bias'[14].

Notice that we are interested here in three distinct relationships: between candidates and votes (first preferences): between candidates and seats; and between votes and seats (electoral representativity proper). Aggregating five Irish general elections (1981-89), Table 5 confirms that higher placed excumbent candidates attracted disproportionately more first preferences (V%/C% decreasing, from 1.18 for A-C surnames, to 0.89 for N-Z surnames).

Table 5: Excumbent Candidate Surname/Forename Representativity Index Irish Republic general elections 1981-89

Name	Initial letter	Vote/ Candidate =V%/C%	Seat/ Candidate =S%/C%	Seat/ Vote =S%/V%
Surname	A-C	1.18	1.20	1.02
	D-J	0.99	1.12	1.12
	K-M	0.96	0.80	0.84
	N-Z	0.89	0.91	1.01
Forename	A-E	0.95	0.95	1.01
	F-K	1.08	1.02	0.94
	L-P	1.06	1.32	1.24
	Q-Z	0.88	0.60	0.68

Source: Dáil Éireann[4]

Consequently, excumbent TDs over-represented candidates with A-C surnames by 20% (S%/C% = 1.20); under-representing those with K-M surnames by 20% (S%/C% = 0.80). However, notice something else: excumbent candidates with L-P forenames were even more over-represented (S%/C% = 1.32); leaving those with Q-Z forenames even more under-represented (S%/C% = 0.60). All the more remarkable, considering that forenames are *not* ordered alphabetically on ballot-forms; and perhaps voters' preferences for surnames were not positional, after all!

Relating excumbent first preferences to seats (electoral representativity proper), D-J surnames and L-P forenames were over-represented (by 12% and 24%, respectively); while K-M surnames and Q-Z forenames were under-represented (by 16% and 32%, respectively). How should we compare surname with forename representativities overall?

We could treat every single name-initial letter of the alphabet like a party (N=22), and calculate both *Surname Total Representativity* (STR) and *Forename Total Representativity* (FTR). Aggregating five Irish general elections again gives Table 6, comparing STR with FTR by district magnitude and cumbency.

Table 6: Excumbent Candidate Surname/Forename Total Representativity Index by district magnitude in Irish Republic general elections 1981-89

	District magnitude			
Cumbency	All	3	4	5
All	95.5/94.8	93.3/91.7	91.8/89.7	93.1/95.6
Incumbent	97.1/96.4	96.0/95.3	94.4/93.4	96.1/97.2
Excumbent	90.7/88.6	75.3/80.8	83.8/85.0	85.2/88.1

Source: Dáil Éireann[4]

Overall, first preferences for surnames and forenames were represented with comparable fidelity (STR = 95.5%: FTR = 94.8%); again, with little difference by district magnitude. Among excumbent candidates, TDs represented surnames slightly more faithfully than forenames (STR = 90.7%: FTR = 88.6%) overall; by district magnitude, somewhat less. Altogether a muddy picture, without obvious implications for ordering candidates' names on ballot forms (surname — alphabetical or random).

Conclusions

Considering the quantitative notion of PR, the measurement of electoral representativity remains curiously neglected. The simplest measure of overall party disproportionality, the Loosemore-Hanby Index (LHI), complements the Rose Index, or Party Total Representativity (PTR). Indeed, PTR may be construed as the degree to which any given election — from a national aggregate down to a single member constituency — achieves PR (rarely 100%).

Single member STV (Alternative Voting) hardly mediates PR, even at the national level (as in Australia[2]). In the four European Parliamentary elections in the Irish Republic, even multi-member STV has only achieved PTRs ranging from 76% to 87%: scarcely more representative than First-Past-the-Post in Britain: ranging from 70% to 79%.

However, the last six Irish general elections (1981-92) have proved considerably more representative, PTR ranging from 90% to 97%. Thus multi-member STV alone mediates quasi-PR[15]; requiring a few additional members to guarantee PR (eg. final count best losers: 'STV-plus', as in Malta[7]).

More remarkably, Irish three and five member STV constituencies have proved comparably representative. That is good news for voters, oppressed by the lengthy ballot-forms characterising larger STV constituencies (perhaps listing 20 names). It is equally good news for reformers, dismayed at the prospect of anonymously vast STV constituencies, electing as many as seven MPs (eg. representing all three London boroughs of Greenwich, Lewisham and Southwark[13]).

The concept of *Total Representativity* proves a most versatile tool, even beyond party considerations. In respect of cumbency, multi-member STV remains invariably non-PR; with Cumbency Total Representativity ranging from 82% to 88%. On the other hand, first preferences for women candidates have been represented near-proportionally; with an aggregate female-to-male S%/V% ratio of 0.94 overall. Nonetheless, three member STV constituencies over-represented votes for women, under-represented in five member constituencies.

Aggregating five Irish general elections also confirmed that excumbent candidates listed higher on ballot-forms tended to attract disproportionately more first preferences; thereby over-representing candidates with A-C surnames, and under-representing K-M surnames(S%/C%). Yet TDs over/under-represented candidates with L-P/Q-Z forenames even more steeply. Moreover, first preferences for both surnames and forenames were represented with comparable fidelity. It may not be so important to randomise ballot-forms after all: another relief for preferential voters accustomed to alphabetic order!

At best therefore, in mediating party first preferences (the main consideration), multi-member STV alone is not quite a form of PR. Nonetheless, in national parliamentary elections, Irish STV has proved far more representative than British FPP. That conclusion may be brought even closer to home, by calculating another measure (perhaps user-friendlier than PTR).

Under both AV and FPP, around half of all voters elect candidates; whereas under multi-member STV, nearly 90% of voters elect at least one representative of their preferred party. In Irish general elections, this *Constituency-Represented Party Vote-fraction* (CRPV) has also proved conspicuously invariant with district magnitude, as shown in Table 7.

Maximising each CRPV, multi-member STV minimises vote-wastage. Thus quantifying STV's principal virtue, CRPV should allay the concern over STV — apart from its complexity — expressed by the Plant Working Party on Electoral Systems[8]. Of course, STV enjoys other virtues!

Table 7: Constituency-Represented Party Vote-fraction by district size for Irish Republic General Elections 1981-89

Date	District size			
	All	3	4	5
1981	90.8	91.3	86.9	93.0
1982 (Feb)	92.6	94.0	91.3	92.8
1982 (Nov)	92.3	94.6	89.4	93.4
1987	83.6	85.3	81.0	84.8
1989	87.3	90.9	84.9	86.7
1992	86.7	85.7	84.9	88.7

Source: Dáil Éireann[4]

References

1. V Bogdanor. Introduction, in *Democracy and Elections*, edited by V Bogdanor and D Butler. 1983. pp 1-19.

2. V Bogdanor. *What is Proportional Representation?* 1984.

3. P Cole. Measuring Proportionality and Presenting the Evidence, *Representation*, Vol 30, pp 66-67, 1992.

4. Dáil Éireann. *Election Results and Transfer of Votes in General Election*. June 1981 - November 1992.

5. P Dunleavy, *et al*. *Replaying the 1992 General Election*. 1992.

6. M Gallagher. The Political Consequences of the Electoral System in the Republic of Ireland, *Electoral Studies*, Vol 5, pp 253-275, 1986.

7. S Howe. The Maltese General Election of 1987. *Electoral Studies*, Vol 6, pp 235-247, 1987.

8. Labour Party. *Second Interim Report of the Working Party on Electoral Systems*. 1992.

9. A Lijphart. *Electoral Systems and Party Systems*. 1994.

10. J Loosemore and V Hanby. The Theoretical Limits of Maximum Distortion: Some Analytic Expressions for Electoral Systems. *British Journal of Political Science*, Vol 1, pp 467-477. 1971.

11. P McKee. The Republic of Ireland, in *Democracy and Elections*, edited by V Bogdanor and D Butler. 1983. pp 167-189.

12. T Mackie and R Rose. *The International Almanac of Electoral History* (3rd edition). 1991.

13. R A Newland. *Electing the United Kingdom Parliament* (3rd edition). 1992.

14. C Robson and B Walsh. The Importance of Positional Voting Bias in the Irish General Election of 1973. *Political Studies*, Vol 22, pp 191-203. 1974.

15. R Taagepera and M S Shugart. *Seats and Votes*. 1989.

Changes to *Voting matters*, as recommended by the Technical Committee of the ERS are as follows:

1. You can see that a subtitle now appears. The reason for this is that some readers did not appreciate the technical nature of the publication.

2. As Editor, I will try to avoid excessively technical jargon. I will attempt to ensure that terms like *monotonicity* are explained (even though that has been defined in a previous issue).

3. The main publication of ERS, *Representation*, is being asked to reproduce the contents list of our issues, so that those interested will be aware of *Voting matters*.

Monotonicity and Single-Seat Election Rules

D R Woodall

Douglas Woodall is Reader in Pure Mathematics at Nottingham University.

1. Introduction

This article investigates the monotonicity properties of preferential election rules for filling a single seat. Section 2 lists the properties of interest, which form a subset of those introduced in Woodall[4]. Section 3 describes several known election rules and two new ones (QLTD and DAC), whose properties are tabulated in Table 1. Section 4 describes a number of impossibility theorems, which are also represented symbolically in Table 1. These theorems say that certain combinations of properties cannot hold simultaneously, because the properties in question are mutually incompatible. In Section 5 I attempt to summarize the current state of knowledge and indicate what remains to be done.

Throughout this article I consider only the single-seat case. This does not reduce the force of the impossibility theorems in Section 4. We are interested in universal election rules, which will work for filling any number of seats. If certain properties are mutually incompatible even in the single-seat case — that is, there is not even a single-seat election rule with all these properties — then it is almost inconceivable that there will be an election rule with all these properties that works for any other number of seats, and there certainly cannot exist a universal election rule with them all. So, in practice, an impossibility theorem for single-seat election rules is as good as one that considers multi-seat elections as well. But in the case of the examples in Section 3, considering only single-seat elections is a real limitation, and I have resorted to it only because I have found the multi-seat case too hard to handle. There are many election rules that possess properties in the single-seat case that they do not possess in the multi-seat case, and there are many single-seat election rules that cannot apparently be extended to multi-seat elections in any sensible way, and so the multi-seat case is much harder to analyze.

I think the most important problems facing mathematicians who are interested in STV are, first, to discover which monotonicity properties are compatible with DPC (the Droop Proportionality Criterion)[4], or with majority (the property that DPC reduces to in single-seat elections—see Section 2 below); and then to find an election rule that satisfies DPC and as many monotonicity properties as possible. In the case of single-seat elections, I have found a rule (DAC) that satisfies majority and many monotonicity properties, which I would be prepared to recommend as preferable to the Alternative Vote (AV). Admittedly it fails to satisfy one important property of AV, that later preferences should not count against earlier preferences, but in return for this it gains five properties that AV does not possess. However, at the moment I have not been able to extend DAC in any sensible way to multi-seat elections, and I do not know whether this will prove to be possible, or whether it will be necessary to start afresh with a new idea.

2. The properties

These properties were all introduced in Woodall[4], where they were discussed in more detail, and so I shall only state them briefly here. Of the seven global or absolute properties mentioned there, three are of interest to us now:

Plurality. If some candidate *x* has strictly fewer votes in total than some other candidate *y* has first-preference votes, then *x* should not have greater probability than *y* of being elected.

Majority. If more than half the voters put the same set of candidates (not necessarily in the same order) at the top of their preference listings, then at least one of those candidates should be elected.

Condorcet. If there is a Condorcet winner (that is, a candidate who would beat every other candidate in pairwise comparisons), then the Condorcet winner should be elected.

Of these three properties, **majority** is by far and away the most important. **Plurality** is also important, but it is much less likely to be violated: every reasonable electoral system seems to satisfy it, whereas many systems proposed or actually used, such as first-past-the-post, point-scoring systems and approval voting, fail **majority**. **Condorcet** is a very attractive property, but, as we shall see in Section 4, it leads to problems with monotonicity. My aim is to find a system that satisfies **majority** and as many of the monotonicity properties as possible.

Among the local or relative properties introduced in Woodall[4] we shall consider seven of the nine versions of monotonicity, together with **participation**, **later-no-help** and **later-no-harm**. The remaining two versions of monotonicity, **mono-append** and **mono-add-plump**, are omitted because they hold for all the election rules discussed in Section 3 and do not feature in any of the impossibility theorems in Section 4.

Monotonicity. A candidate *x* should not be harmed if:

(**mono-raise**) *x* is raised on some ballots without changing the orders of the other candidates;

(**mono-raise-delete**) *x* is raised on some ballots and all candidates now below *x* on those ballots are deleted from them;

(**mono-raise-random**) *x* is raised on some ballots and the positions now below *x* on those ballots are filled (or left vacant) in any way that results in a valid ballot;

(**mono-sub-plump**) some ballots that do not have *x* top are replaced by ballots that have *x* top with no second choice;

(**mono-sub-top**) some ballots that do not have *x* top are replaced by ballots that have *x* top (and are otherwise arbitrary);

(**mono-add-top**) further ballots are added that have *x* top (and are otherwise arbitrary);

(**mono-remove-bottom**) some ballots are removed, all of which have *x* bottom, below all other candidates.

Participation. The addition of a further ballot should not, for any positive whole number *k*, reduce the probability that at least one candidate is elected out of the first *k* candidates listed on that ballot.

Later-no-help. Adding a later preference to a ballot should not help any candidate already listed.

Later-no-harm. Adding a later preference to a ballot should not harm any candidate already listed.

3. Examples of election rules

First-Preference Plurality (FPP), or First-Past-the-Post, elects the candidate with the largest number of first-preference votes. This rule behaves extremely well with regard to all the local properties (although it satisfies **later-no-harm** only if second and subsequent preferences are ignored totally, and are not used to separate ties). However, it does not satisfy **majority** or **Condorcet**: in Election 1 below, FPP elects *c*, but majority requires that *a* or *b* should be elected, and *a* is the Condorcet winner.

	ab	30
Election 1:	*ba*	25
	c	45

Point Scoring (PS) methods are those where each candidate is given a certain number of points for every voter who puts them first, a certain (smaller) number for every voter who puts them second, and so on, and the candidate with the largest total number of points is elected. These methods have very similar properties to FPP, although later preferences can now count against earlier preferences, so that **later-no-harm** fails, and **mono-raise-random** and **mono-sub-top** also fail in most cases. To see that PS systems do not satisfy **majority** or **Condorcet**, suppose that just over half the voters list three candidates in the order *abc*, and just under half list them in the order *bca*. Then both **majority** and **Condorcet** require that *a* should be elected, but any PS method will choose *b*.

Table 1

	Properties of specific election rules						Impossibility theorems		
	FPP	PS	AV	C-PS	QLTD	DAC	1	2	3
Plurality	√	√	√	√	√	√	•		
Majority	×	×	√	√	√	√			•
Condorcet	×	×	×	√	×	×	•	•	
Mono-raise	√	√	×	√	√	√			×
Mono-add-top	√	√	√	×	×	√	×		
Mono-remove-bottom	√	√	×	×	√	√			×
Participation	√	√	×	×	×	√		×	×
Mono-raise-random	√	×	×	×	×	×		×	×
Mono-sub-top	√	×	×	×	×	×		×	×
Mono-raise-delete	√	√	×	×	√	√		×	×
Mono-sub-plump	√	√	×	×	√	√		×	×
Later-no-help	√	√	√	×	√	√		×	•
Later-no-harm	√	×	√	×	×	×		×	•

The thick box delimits those properties that make sense even if truncated preference listings are not allowed. The top three properties are global while the others are local or relative.

The Alternative Vote (AV) was discussed at length in Woodall[4] and so I shall content myself here with tabulating its properties in Table 1. Unlike FPP and PS, it satisfies the all-important **majority** property, but it behaves rather badly with respect to monotonicity.

There are many known election rules that satisfy Condorcet's principle; for example, nine such rules are discussed by Fishburn[1]. In the present context (looking for a more monotonic substitute for AV) we are really only interested in rules that satisfy **majority**. Among such rules, the one with the largest number of other properties seems to be one that is not among the nine considered by Fishburn, namely to use a point scoring method to select a candidate from the Condorcet top tier. This method is described as C−PS in Table 1. It satisfies all three of the global properties that we are considering, but it behaves badly with respect to the local properties.

My first serious attempt to find a rule that would rival AV resulted in what I call Quota-Limited Trickle-Down (QLTD). Although this has now been superseded by DAC, I describe it here because it is simpler. One starts by crediting every candidate with all their first-preference votes. If no candidate exceeds the quota (of half the number of votes cast), then one gradually adds in the second-preference votes, then the third-preference votes, and so on, until some candidate reaches the quota. For example, it may be that if one credits every candidate with all their first-preference votes, all their second-preference votes and 0.53 times their number of third-preference votes, then exactly one candidate is brought up to the quota; that candidate is then declared elected.

	abcdef	12
	cabdef	11
Election 2:	*bcadef*	10
	def	27

It is easy to see that this rule satisfies **majority**. At first I thought it satisfied all the most important monotonicity properties as well. However, I now realize that it fails **mono-add-top**. This can be seen from Election 2 above. Here the quota is 30, and if one gives every candidate all their first and second-preference votes, plus 0.7 of their third-preference votes, then *a* gets 30 votes, *b* 29.7, *c* 29.4, *d* 27, *e* 27 and *f* 18.9; thus *a* is elected. However, if one adds six extra ballots marked *ad*, then the quota goes up to 33, but now *d* reaches the quota on first and second preferences alone: the count is *d* 33, *a* 29, *b* 22, *c* 21, *e* 27 and *f* zero. In Election 2 itself, *a* is behind *d* (by 23 to 27) on the basis of first and second-preference votes, but *a* overtakes *d* when the third-preference votes are added in. Adding six extra *ad* ballots increases *a*'s and *d*'s first and second-preference votes by the same amount, and this causes *d* to reach the quota: *a* would again overtake *d* if the third-preference votes were added in, but this does not happen because the election has already ended.

Election 3		Acquiescing Coalitions			
ab	11	{*a*, *b*, *c*}	30	{*c*}	12
b	7	{*b*, *c*}	19	{*a*}	11
c	12	{*a*, *b*}	18	{*b*}	7
		{*a*, *c*}	12		

My most recent attempt to find a substitute for AV has resulted in what I call the method of Descending Acquiescing Coalitions, or DAC, which is the first election rule that I am really happy with. The *coalition acquiescing* to any set of candidates comprises all voters who have not indicated that they prefer any candidate not in that set to any candidate in that set. For example, in Election 3 above, there are 19 voters who acquiesce to *b* and *c*, namely, the 7 who voted *b* and the 12 who voted *c*; none of them actually voted for both *b* and *c*, but none of them have said that they prefer *a* to either of these candidates, and so they are said to acquiesce to this set of two candidates. Similarly, the 18 voters who acquiesce to *a* and *b* comprise the 11 who voted *ab* and the 7 who voted *b*. The 12 voters who acquiesce to *a* and *c* are exactly the same as those who acquiesce to *c*, namely, the 12 *c* voters. And so on.

In DAC, one first lists the sizes of all the acquiescing coalitions in decreasing order, as I have done above, and then works down the list from the top, eliminating candidates until only one is left. The largest acquiescing coalition always contains every voter, since every voter acquiesces to the set of all candidates; this does not help towards deciding who should be eliminated. In the above example, the next largest acquiescing coalition comprises 19 voters, for {*b*, *c*}; the fact that *a* is not included in this set means that *a* is the first candidate to be eliminated. The next acquiescing coalition comprises 18 voters, for {*a*, *b*}. Since *c* is not included in this set, *c* is next to be eliminated. This leaves only one candidate not eliminated, namely *b*, and so *b* is declared elected. (Note that AV would exclude *b* first and then elect *c* in this example.)

Election 4		Largest Acquiescing Coalitions			
adcb	5	{*a*, *b*, *c*, *d*}	30	{*a*, *c*}	8
bcad	5	{*a*, *b*, *c*}	13	{*b*, *c*, *d*}	8
cabd	8	{*d*}	12	{*b*, *d*}	8
dabc	4	{*a*, *d*}	9	{*c*}	8
dbca	8				

Sometimes several candidates can be eliminated at once. For example, in Election 4, the largest acquiescing coalition not containing all voters comprises 13 voters, for {*a*, *b*, *c*}; thus *d* is the first candidate to be eliminated. The next largest acquiescing coalition is for {*d*}, and so it appears that *a*, *b* and *c* should all be eliminated at once, leaving no candidate remaining uneliminated. In this case one simply ignores this coalition: it does not help in distinguishing between the remaining three candidates. The next coalition is for {*a*, *d*}, and this causes *b* and *c* to be eliminated, so that *a* is elected.

Election 5		Largest Acquiescing Coalitions	
acbd	6	{*a*, *b*, *c*, *d*}	25
adbc	3	{*a*, *b*, *c*}	14
adcb	3	{*a*}	12
bcad	4	{*a*, *c*}	10
cabd	4	{*a*, *d*}	6
dbca	5		

It is not difficult to see that DAC satisfies **majority**, since if more than half the voters put the same set of candidates (in various orders) at the top of their preference listings, then every other candidate will be eliminated before any candidate in that set. With slightly more difficulty, it can be proved that DAC satisfies all the other properties ticked in Table 1. However, it does not satisfy **mono-raise-random** or **mono-sub-top**: if two of the four *dabc* ballots in Election 4 were replaced by *acbd* then *c* would be elected instead of *a*. Also, it does not satisfy **Condorcet**: in Election 5, DAC elects *a*, but *c* is the Condorcet winner. And it does not satisfy **later-no-harm**: if the seven *b* voters in Election 3 had voted *bc* instead, then *c* would have been elected instead of *b*. We shall see in the next section that there cannot exist any election rule satisfying **Condorcet** or **later-no-harm** as well as all the properties of DAC; but it is not clear whether there is any rule that satisfies **mono-raise-random** or **mono-sub-top** as well as everything that DAC does.

4. Impossibility theorems

Of the three theorems summarized symbolically in Table 1, the one of greatest interest in the present context is Theorem 3. However, it is also the most difficult to prove, and so I shall discuss the two simpler theorems first.

Theorem 1 says that if **plurality** and **Condorcet** hold then **mono-add-top** cannot hold; that is, there is no election rule that satisfies all three of these properties. This is easily seen by considering Election 3. Which candidate would such a rule elect? Since *c* has more first-preference votes than *a* has votes in total, *a* cannot be elected, by **plurality**. But adding two *ba* ballots would make *a* the Condorcet winner, and so *b* cannot be elected, by **Condorcet** and **mono-add-top**. And similarly *c* cannot be elected, because adding five *cb* ballots would make *b* the Condorcet winner. Thus, whichever candidate was elected, at least one of the three properties would be violated! (Of course, our rule could declare the result of Election 3 to be a tie; but this would lead to a contradiction in a similar way.)

It seems that most of the Condorcet-based properties discussed in the Social Choice literature would in fact elect *a* in Election 3, and so violate **plurality** (whereas AV elects *c* and DAC elects *b*). How seriously one regards the failure of plurality depends on how one interprets truncated preference listings, and that in turn may depend on the rubric on the ballot paper. If the 12 *c* voters are merely expressing indifference between *a* and *b* and not hostility to them, and so can be treated in exactly the same way as if half of them voted *cab* and half voted *cba*, then the violation is not too serious. But if, by plumping for *c*, these voters are not just saying that they prefer *c* to *a*, but that they want *c* and definitely do not want *a* (or *b*), and if the seven *b* voters also definitely do not want *a* (or *c*), then it is clear that *c* has more support than *a* and so *a* should not be elected.

	abc	3	*acb*	2
Election 6:	*bca*	3	*bac*	2
	cab	3	*cba*	2

Theorem 2 says that if an election rule satisfies Condorcet's principle, then it cannot possess any of the seven properties that are crossed in the column headed 2 in Table 1. This is a lot to prove. Fortunately most of it can be proved by considering variants of Election 6 above. The only bit that cannot is the incompatibility of **Condorcet** with **participation**; this is proved by Moulin[2], and I shall not attempt to reproduce his proof here. The following proof of the rest of Theorem 2 invokes the axioms of symmetry and discrimination, for a precise statement of which see Woodall[4].

So suppose we have an election rule that satisfies **Condorcet**. By symmetry, the result of this rule applied to Election 6 above must be a 3-way tie. But by the axiom of discrimination, there must be a profile *P* very close to the one in Election 6 (in terms of the *proportions* of ballots of each type) that does not yield a tie. So our election rule, applied to profile *P*, elects one candidate unambiguously; and there is no loss of generality in supposing that this candidate is *a*. However, there are ways of modifying the profile *P* so that *c* becomes the Condorcet winner, so that our election rule must then elect *c* instead of *a*. This happens, for example, if all the *bac* ballots are replaced by *a*; and the fact that this causes *c* to be elected instead of *a* means that our election rule does not satisfy **mono-raise-random**, **mono-raise-delete**, **mono-sub-top** or **mono-sub-plump**. It also happens if all the *abc* ballots are replaced by *a*, and this shows that our election rule does not satisfy **later-no-help**.

To prove that our election rule does not satisfy **later-no-harm**, it is necessary to consider a slight modification of the profile in Election 6, in which the second and third choices are deleted from all the *abc*, *bca* and *cab* ballots. Again, our election rule, applied to this profile, must result in a 3-way tie. But again, there must be a profile *P'* very close to this (in terms of the *proportions* of ballots of each type) that does not give rise to a tie, and we may suppose that our election rule elects *a* when applied to profile *P'*. But if we replace the *a* ballots in *P'* by *abc*, then *b* becomes the Condorcet winner, and so must be elected by Condorcet's principle; and this shows that our election rule does not satisfy **later-no-harm**.

Together with the result of Moulin[2] already mentioned, this completes the proof of Theorem 2, that an election rule that satisfies **Condorcet** cannot satisfy any of the seven properties crossed in the column headed 2 in Table 1.

Theorem 3 is a result that looks superficially similar to Theorem 2, and the proof is similar in character but much harder. The theorem says that if an election rule satisfies **majority**, **later-no-help** and **later-no-harm** then it cannot possess any of the seven properties crossed in the column headed 3 in Table 1. This is a substantial improvement on the result sometimes known as “Woodall's impossibility theorem”[3], which asserts that there is no election rule that satisfies **plurality**, **majority**, **later-no-help**, **later-no-harm** and **mono-sub-top**. In obtaining the improvement, I have needed to adopt an axiom of discrimination that is somewhat stronger than the one stated in Woodall[4], although one that must surely still hold for any real election rule. I am also grateful for help from my research student, Ben Tarlow.

A1	A2		A3	A4		A5	A6	
a	*ab*	0.34	*a*	*ab*	0.34	*ab*	*ab*	0.3
b	*b*	0.33	*b*	*b*	0.3	*b*	*ba*	0.3
c	*c*	0.33	*c*	*c*	0.36	*c*	*c*	0.4

Because the proofs of the different parts of Theorem 3 are quite complicated, I shall just sketch the proof of the easiest part, which is that there is no election rule that satisfies **majority**, **later-no-help**, **later-no-harm** and **mono-sub-plump** (or **mono-sub-top**). Suppose, on the contrary, that we have a rule that satisfies these four properties. The first part of the proof is to show that it must elect *a* in election A1 and *c* in election A3 in the above table. This is not too difficult to prove, using symmetry and **mono-sub-top**, provided that neither of these elections results in a tie. However, although it may seem highly implausible that either of them should yield a tie, I cannot see any way of proving that this is impossible. Instead, I have used the strong form of the axiom of discrimination in order to show that, if it does happen, then one can vary the proportions 0.34, 0.33, 0.3 and 0.36 in these profiles by very small amounts in a consistent way so as to obtain very similar profiles in which it does not happen.

The rest of the proof is much easier to explain. Let us write $X \to x$ to mean that x is definitely elected in Election X (that is, with probability 1), and $X \not\to x$ to mean that x is definitely not elected in Election X (that is, x does not even tie for election in Election X). Also, $\Rightarrow$ is used to mean "implies that". Therefore

$\mathrm{A1} \to a \Rightarrow \mathrm{A2} \to a$ by **later-no-harm**,

$\mathrm{A2} \to a \Rightarrow \mathrm{A2} \not\to b$ (clearly),

$\mathrm{A2} \not\to b \Rightarrow \mathrm{A4} \not\to b$ by **mono-sub-plump**,

$\mathrm{A3} \to c \Rightarrow \mathrm{A3} \not\to a$ (clearly),

$\mathrm{A3} \not\to a \Rightarrow \mathrm{A4} \not\to a$ by **later-no-help**,

$\mathrm{A4} \not\to a$ and $\mathrm{A4} \not\to b \Rightarrow \mathrm{A4} \to c$,

$\mathrm{A4} \to c \Rightarrow \mathrm{A5} \to c$ by **mono-sub-plump**,

$A5 \to c \Rightarrow A5 \not\to b$ (clearly),

$A5 \not\to b \Rightarrow A6 \not\to b$ by **later-no-help**.

However, majority requires that A6 should result in the election of either *a* or *b*, and the axiom of symmetry therefore requires that *a* and *b* should tie for election in A6, each with probability ½. This contradiction shows that there can be no election rule satisfying the four properties described.

The details of this proof, and the proof of the rest of Theorem 3, can be found in Woodall[5], which is not yet published but is available from the author at the Department of Mathematics, University Park, Nottingham, NG7 2RD, email drw@maths.nott.ac.uk .

5. Conclusions

In attempting to find a single-seat preferential election rule that satisfies majority and is generally monotonic, I have come up with only one rule, DAC, that I would be prepared to recommend as preferable to the Alternative Vote, and then only when the count is carried out by computer. DAC is much more complicated than AV, and I have not given great thought to how one would implement it on a computer, but I do not think there would be any great difficulty unless the number of candidates was unrealistically large. DAC admittedly fails to satisfy one important property of AV, that later preferences should not count against earlier preferences, but in return for this it gains five monotonicity properties that AV does not possess, including the very strong **participation** property, and so I would regard it as preferable.

However, DAC only works for filling a single seat, and I have not so far found any sensible way of extending it to multi-seat elections. The major remaining problem seems to me to be to find a multi-seat preferential election rule that satisfies the Droop Proportionality Criterion and is generally monotonic. It is not clear whether one can do this by modifying DAC, or whether it will be necessary to start afresh with a new idea.

From the mathematical point of view, there is still a great deal of work to be done on single-seat elections. The general problem is to determine which sets of the properties listed in Table 1 are mutually compatible. The examples discussed in Section 3 and the impossibility theorems in Section 4 give some answers. For example, Theorems 2 and 3 show that both FPP and AV possess maximal compatible sets of these properties, and that moreover these are the only two maximal compatible sets of properties that include both **later-no-help** and **later-no-harm**. Surprisingly, I have not been able to prove that the properties possessed by DAC form a maximal compatible set; Theorems 2 and 3 show that one cannot add either **Condorcet** or **later-no-harm** to these properties, but I cannot prove that one cannot add **mono-raise-random** or **mono-sub-top** (although this seems unlikely, since these last two are extremely strong properties, which hardly any election rules seem to possess). Another problem of this type is to determine whether there is any rule that satisfies **majority**, **Condorcet** and either **mono-add-top** or **mono-remove-bottom**. While problems of this type may seem to have little direct relevance to STV, the ideas generated by attempts to solve them may turn out to be more relevant than at first appears, and in any case we cannot afford to know less about such questions than our opponents do.

References

1. P C Fishburn, Condorcet social choice functions, *SIAM Journal on Applied Mathematics* 33 (1977), 469-489.

2. H Moulin, Condorcet's principle implies the no show paradox, *Journal of Economic Theory* 45 (1988), 53-64.

3. D R Woodall, An impossibility theorem for electoral systems, *Discrete Mathematics* 66 (1987), 209-211.

4. D R Woodall, Properties of preferential election rules, *Voting matters* Issue 3 (1994), 8-15.

5. D R Woodall, Monotonic single-seat preferential election rules, submitted.

Some comments on *Replaying the 1992 general election*

I D Hill

David Hill is Chairman of the ERS Technical Committee

At the time of the 1992 general election, Patrick Dunleavy, Helen Margetts and Stuart Weir conducted research designed to indicate how Britain would have voted under alternative forms of voting. Their report [1,2] states that the result "poses a problem for STV advocates" in that the allocation of seats is far from proportional by first preferences and severely disadvantages the Conservatives. They are very forthright in their claims that the study shows what would actually have happened. A subsequent letter[3] hoped that the Electoral Reform Society would "address the problems for STV that our ... study identified". It is, of course, not possible fully to address such problems without the data, and I am grateful to the authors for letting me have a copy.

In the comments that follow, I have concentrated entirely on the STV part of the document, ignoring the work that they also did on Alternative Vote, Additional Member, and List PR systems.

The data were obtained from a sample of 9614 people across 13 regions of the UK (excluding Northern Ireland). The sampling and interviewing was done by ICM, using their professional experience of getting a representative sample within each region. The interviewees were given a ballot paper of 17 candidates, in sections by party, their names being those of actual candidates in the general election in that region. They always consisted of 4 each from the 3 main parties, plus 5 others who included the nationalist parties in Scotland and Wales. Within this pattern, the aim was to give a mix of well-known and lesser-known, of men and women, etc.

The country was divided into 5-member constituencies so far as possible, but with some 4-member ones, by combining the actual single-member constituencies within each region. There were 133 such constituencies, consistently misquoted as 123 in their reports.

Much trouble was taken to get representative samples, but for analysis the regional results were reweighted for each multi-member constituency “to produce distinctive local profiles”. I do not doubt that this was done with good intentions but, so far as I can see, the anomalous results that “pose a problem for STV advocates” result almost entirely from this reweighting.

My analysis has necessarily had to be slightly incomplete because of some missing files. I am told that some computer discs have become unreadable and these files are going to be difficult to retrieve, so it seems better to go ahead with reporting what I can without them. Those missing concern all four of their East Anglian constituencies, 5.4% of the total data, and three of the Greater London ones (those that they call Richmond and Kingston, Hillingdon, and Central London). We can derive the regional Greater London results, from the other constituencies in the region, but not the reweighting for each of these three missing constituencies.

Of the available files, there are some that show trouble in the data in that some spurious figure zeros appear, that lead much of the data to be ignored by my STV computer program that was used. Luckily only one of these instances leads to a different result by political party from what they found, but there are also four others not suffering from this particular trouble, where the results by party seem to have been incorrectly reported. For the reweighting of the data, the authors say that “the 1992 general election results provide a complete picture of people's first preferences” so they use those to reweight the voting patterns. Even if this actually gave an improvement, I completely disagree with the beliefs behind it. A very important reason for wanting electoral reform is that election results at present do *not* show people's true preferences. Common observation shows that vast numbers of people vote tactically, not for what they would most like to see but for candidates who, they think, have some chance of success, and trying to keep out the party they most dislike. The *squeeze* of the third candidate in by-elections is notorious and a similar effect in general elections, to a lesser extent, certainly exists. Whether a better electoral system would make much change in voters' stated preferences or not we simply do not know; until we try the real thing the evidence is not available.

Having done the reweighting, for better or for worse, they report (in their Table 11):

	Con	Lab	L/D	others
Pure proportionality	273	222	114	25
STV	256	250	102	26

and this is what they say that STV supporters have to ponder. If we do the analysis by what appear to be the original data for each region, without such reweighting, it means using the same voting pattern for every multi-member constituency within the region, which is a crude model and often unrealistic but is probably the best that can be done with the data available. The results (with some assumptions for missing files) are:

	Con	Lab	L/D	others
Pure proportionality	273	222	114	25
STV	274	230	108	22

I think that it is they who have some pondering to do.

I cannot see how the numerical values of their reweighting were derived, but my requests for clarification have not been successful. If, as I believe, it was intended to bring the first preferences, by party, closer to the general election votes, it does not seem to have done so. The results are in the large table.

If anything the results after reweighting seem further from the general election results than do the raw ones, and certainly the Conservatives have been marked down.

It might be claimed that it is the individual multi-member constituency figures that matter rather than these overall ones, so I have looked at one constituency in detail to see whether that improves the picture. I chose to do this for my home constituency which, in their scheme, would be the combination of the present constituencies of Herts SW, Herts W, Hertsmere, Watford and St Albans. As an example of their reweighting, in this constituency every vote in the raw data with a Conservative first preference has been treated as 98 identical votes, every vote with a Labour first preference as 121 identical votes, every vote with a Liberal Democrat first preference as 87 identical votes and every other vote as 94 identical votes.

For this constituency I find:

	Con	Lab	L/D	other
General election	53.3%	25.1%	20.3%	1.3%
STV (raw)	52.7%	23.0%	20.0%	4.3%
STV (reweighted)	51.2%	27.6%	17.3%	4.0%

	General election				STV (raw figures)				STV (after reweighting)			
Region	Con	Lab	L/D	other	Con	Lab	L/D	other	Con	Lab	L/D	other
East Anglia	48.7	37.1	12.7	1.5	-	-	-	-	-	-	-	-
East Midlands	48.9	35.2	14.8	1.1	46.5	39.2	12.0	2.3	46.3	39.5	12.0	2.3
Greater London	45.2	35.1	18.0	1.7	41.0	37.1	18.1	3.8	*40.3*	*39.5*	*16.0*	*4.2*
North West	39.4	45.1	14.2	1.3	37.6	46.3	13.9	2.1	36.7	47.3	13.8	2.2
Northern	29.6	55.1	15.2	0.2	26.0	56.8	15.2	2.0	26.1	56.8	15.1	2.0
South East	54.6	20.4	23.7	1.2	52.7	23.0	20.0	4.3	52.7	23.3	19.8	4.3
South West	48.1	17.2	32.8	1.9	49.3	21.7	24.8	4.1	49.2	21.9	24.8	4.1
West Midlands	49.4	34.1	15.7	0.7	45.2	40.5	10.7	3.6	44.3	41.6	10.4	3.7
Yorks & Humber	37.5	45.7	15.5	1.3	39.0	44.2	12.8	4.1	38.3	45.0	12.5	4.2
Highlands	38.0	11.0	20.1	30.9	23.2	28.0	17.4	31.4	22.0	28.5	18.9	30.5
Strathclyde	19.8	49.9	7.8	22.5	24.1	47.2	5.2	23.5	23.3	48.0	5.2	23.6
East Central Scotland	29.9	34.1	14.8	21.2	28.0	36.0	18.3	17.7	28.4	35.2	18.7	17.6
Wales	26.0	50.8	11.8	11.4	32.0	43.7	13.1	11.2	31.1	44.1	13.2	11.6

Percentage share of votes by region

East Anglia missing, italic figures approximate due to missing files

Apart from slightly reducing the *others* figure, which is far too big nevertheless, has the reweighting helped? I doubt it.

I am well aware that it is much easier to criticize such a study than to perform one, but it does seem to me that a better scheme would have been to take their 9614 interviews equally from each of their 133 multi-member constituencies, i.e. about 72 per constituency, and then use the results in raw form. It might be argued that 72 votes are not many for electing to 5 seats, but that is all you get with a total of 9614. You do not get any more actual information by using the same figures many times with reweighting.

The authors also comment on “some apparently extraordinary results — as with the election of 5 Green MPs in the south east region”, and that only 2 of the 5 would survive if Meek rules were used (I make it only 1 of the 5 actually). In interpreting this we need to remember that it is the same set of votes being analysed over and over again, and the identical person as Green candidate, merely with different reweighting for each constituency in the region. That may have been the best that could be done in the circumstances, but I wish they would not claim that this is what would actually happen in practice. Again it is the reweighting that has produced the odd effect — no Green is elected if the original, unmodified, observations are used.

They seem to think it a disadvantage of STV that it can react with different results when the votes change only slightly. I think it an advantage that most constituencies become marginal for their final seat. At present it is only the marginal constituencies that have any real effect on who wins a general election. Under STV nearly every voter can feel that it is worth voting as it could make a difference.

They also make a point of the fact that “STV is only contingently proportional” if comparing seats with first preferences by party. So it should be. It often helps to explain a point such as this by using an exaggerated example. To repeat one that I have used elsewhere, if we have 9 candidates for 3 seats, A1, A2 and A3 from party A; B1, B2 and B3 from party B; C1, C2 and C3 from party C and the votes are

```
20   A1   B1
20   A1   C1
```

a party-based proportional system would have to elect A1, A2 and A3 as all first preferences were for party A, whereas STV will elect A1, B1 and C1 and appear to do badly if one insists on comparing seats with first preferences by party, but it has done what the voters have asked for, and that should be the aim of an electoral system.

What is more their data show that, of all interviewees who selected at least two preferences with each of their first two from the three main parties, only 79% of them chose the same party for both choices. If this is nothing like what would happen in practice, then the exercise cannot be quoted as meaningful in this respect. Their report claims strongly that their figures do represent what would happen in practice, but they cannot have it both ways; if they are right in that, then the authors' wish to see party proportionality by first preferences is not shared by the electorate. I believe that the wishes of the electors are what matter.

My overall conclusion is that it has been claimed that STV advocates have some problems to deal with, but in fact it is the authors of the study who need to deal first with the problems that they have created.

References

1. P Dunleavy, H Margetts and S Weir. *Replaying the 1992 general election* (LSE Public Policy paper number 3), 1992.

2. P Dunleavy, H Margetts and S Weir. Proportional representation in action: a report on simulated PR elections in 1992. *Representation*, vol 31, issue 113, pages 14-18.

3. S Weir. Letter to the editor. *Representation*, vol 32, issue 117, page 20.

Added in this reprinting

Brian Lawrence Meek, M.Sc, FRAS, C.Eng, FBCS.
1934 - 1997.

Brian Meek died on 12 July 1997. He was a member of the Electoral Reform Society for many years, and in the annals of the Single Transferable Vote his name will surely be immortal. Alongside the three pioneers Hill, Andrae and Hare, the other great names are Droop, Gregory and Meek. Various others have made improvements from time to time. This is not intended as any disparagement of them — fine-tuning of the system is not to be despised; it all helps if well done. It was Meek though who re-thought the system from scratch for the age of the computer and put it upon a proper mathematical basis. It should be recorded that a major part of the Meek system was also devised, quite independently, by Douglas Woodall a little later.

It is a pity that, although Meek's system is simpler in principle and easier to understand than other versions of STV, it is too long-winded if tried by hand. A computer is necessary, and since not everyone is willing to use computers for counting all elections, it will be necessary for a number of years yet to keep the approximate methods, suitable for hand-counting, available too. However, for any organisation that is willing always to use computers for its elections it would be madness to continue with approximate methods when Meek's method is available.

We have lost a man who did something really great in this field. One day that fact will be common knowledge for all proponents of STV.

I.D. Hill

Issue 7, September 1996

Editorial

This issue contains five articles within the tradition that has now been established. This concentrates upon the properties of various STV algorithms as seen from examples or computer simulation.

In the first paper, Hugh Warren illustrates a counter-intuitive case of the application of STV where two halves are not the same as one whole.

My own article provides the results of a computer simulation of 'large' STV elections which casts doubt on the use of the hand-counting rules in that case.

David Hill provides a simple comparison between the hand-counting rules and the computer method due to Meek. In a separate article, he shows how one can compute with Meek how one's vote has contributed (or otherwise) to the elected candidates.

In another paper by Hugh Warren, an example is provided in which equality of preference does not have a property that one might reasonably expect. David Hill responds to this in the final paper of this issue.

On reviewing this material, I conclude that I should appeal for a broader spectrum of papers. STV is not just a minority interest. I am a member of the *John Muir Trust* which aims to preserve wild places in Scotland. The trustees are elected annually by the membership by STV using the Meek algorithm. (Nothing to do with me.) I have been given an impressive list by Eric Syddique of organisations known to ERS that use STV (107 in total, but omitting the *John Muir Trust*). Can I appeal to readers to send details of other organisations so that I can publish the list in a subsequent issue of *Voting matters*?

Brian Wichmann

On the lack of Convexity in STV

C H E Warren

Hugh Warren is a retired scientist

If the voters in a constituency are divided into two districts and the ballots are processed separately and the results in the two districts are the same, then there is said to be convexity if processing the ballots of all voters together gives the same result.

As Woodall[1] has pointed out, quoting an example of David Hill's, STV does not satisfy convexity. We give here a further example, in which the lack of convexity arises, not from the elimination of candidates as in David Hill's example, but from the transfer of surpluses. We assume that these transfers are made by the method currently recommended by the Electoral Reform Society, and which the Electoral Reform Society uses for its own elections — the Meek[2] method.

There are four candidates A, B, C, D, and three seats to be filled. The voting is as follows:

	District 1	District 2	Constituency
ABD	10	-	10
BAD	-	10	10
AC	-	8	8
AD	-	1	1
BC	8	-	8
BD	1	-	1
D	1	1	2
Totals	20	20	40
Elected	A,B,C	A,B,C	A,B,D

This further example reinforces Woodall's comment in the article quoted that ... *sadly, convexity is of no use to us*, as this seemingly ideal property conflicts with a more desirable property.

References

1 D R Woodall, Properties of Preferential Election Rules, *Voting matters*, Issue 3, pp 8-15, December 1994.

2 B L Meek, A new approach to the Single Transferable Vote, reproduced in *Voting matters*, Issue 1, pp 1-10, March 1994.

Large elections by computer

B A Wichmann

Introduction

By a large election, in this article we mean elections in which there are a large number of candidates, say over 100. Such an election was reported in reference 1, in which the periodicals to be retained in a library were to be decided. In that case, the Meek algorithm was used[4], but on re-running the same data with the Newland-Britton (ERS) rules[5], a disturbing fact was noted. Towards the end of the count, none of the remaining candidates were credited with any votes at all, so that the last few ‘seats’ were filled at random from the remaining candidates. This was quite inappropriate, since the number of journals that received some support in the votes was more than enough to fill all the places. Hence Woodall has defined the property No-support in reference 2 to cover this issue.

In this paper, we are concerned not with the limiting case of the ERS rules electing candidates without support, but with other large elections in which some candidates are elected despite having less than half the quota. In such situations, it might appear that the ERS rules might elect the ‘wrong’ person. Unfortunately, it is not easy to devise a means of determining the ‘right’ choice. Here we use random ballot papers with some characteristics of a real election.

UKCC

The United Kingdom Central Council for Nursing, Mid-wifery and Health Visiting (UKCC) election is the largest one conducted by Electoral Reform Ballot Services Ltd (at least, using STV). It is possible that other such elections could arise of this type if multi-national organisations undertake employee-council elections to satisfy the ‘Social Chapter’.

The data from the last UKCC election is impressive: 129 candidates for 7 seats with 62,216 ballot papers. The election is conducted to the ERS rules assisted by Rosenstiel's program.

Mr Wadsworth of ERBS has kindly given me the information above and also a print-out from the Rosenstiel program which gives for the seven elected candidates:

Candidate	First Preference Votes/ Quota	Stage when elected	Votes when elected/ Quota
A	112.8%	1	100.0%
B	17.3%	122	47.7%
C	19.8%	121	53.2%
D	21.0%	121	50.5%
E	16.0%	123	34.9%
F	11.1%	123	36.4%
G	11.0%	123	42.6%

The concern here is that since one candidate was elected on only about one third of the votes that had to be retained by the most popular candidate, can one be sure the correct choice was made? The result of that particular election is not being questioned, but the choice of algorithm for elections of this type.

Computer processing

Since the computer programs to conduct elections are not used for the large public elections, there is no experience in using these programs for very large elections. As noted above, Rosenstiel's program was used for UKCC, but this program is for assisting a manual count, and could not be used for the Meek algorithm (for instance). Although the programs for the ERS rules (by I D Hill) or that for Meek do not have hard limits, it is not immediately obvious that they could be used for elections as large as that for UKCC.

To determine the feasibility of using these programs on a PC (personal computer) for elections like UKCC, a program was written to construct a large number of random ballot papers. Of course, real ballot papers were not available, and even if they were, the data preparation problem would be formidable.

At this point, a major problem arose. Both Meek and the ERS computer programs allow for the storage of the complete set of preferences. If this information is written to temporary disc storage, then the programs will run quite slowly. However, the total storage for UKCC-like elections is around 8Mbytes, which is only just within the reach of current PCs. The obvious solution was to undertake modifications to both programs to take advantage of the fact that only a small fraction of the total possible number of preferences would be specified. In fact, the modification to Meek was very easy and undertaken, but that for ERS (which is much more complex in computer terms) was too difficult. In any case, both programs were successfully run with random data on my home computer.

The conclusion from this study was that running the Meek or ERS rules on a modern PC would be possible for large elections. However, program modifications would be desirable to ensure that the programs kept within system limits. It was also observed that both programs produced result files which were excessive in size (and too big to print with convenience). With the preferences kept in main store, the time both programs took to execute was limited by the speed of processing; moreover, it was linear in the number of ballot papers. The time taken for the programs on my home computer was about 500 seconds per 10,000 papers for Meek and about ten times faster than that for the ERS rules. These times are clearly minor compared with the data preparation overheads in undertaking such counts.

Random UKCC-like data

Having determined that it is feasible to undertake UKCC-like elections on a computer with either Meek or the ERS rules, we now wish to see if there is a significant risk of either algorithm producing the 'wrong' result.

For this part of the study we use simulated data with only 1,000 papers, rather than the 62,000 that were actually recorded for UKCC. The reason for this reduction is to save on the computer time required, since many elections must be analysed (in fact, 100 elections were used). However, to have a realistic chance of determining the effect of using either algorithm, it is clear that the ballot papers must adhere to some of the characteristics of the real data.

The method used to construct the papers was to use a random number generator, but to use some of the characteristics of the UKCC election to determine the distribution functions used. The two major parameters are the popularity of each candidate and the length of each ballot paper. We can estimate the popularity of each candidate in the real election by means of their (known) number of first preference votes. Hence the popularities of the candidates in the simulated elected were adjusted so that the leading candidate had more than the quota of first-preference votes, candidates numbered 2 to 20 had reducing popularity of 95% of the previous candidate, and the remaining candidates had a constant popularity of 95% of the 20th candidate. The reason for this constant tail is that if the 95% rule was carried on, it was observed that the lower candidates had virtually no votes at all.

The distribution of the length of preferences chosen was as follows: For those expressing a single preference: 8.0% of the papers; for two preferences, 8.7%; for 3: 9.4%; for 4: 10.1%; for 5: 10.9%; for 6: 11.6%, for 7: 12.3%, and for 8 to 11 preferences: 7.2%. This distribution increases linearly to 7, the number of candidates then drops to a constant amount.

We can now compare a randomly produced set of papers with those above from UKCC. In this case with random ballots, the quota becomes 1000/8=125, instead of 62216/8=7777 for the real election. The table entries below and for the comparative table for UKCC are expressed in proportion to the quota to give directly comparable data.

Candidate	First Preference Votes/ Quota	Stage when elected	Votes when elected/ Quota
A	129.6%	1	100.0%
B	16.8%	119	53.3%
C	10.4%	121	46.1%
D	13.6%	121	46.9%
E	12.0%	121	45.3%
F	8.8%	121	44.1%
G	9.6%	121	42.2%

The pattern is clearly similar. We need not be concerned about minor differences, since the study is of elections of this general type. To generate each set of ballot papers merely requires as input the three integer seeds for the random number generator. In consequence, all the data presented here which is based upon a set of 100 elections can be recomputed from 300 integers. The seeds for the election in the above table were 1, 1 and 18.

Comparative tests: Meek versus ERS

We now have the ability to generate large election data and process the results with two algorithms: Meek and the ERS rules. The remaining problem is to determine characteristics of the results which would decide between the two. In fact, four different tests were applied as follows:

Non-transferables: In this test, the number of non-transferable votes of each algorithm are compared. The 'better' algorithm is the one which gives the lower figure.

Condorcet: In this test, we take those elections produced in which the two algorithms elected different candidates. We then compare the first candidate elected by Meek who was not elected by ERS with the first candidate elected by ERS who was not elected by Meek. The comparison is by Condorcet. Since there is no correlation between the votes for different candidates, the winning algorithm for this test is the one which has the higher number of Condorcet winners.

No-hopers: In this test, we eliminate the candidates with no realistic hope of being elected, namely the candidates numbered 21-129, so there are 20 candidates. Again, since there is no correlation between the votes of different candidates, the winning algorithm for this test is the one for whom this change makes the least difference. In other words, one is expecting the removal of the no-hope candidates to make no difference.

Steadiness: This test is that specified by I D Hill in reference 3. The test is applied when there is only one pair of candidates elected differently by Meek and ERS. The election is then re-run with only 8 candidates. The winner is the algorithm for which this makes the least difference to the result.

At this point, the author thinks that readers should reflect upon the tests above. If the results are against your favourite algorithm, will you be convinced that your algorithm should not be used for such elections?

We now consider the results of each of these tests:

Non-transferables: There is a consistent pattern with the number of non-transferable votes with each algorithm which can be summarised as follows:

Meek 559.0 (±18.8); ERS 482.6 (±13.7); Meek/ERS 1.159 (±0.031); where the range represents two standard deviations. Hence Meek consistently gives 16% more non-transferable votes.

Condorcet: Out of the 100 elections constructed with the random ballot papers, 30 produced a different result. Hence for these 30, the Condorcet test could be applied. The results were that for 24 cases, Meek elected the Condorcet winner, and for 6 cases, ERS elected the Condorcet winner.

No-hopers: In this test, we wish to know if the elimination of the no-hope candidates changed those that were elected. For Meek no change occurred for any of the 29 cases examined, but there were changes for all but three cases with ERS. Hence Meek is a clear winner here.

Steadiness: This test is applied to the 29 cases in which there was one difference between the two algorithms. To pass the test, the result of the election with just eight candidates must be the same as for the full election. Meek passes the test for all of the 29 cases, and ERS 6 times (and failed 23 times). Again, Meek is the clear winner of this test.

The above analysis understates the differences between the two algorithms. Of the 29 cases that can be compared for steadiness, the following table indicates how the results compared in 20 cases:

	Meek Elects	ERS Elects
129 candidates	[S]+A	[S]+B
20 candidates	[S]+A	[S]+A
8 candidates	[S]+A	[S]+A

Here `[S]` represents a set of six candidates and `A` and `B` are different candidates not in the set `[S]`. In other words, ERS reverts to the Meek result when the no-hope candidates are removed, and this reversion is retained when only 8 candidates are considered. This is clearly strong evidence that the full election using the ERS rules produces the 'wrong' result.

Conclusions

The study indicates that it is feasible to use computer algorithms such as Meek on a PC for elections as large as that for UKCC (although the data preparation problem has not been considered). Furthermore, a comparison between Meek and ERS shows that Meek is superior except for the number of non-transferable votes. The increased number of non-transferable votes is clearly secondary to producing the 'correct' result, and from that perspective, the Meek algorithm appears to be superior. The fact that random papers from no-hope candidates can change the result is strong evidence against the ERS rules.

Of course, this study only relates to elections with a large number of candidates. It can hardly be considered a criticism of Newland and Britton, since it is doubtful whether they ever conducted an election of the size considered here. [Added in this printing: See Issue 8, page 3].

Acknowledgements

I should like to thank Joe Wadsworth of ERBS for providing me with a summary of the UKCC election result 'sheet'. Also, David Hill and Douglas Woodall provided many comments on a draft of this paper which I hope has resulted in a clearer presentation here.

References

1. B A Wichmann. Two STV Elections. *Voting matters*, Issue 2, pp7-9, September 1994.

2. D R Woodall. Properties of Preferential Election Rules. *Voting matters*, Issue 3. pp8-15, December 1994.

3. I D Hill. The comparative steadiness test of electoral methods. *Voting matters*, Issue 3, p5, December 1994.

4. B L Meek. A New Approach to the Single Transferable Vote. Reprinted in *Voting matters*, Issue 1, pp1-10, March 1994.

5. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote, second edition, ERS, 1976.

Meek style STV − a simple introduction

I D Hill

Until recently, David Hill was Chairman of the ERS Technical Committee

For its 1996 Council election, ERS used the Meek counting rules, instead of the Newland and Britton rules that are suitable for counting by hand. Now that there is sufficient availability of computers, I believe that ERS owes it to itself and to its members to use the best rules of which we are aware.

However many people seem to be muddled as to what this involves and some seem to be sadly misinformed. It is therefore desirable to have available a simple listing of what is the same and what is different in these systems.

It needs to be said clearly that there is no intention of abandoning STV. The system adopted (taking its name from B L Meek who first proposed it) retains all the essential features and aims of STV, but uses the power of modern computers to get a closer realisation of the voters' wishes, better meeting all the traditional STV virtues.

Some of the main changes were mentioned by Robert Newland in *Comparative Electoral Systems*, section 7.8(c). He wrote that these further refinements ‘which would be likely rarely to change the result of an election but which greatly lengthen the count, are not recommended’. At the time, that was probably a reasonable judgement but information gained since then has shown it to be untrue that the result would rarely change, whereas lengthening the count is unimportant when counting is by computer where, either way, the counting time is trivial compared with the effort needed to input the data.

Meek style STV - what is the same?

1. Each voter votes by listing some or all of the candidates in order of preference.

2. Each voter is treated as having one vote, which is assigned initially to that voter's first-preference candidate.

3. A quota is calculated, as the minimum number of votes needed by a candidate to secure election.

4. If a candidate receives a quota of votes or more, then that candidate is elected, and any surplus votes (over the quota) are transferred to other candidates in accordance with the later preferences expressed by the relevant voters.

5. If, at any stage of the count, no surplus remains to be transferred, but not all seats are yet filled, then the candidate who currently has fewest votes is excluded. Votes assigned to that candidate are then transferred to other candidates in accordance with the later preferences of the relevant voters.

Meek style STV - what is different?

6. All surpluses are transferred simultaneously instead of in a particular order.

7. Surpluses are taken, in due proportion, from all relevant votes, not only from those most recently received.

8. To make that work properly it is necessary to give votes to already-elected candidates and not "leap frog" over them. This does not waste votes as the same number are transferred away again, but now in due proportion to all relevant votes.

9. Whenever a candidate is excluded, the count behaves as if that candidate had never existed (except that anyone previously excluded cannot be reinstated).

10. Whenever any votes become non-transferable, the quota is re-calculated, based on active votes only. This lower quota then applies not only for future election of candidates, but also to already-elected candidates giving them all new surpluses.

11. No candidate is ever elected without reaching the current quota.

12. For surpluses, every relevant vote goes to the voter's next choice, at fractional value. If there is no next choice, the fraction becomes non-transferable.

13. At an exclusion all the relevant votes are dealt with at once. There is no doing one little bit at a time.

14. The only disadvantage is that it is too tedious to do by hand, but has to be by computer.

Examples

1. A very simple, though artificial, example of the superiority of the Meek method is seen in 4 candidates for 3 seats. If there are only 5 voters and the votes are: 2 ABC, 2 ABD, 1 BC it is obvious to anyone, whether knowing anything of STV or not, that the right solution must be to elect A, B and C, as the Meek method does, yet traditional hand-counting rules elect A and B but declare the third seat to be a tie between C and D.

2. In a real election held recently, I shall call 4 of the

candidates A, B, C and D of whom at the last stage, A and B had each been elected with a surplus, C had been excluded and D was still continuing, to be either the last elected or the runner-up. Four of the votes gave preferences as ABCD, ACBD, CABD and ABD. As C had been excluded, these became identical votes, each now having A as first preference, B as second and D as third. The Meek method would have treated them identically, but the rules actually in use gave D wildly different portions of these votes, as follows:

Vote	Rules as used				Meek rules			
	Portion of vote assigned to				Portion of vote assigned to			
	A	B	C	D	A	B	C	D
ABCD	0.72	0.28	-	-	0.471	0.285	-	0.244
ACBD	0.72	-	-	0.28	0.471	0.285	-	0.244
CABD	-	-	-	1.00	0.471	0.285	-	0.244
ABD	0.72	0.28	-	-	0.471	0.285	-	0.244

The variation between all of the vote going to D, and none of it doing so, is really startling.

How was my vote used?

I D Hill

If an election has been conducted by STV using Meek counting, and the final keep values have been published (as I think that they should be), any voters who remember their preference orders can work out how their votes were used, as follows.

Suppose you voted for Bodkins as first preference, for Edkins as second preference, etc., where their final keep values were published as 0.310, 0.772, etc., as shown in the table below. The first thing to do is to make such a table with the order of preference that you actually used for the real candidates and fill in their published final keep values in column (3).

Always start with 1.000 as the first item, one line above your first candidate, in column (6), and then in each row in turn, fill in columns (4), (5) and (6) using the rules shown.

Preference	Candidate	Final keep value	Previous vote remaining	Vote kept	Vote remaining
(1)	(2)	(3)	(4)	(5)	(6)
			previous (6)	(3) × (4)	(4)−(5)
					1.000
1	Bodkins	0.310	1.000	0.310	0.690
2	Edkins	0.772	0.690	0.533	0.157
3	Atkins	0.000	0.157	0.000	0.157
4	Dawkins	0.702	0.157	0.110	0.047
5	Firkins	1.000	0.047	0.047	0.000
6	Gaskins	0.570			
7	Catkins	0.978			

When an excluded candidate appears, such as Atkins above, the keep value is 0.000, so no part of the vote is kept. When a candidate was either the runner-up or the last to be elected, such as Firkins, the keep value is 1.000, so that candidate keeps everything received and later preferences get nothing.

Column (5) tells how the vote was used. 0.310 of it went to help elect Bodkins, 0.533 of it went to help elect Edkins, 0.110 of it went to help elect Dawkins and the remaining 0.047 went to Firkins and, if Firkins was runner-up, was unused.

I have been asked by someone who has seen the above to produce something similar for traditional-style STV (and, in particular, for Newland and Britton rules, second edition). Having had a look at the problem, I have concluded that, for anyone who really understands what is going on, the information can be derived from the result sheet in an *ad hoc* way, but that it is not possible to do anything as general, or as simple, as the above.

This should be offered as an exercise for those who think the traditional rules simpler than the Meek rules. Let them do it. I do not deny, of course, that the traditional rules are less long-winded for making a hand-count, but in every other way, in principle and in practice, the Meek rules are much the simpler.

STV and Equality of Preference

C H E Warren

The Single Transferable Vote is a preferential voting system, in which the voter has to list the candidates in the order in which he prefers them.

One of the questions which is asked is whether a voter should be permitted to express an equality of preference between two candidates whom the voter assesses as equal in his judgement. My view is that the expression of equality of preference should be permitted in principle, although of course it would complicate both the voting and the subsequent count.

If a voter does express an equality of preference between two candidates A and B, then it is assumed that this is tantamount to his expressing two half-votes with non-equal preferences, one half-vote for A followed by B, and the other half-vote for B followed by A, but the half-votes otherwise identical.

However, Bernard Black is concerned that, if equality of preference is permitted, a voter may see neither of his equal preferences elected, whereas if the voter had given one of his two a clear preference then at least he would have got that one elected.

The following example of an election for 3 seats from 6 candidates by 30 voters, for which the quota is 7.5, exemplifies Black's concern. 29 of the voters vote as follows:

```
1 AB
1 BA
9 CAB
1 CEF
9 DBA
1 DEF
3 EF
4 F
```

The thirtieth voter is undecided between A and B. If this thirtieth voter votes AB, or votes BA or expresses an equality of preference between A and B, then the votes after the surpluses of C and D have been transferred are:

AB		BA		½AB + ½BA	
A	4.25	A	3.25	A	3.75
B	3.25	B	4.25	B	3.75
C	7.5	C	7.5	C	7.5
D	7.5	D	7.5	D	7.5
E	3.5	E	3.5	E	3.5
F	4	F	4	F	4

We see that if the voter gives a clear preference for either A or B, then that one gets elected, because the other one is now eliminated and his votes then transferred to the preferred one. However, if the voter expresses equality of preference, then E is now eliminated, and E's votes then transferred to F who is elected, so that neither A nor B is elected. Hence Black's concern is justified.

The main benefit that is likely to arise from permitting equality of preference, as Douglas Woodall has said, is not for voters who are undecided between their top preferences, but for voters who want to put certain candidates as their bottom preferences, below a whole lot of candidates whom they do not know much about, but for whom being able to give equality of preference would be ideal.

David Hill has shown, in an unpublished paper, that, in a real election, this middle group of candidates whom the voter does not know much about is more likely to be of relevance with Meek[1] counting than with Warren[2] counting, because with Warren counting the count does not extend down to this middle group of candidates.

References

1 B L Meek. A new approach to the Single Transferable Vote. Reproduced in *Voting matters*, Issue 1, pp1-10. March 1994.

2 C H E Warren. Counting in STV Elections. *Voting matters*, Issue 1, pp12-13. March 1994.

Equality of preference − an alternative view

I D Hill

In the preceding paper[1], Hugh Warren states 'Hence Black's concern is justified', but the example from which he derives this opinion is not convincing. It really concerns the question of how a tie is to be resolved, since in each of his three cases the AB supporters have 7.5 votes and the EF supporters have 7.5 votes. This makes it critically dependent on using a version of STV in which the quota is defined to give precisely 7.5 as in Newland and Britton, second edition[2] and not 7.5 plus a minimal amount as in most versions of STV, such as Newland and Britton, first edition[3], for example. It also depends on the rule that anyone reaching the quota is to be deemed elected at once even though some other candidate could catch up if the process were continued.

I am not objecting to those features, but if we are prepared to base conclusions on examples that depend critically on them, it is easy enough to construct one that points to the opposite conclusion. Consider 4 candidates for 3 seats with an odd number, n, of voters who support A and B, and an equal number, n, who support C and D. The quota will be $n/2$ and if the AB party do not use equality, no matter how they arrange their votes between saying AB and saying BA, one of their candidates will have more than a quota, and the other less than a quota, on the first count. If the CD party all put C and D as equal first, each of their candidates will have exactly a quota on the first count and consequently either ACD or BCD will be elected.

It follows that Black's concern is *not* justified. In these extreme cases use of equality could either harm or help and it is not possible to know which. In reality such extreme cases rarely, if ever, occur. What would normally happen if equality were used would be for one of the two candidates to go out (either as excluded or elected) at some stage and then the relevant part of the vote would be transferred to the other candidate, so nothing would be lost.

References

1. C H E Warren. STV and Equality of Preference. *Voting matters*, Issue 7, p6, September 1996.

2. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote, second edition, ERS, 1976.

3. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote, first edition, ERS, 1973.

Issue 8, May 1997

Editorial

In this issue, a new format is being used, but without any change to the content or type of material being published.

It is hoped that future issues of *Voting matters* will be made available via the Internet. However, printed copies will continue to be made which can be ordered from ERS. Due to some limitations of the most straightforward means of producing material on the World Wide Web, the printed copies will be the master ones, and presentation on the Web may have some defects.

The first article which lists those organisations known to use STV is an example of material which should be available on the Web anyway. Given this, then updating the list can more easily be undertaken.

As before, I am concerned about the lack of variety in the authors of material. Electronic publication could easily encourage contributions from other countries.

Brian Wichmann.

Organisations using STV

The following is an alphabetical list of organisations known to use STV in the UK or the Republic of Ireland. In the interests of brevity, local organisations are not always included.

3M plc
Aberdeen University SRC
Adlerian Society of the UK
Allied Dunbar
Amnesty International
Association for Jewish Youth
Association of Municipal Engineers
Association of University Teachers (AUT)
Association of Teachers & Lecturers (ATL)
Association of Logic Programming
Automobile Association
Avon Cosmetics
Bar Council
Bardsey Island Trust
Bass plc
Beechlawn School
Birmingham Labour Group
Birmingham University
Bow Group
British Airports Authority
British Dental Association
British Psychological Society
British Association of Colliery Management
British Association of Dermatologists
British Association of Counselling
British Computer Society
British Council
British Humanist Association
British Medical Association
British Mensa Ltd
British Union of Anti-Vivisection
Brittle Bone Society
BUPA plc
Cambridge University Student Union
Campaign for Homosexual Equality
Cardiff Union Services
Celtic Film and TV Association
Church of England
Church of Wales
City Literary Institute
Committee of Vice Chancellors & Principals
Consumers' Association
Coopers & Lybrand
Crosslinks
Derbyshire E. R. Group
Drake & Scull Engineering Ltd
Du Pont UK Ltd
Eastern Electricity
East Midlands Electricity
Engineering Council
Electronic Data Systems
Express Newspapers Pension Ltd
Faculty of Public Health Medicine
Family Law Bar Association
Gateshead & South Tyneside LMC
General Dental Council
General Medical Council
Gilberd School
Glasgow Caledonian University
Gouldens
Greater Manchester Police
Greater London Unison
Guild of Hospital Pharmacists
Headmasters' Conference
Hoechst UK Ltd
ICL plc
Imperial Tobacco
Institute of the Motor Industry
Institute of Chartered Accountants
Institute of Civil Engineers
Institute of Electrical Engineers
Institute of Linguists
Institute of Management Services
Institute of Mechanical Engineers
Institute of Public Relations
International Association of Teachers of English as a Foreign Language
John Muir Trust
King's College London Students' Union
Leeds University Union
Lewisham & Kent Islamic Centre
London Borough of Sutton
London Electricity plc
London School of Economics Students' Union
Liberal Democratic Party
Liberty
Logica plc
Manweb plc
Mercury Communications
Methodist Conference
Midland Bank plc
Mountain Bothies Association
National Association of Teachers in Further & Higher Education (NATHE)
National Citizens' Advice Bureaux
National Federation of Housing Associations
National Freight Consortium
National Grid plc
National Westminster Group
National Power
National Union of Journalists
National Union of Mineworkers
National Union of Rail, Maritime & Transport Workers (RMT)
National Union of Students
National Union of Teachers
Neural Computing Applications Forum
News International Newspapers Ltd
Northeast Fife D.C.
Northern Electric plc
Northern Sinfonia
Norweb plc
Pensions Management Institute
Pensions Trust
Pharmaceutical Society of Great Britain
Powergen
Price Waterhouse
Professional Association of Teachers
Prudential Assurance
Royal College of General Practitioners
Royal College of Midwives
Royal College of Nursing
Royal College of Pathologists
Royal Statistical Society
Royal Town Planning Institute
Scottish Nuclear
Secondary Heads Association
Shantiniketan Centre, Southall
Shell UK
SeeBoard plc
Smith & Nephew plc
Solicitors Family Law Association

South Oxfordshire D.C.
Stoneham Housing Association
Southern Electricity plc
South East Electricity plc
South Wales Electricity plc
South West Electricity plc
Telegraph Newspapers
Total Oil Ltd
Theatrical Management Association
UK Central Council for Nursing, Midwifery and Health Visiting (UKCC)
UK Council for Graduate Education
University of Bristol
University of Wales Swansea Students' Union
University of Ulster Students' Union
Union of Democratic Mineworkers
Union of UEA Students
Yorkshire Housing Association
Yorkshire Water
Zionist Federation of Great Britain

The various companies named above will not be using STV to elect their Boards of Directors which are usually Yes/No ballots, but to elect Pension Fund Trustees. The accountancy partnerships of Coopers & Lybrand and Price Waterhouse use it to elect their Executive partners. These particular elections are unique in that, apart from partners retiring during the year, all partners are automatically candidates.

Quotation Marks

Dear Sir,

There are one or two matters I would like to comment upon.

In his article *Large Elections by Computer*, Dr Wichmann says there is strong evidence that the traditional method of STV counting produces the 'wrong' result. I would suggest that even with the use of quotation marks this is an unfortunate comment. The result is surely correct within the rules which have been used, and to suggest otherwise is to imply that there is something inaccurate, or wrong, with the count. It might lead to defeated candidates thinking they were defeated as a result of some procedural error by the Returning Officer, which would not be the case. It would be wiser to say that the election result might be different. I do not think we would wish to appear to cast doubts upon our own ballot organisation to count an election by STV.

The real problem with elections of this kind is the proportion of candidates to the number of places to be filled. In the UKCC example there were over 18 times more candidates than the 7 places to be filled. 129 candidates appears to offer those voting the widest possible choice, but the choice is unreal. Unfortunately few of the voters have sufficient knowledge about the candidates to be able to put more that a small number of candidates in preferential order. The candidates are allowed to provide information about themselves but there is still a great deal of information to read. One answer might be to re-examine the nomination process, with a view to there being more assentors to the nominations. The organisation may, of course, not wish to do this because it might create an unreasonable hurdle to nomination.

Dr Wichmann is under a misapprehension when he says *It can hardly be considered a criticism of Newland and Britton, since it is doubtful whether they ever conducted an election of the size considered here*. Major Britton and Mr Newland were closely involved with drafting the electoral arrangements for the UKCC and took a very close interest in the first two elections at Chancel Street to see how the counts went. The report of the first UKCC election records that 441 candidates were nominated and that 61,715 people voted. Therefore it would appear that there has been a decline in the number of candidates nominated. I can recall both Major Britton and Mr Newland being somewhat concerned at the number of candidates nominated for the first election, but thought the number would decline when it was realised that most of those nominated had no real hope of being elected. At the first two elections no candidate achieved the quota, the whole election consisting of exclusions, candidates being elected with a reduced quotient as votes became non-transferable. It was not until the third or fourth UKCC election that I recall being told that for the very first time a candidate had attained the quota during the count. My recollection is that Major Britton and Mr Newland would probably have recommended that the nomination procedure be amended if the number of candidates had not declined to a more manageable number for the voters.

E M Syddique, ERS

A reply

I owe readers an apology if they were under any misapprehension on the use of the quotation marks. Of course, there was no implication that the rules were not correctly applied; indeed the simulations I made assumed that. It is also worth noting that since I used artificial ballot papers, the implications for any specific election (like the last UKCC one) are unclear.

I cannot apologise to Major Britton and Mr Newland for not realising their involvement in the early UKCC elections which were even bigger than the one I analysed.

I believe that a major contributory factor to the results I obtained was that not only was one candidate elected with the quota, but that the others were elected with very much less than the quota. The re-computation of the quota undertaken by the Meek method therefore makes a bigger difference than would typically be the case.

Lastly, it seems to me that an advantage of STV should be that many candidates can compete. Hence introducing barriers to nominations seems against the spirit of STV.

B A Wichmann.

Are non-transferables bad?

I D Hill

Brian Wichmann[1] put forward four different tests of whether one vote-counting algorithm had done better than another and invited readers, before reading on, to consider whether they would regard failure on each test as a serious matter.

I did not cheat, but made the requested consideration before reading on. I concluded that I accepted his tests called Condorcet, No-hopers and Steadiness but I totally rejected his test called Non-transferables. I then found, not much to my surprise, that he had found Meek's method to be a clear winner (on his particular data) on the three tests that I accepted as valid, while Newland and Britton (2nd edition) rules had done ‘better’ on the Non-transferables test which I had rejected, so I think it important to explain just why I had rejected it.

My view is that everything should always be in accordance with what the votes say, in proportion to their numbers and, if some votes, in whole or in part, are entitled to transfer and do not indicate a wish to be transferred anywhere, then it is morally wrong not to make them non-transferable, in whole or in part as the case may be.

That being so we cannot say which of two methods is better on the basis of the number of non-transferables, until we know the cause of the difference. If method 1 shows more than method 2, we must ask whether this is due to method 1 making some votes non-transferable unnecessarily, or to method 2 failing to make votes non-transferable when they should be. With methods of which we know nothing except the outcome of this particular test, we can really say no more than that.

In the actual case, however, we do know the methods in detail and are aware that Meek's method never makes anything non-transferable except when it is right to do so. It follows that, if the Newland and Britton rules get a smaller number, it is they that are failing to do the right thing.

Reference

1. B A Wichmann. Large elections by computer. *Voting matters*, 1996, issue 7, 2-4.

Some Council Elections

B A Wichmann

Introduction

This paper is an analysis of some Council elections based upon computer simulation in a similar manner to two previous papers [1,2]. The analysis starts with (five) result sheets, since they are the publicly available record of the elections. The first stage consists of using a computer program to produce a set of ballot papers which reproduces the result sheet (or gets very close to that). The second stage consists of running a number of experiments based upon elections which select a random subset of the ballot papers. The third stage is a further analysis of the results.

This paper is concerned with STV elections in which there are no ‘party’ affiliations. Hence the voting patterns are different from those which applied in the Irish elections analysed in the first reference. The identity of the actual council elections used for this study is not stated here, since this is irrelevant and could detract from the conclusions which are thought to be relevant for all elections for several seats in which there are no parties involved.

Constructing ballot papers

Given a result sheet, then it is possible to construct a set of papers which would produce the same results. In producing such a set by hand, the obvious method is to work forward stage by stage. However if no transfers occur from candidate A (say), such a method will give preferences that, if A appears, stop at that point. In other words, preferences that are not required to produce the results as given in the result sheet are not given. Clearly, the voter will not necessarily do this, and more significantly, other algorithms may use subsequent preferences. Hence a more general means is required of producing ballot papers.

The program used in this study works as follows. The program computes transfer rates from A to B if candidate A was eliminated or had a surplus to transfer (and B was available for transfers). If no such transfer occurred, then an estimate is used based upon the first preferences for B.

Ballot papers are now constructed using a random number generator with an exact match for the first preferences. This set is then used as the starting point of an iterative process, working stage by stage, to obtain a very close fit to the actual result sheet. The program cannot necessarily obtain a perfect match when transfers of surpluses are involved. Experiments showed that the starting position which was dependent upon the seeds for the random number generator did not have a large effect on the accuracy of the final fit to the actual election.

An example

To give a fuller explanation of the method of constructing ballot papers from a result sheet, we give a simple example. Consider an election in which the votes for electing one candidate from 4 was:

```
10 AB
5 BCD
6 BAD
6 CDA
1 C
8 DAB
```

The result sheet from these ballot papers using Newland-Britton is:

	Stage 1	Stage 2	Stage 3
A	10	10	0
B	11	11	21
C	7	0	0
D	8	14	14
Non-T	0	1	1

Since we are concerned with a council election without parties, we consider each candidate in the same way. We can judge the overall popularity of each candidate from the first preference votes. We now construct a matrix to represent the probability of X being followed by Y in any preference (X could clearly be the last preference given, so Y is allowed to be the Non-Transferable option). For instance, given candidate D, then the preference specified after D is assumed to be A, B or C in the ratio 10:11:7 (since these are the ratios of the votes on the first preferences).

We can make a better estimate of the transfer probabilities, since we do have a limited amount of information from the result sheet. In this case, for stage 2 in which C is eliminated, we know that the next preferences were either D or non-transferable in the ratio 6:1, respectively. Hence, we can adjust our matrix accordingly. For stage 3, in which A is eliminated, the transfers were entirely to B, but the papers could have had a preference to C which would have been ignored. This clearly reflects the adjustments made to the matrix. The final matrix, based upon one hundredths of a vote, in this case becomes:

			TO		
FROM	NT	A	B	C	D
STRT	-	278	306	194	222
A	0	-	1000	194	222
B	0	278	-	194	222
C	143	278	306	-	857
D	0	278	306	194	-

The program now computes a trial set of ballot papers with an exact match on the first preferences, but using a pseudo-random number generator and the above matrix to produce the remaining preferences. Finally, adjustments are made to the papers to obtain a better match to the result sheet. The root mean square error is computed over the entries in the result sheet, which gives 0 in this case for the 3×5 entries, since we have a perfect match.

The ballot papers produced in this case (which depends upon the seeds used for the random-number generator) were:

```
2 ABCD
7 ABDC
1 ACBD
3 BADC
4 BCDA
2 BDAC
1 BDC
1 BDCA
1 C
2 CDAB
4 CDBA
1 DABC
4 DBAC
1 DBCA
1 DCAB
1 DCBA
```

There are clearly many differences between the initial ballot papers and the above. However, since there are 64 ways of voting, it is quite unlikely that 10 ballot papers would be identical as with the initial papers (and in this sense, the final set must be regarded as more likely than the starting set). The construction method in this case gives very few papers with incomplete preferences, since the result sheet had few non-transferables.

Five real elections

The results of running the program for the five elections are given in Table 1. The result sheets were from the application of Newland-Britton[3]. A very close fit was obtained in all cases. The entry *Next* gives the difference in the number of votes between the last candidate elected and the next highest. This figure is also divided by the number of votes to give a numeric indication of how close the choice of the last elected candidate is. For election B, the result was very close since this difference was a mere 14 votes (from 8739, ie 0.16%). In performing both Newland-Britton and Meek upon the ballot papers constructed by the program, only one result was obtained which was different from the actual result. For election B, Meek produced a result different from the actual election, but this is hardly surprising, due to the closeness of the final candidate elected.

The experiment

The experiment concerns the influence of candidates with no realistic hope of being elected upon the result. With the UKCC analysis[2], it was observed that such no-hopers had a bigger influence with Newland-Britton rules than with

Table 1: Five Council Elections

Election	A	B	C	D	E	Total
Candidates	17	17	16	13	12	75
Seats	4	4	4	4	6	22
Votes	5764	8739	9364	8486	1669	34022
Stages	13	15	12	10	10	
RMS error(votes)	0.05	0.04	0.41	0.06	0.28	
Next	128	14	221	75	7.46	
Next/Votes	0.0222	0.0016	0.0236	0.0088	0.0045	
Actual=New-Br	yes	yes	yes	yes	yes	
Actual=Meek	yes	no	yes	yes	yes	

Meek. In this case, 100 elections were run by selecting 200 ballot papers at random (repeated five times for each actual election). For these 100 elections, both Newland-Britton and Meek were run. The second row in Table 2 gives the number of times out of the 100 that the results from Newland-Britton and Meek were different. For the 500 elections the result was different for 88 cases, which implies that 4% of the candidates were treated differently.

The first row in Table 2 gives the number of candidates which were never elected in any of the 100 elections, called *no-hopers*. It would seem that this is not an unreasonable definition of those that have no chance of election, since we know that the number of first-preference votes is not always a good indication.

The 88 elections in which Newland-Britton/Meek gave a different result were now re-run with the no-hopers eliminated. The results of this are recorded in Table 2 in the rows with indented titles. In all but one case, the difference between the two algorithms was just one candidate. However, the result of the re-runs is somewhat confusing except for the simple case in which the elimination of the no-hopers makes no difference. The results in the table are classified as follows:

No change. In this case, the elimination makes no difference and hence these cases are not supportive of either Newland-Britton or Meek.

Revert to Meek. In this case, the result from Meek does not change, but that for Newland-Britton changes to that of Meek. Such a case is taken as supporting the use of Meek.

Revert to Newland-Britton. This is the exact opposite of the previous case and is taken as supporting the use of Newland-Britton.

Meek unchanged. In this case, the result for Meek does not change, but that for Newland-Britton does (but not to that of Meek). This case is regarded as supporting the use of Meek.

Table 2: Results of simulations

Election	A	B	C	D	E	Total
No. No-hopers	7	7	6	3	3	26
New-Br = Meek	78	84	83	86	81	412
New-Br≠Meek	22	16	17	14	19	88
No change	13	8	11	10	9	51
Revert to Meek	6	5	3	2	7	23
Revert to New-Br	1	2	1	1	0	5
Meek unchanged	2	1	0	1	0	4
Both change	0	0	2	0	1	3
Invert both	0	0	0	0	1	1
Other	0	0	0	0	1	1

Both change. In this case, both change to a different result. This is obviously not supportive of either algorithm.

Invert both. In this case, the results of both algorithms change to the previous result of the other one! Clearly not indicative of either Newland-Britton or Meek.

Other. None of the above, and again not supportive of either algorithm.

The overall count from the above classification is that 56 cases are neutral, 27 support Meek and 5 support Newland-Britton.

Conclusions

In appears that realistic ballot papers can be computed from the result sheets. However, it is difficult to validate this process, since at the moment, actual ballot papers are not available from real elections of any size. I would like to appeal for such ballot papers, perhaps in computer format, since such papers could be made available without revealing the source which surely would be satisfactory once the period of elected candidates had finished. All the election data obtained so far is for small elections for which the study above could not be applied.

The first result from this study is that Newland-Britton and Meek produce a different result for about 4% of the seats. The observed rate for the Irish elections in 1969 was 2.8% (3 out of 143) and for 1973 was 4.9% (7 out of 143). The difference between 1969 and 1973 is due to a decline in the party voting and hence is consistent with a figure of 4% given in this study.

Does a difference of 4% matter between two STV algorithms? Obviously, it is reasonable to say this is insignificant against a difference of around 30% when STV is compared to First Past The Post. On the other hand, for the Electoral Reform Society, it is surely unsatisfactory to have such differences. Unfortunately, resolving this issue, as we are all aware, is not easy.

The remaining result is that Meek has more indicative cases in its support than Newland-Britton by about 5 to 1 in the above experiment. Does this matter? Surely, a key advantage of STV is that candidates can enter without upsetting the result if they have no realistic chance of being elected. Providing other hurdles for candidates seems against the spirit of democracy.

References

1. B A Wichmann. Producing plausible party election data. *Voting matters*, Issue 5. pp6-9. January 1996.

2. B A Wichmann. Large elections by computer. *Voting matters*, Issue 7. pp2-4. September 1996.

3. R A Newland and F S Britton, How to conduct an election by the Single Transferable Vote, second edition, ERS 1976.

4. B L Meek, A new approach to the Single Transferable Vote, reproduced in *Voting matters*, Issue 1, pp1-10, March 1994.

Measuring proportionality

I D Hill

When you can measure what you are speaking about and express it in numbers, you know something about it, but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind. Lord Kelvin.

It is important to consider what the problem actually is, and solve it as well as you can, even if only approximately, rather than invent a substitute problem that can be solved exactly but is irrelevant. Anon.

I agree with the first of those quotations but I agree much more strongly with the second one. As Philip Kestelman points out in a recent article[1], if we are to talk of proportional representation, and to claim that one aim of STV is to achieve it, it is desirable that we should have some idea of how to measure it and thus be able to detect the extent to which one system or another is able to achieve it.

Many indices have been proposed for the purpose, of which Kestelman prefers the Rose index, or Party Total Representativity (PTR) as he renames it. While differently formulated, the various indices all seem to have similar effects, usually placing different elections in the same order of merit even if the numbers that they assign are very different. They mostly depend, in one way or another, on the differences between percentages of votes by party and percentages of seats by party. It seems a little odd when considering a multiplicative type of thing, like proportionality, to use an additive type of measure, but this does overcome some difficulties that might otherwise arise when parties get zero seats.

A correlation measure

There is an additional measure that is rather different from all these, mentioned by Douglas Woodall[2] as having been proposed by Dr J E G Farina and depending on the cosine of an angle in multi-dimensional space. This is not a concept with which the general public would feel easily at home, but the measure does turn out to be closely associated with the statistical measure known as the correlation coefficient, and many people seem to feel happy that they know what correlation means (even if, in fact, they do not). However the ordinary correlation will not do, because it measures whether points tend to be grouped around a straight line, but not all straight lines give proportionality.

For example with votes of 200, 400 and 600 and the proportional 2, 4 and 6 seats we get a correlation of 1.0, but the non-proportional 3, 4 and 5 seats equally get 1.0 as those points also fall on a straight line. To get a suitable measure we also need to include the same numbers over again, but negated. Thus 200, 400, 600, −200, −400, −600 with 2, 4, 6, −2, −4, −6 gives a correlation of 1.0 as before, but 200, 400, 600, −200, −400, −600 with 3, 4, 5, −3, −4, −5 gives only 0.983 demonstrating a less good fit.

The fatal flaw

If going for any of these measures, I like the last one best, but they all have one fatal flaw — they depend only upon party representation and only upon first preference votes. It is possible to use them upon features other than formal political parties if there is enough information available on those other features, which usually there is not. Kestelman does so, but this is rarely done, while how to extend them to deal with anything other than first preferences does not even seem to be discussed. They therefore, to my mind, fall within the terms of the second quotation in my heading, as the substitute problem that is irrelevant.

It is true that, in many elections, voting is mainly in terms of party, and that most people's party allegiances will be detectable in terms of their first-preference votes, but I object to those who say that all we need to know about an electorate is to be found in those things. I much more strongly object to any suggestion that voters ought not to vote cross-party if they wish, or even should not be allowed to do so.

It often helps discussion to look at an exaggerated case, even though it is far removed from what normally happens in practice. An example that I have used before concerns 9 candidates: A1, A2 and A3 from party A; B1, B2 and B3 from party B; C1, C2 and C3 from party C. The election is for 3 seats and the votes are, say,

```
50% A1 B1
50% A1 C1
```

If a system elects A1, A2 and A3 the above measures will all say that it has done well — with 100% of the votes for party A and 100% of the seats for party A. Yet nobody actually voted for A2 or A3 at any level of preference. From that election STV would elect A1, B1 and C1, the candidates whom the voters mentioned, yet such measures will all say that it has done badly. While I believe that a measure of proportionality, if we can find a suitable one, would be a good thing I am not prepared to accept as useful any measure that cannot deal sensibly with that case.

Minor parties and independents

A further difficulty with all these measures occurs if there are a number of minor parties (and/or independent candidates), none of which get enough votes to be entitled to a seat. If each of them is put into the formula as a separate entity, you get one answer, but if you put them together as “others” you may get a very different answer because that number of votes for a single party would have been worth a seat (or more). Such minor parties are likely to be so divergent that to elect any one of their candidates to represent all their voters would be quite unsatisfactory.

STV's proportionality

STV's proportionality comes from what Woodall [3] calls DPC for “Droop proportionality criterion”. This says that if, for some whole numbers *k* and *m* (where *k* is greater than 0 and *m* is greater than or equal to *k*), more than *k* Droop quotas of voters put the same *m* candidates (not necessarily in the same order) as their top *m* preferences, then at least *k* of those *m* candidates will be elected. In particular this must lead to proportionality by party (except for one Droop quota necessarily unrepresented) if voters decide to vote solely by party. Anti-STVites may argue that this is not altogether relevant because people may not vote like that, but they cannot have it both ways — if voters are not concerned solely with party, and do not vote solely by party, then measures that assume that only party matters must be wrong.

The STV argument is that it will behave proportionately, as defined above, so long as voters do vote solely by one thing, whether that is party or not, but if (as is usual) voters have a mixture of aims and motives it will adjust itself to match what they do want to a reasonable degree. Looking at how it works suggests that it must do so, but I know of no way of proving it. What I find obnoxious is to find those who oppose it saying that it cannot be guaranteed to do so, and therefore wanting instead some system that does not even attempt it.

Furthermore STV gives the voters freedom to show their true wishes, major party, minor party, independents, sole party or cross-party, by sex or race or religion or colour of socks, or whatever they wish, whether others think that a sensible way of choosing or not. Even if it did not give a reasonable degree of proportionality as well, it would be worth it for that freedom and choice. Party proportionality is a bonus, not the be-all-and-end-all. It may be that “when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind” but can we measure love, or aesthetic pleasure, or scientific curiosity? Perhaps there would be some advantages if we could measure them, but our inability to do so does not in the least affect our conviction that they are things worth having. Let us continue to seek a useful measure, but not be bound by imperfect ones.

First-preference measures unsatisfactory

Even within strictly party voting, the first-preference measures are unsatisfactory. Consider a 5-seater constituency and several candidates from each of Right, Left and Far-left parties. Suppose that all voters vote first for all the candidates of their favoured parties, but Left and Far-left then put the other of those on the ends of their lists. If the first preferences are 48% Right, 43% Left, 9% Far-left, all the measures will say that 3, 2, 0 is a more proportional result than 2, 3, 0. Yet STV will elect 2, 3, 0 and that is the genuinely best result, because there were more left-wing than right-wing voters. There is no escape by comparing with final preferences, after redistribution, instead of first preferences. That is merely to claim that STV has done well by comparing it with itself. Our opponents may sometimes be dim, but I doubt whether they are dim enough to fall for that one.

Conclusion

I remain of the opinion that a measure of proportionality is very much desired if we can find a suitable one, but we know of none, and an unsuitable one may be worse than useless. What do others think?

References

1. P Kestelman. Is STV a form of PR? *Voting matters*, 1996, issue 6, 5-9.

2. D R Woodall. How proportional is proportional representation? *Mathematical intelligencer*, 1986, 8, 36-46.

3. D R Woodall. Properties of preferential election rules. *Voting matters*, 1994, issue 3, 8-15.

Issue 9, May 1998

Editorial

I must apologise for the absence of an issue since May1997, but this has been due to a lack of material. There is no doubt that the primary reason for this lack has been the May 1997 elections and the consequences in terms of the political debate on voting reform which has engaged many potential contributors.

The first article by David Hill considers the vexed question of constraints. After producing an elegant possible solution to the problem, he advocates that constraints should not be used. It seems to me that constraints can be used, but only modestly. For instance, if a Council is to be elected having a treasurer who must be a qualified accountant, then a constraint is better than having a separate election. Also, for national bodies, it is difficult to get young people elected since they are not as well-known which again seems to me to be reasonable grounds for a constraint. What do others think?

The second paper was prepared to submit to the Scottish Office as a result of the paper giving the proposed electoral system for the Scottish Parliament. This has obviously been partly overtaken by events.

The third paper on voter choice and proportionality was prepared as a result of the ERS AGM, and has been submitted to the Electoral Commission.

The last two papers consider a topical issue: how to prepare an ordered list of candidates given preferential voting. In this case, the Liberal Democrats and the Green Party have decided upon different methods which are specified in these two papers. The one point of agreement, which is also supported by others, is that the method of ordering a list given by Newland and Britton should *not* be used for this purpose! In both the second and third edition of Newland and Britton's book, they suggest that the order of election within the STV stages should be used to order the candidates (see section 2.5).

For the tenth issue, I plan to produce a combined index of all the issues to that date. I also hope to produce a volume containing all these issues in one binding, hopefully with good reproductive quality. The intent is to provide a more convenient permanent record.

I also plan to provide Internet access to *Voting matters* via the UK Citizens On-Line Democracy, which has been agreed by ERS. This Internet site provides discussion groups and other informal material on democratic issues. UKCOD has also collected comments on the Government proposals for a Freedom of Information Bill (`http://foi.democracy.org.uk/`). I believe that providing access to *Voting matters* by this means will encourage further international contributions. The printed version will be the authoritative one, since it is not possible to control layout precisely on the Internet, nor can references be guaranteed to remain valid over a long time-frame. The Internet version may also be delayed by the conversion effort required.

Brian Wichmann

STV with constraints

I D Hill

David Hill is the author of the computer program certified for use in the Church of England elections.

Introduction

Elections sometimes include constraints such as, for example, saying that those elected must include at least a given number of each sex. How is it to be done?

The traditional way is set out, for example, in the ERS booklet by Grey and Fitzgerald, that preceded the later rules by Newland and Britton. I have sometimes heard their method referred to rather rudely as “the naïve rules”. Basically they are the same as those in the Church of England's 1981 regulations and say: (1) that if a point is reached in the count where a specified maximum number of candidates of the constrained type has been elected, then any other candidate of that type must be excluded as soon as possible; (2) that if a point is reached where the number unexcluded of the constrained type equals a specified minimum, then any such candidate not yet elected must be guarded, such that when choosing a candidate for exclusion at any later point, the lowest non-guarded one must be chosen.

Multiple constraints

Grey and Fitzgerald make no mention of the possibility of more than one constraint or how such is to be handled. The Church of England's 1981 regulations, however, specified that the same rules should be applied to each constraint independently. It was pointed out that this could lead to trouble because two constraints may interfere with each other. The example used was: suppose there are 3 seats to be filled, and one constraint requires at least 2 women, while another requires at least 2 black people. If the available candidates are 2 black men, 2 white women and 1 black woman, where no-one has a quota and the last-named has fewest votes, she would be excluded by looking at each constraint separately, whereas that exclusion makes it impossible for the constraints to be met. It might be objected that such requirements are unlikely but: (a) regulations must allow for all possibilities; (b) however unlikely for a complete election, such a thing could easily arise at a late stage of something larger.

In consequence, the Church of England's 1990 regulations gave no specific rule for handling multiple constraints but left it to the presiding officer to do as seemed right at the time.

An alternative for a single constraint

An alternative approach has been devised by Colin Rosenstiel and colleagues for use by the Liberal Democrats in their internal elections, where STV is to be used with a constraint on the minimum number of each sex to be elected. Their method is (a) to run STV with the correct number of seats and no constraints. If more of one sex are found to be required, then (b) to rerun with more and more seats until it is found which extra candidates of that sex to elect, and (c) to rerun with fewer and fewer seats until it is found which candidates of the other sex to exclude. There are some difficulties, but on the whole this seems at first glance to be an elegant solution for a single constraint, though it is not feasible unless the count is to be by computer and it is not easy to see how it could cope with more than one constraint. It should be noted, however, that it is incompatible with any promise to voters that their later preferences cannot upset their earlier ones.

Although attractive at first sight, I have now come to the opinion that this method is wrong in principle. Indeed this opinion relates to any scheme that starts with ordinary STV and says that, if that produces a result that meets the constraints, it should be accepted. Such a method is always wrong. This opinion may seem odd; does it mean that there is something wrong with ordinary STV? Yes, of course there is. We know well that a perfect electoral method is impossible. The main fault with ordinary STV lies in its “exclude the lowest” rule, which can lead to unjustified exclusion on occasions. The justification of the rule is that it seems to be impossible to find a better one without violating the promise that a voter's later choices cannot upset that voter's earlier choices. It is generally thought to be better to accept the fault than to violate that.

Excluding the lowest is on the grounds that we must exclude someone and that candidate looks, on current evidence, the least likely to succeed. But if we have a constraint that makes it totally impossible for some other candidate to succeed, it is plain daft not to exclude that candidate first.

A simple example can explain the point clearly. Suppose 4 candidates for 2 seats. A and B are men, C and D are women. The votes are:

```
19 ABD.
 8 CD..
 3 DC..
```

giving a quota of 10. A is elected at once and passes his surplus to B, but with no further surplus someone must be excluded and, without constraints, it is most sensible to exclude D, who looks the least likely to succeed, and make a fair fight between B and C for the second seat, which C wins.

Suppose, however, that there is a constraint to say that 1 man and 1 woman must be elected. A and C, as by plain STV, are 1 man and 1 woman, but the reasoning by which they were chosen is now quite inadmissible. It cannot be said that D “looks the least likely to succeed” because, no matter what happens, it is absolutely certain that B cannot succeed and a fair fight between B and C is impossible. The remaining seat must go to a woman and a fair fight between C and D is what is necessary. Excluding B, D beats C.

So, by plain STV, A and C are elected, 1 man and 1 woman. Yet, with the constraint of 1 man and 1 woman, it is right to elect A and D instead. This may seem remarkable, but if there is any flaw in the logic I should like to hear of it. The conclusion must be that the title “naïve method” has been wrongly ascribed.

Tackling multiple constraints

How then should multiple constraints be tackled? I believe that the traditional way for a single constraint is right but it needs to be extended to deal with multiple constraints as such, not with each single constraint independently. This is not easy, but if people will introduce multiple constraints, the difficulties are their fault.

The Grey and Fitzgerald rules must be extended to say that whenever a situation is reached such that certain candidates must be elected, or must fail to be elected, if all constraints are to be met then the appropriate action is required. It should be noted that such situations can be met even before vote counting starts, and it may even be that no solution is possible. Regulations need to deal with such cases.

It is superficially attractive to look at each possible set of candidates that could be elected and enquire of each set whether it meets all the constraints, classifying each set as positive or negative. At every stage, each set ruled out as inconsistent with those now elected or excluded would be reclassified as negative, any candidate appearing in every positive set would be marked as “guarded” (i.e. not to be excluded, but still to receive votes until reaching a quota), while any candidate appearing in no positive set would be at once excluded. However, if thoughtlessly implemented, this scheme could easily lead to a combinatorial explosion. For example to elect 20 candidates out of 40, there would be over 10,000,000,000 sets and if a computer could examine 1000 sets per second to classify them, it would take over 4 years merely to go through them once.

A more practicable scheme is to note that the candidates can be grouped according to which constraint features they possess. Usually there are many identical in such respects and looking at them individually is not necessary but only at the number in each group. With that simplification it has been found possible to implement a solution, but it remains sufficiently complicated that to try to do it without computer help is inadvisable. By hand and eye it is all too easy to miss the vital moment when constraints need to be applied, and if missed, disaster can ensue later.

A (disguised) real election

An example of this can be seen in an election that actually occurred though, for obvious reasons, I shall disguise it. I shall also simplify it a little.

Suppose an election in which there are 28 candidates for 14 seats. The candidates, with two-letter code-names for the groups are

```
 4 English men    (EM)
 7 English women  (EW)
11 Scottish men   (SM)
 3 Scottish women (SW)
 2 Welsh men      (WM)
 1 Welsh woman    (WW)
```

Constraints say that those elected have to be 7 English, 6 Scottish, 1 Welsh and additionally 7 men, 7 women.

Suppose that the first to be elected is a Welsh man. Anyone would at once see that the other Welsh man and the Welsh woman cannot now succeed so it is right to exclude them at once to let their supporters move elsewhere.

Suppose that the next to be elected are 2 English men and 2 English women, and that the next step after that is to exclude a Scottish woman. How many people would notice that this is a critical point, where everything will go wrong later unless constraints are applied? I think that few people would; without careful analysis it is hard to notice.

The point is that only 2 Scottish women remain, we have to elect 6 Scottish altogether, and have elected none as yet. Therefore we must elect at least 4 Scottish men. But we are restricted to 7 men in total and we have already elected 3. It follows that we must elect exactly 4 Scottish men, and that means that the 2 remaining Scottish women must now be guarded, and that the 2 remaining English men must be excluded as soon as possible, as they cannot now succeed without upsetting the constraints.

If such an election has to be carried out by hand, the best way is to prepare in advance, preferably with a computer to help, a list of all the possible ways, by groups not by individuals, in which a conformant result could be obtained. This can be done as soon as the candidates are known, when there is time to devote to it before the count. In the present instance, there are 8 possibilities:

EM	EW	SM	SW	WM	WW
0	7	6	0	1	0
1	6	5	1	1	0
1	6	6	0	0	1
2	5	4	2	1	0
2	5	5	1	0	1
3	4	3	3	1	0
3	4	4	2	0	1
4	3	3	3	0	1

With such a list at hand during the count, its lines can be deleted as soon as they become impossible. Thus as soon as the first to be elected is found to be a Welsh man, any line with WM set to 0 goes out, leaving just

EM	EW	SM	SW	WM	WW
0	7	6	0	1	0
1	6	5	1	1	0
2	5	4	2	1	0
3	4	3	3	1	0

The election of 2 English men and 2 English women leaves just

EM	EW	SM	SW	WM	WW
2	5	4	2	1	0
3	4	3	3	1	0

and the second of these becomes untenable when only 2 Scottish women remain. Knowing that the first line is now the only way to meet the constraints shows the steps necessary much more clearly than could be seen without it. With a bit of practice, to follow such a list, as an indication of the interaction of the constraints with the STV count, becomes a little easier. However, it can never be really easy.

In case anyone should suggest that such a complicated example is implausible, I should repeat that it did actually occur except that I have disguised it and *simplified* it.

Conclusions

I believe that the approach given above is the best way, within STV, to implement constraints but that they should not be employed unless it cannot be avoided.

The mechanisms of STV are already designed to give voters what they want, so far as possible, in proportion to their numbers. It should be for the voters to decide what they want, not for anyone else to tell them what they ought to want.

The magazine *Punch* in 1845 included “Advice to persons about to marry — Don't”. My advice on constraints is similar.

Comments on the Scottish Electoral proposals

I D Hill, R F Maddock and B A Wichmann

Co-incidentally, all three authors have been members of the British Standards Institution programming language standards committee at various times.

It is clear than the proposal (made in July by the Government[1] in advance of the September 1997 Referendum) is an incomplete draft. Nevertheless, it seems appropriate to list the logical problems which are in this draft, since it is unclear how a complete proposal would rectify the flaws. In some cases, aspects which are undefined could be resolved by taking the proposals made at the Scottish Constitutional Convention, but this is something to be submitted to a referendum to authorise a constitutional change. No matter how worthy that body, it would be absurd to regard its proposals as being in any way definitive for such a purpose.

Why admit the existence of parties?

Although the existence of parties is a key aspect of the proposals, we feel bound to query this for the reasons below.

1. To formally acknowledge the existence of political parties is not currently part of the UK electoral framework. Surely such a significant step should be justified by showing that the general objectives can only be satisfied by this step.

2. Who is to be entitled to register a party? How are the names of such parties to be resolved to avoid confusion? Several names could cause confusion: The New Labour Party or The Tory Party or even just Liberal.

3. The proposal appears to suggest that the stated objective is to attain proportionality of party representation within the Scottish Parliament. However, the UK already accepts that proportionality can be attained without formal recognition of parties by means of the Single Transferable Voting system used for the Northern Ireland European Elections.

4. The proposal has several logical flaws, most of which arise from the party identification (see below).

The whole process appears to have been designed to give as much power as possible to party organisations and as little as possible to the electorate, making a mockery of what democracy should be.

On the other hand, the recent case of the Literal Democrats indicates that standardized party labels have some benefits.

Can one have independent MSPs?

It is clear that independent MSPs could not be elected from the Party lists, but for the constituency MSPs this appears to be possible. After all, we have such an MP for Westminster and therefore the question is not academic. The basic right for anyone to seek election should not be unreasonably restricted and therefore one must assume that those seeking election as a constituency MSP need not have a party affiliation.

Can a 'rejected' MSP be elected?

This can happen under the German system and results in the electorate being very sceptical about elections. This happens as follows:

A candidate who is seeking re-election is both a constituency candidate and is on a party list. If the candidate fails to obtain election for the constituency, the person can nevertheless be elected via the party list. If the person concerned was overtly unpopular and lost by a significant swing, then to be subsequently elected is perverse.

No electoral system should give rise to anomalies as gross as the above, since it can seriously damage the electoral process in the eyes of the electorate. (However, we know that 'perfection' is not possible for electoral systems which implies that minor anomalies cannot be avoided.)

One party list or many?

It is not clear if there is a single party list for each party, or one for each European Constituency. Note that the rules appear to allow for a party which is already over-represented to obtain additional seats due to being under-represented within one European Constituency (thus increasing the lack of proportionality).

Better proportionality would be obtained for a single list allocated on the basis of the entire Scottish vote. If the aim is to elect on the basis of European Constituencies, then why not STV for each such constituency?

Some problems

A list is made here of the main flaws that we have noticed. We cannot guarantee that the list is complete.

1. Who specifies the party lists? In practice, a good fraction of the MSPs are not determined by the electorate but by those who draw up the lists. In consequence, it is most important that the mechanism for producing these lists should be well-defined (or even an explicit statement that the party organisations determine the list by means of their own choosing). If the list is specified by the party organisations without any electoral process, then it is clear that this aspect is less democratic than any other mechanism currently in use within the UK.

2. When are the party lists published, and by whom? Is the list on the ballot paper? Surely the lists have to be published by the returning officers, but what restriction, if any, is placed upon the lists? (One could allow 'cross-benchers' to appear, as in the Lords. We assume that the lists are published before polling day!) The Scottish Constitutional Convention proposals appear to suggest that the list is just that, with no 'party' as such, which leaves open how parties are linked to constituency MSP's to determine the number of additional members.

3. Can a (previously) sitting MSP also be on a party list? If this is allowed, then the German problem arises, as noted above. In consequence, it seems best to exclude this. Obviously, if an MSP is elected as a constituency member, then one must assume that his/her name is deleted from the party list. This might present a practical problem if the MSP appeared on a different list from his/her own European Constituency.

4. What duplicates can appear on the party lists? If a person could appear on the party list for more than one European constituency, then logical problems arise due to the coupling of the voting between the European Constituencies. In particular, the result would depend upon the order in which the European Constituencies were considered.

5. The dependence of the proposals on the European Constituencies seems odd since the government has indicated its intention that the next European election, which will occur before the elections to the Scottish parliament, will use a regional list system, and thus the current European constituencies will no longer exist. The white paper does say that if the European constituencies are changed the boundary commission will make "appropriate arrangements for the Scottish Parliament".

6. A popular MSP could stand as an 'independent' so that his/her seat would not count for his/her party, thus increasing their additional members by one.

7. In a somewhat similar position to the last problem, a party could have a different label for its constituency candidates than for its party list. This would make the party list label appear under-represented (no seats), thus being eligible for additional members.

8. Apart from the voting system, we regard it as quite wrong that Scottish MPs will apparently be allowed to continue to vote at Westminster for what is to happen in England on the devolved issues.

9. The statement that the number of Scottish seats [in Westminster] will be reviewed begs more questions than it answers. The number of seats could even be increased! (However, Donald Dewar, introducing the white paper in the Commons, indicated that the number of Westminster constituencies was likely to be reduced at the next boundary review, and the white paper says that such changes would lead to corresponding changes in the number of both constituency and additional members in the Scottish parliament.)

10. It has been noted in New Zealand that a result of a mixed system of constituency members and party lists is a potential conflict between local party workers (who want to get their constituency member elected) and the party organisation (who might prefer the next person on the list instead).

11. The proposals call for 129 members which appears to be a consequence of the constituency numbers with the need for 56 additional members to obtain proportionality. Contrast this with STV for each of the 7 European Constituencies which could obtain the same degree of proportionality with around half the number of MSPs. (The cost saving would be very significant, and the body might well be more effective.)

12. Candidates must be *resident in the UK*, including therefore resident outside Scotland, which is different from most local elections in Britain, where the candidate must reside in the area administered by the assembly in question.

13. Can a Westminster MP simultaneously be an MSP? Nothing is mentioned about this, so one assumes the answer is yes, as it is for MEP, MP, county councillor, district councillor, parish councillor,... However, the proposals made by the Scottish Constitutional Convention state that being an MSP is *a full-time appointment* and thus excluding such roles (except perhaps being a Peer).

14. The arrangements for by-elections are not stated, although proposals were made by the Scottish Constitutional Convention, which we assume apply (namely, a conventional by-election for constituency MSPs, and the next on the party list for the additional members).

15. It is not specified what happens if a party list is exhausted.

16. If an MSP, elected from the party list, resigns from the party or is expelled from it, is resignation as an MSP to be required?

Reference

1. The Internet Scottish Office pages, and those from the Scottish Constitutional Convention.

The above paper records our comments at the time that it was written. We recognise that some of its queries have now been answered.

Voter Choice and Proportionality

B A Wichmann and R F Maddock

At the Electoral Reform Society 1997 AGM, Hugh Warren produced an eye-catching diagram in which several electoral systems were plotted on a diagram in which the two axes were voter-choice and proportionality. The diagram was not intended to give precise measures of the characteristics of each electoral system, but merely their relationship. However, for (party) proportionality, the Rose Index is a reasonable approximate measure. For voter-choice, no existing measure appears to be available which would be necessary to provide a more accurate representation of the diagram.

A possible measure of voter choice is the information-theoretic value of the result of an election, which appears to be new. For instance, in a dictatorship which has mock elections, the result is known beforehand, and therefore the information-theoretic value is zero. On the other hand, if the electorate is given a choice between three candidates then, assuming that each outcome is equally likely, the information-theoretic value is $\log_2(3)=1.58$. As the number of possible outcomes increases, so does this measure of voter choice.

For values of the Rose Index, Kestelman[1] gives values for the major electoral systems. It must be acknowledged that the Rose Index as a measure of party proportionality, may not be appropriate for STV elections, as pointed out by David Hill[2].

We compute the values for a hypothetical election for a 600 seat assembly in which there are three parties. For the use of STV, we take 120 constituencies each electing 5 members. For the regional list, we take 10 regions electing 60 candidates each. For the additional member system, we assume 300 seats elected directly and 300 added by proportionality. Note that if n seats are to be filled with 3 parties, then the number of ways to do this is $n^2/2+3n/2+1$. We assume that all possible outcomes are equally likely. The entries in the diagram are as follows:

First Past The Post (**FPTP)**: Rose Index 70% (UK), voter choice is 600 $\log_2(3)=951$.

Alternative Vote (**AV**): Rose Index 72% (Australia), voter choice is 600 $\log_2(3)=951$.

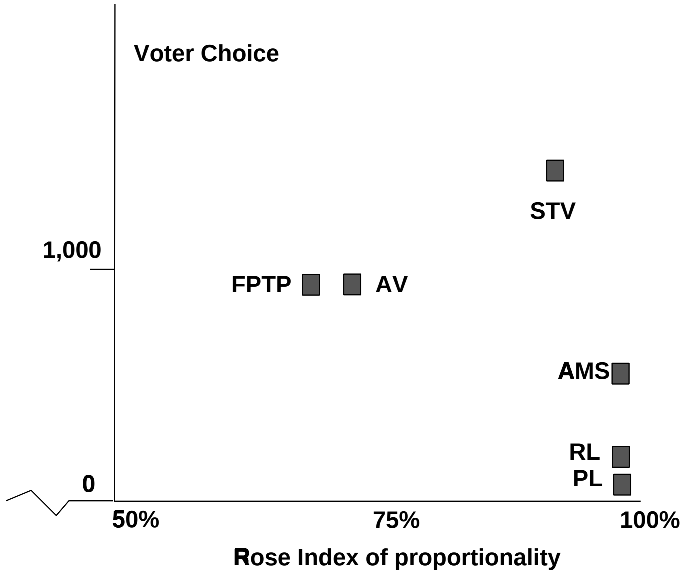
Single Transferable Vote (**STV**): Rose Index 92% (Ireland), voter choice is 120 $\log_2({}^5C_{15})=1386$. (We are assuming each party has five candidates and therefore could theoretically obtain all five seats; hence the number of possibilities is the number of ways of selecting 5 from 15.)

Additional Member System (**AMS**): Rose Index 98% (estimated), voter choice

$300 \log_2(3)+\log_2(300^2/2 +3\times300/2+1)=491$.

Party List (**PL**): Rose Index 98% (estimated), voter choice is $\log_2(600^2/2+3\times600/2+1)=17.5$.

Regional party Lists **(RL)**: Rose Index 98% (estimated), voter choice is 10 $\log_2(60^2/2+3\times60/2+1)=109$.



It is important to note that this diagram will change if the underlying assumptions are changed, for instance, if the number of parties was increased from 3 to 4. An alternative way to compute voter choice values would be to take into account the probability of the various outcomes, based upon appropriate statistical data. This was considered initially but rejected due to the difficulty of the calculation and the problems in finding appropriate statistical data. If the voting system was changed, then one can only guess at the future statistical data. (The diagram here has the *x*-axis reflected from Hugh Warren's version so that the Rose Index is increasing.)

The conclusion from this diagram is hardly unexpected: party lists do not give voter choice, and FPTP/AV do not give party proportionality, while STV can claim, to a reasonable degree, to provide both.

References

1. P Kestelman. Is STV a form of PR? *Voting matters*. Issue 6. p5-9.

2. I D Hill. Measuring proportionality. *Voting matters*. Issue 8. p7-8.

Producing a Party List using STV

C Rosenstiel

Colin Rosenstiel is a member of the Council of ERS and the author of a computer program for STV.

With some of the current proposals for electoral reform, parties will be required to produce a list from whom candidates will be elected in order from the top. STV can be used to construct the ordered list, given a preferential ballot of all party members.

The conventional use of STV to elect *n* members gives members of equal status, since the order in which STV elects does not necessarily determine the strength of their support. Repeated use of STV elections can be used to determine an order as follows:

Given a total list of 10 (say), then the first step is to determine those on the list (without an ordering) by running an STV election with all the candidates and 10 seats to fill. The next step is to run an STV election for 9 seats with 10 candidates being those previously elected (using the same ballot papers). The eliminated candidate is then placed last on the list. Next, an STV election if run with the remaining 9 candidates with 8 seats to determine the next lowest candidate, and so on.

This process might sound tedious, since so many STV elections are run, but if a computer is used, it is straightforward. Note that the above process will not work in reverse, i.e. selecting the top candidate first. The reason for this is that when electing two candidates, it can happen that neither of those elected is the previously selected ‘top’ candidate.

Two elections were taken in which there was more than ten candidates to which we have applied the algorithm above to order the top 10 candidates. The results obtained were as follows:

```
                     Election 1      Election 2
This algorithm    :  ABCDEFGHIJ      ABCDEFGHIJ
Order of election :  CABDEFGHJI      CBAFEHDGIJ
```

As expected, it can be seen that the order of election does not give the same result as successive elimination. Hence this algorithm is recommended in producing party lists.

Editorial comment

It has been suggested to me that if the Meek method is used, then just one election would suffice (to determine the order of the 10 candidates). Their order can be found from the retention factor in the final table of the election results — the smallest retention factor implying the strongest candidate since that candidate required the smallest proportion of the votes retained to get the quota. These values do give a measure of their relative support, unlike the order of election. In the elections above the Meek results were:

```
                     Election 1      Election 2
Meek 'keep' factor:  ABCDEFGJHI      ABCEFDHGIJ
```

This would appear to indicate that the methods of ordering of the candidates produce a similar result. In practice, both methods would need to use a computer and hence there seems to be little to choose between them.

Ordered List Selection

J Otten

Joseph Otten is the Green Party Policy Co-ordinator.

Rationale

The electoral system to be used for the next European Elections requires ordered lists of candidates from each party. It was felt that the advice in the ERS booklet [1] that *If an order is desired, this is provided by the order of election* (2.5) was inadequate — it would effectively lead to a First Past the Post contest for the top place on the list.

Were we to know in advance that we would win, say, *n* seats in a region, then it would be straightforward to use STV to select *n* candidates from the potential candidates and put them in the top *n* places in our list. If we don't know *n* in advance (which we don't!) then we can perform this operation for every possible *n*, i.e. from 1 up to the number of seats available in the region, and attempt to construct a list whose top *n* candidates are those victorious in the *n*th selection ballot. (There is really only 1 ballot — the division into *n* ballots is notional.)

This ideal solution fails when a candidate elected for one value of *n* is not elected for a larger *n*. In such cases either the STV result for a smaller *n* must constrain that for the larger (top-down) or vice versa (bottom up). Reasoning that the Green Party would be unlikely to win large numbers of seats in any region, we opted for top-down.

Algorithm

Each count is conducted by ERS rules [1] with the following alterations. We start with the count for the first place (*n*=1) and work down.

5.1.6 Calculate the quota

Divide the total valid vote by one more than the ordinal number of the count. Eg for the third count, divide the total valid vote by 4. If the result is not exact, round up to the nearest 0.01.

5.2.5 Excluding Candidates

Do not exclude any candidate in one count if they have already been elected to the list in an earlier count. This may introduce distortions to the results of later counts, but is necessary to preserve the integrity of the earlier counts.

If a count is proceeding identically to an earlier count, and an exclusion by lot is required, then the result of the earlier lot should be taken as read. Otherwise the lot must be recast. (cf 5.6.3)

5.3.3 and 5.4.2

For the purpose of these rules (i.e. receiving transfers), a candidate elected in a previous count (not stage) should be treated as a continuing candidate for purpose of receiving transfers during the count, until they are deemed elected again.

5.5.2 Completion of the Count

For the purpose of this rule, any candidate elected to the list in a previous count shall be deemed elected. Therefore the count may stop as soon as a single candidate is deemed elected, who was not elected in a previous count. In exceptional circumstances it is possible that two candidates, not previously elected may exceed the quota in the same stage. Only one may be elected. Resolve as follows (in order of priority):

1. If more than one value of papers is transferred during that stage, and only one candidate is elected as a result of the transfer of an earlier (i.e. higher valued) batch, then that candidate is deemed elected.

2. If both exceed the quota during the transfer of the same batch, then elect the one with the higher vote.

3. According to 5.6.2

4. By lot.

Other deviations

My apologies to the Electoral Reform Society for these, but they do seem to be popular in some quarters.

Where regional parties have agreed to adopt gender balance constraints, then the usual constraint rules shall be used. This usually means excluding all the candidates of a particular sex at the beginning of an even-numbered count.

Each region was free to determine its own gender balance formula. For example one region might require a list of half men and half women with no constraints on position, and another region might require that the top two candidates were a man and a woman with no constraints on the other candidates. Whatever formula was chosen, this was applied within the system by excluding any ineligible candidates at the beginning of a round. Hence the top place on each list would be open to both sexes, and subsequent places would only be constrained in the event of an imbalance. Notably the London region decided not to impose a gender balance formula, and the top three candidates are all women.

On each ballot form there is a notional candidate called “Re-Open Nominations” (who is of indeterminate sex). If Re-Open Nominations is elected to the list, then there must be a fresh election for that place and lower places on the list. This is a distortion of STV which could be used by a majority to deny minority representation, although there is no evidence of this happening. STV, rightly in my view, omits this sort of negative voting, but it is popular in the real world outside public elections, such as in student unions.

Conclusions

Although the justification for starting at the top of the list and working down, as opposed to starting at the bottom or even in the middle, is not particularly strong, this system is a reasonable solution to the question of seeking an ordered list. In particular it ensures that however MEPs are elected in any region from the party, they are as proportionally representative of the range of opinion in the party as their number allows.

Reference

1. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote. ERS 3rd Edition. 1997.

Issue 10, March 1999

Editorial

The publication of the Jenkins Commission report has presented ERS with a dilemma. The role of STV is minimal and on any reasonable measure, the degree of voter choice is not on the same scale that STV would provide. However, the report in my mind raises a technical challenge. If one accepts as a political imperative that 80% of the seats must be from single-member constituencies, can one devise a scheme with an increased voter choice which is simple to understand?

I certainly believe that this is possible. Moreover, I think that one should accept the significant support for First Past The Post (FPTP).

The Jenkins proposal of having two votes seems to me to be basically flawed since it then requires a mathematical adjustment to correct the mis-representation from the single member constituencies. Why not have just one vote, which either elects your chosen candidate for the single-member constituency or is transferred to the ‘county’ vote?

With an additional 20% to be elected at the ‘county’ level, and STV being the electoral mechanism, one needs about 15 single member constituencies to be merged into counties. These counties would therefore elect three members, giving useful voter-choice and good proportionality.

What would such a scheme look like from the point of view of the voter? The single-member voting would retain FPTP and hence would correspond to the existing system apart from a 20% increase in the size of these constituencies.

This implies that votes which would undeniably be ‘wasted’ under the present system would now be transferred to the county vote. Here the voter has a bigger choice, but more difficult decisions to make. With perhaps 12-20 candidates to rank in order to elect 3 people, the situation would be very similar to that of the voter in the Irish Republic. The key difference is that this STV vote would only apply to those voters who did not select a winning candidate at the single-member constituency level. Surely this scheme would end the need for strategic voting. The use of the Alternative Vote, as proposed by Jenkins, would therefore be unnecessary.

If voter choice is to mean anything at all, surely the voter must be able to choose between candidates of the same party. By having STV with three seats, such a choice would be effective. Increasing the size of the STV areas would have some advantages in terms of proportionality, but would probably give a ballot paper that was too cumbersome (compared with current practice).

Brian Wichmann.

The Handsomely Supported Candidate Ploy

C H E Warren

Hugh Warren is a retired mathematician

There is an electorate of 1400, who have to elect candidates to fill 6 seats, so clearly the quota is 200. The electorate is made up of 418 members of the Labour Party and 982 members of the Conservative Party. Labour should, therefore, get 2 seats, and the Conservatives 4.

The Conservatives put up 5 candidates — L, A, B, Z and Y. Candidate L is the Party Leader, and is handsomely supported because of his ability to hold the party together, despite its Europhile and Europhobe wings. Candidates A, B are on the Europhile wing, and Candidates Z, Y on the Europhobe wing. If all the Conservatives voted sincerely their voting pattern would be as follows:

```
503 LAB
479 LZY
```

Whether the count is done by Newland & Britton[1], Meek[2] or Warren[3], 4 Conservatives would be elected — L, A, Z and B. Not surprisingly the Europhiles get one more seat than the Europhobes because they are the slightly larger faction. 182 Conservative votes would be ‘wasted’, as would 18 Labour votes, thereby making up a quota of 200 votes in total which are perforce ‘wasted’ in any STV election.

However, the Europhobe Conservatives adopt the Handsomely Supported Candidate Ploy. Above everything else they want to see their leader, Candidate L, elected. But they argue that their support of 479 voters should be enough to ensure that Candidate L is elected if they insincerely give him their second preference only, even if those Europhiles are even more insincere and don't give Candidate L a preference at all!

In practice the Europhiles vote sincerely, so the voting pattern turns out to be:

```
503 LAB
479 ZLY
```

If the count is done by Newland & Britton or Meek, the Europhobes' ploy pays off, because the 4 Conservatives elected are L, A, Z, Y. So, by their ploy, the Europhobes have ‘captured’ the fourth Conservative seat for the Europhobes.

Of course one can not guarantee that one will always gain an advantage by adopting the ploy, but it is always worth trying on, for one can not lose provided it is done prudently, as in the example here, by not relegating a handsomely supported candidate to a preference where one has not the support to get him elected no matter what other voters do.

The Handsomely Supported Candidate Ploy, if practised by a group, can lead to a discernible gain, as just demonstrated. However, the principle, that one can gain an advantage by not giving one's first preference to a handsomely supported candidate, holds even for voters voting individually.

Consider an election for nine seats by 100 voters, so the quota is 10, in which the voting pattern is as follows:

```
39 H...
19 M...
41 ....
 1 HM...
```

H is clearly a handsomely supported candidate, and M a moderately supported candidate. These two candidates do not figure in the voting pattern other than in the places shown.

If the count is done by Meek the individual voter HM... will have 0.37025 of a vote to pass on to his third preference after H and M have retained just the votes necessary to attain the quota.

However, if the individual voter decides not to give his first preference to the handsomely supported candidate H, who would be his sincere first preference, but instead to vote MH..., then he finds that he has 0.37342 of a vote to pass on to his third preference.

It is the principle that is salient from this example — that one can get more out of one's single vote by not giving one's first preference to a handsomely supported candidate.

References

1. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote. ERS 3rd Edition. 1997.

2. B L Meek, A new approach to the Single Transferable Vote, reproduced in *Voting matters*, Issue 1, pp1-11, March 1994.

3 C H E Warren. Counting in STV Elections. *Voting matters*, Issue 1, pp12-13. March 1994.

An example of ordering elected candidates

C H E Warren

Colin Rosenstiel has proposed that elected candidates can be ordered by successive elimination[1]. In an unpublished note of the same date (May 1998), Eric Syddique has proposed essentially the same method. However, in Newland & Britton[2], the method proposed is to take the order of election. The purpose of this paper is to show that these two methods can produce very different results.

Consider the following election of 4 candidates from 7 contenders by 600 voters, for which the voting profile is:

```
 50 AC
 70 AD
115 BED
100 CD
115 D
 65 ED
 50 FCD
 35 GBED
```

Since the quota is 120, we obtain the following result sheet from the ascription of the Newland & Britton principles, avoiding the rounding errors which the practical application of their method as given by them introduces.

A	120		120		120		120		120
B	115	+35	150	-30	120		120		120
C	100		100		100	+50	150	-30	120
D	115		115		115		115	+30	145
E	65		65	+30	95		95		95
F	50		50		50	-50	0		0
G	35	-35	0		0		0		0

Hence the order of election is A, B, C and then D.

With the Rosenstiel/Syddique method of successive elimination, with E, F and G eliminated the votes are:

```
120 A
150 B
150 C
180 D
```

B, C and D are selected and A is placed fourth and eliminated henceforth. The votes are then:

```
150 B
200 C
250 D
```

C, D are selected, and B is placed third and eliminated henceforth. The votes are then:

```
200 C
400 D
```

C is now placed second and D first. To summarise, the order is D, C, B and then A.

Hence the two methods produce ordering which is exactly the opposite of each other.

References

1. C Rosenstiel. Producing a Party List using STV. *Voting matters*, Issue 9, pp7-8, May 1998.

2. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote. ERS 3rd Edition. 1997.

STV with constraints

Earl Kitchener

Hill[1] describes a useful way of dealing with constraints, but then says that “It should be for the voters to decide what they want, not for anyone else to tell them what they ought to want”. If, as is normally the case, the rules for elections have been set by the voters, there *is* no-one else, because it is the voters themselves who have decided in advance that they want, say, at least one new member and at least one sitting member. I feel that the ERS should encourage the use of constraints, so that we can learn whether they turn out to be helpful.

When voting is on party lines, it may be desirable to ‘entrench’ the rules by only allowing them to be changed by more than a simple majority. This is because a party in power can often find alterations whose only merit is that they would favour it. In some cases this would force constraints on unwilling voters.

Reference

1. I D Hill. STV with constraints, *Voting matters*, Issue 9, pp2-4, May 1998.

Response by I D Hill

I am grateful to Lord Kitchener for his courtesy in letting me see his paper in advance and for having no objection to my putting a reply in the same issue.

Although it is true that constraints would sometimes have been set by the voters themselves, it is by no means always so. For example, some Church of England elections are subject to constraints that have been set by Act of Parliament. Even where the voters have set them, it will usually be an earlier set of voters who have done so, constraints being set in the bye-laws of the organisation and continuing to exist for many years; the actual voters have no opportunity to alter them at the time of an election.

There is much to be said in favour of rules specifying that at least a certain number of people of particular types must be among those nominated as candidates, but it should be for the voters to decide whether they wish to elect them or not. As soon as they are forced to elect some, the whole election can become distorted by that fact. So I stick to my point of view that, in general, constraints within STV are a bad thing and should be avoided if at all possible. If there is no avoiding their use, the method employed should be as in my article.

A problem for Andrae and Hare

I D Hill

David Hill's great-great-great-grandfather is the Thomas Wright Hill mentioned in this article.

With any form of STV there is a question about the best way to transfer surpluses when they arise. Some people seem to think that provided the right number are taken, and no vote is specifically misused, it does not much matter how it is done. Others claim that such conditions are not sufficient, and that methods should be used that correctly interpret the wishes of the relevant voters as a whole.

The argument turns up interestingly in a fascinating book, to be found in the McDougall Library *Andrae and Proportional Representation* by Poul Andrae, son of Carl Andrae who introduced STV to Danish elections in 1855. The book is partly a complaint that Thomas Hare gets all the credit for the invention of STV and his father very little. Hare first suggested STV in 1857, whereas Andrae actually introduced it in 1855. The complaint appears to be justified and it looks as though perhaps Hare himself did not really want to know about Andrae, but it is always dangerous to judge something like this after hearing argument from one side only. The author of the book is evidently totally unaware of what Thomas Wright Hill did in 1819.

Andrae's system was simply to shuffle the voting papers and then count them just once, allocating each to its earliest preference who had not already attained a quota, and finally elect all those with a quota, plus the highest of those with less, to give the right number to fill the seats. There was no system of exclusions, with redistribution of those votes. Hare's earliest versions were somewhat similar to this.

On the question of how to redistribute a surplus, there is in the book a problem that was put to Andrae, of a case where it was said that his system could give an absurd answer. Andrae, in reply, points out that one of the rules of his system is that the voting papers are to be thoroughly shuffled before counting and, if that rule is obeyed, the probability that they are counted in the particular order on which the absurd result depends is so small that it can be ignored. In this he is correct (and he calculates the probability correctly too).

However the problem was also put to Hare, and Hare's reply is to try to justify the absurd answer as reasonable. I wonder whether any STV supporter nowadays would agree with Hare.

The problem concerns 5 candidates for 3 seats, and votes:

```
299 ABD
200 ACB
101 ACE
```

Hare and Andrae agree that the quota should be 600/3 = 200 and for present purposes let us not dispute that, even though we think that Droop's quota is preferable. The problem says: suppose the votes are counted in the order as given, using Andrae's system. Then, of the first 299, 200 go to A and 99 to B, the next 200 all go to C (leap-frogging A) and the final 101 to E (leap-frogging both A and C). As the system does not use exclusions, the final seat is awarded to E, because 101 exceeds 99 even though nowhere near a quota.

Andrae's correct reply is that, even in the unlikely event of such votes being made, the probability merely that all the 299 come out before any of the 200 is $1/q$ where q is a number of 117 digits, without even taking account of the fact that all those have to come before the final 101. This is certainly a remote enough probability to be ignored.

He does not mention that a similarly silly answer could result from

```
2 ABD
2 ACB
2 ACE
```

where the probability is as high as 1/90, but I feel sure that he would have said that his system was designed for big elections, not such tiny ones, though to my mind a good system ought to work sensibly for any size of election.

Hare, however, according to the book, wrote

> I am willing ... to adopt the result, which I believe is perfectly reconcilable with the principle that is at the foundation of this method of voting, and also reconcilable with justice. The object is to give the elector the means of voting for the candidate who most perfectly attains his ideal of what a legislator should be, but it does not contemplate giving him the choice of more than one ...

> The primary purpose of giving the voter the opportunity of adding to his paper the second, third, or other names for one of whom his vote is to be taken on the contingency of the name at the head not requiring it, is not to add greater weight to his vote, but to prevent it from being thrown away or lost owing to a greater number of voters than is necessary placing the same popular candidate at the head of their papers ...

> Thus the first 200 voters, whose voting papers are appropriated to A, have no ground of complaint (because of the non-election of B), for their votes have been attended with entire success ... Still less have the second 200 voters, whose votes were appropriated to C, any reason to complain, for they also have not only elected a favourite candidate of their own, but, equally with the first 200, they are gratified by the triumphant success of A. The 99 voters for B have also the latter satisfaction, and if they failed to return their next favourite candidate, it is simply because 101 are more than 99.

I should have to change my mind about supporting STV, if that were good STV reasoning, but I do not accept that it is. I agree that it is right to allow each voter just the one vote, but if 299 say AB whereas 301 say AC, to pass A's surplus as 301 to C and only 99 to B, instead of dividing it out in proportion to the voters' wishes, is grotesque.

It is extraordinary that Hare thinks it just and reasonable to elect E even though the total number of voters mentioning E at any level of preference is far less than a Droop quota. Any modern STV system would take the quota as 600/4 = 150, elect A with a surplus of 450 to be divided almost equally between B and C, who then each have more than a quota and all seats are filled.

Even if the votes had been merely

```
200 ACB
101 ACE
```

to elect ACE rather than ACB would be obviously absurd. With the additional 299 ABD votes it becomes even more so. Does any reader think that Hare was talking sense?

A review of the ERS97 rules

B A Wichmann

Recently, I was asked to interpret the Newland and Britton 3rd edition rules[1] (referenced as ERS97) with some specific examples and therefore read the rules carefully for the first time. I think I was largely successful in interpreting the rules correctly, but was surprised at a number of features of their presentation.

Over the last 20 years, I have been involved with the specification of programming languages for the International Standardization Organization (ISO). The requirements here are again for precision and clarity. ISO have adopted drafting rules for standards which I think are very helpful and are not far removed from the style of the presentation of section 5 of ERS97. There are a number of detailed differences in which I prefer the ISO approach. These differences are as follows:

1. Separation of normative (requirements of the standard) and non-normative text. In ISO, the model election would be a non-normative annex. In fact only sections 5 and 6 are normative.

2. In ISO, defined terms would appear before the main text. In ERS97, the Glossary in section 6 appears after their usage in section 5.

3. In ISO, notes are non-normative and laid out in a manner to make this clear. The note in 5.6.2 in ERS97 is clearly normative (and uses *shall*, as in ISO standards).

It seems to me that the adoption of the ISO drafting rules would be a worthwhile undertaking if any revision of the rules was contemplated. Indeed, I see no reason why a suitable revision should not be proposed to ISO as a standard, since it would allow other organisations (in any country) to use it by reference. Currently, many organisations contain rules for STV in their constitution which is unsatisfactory when the rules themselves are very old — a method of reference would be useful in such contexts.

My major and perhaps controversial comments on the rules arise from my desire to see it formulated more closely as an algorithm, rather than as a description. In trying to interpret the rules, one is necessarily performing a function like that of executing a computer program. Since the main purpose of the rules is surely to aid Returning Officers, then the computer program approach is helpful. Of course, I am not suggesting that computer terminology should be used, but merely that the style should allow for conversion into a program in an obvious manner.

My specific points arising from the above computer perspective, and from other analysis are as follows:

1. There is no provision for conducting a count with the aid of a computer or by an entirely automatic process. Since computer programs of both types are routinely used for this purpose, this is a major fault. Note that the Church of England rules[2] make specific provision for this, including the certification of appropriate programs by ERS. Breaking ties by lot needs a different wording allowing for the use of a pseudo-random number generator.

2. I think the wording associated with checking and records should be separated by being in a paragraph after the corresponding actions. (This is not straightforward as some paragraphs are a mixture.)

3. As I see it, only those paragraphs which are needed for reference purposes need be numbered. This would reduce the apparent complexity of the rules. Currently, the whole of section 5 needs to be read to determine what use is made of each part of the rules.

4. Section 5.5 (completion of the count) is not referenced at all, since it is invoked when appropriate conditions are satisfied. This is not algorithmic in the conventional sense, indeed, in computer terms could be seen as 'interrupt-driven'. I think this section should be used explicitly.

5. The calculation of the quota and the recording of transfers appears to give the impression of undertaking computations to one hundredth of the vote. However, this is not achieved, since that accuracy requires that the transfer values are computed to a greater accuracy. Indeed, if p votes are transferred, then there is a truncation error of at most $p/100$, which implies that transfer values should be computed to about *(number of digits in total votes)*+1 digits. I do not believe that an arithmetic approximation which can lose a whole vote is acceptable since the voter could reasonably equate the loss to his/her vote. Unfortunately, the rules depend upon (*number of papers*)×(*transfer value*) in hundredths of a vote, so it is difficult to increase the accuracy without complexities elsewhere. Hence I conclude that this problem is inherent in this type of rule and could be seen as a defect in ERS97.

6. The rules mention coloured forms, but the colouring is not apparent in the copy of the forms in the example — the solution is to print the 'beige/blue/green/white/pink/yellow' on the forms, so that photocopying them retains the information (or so they can be photocopied onto the correct colour paper).

7. Not all uses of the defined terms appear in bold in the rules. I would suggest that the uses of a defined term uses a different font (say, italic).

8. Paragraphs used in more than one place should be given a name and referenced by name (as with the sections 5.3 and 5.4).

9. A batch is a set of bundles each having the same transfer value, not a type of bundle as given in 6.1.

10. The definition of stage of the count is ambiguous, or perhaps depends upon the layout to parse.

11. The definition of transfer value should have 'deemed elected' rather than 'elected'.

12. The statement that for small elections counting slips are not required should be made once at the start, rather than each time slips are mentioned.

13. The second sentence of 5.6.4 is confusing. Surely the point is that an auditable record of the count should be kept? If all recording forms are optional, then why are counting slips specifically mentioned in 5.1.3, 5.3.12 and 5.4.3?

14. The term 'formally excluded' (in 5.5.2) clearly means exclusion without the application of the rules associated with exclusion, although this is not explicitly stated.

I have attempted to reformulate the rules along the lines that ISO would use, but I do not regard the result as at all satisfactory. My attempts were based upon a minimal change to the wording, but it appears that a more radical approach is needed.

A few issues have been noted by others that I should also add for completeness:

a) Conventional practice appears to be that the transfer values are not included in the result sheet. I do not like this, since the values are hard to reconstruct and are available.

b) The handling of withdrawn candidates is not mentioned in section 5 of ERS97, although it is surely a possibility with all elections (and is noted in section 2.2).

c) A minor ambiguity has been noted in the rules. (I hope to report fully on this in the next issue of *Voting matters*.)

Conclusions

Is any ‘improvement’ to the wording needed? I think the rules should be readily usable just from the booklet. In this regard, the model election and examples given are very helpful. However, they do not cover all the situations that can arise. Moreover, for the model election, the actual papers are not included (not unreasonable for 785 voters, but this means that this single long example cannot be re-worked completely by the reader). Also, the explanations given are not always adequate. For instance, in Section 8.2 it is said that, because the surplus could change the order of the last two, it ‘must be transferred’, without any hint being given that it is required to look at whether the next two or more to go out are definite, in which case it must not be transferred. In the particular instance the action taken is correct, but that is not the point. How to decide that it is correct is not fully stated as it should be[3].

Of course, the fact that ERS runs courses in conducting an STV election is very helpful as is the large number of people that have had such training and can pass on their skills to others.

Hence I conclude that improving the wording is not that vital, but it would be a shame not to consider the ISO approach if a revision was produced in the future.

References

1. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote. ERS 3rd Edition. 1997.

2. Church of England, General Synod. Regulations for the conduct of elections by the method of the single transferable vote. GS930. 1990.

3. I D Hill. Private communication.

Quantifying Representativity

P Kestelman

Philip Kestelman is keen on measuring electoral representativity, and works in family planning.

Introduction

What is a Proportional Representation (PR) electoral system? Seriously begging that question, Gallagher (1991: 49) argued that “Each method of PR minimizes disproportionality according to the way it defines disproportionality, and thus each in effect generates its own measure of disproportionality”.

However, Gallagher overlooked Single Transferable Voting (STV); an omission repaired by Hill (1997), invoking a ‘Droop proportionality criterion’ (DPC: Woodall, 1994: 10): “If, for some whole numbers *k* and *m* (where *k* is greater than 0 and *m* is greater than or equal to *k*), more than *k* Droop quotas of voters put the same *m* candidates (not necessarily in the same order) as their top *m* preferences, then at least *k* of those *m* candidates will be elected. In particular this must lead to proportionality by party (except for one Droop quota necessarily unrepresented) if voters decide to vote solely by party”.

Thus defined, PR systems include Alternative Voting (AV: *k*=1); though over half the voters may be unrepresented! According to the 1937 Irish Constitution, not only parliamentary deputies (multi-member STV), but also the President (AV), shall be elected “on the system of proportional representation by means of the single transferable vote”.

Yet nobody regards AV as a PR electoral system. In fairness to Woodall (1994: 10), “Any system that satisfies DPC deserves ... to be regarded as a system of proportional representation (within each constituency)”. At that level, Hill's “exaggerated case” (three-member STV) is persuasive; however disproportional to Party First Preferences. Nonetheless, *constituency* level ‘PR’ (including AV) is not enough for PR as normally construed.

Hill (1992) reasoned that, if voters vote solely by party, each nominating sufficient candidates, “then STV will produce splendid proportionality, ... , while any discrepancy due to fractions of quotas can be expected to even out over a number of multi-member constituencies”. Indeed, the main political question is how faithfully total seats reflect Party First Preferences overall (regionally and/or nationally).

Party Total Representativity

In parliamentary elections, the simplest measure of total disproportionality is the overall deviation between over-represented Party Seat-fractions and Vote-fractions: the *Loosemore-Hanby Index* (LHI),

$$LHI = \tfrac{1}{2} \Sigma \text{ ABS (S\% - V\%)},$$

where S% = Party Seat-fraction (percent);
V% = Party Vote-fraction (percent); and
Σ ABS = Sum of magnitudes (over all parties).

LHI complements the *Rose Index of Proportionality* (RIP); for which I prefer the more explicit term, *Party Total Representativity* (PTR).

Table 1 below demonstrates the calculation of PTR = 100%-LHI for the 1997 Irish General Election, which proved unprecedentedly disproportional to Party First Preferences.

Table 1: STV Party (First Preference) Votes and Seats: Numbers, Fractions and Deviations: General Election, Irish Republic, 1997

Party (Constituency)	Number Votes (V)	Seats (S)	Fraction (%) Votes (V%)	Seats (S%)	Deviation (S%-V%)
Total	1,788,985	166	100.0	100.0	0.0
Fianna Fáil	703,682	77	39.3	46.4	+7.1
Fine Gael	499,936	54	27.9	32.5	+4.6
Labour	186,044	17	10.4	10.2	-0.2
Progressive Democrats	83,765	4	4.7	2.4	-2.3
Green	49,323	2	2.8	1.2	-1.6
Sinn Féin	45,614	1	2.5	0.6	-1.9
Democratic Left	44,901	4	2.5	2.4	-0.1
Socialist	12,445	1	0.7	0.6	-0.1
Lowry (Tipperary N)	11,638	1	0.7	0.6	-0.0
Blaney (Donegal NE)	7,484	1	0.4	0.6	+0.2
Healy-Rae (Kerry S)	7,220	1	0.4	0.6	+0.2
Gildea (Donegal SW)	5,592	1	0.3	0.6	+0.3
Fox (Wicklow)	5,590	1	0.3	0.6	+0.3
Gregory (Dublin C)	5,261	1	0.3	0.6	+0.3
Unrepresented	120,490	0	6.7	0.0	-6.7
Over-represented	1,234,765	136	69.0	81.9	+12.9*
Under-represented	554,220	30	31.0	18.1	-12.9

* *Loosemore-Hanby Index* (LHI) = 12.9 percent = Overall deviation between over-represented Party Seat-fractions and Vote-fractions: complementing *Party Total Representativity* (PTR) = 87.1 percent.
Source: Dáil Éireann (1998).

The Independent Commission on the Voting System (Jenkins, 1998: 47) gave a 1997 Irish General Election LHI of only 9.8 percent (their DV or ‘deviation from proportionality’: Dunleavy *et al*, 1997: 10). However, the two main parties (Fianna Fáil and Fine Gael) alone received 11.6 percent more Seats than Votes (First Preferences); and exact LHI=12.9 per cent (Table 1). LHI (and hence PTR) are often miscalculated.

Other Measures

McBride (1997: 9) invoked “O'Leary's index of proportionality”: the ratio of each party's Seat-fraction to its First Preference Vote-fraction (S%/V%). However, the problem is how to combine such party-specific ratios (or deviations, S% - V%: see Table 1 above) into some measure of overall disproportionality. O'Leary (1979: 100) favoured the *Rae Index of Disproportionality* (RID), measuring party average disproportionality (contrast LHI above):

$$RID = 1/N \, \Sigma \text{ ABS (S\% - V\%)},$$

where N = Number of parties exceeding 0.5 percent of votes.

The palpable arbitrariness of this average disproportionality per party (why not a cutoff-point of 0.1 percent, or 5.0 percent of votes, for that matter?) may be redeemed somewhat by defining N as the ‘effective number of parties’ (Laakso and Taagepera, 1979):

$$N_1 = 1 / \Pi P^P \quad \text{or} \quad N_2 = 1 / \Sigma P^2,$$

where P = Party Vote-fraction or Seat-fraction;
and Π = Product (over all parties).

Taagepera and Shugart (1989: 260) preferred N_2 on practical grounds; though (entropy-based) N_1 enjoyed “equally good conceptual credentials”.

Gallagher (1991) argued that RID was “too sensitive to the number of parties”; to which LHI was “much too insensitive”. Accordingly, he proposed a “least squares index”: the *Gallagher Index of Disproportionality*,

$$GID = \left(\tfrac{1}{2}\Sigma \text{ (S\% - V\%)}^2\right)^{1/2}.$$

Nevertheless, Gallagher (1991: 47) considered “probably the soundest of all the measures” the *Sainte-Laguë Index*,

$$SLI = \Sigma \text{ (S\% - V\%)}^2 / V\% = (\Sigma \, S\%^2 / V\%) - 100 \, \%.$$

Unfortunately, SLI ranges theoretically from zero to infinity; which Gallagher acknowledged was “less easily interpreted” than LHI or GID (ranging 0 - 100 percent). Thus in the 1997 Irish Presidential Election, AV First Count LHI = 55 percent (complementing PTR = 45 percent: President McAleese's First Preference Vote-fraction: Table 2 below); whereas SLI = 121 percent!

Table 2: AV Party Vote-fractions, Seat-fractions and Deviations, by Count: Presidential Election, Irish Republic, 1997

Candidate (Party)	Vote-fraction (V%) First	Vote-fraction (V%) Final	Seat-fraction (S%)	Deviation (S%-V%) First	Deviation (S%-V%) Final
Total	100.0	100.0	100.0	0.0	0.0
McAleese (FF)	45.2	55.6	100.0	+54.8	+44.4
Banotti (FG)	29.3	39.2	0.0	-29.3	-39.2
Scallon (Ind)	13.8	0.0	0.0	-13.8	0.0
Roche (Labour)	7.0	0.0	0.0	-7.0	0.0
Nally (Ind)	4.7	0.0	0.0	-4.7	0.0
Non-transferable	0.0	5.2	0.0	0.0	-5.2
Over-represented	45.2	55.6	100.0	+54.8*	+44.4†
Under-represented	54.8	44.4	0.0	-54.8	-44.4

* First Count LHI = 54.8 percent: PTR = 45.2 percent.

† Final Count LHI = 44.4 percent: PTR = 55.6 percent.
Source: Irish Times, 1 November 1997.

Lijphart (1994: 60) preferred GID as steering “A middle course between the Rae and Loosemore-Hanby indices. Its key feature is that it registers a few large deviations much more strongly than a lot of small ones”; and contrasted two hypothetical elections (abstracted in Table 3 below).

Without defining any ‘Lijphart Proportionality Criterion’, he maintained that Election 1 was “highly disproportional” (GID = LHI = 5.0 percent); whereas Election 2 was “highly proportional” (GID = 2.2 percent; but LHI = 5.0 percent). Ironically, his intuitively “much more proportional” Election 2 yielded the higher SLI, considered by Gallagher (1991: 49) “the standard measure of disproportionality” !

Woodall (1986: 45) preferred the *Farina Index*,

$$FI = \cos^{-1} (\Sigma S\% V\% / [\Sigma S\%^2 \Sigma V\%^2]^{½}).$$

FI is the angle between two multi-dimensional vectors, whose coordinates are Party Seat-fractions and Vote-fractions: theoretically ranging between 90° (cos FI=0) and zero degrees (parallel vectors: exact PR). As a fraction of a right angle, FID = FI/90°; so ranging 0 - 100 percent (instead of 0 - 90°).

In Table 3, FID (like RID and GID) evaluates Election 1 as more disproportional than Election 2. However, as Hill (1997) recognised, FID also poses problems of interpretation; remaining a far cry from the pristine simplicity of LHI.

Hill (1997) reproached PTR and other measures (their “fatal flaw”) as confined to Party First Preferences. Nonetheless, he acknowledged that the concept of Total Representativity may be generalised (e.g. from Party to ‘Cumbency’, Gender and Name: Kestelman, 1996); and extended beyond the STV First Count. Yet he regarded Final Count PTR as merely comparing STV with itself!

Table 3: Five Measures of Overall Disproportionality: Two Hypothetical Elections

Party	Election 1 Votes (V%)	Election 1 Seats (S%)	Election 2 Votes (V%)	Election 2 Seats (S%)
Total	100	100	100	100
A	55	60	15	16
B	-	-	15	16
C	-	-	15	16
D	-	-	15	16
E	-	-	15	16
F	45	40	5	4
G	-	-	5	4
H	-	-	5	4
J	-	-	5	4
K	-	-	5	4
Disproportionality Index (percent)*				
LHI	5.0		5.0	
RID	5.0		1.0	
GID	5.0		2.2	
SLI	1.0		1.3	
FID	6.2		4.9	

*As defined in the text above.

Minor / Micro-Parties

As Hill (1997) implied, minor parties and independents (‘microparties’ — representing nobody but themselves) may need disaggregating before calculating overall measures of disproportionality. Exact LHI necessitates disaggregating the votes for every represented party (and elected independent) from unrepresented parties; as in Table 1 above. SLI may also be calculated without disaggregating unrepresented parties.

On the other hand, exact GID requires disaggregating even unrepresented party votes. Moreover, in evaluating a few large deviations (S% - V%) as more disproportional than many small deviations, with the same total deviation (and hence LHI), GID implies that, the more fissiparously people vote, the more they *deserve* to be under-represented. In contrast, LHI consistently measures the total under-representation (Σ S%-V%) of all under-represented voters.

Conclusions

Gallagher (1991: 33-34) lamented that “There is remarkably little discussion of what exactly we mean by proportionality and how we should measure it ... how do we decide which is closer to perfect proportionality?” — when comparing different elections. Notice already two different senses of the term ‘proportionality’ here! Hence my preference for the term ‘representativity’ for measures admitting matters of degree to the relationship between votes and seats.

Gallagher (1991: 46) reported that, at 82 national elections in 23 countries (1979-89), LHI, GID and SLI (but not RID) proved impressively correlated with each other: so why complicate matters? Besides, measuring average disproportionality (RID) necessitates counting parties — a rather moveable feast — and there seems little virtue in quantifying some hybrid between the distinct concepts of total and average disproportionality.

There remains legitimate scope for debating the relative merits of STV first or final preference representativity, in national aggregate or constituency average, respecting party or other considerations. In evaluating the representativity mediated by different electoral systems, no measure is perfect.

A generation after its introduction (Loosemore and Hanby, 1971), LHI survives relatively unscathed. I remain peculiarly susceptible to the complement (PTR) of that simplest LHI; doubting whether more complex measures of overall disproportionality would materially affect electoral comparisons (for example, STV representativity by District Magnitude: Kestelman, 1996).

References

1. Dáil Éireann (1998): 28th Dáil General election, June 1997. Election Results and Transfer of Votes. The Stationery Office, Dublin.

2. Dunleavy, P *et al* (1997): Making Votes Count. Replaying the 1990s General Elections under alternative electoral systems. Democratic Audit, Colchester.

3. Gallagher, M (1991): ‘Proportionality, Dis-proportionality and Electoral Systems’: *Electoral Studies* 10, 33-51.

4. Hill, I D (1992): ‘STV and Party Proportionality: Another View’: *Representation* 31: 113, 19.

5. Hill, I D (1997): ‘Measuring proportionality’: *Voting matters* 8, 7-8.

6. Jenkins, R (1998): The Report of the Independent Commission on the Voting System. Cm 4090-I. The Stationery Office, London.

7. Kestelman, P (1996): ‘Is STV a form of PR?’: *Voting matters* 6, 5-9.

8. Laakso, M and Taagepera, R (1979): '"Effective" Number of Parties: A Measure with Application to West Europe': *Comparative Political Studies* 12, 3-27.

9. Lijphart, A (1994): Electoral Systems and Party Systems. Oxford University Press.

10. Loosemore, J and Hanby, V (1971): ‘The Theoretical Limits of Maximum Distortion: Some Analytic Expressions for Electoral Systems’: *British Journal of Political Science* 1, 467-477.

11. McBride, J P (1997): ‘Positive Aspects of PR-STV for the Irish Political System’: *Representation* 34:1, 7-12.

12. O'Leary, C (1979): Irish Elections, 1918-1977: Parties, Voters and Proportional Representation. Gill and Macmillan, Dublin.

13. Taagepera, R and Shugart, M S (1989): Seats and Votes. Yale University Press, New Haven.

14. Woodall, D R (1986): ‘How Proportional is Proportional Representation?’: *Mathematical intelligencer* 8:4, 36-46.

15. Woodall, D R (1994): ‘Properties of Preferential Election Rules’: *Voting matters* 3, 8-15.

All correspondence regarding *Voting matters* should be addressed to:

B A Wichmann, The Editor, *Voting matters*, The Electoral Reform Society, 6 Chancel Street, London, SE1 0UU.

or using e-mail to Brian.Wichmann@freenet.co.uk.

To aid production, the Editor would welcome contributions on IBM-PC discs (with a printed copy as well) or to Brian.Wichmann@freenet.co.uk.

Voting matters

for the technical issues of STV

The Electoral Reform Society

Issue 11 April 2000

Editorial

This publication has now entered the Internet age! Future issues will appear on the ERS Web site (http://www.electoral-form.org.uk). Those who have no access to the Internet and do not wish to do so, need not be concerned, since a printed copy will continue to the available from ERS as before.

As Editor, I will ensure that those without Internet access are at no disadvantage. On the other hand, I would be happy to receive articles by e-mail to Brian.Wichmann@freenet.co.uk. Material should be in standard formats, such as HTML or PDF (and also RTF), rather than in proprietary word-processor formats.

Although the ERS Web site will have the current issue, there will be some delay in conversion and checking before it will be available there. Hence the printed copy should be available first, and that version should be regarded as the authoritative source (due to conversion and presentation problems with HTML).

I hope to arrange for all the back issues to be available on the Internet via a suitable Archive site. I have prepared a ‘combined’ issue for all of Issues 1-10, which is available from me in electronic format (HTML and PDF). Unfortunately, since this combined issue amount to 112 pages, it has not been possible for it to be professionally printed, since the cost is excessive for the likely sales.

The delay in this issue indicates the continuing problem of the lack of material from a small authorship. I am hoping that exposure of the material to international access via the Internet will encourage other parties to contribute in the future.

In the first article, Hugh Warren suggests a way of merging STV with FPTP, at least as far as the ballot itself is concerned. Could this encourage STV counting? Comments are welcome.

Philip Kestelman provides another article on proportionality with reference to the Jenkin's proposals.

Earl Kitchener makes a suggestion that Borda scores should be used to break tie rather than relying on a random choice (at least in the first instance).

My own article on checking two STV computer programs has proved controversial due to the issue of quota-reduction which is one of the new features in the 1997 edition of the ERS hand-counting rules. This issue is explained in the following article by David Hill; and Colin Rosenstiel, as co-author of the new rules, provides a response. Readers should judge for themselves whether a revision to the rules is required to ensure that no ambiguity exists.

Brian Wichmann.

Incorporating X-voting into Preference voting by STV

C H E Warren

Hugh Warren is a retired mathematician

1. Introduction

One of the thing said by many people, particularly by those who have used the X-voting system for many years, and by journalists, is that preference voting by STV is difficult to understand. However much advocates of preference voting by STV may find this view unjustified, and itself difficult to understand, they must accept that it is a view that is expressed, and no doubt genuinely held by a lot of people.

The purpose of this paper is to make the point that, instead of trying to win over the X-voting enthusiasts to the STV way of voting, consideration should be given to allowing the X-voting enthusiasts into the preference voting by STV system.

2. The Basic Idea

The basic idea is that, in addition to those who wish to vote in the STV way by showing preferences 1, 2, 3, .. in the recognized way, those who wish to vote by putting an X against the candidates they wish to see elected should be allowed to do so, provided of course that they do not put an X against more candidates than the number to be elected.

3. Interpretation of the Ballot Paper

With some ballot papers marked in the STV way by preferences 1, 2, 3,.. and some marked by an X against a number of candidates, the way in which it is suggested that the two may be accommodated is to treat the X votes as equal preference for a first preference candidate.

The allowing of equal preferences in the STV system is a matter which has been talked about in the past, but usually ruled out on the grounds that it would make an already complicated system more complicated. However, to allow equality of preference to be exercised on the first preference only should not lead to seriously greater complexity.

4. The count

The count is not of course a matter with which the voters have to concern themselves, provided that they can be assured that it is being done in a fair way.

If there are, say, 10 candidates to be elected, then at the first stage of the count, each candidate will have a number of votes of value 1 from the preference votes, and a number of votes of value 0.1 from the X-votes.

From this point onwards the count can proceed just as if it were a regular STV count, except that, of course, when surpluses have to be transferred, it will only be the preference votes for which the amount retained will be reduced, thereby allowing some of the vote to be transferred to the next preference.

5. Conclusion

The advocates of preference voting by STV have been trying for over 100 years to beat the advocates of X-voting. There is an adage which says *If you can't beat them, join them*. What is proposed here is not so much a case of joining them as incorporating them.

It is possible that, in the course of time, the X-voters will see that their interests could be better served by going across to preference voting, but the proposal is not to try and force STV on them.

Editorial Comment

The above proposal effectively merges the voting methods of First Past The Post and STV, so that the user can choose which method to employ. However, given that an STV-style count is to be undertaken, it seems logical to make an extension to Warren's proposal as follows: Allow the voter to place any number of X's on the ballot paper. Each X counts as a first-preference value of $1/n$, where n is the number of X's. With this proposal, an election for a single candidate in which the voter judges two candidates as of equal merit and no others of interest, two X's can be used, counting as 0.5 for each. More significantly, in my own experience for some elections, one can have, say, 6 seats to fill, but one has knowledge of only, say 3 candidates. Under conventional X-voting (and Warren's proposal) one could place 3 X's and loose half of ones voting power. Under this suggestion, 1/3 of a vote would go to each candidate and there would be no loss of voting power.

AV-plus, PR and Essential AMS

Philip Kestelman

Nomenclature

Much like Proportional Representation (PR), Single Transferable Voting (STV) is not an electoral system but a *principle*. There are various forms of STV: single-member STV, better known as Alternative Voting (AV); and multi-member STV, using various counting procedures (with potentially different results).

In October 1998, the Independent Commission on the Voting System (ICVS) recommended AV-plus for electing 659 UK MPs: mostly in around 543 AV constituencies, with 15-20 percent compensatory MPs, in 80 relatively small Top-up areas (electing 4 - 11 total MPs per area, including one or two Additional Members). Compensating parties under-represented by Constituency MPs (AV), d'Hondt allocation of Top-up MPs would render total MPs semi-proportional to Second / Party Votes, with choice of candidate within party (Open List PR[4]).

Is AV-plus a form of PR? Is AV-plus an Additional Member System (AMS)? Indeed, is AV-plus a form of multi-member STV? Answers to all three questions depend on what you mean by PR, AMS and STV, respectively!

Proportional Representation

Ritchie (*Tribune*, 11 June 1999) has argued that

"The Jenkins Committee's recommendations have much to recommend them, but there is little more chance of them delivering a proportional result than there is under the present system".

His introduction of a probabilistic element is welcome: here comparing AV-plus with so-called 'First-Past-the-Post' (FPP).

Jenkins[4] estimated that, in the 1997 UK General Election (FPP), AV-plus would have reduced the "DV score" from 21 percent to 13.2 percent. Measuring Deviation from Proportionality, DV = Loosemore-Hanby Index = LHI[8]. LHIs of 4 - 8 percent represent practically "full proportionality"; and for AV-plus, Jenkins[4] claimed only 'broad proportionality'.

Compare other d'Hondt systems. In the May 1999 Scottish Parliamentary Election (FPP-plus: seven Top-up MSPs per Region × 8 = 56 / 129 = 43 percent), the Second / Party Vote LHI was 10.5 percent. Ironically, total MSPs proved more representative of First / FPP Votes (LHI = 5.4 percent)! In the May 1999 Welsh Assembly Election (FPP-plus: four Top-up MWAs per Region × 5 = 20 / 60 = 33 percent), the Party Vote LHI was 11.2 percent (*Guardian*, 8 May 1999).

In Britain, the June 1999 European Parliamentary Election LHI reached 14.1 percent (Closed List PR: 84 MEPs: 4 - 11 per Region: *Guardian*, 15 June 1999): 'broad proportionality'. Such pure d'Hondt seat allocation favours larger parties, proving considerably less representative than Largest Remainder (which would have yielded LHI = 6.1 percent).

Over the last 10 Irish general elections (multi-member STV, 1969-97), aggregate First Count LHI averaged 7.0 percent (ranging 3.4 - 12.9 percent between elections: from 'full PR' down to 'broad PR' in 1997). Between three- and five-member STV constituencies (averaging 7.0 and 7.4 percent, respectively), LHIs differed insignificantly[7]. In the June 1998 Northern Ireland Assembly Election (six-member STV), First Preference LHI was 6.6 percent (*Irish Times*, 29 June 1998).

Additional Member Systems

Now used in Germany, New Zealand, Scotland and Wales, FPP-plus is frequently referred to misleadingly as *the* AMS. Thus Bogdanor[2]:

"the additional member system is, conceptually, a 'closed' list system ... it combines many of the faults of the first-past-the-post system with many of the defects of list systems of proportional representation".

Confusingly, Bogdanor was alluding to "a variant of the German system", recommended by the Hansard Society Commission on Electoral Reform: FPP without separate party voting, topped-up regionally with FPP 'best losers' (25% of all MPs[1]).

At the 1994 German General Election, 328 Constituency MPs were elected by FPP (First Votes); d'Hondt allocating 328 Top-up MPs, in 16 Regions, according to Second Votes (Closed List PR[9]). However, Second Votes may indicate voters' *second* preference parties[5]; as suspected in the 1999 Scottish and Welsh elections (*Times*, 8 May 1999):

"All electors then had a second vote. This should have been used to indicate their favourite political party. There is widespread confusion on this point and the fear that some people thought that they were being asked for their second preference".

Voting separately for constituency MPs and parties — One Voter *Two* votes — may well encourage tactical (insincere) voting. Especially in areas safe for the most-favoured party, a Second Vote for that party would elect no Top-up MP (and thus be wasted); and it would be more rational to vote for a less-favoured party, against a least-favoured party[4].

The average area represented by a German MP under FPP-plus in 1994 was over 20 times that of 656 FPP constituencies. In contrast, the mean area covered by each MP under AV-plus, with two Top-up MPs per area, would be only three times that of 659 FPP constituencies — just like three-member STV!

STV-plus

It is not widely realised that, in Malta since 1987, five-member STV has operated with a conditional AMS[6]. At the 1981 General Election, the Nationalist Party received an absolute majority of First Preferences (50.9 percent), but a minority of STV seats (31 / 65 = 47.7 percent).

Public outrage forced a constitutional amendment, guaranteeing a bare parliamentary majority to a party exceeding half of all STV First Preferences. At the 1987 Maltese General Election, the Nationalists won the same majority of First Preferences (50.9 percent), and minority of STV seats (47.7 percent); and therefore received four additional seats (totalling 35 / 69 = 50.7 percent of all MPs).

The 1992 General Election required no compensatory seats. Yet at the 1996 General Election − with fine impartiality − the Maltese Labour Party won 50.7 percent of First Preferences, but only 47.7 percent of STV seats! Accordingly, for a bare parliamentary majority, Labour received four additional seats (again totalling 50.7 percent of all MPs).

These few compensatory seats (4 / 69 = six percent) were occupied by STV Final Count ‘best losers’: runners-up for the party under-represented by STV alone. Thus Additional Members both stood for election and retained their constituency links.

The Maltese AMS (STV-plus) neatly solved an acute political problem. Incidentally, Malta remains a two-party polity, despite the opportunities for party fragmentation afforded by multi-member STV.

In the June 1998 Northern Ireland Assembly Election, the Social Democratic and Labour Party won more STV First Preferences than the Ulster Unionist Party (177,963 / 172,225 votes); but fewer Members (24 / 28 seats). That owed little to vote-transfers (*Irish Times*, 26 June 1998): even SDLP final ‘preferences’ exceeded those for the UUP (191,091 / 185,560 votes). The SDLP deserved five Additional Members (29 / 28 total seats proportionating SDLP to UUP).

STV-plus could well be generalised to British conditions; and would remedy the corruption of Party Vote Management — a form of tactical voting which disfigures Irish STV[3]. Party Vote Management involves a party’s supporters spreading their First Preferences evenly among its candidates: intended to keep them in the STV count for as long as possible (hoovering up stray transfers). In addition, each party nominates one more candidate than it expects seats; avoiding premature elimination through spreading its votes too thinly (‘over-nomination’).

Proportionating total (Constituency + Compensatory) seats to Party First Preferences, STV-plus could also reconcile the main parties (fearing the spectacle of disunity) to multi-member STV's wider choice of candidate. With each party's candidates competing for the voters’ affections, their First Preferences would *complement* each other in determining parliamentary party strengths under STV-plus. AV-plus could be redeemed likewise.

Essential AMS

AV-plus clarifies that AMS is not essentially FPP or Closed Party Lists. Both STV-plus (e.g. Malta), and the Hansard Society Commission variant of the German AMS, show that separate voting for Constituency Members and Parties is equally inessential. Anxious to avoid “all traces of a party list”, the Hansard Society Commission recommended that all candidates should stand in constituency elections[1].

Likewise, the ICVS stressed “open as opposed to closed lists for Top-up members”: Second / Party Votes offering a choice of candidate[4]. However, with three candidates per major party, preferential (rank-ordered, numbered) Second Votes are clearly better than categorical (single choice, X-marked) voting.

In that case, why not simply integrate First / AV with Second / Party votes: semi-proportionating total (AV + Top-up) MPs to AV First Preferences; with AV Final Count ‘best losers’ as Top-up MPs? Aiming to maximise AV First Preferences (and hence total MPs), each party would become highly motivated to nominate more than one candidate per constituency.

Thus could an improved AV-plus increase voter choice, both within and between parties. With a transferable choice of candidate within party, Party First Preferences are most sincere.

The ICVS argued that separating Constituency from Party votes would liberate voters from unwanted candidates of preferred parties; and that transmuting Constituency ‘best losers’ into Top-up winners would be hard to explain[4]. Valid against FPP-plus, both objections are much attenuated by more than one AV candidate per Constituency Party.

One Voter One Vote could then become far less wasteful than One Voter Two Votes. In both Scottish and Welsh elections, around half of both First and Second votes elected nobody (*Guardian*, 8 May 1999).

Moreover, the ICVS version of AV-plus (switching between preferential and categorical voting) is even more complicated for voters than multi-member STV. Indeed, it has been argued — rather cruelly — that its very complexity would favour that next step!

Conclusions

ICVS-proposed AV-plus is an Additional Member System (AMS), mediating semi-PR ('broad proportionality'). AMS is confined neither to FPP-plus nor to separate Constituency and Party List voting.

AV-plus would be simplified by integrating Constituency with Party voting, each party nominating more than one AV candidate per constituency; rendering total MPs semi-proportional to First Preferences; and exploiting the rich crop of Final Count 'best losers' as Top-up MPs. AV-plus could thus achieve much towards multi-member STV (which may also benefit from some mild topping-up: STV-plus).

It remains unclear why the Scottish Parliament includes more Top-up Members (43%) than the Welsh Assembly (33%): both more than the ICVS-proposed House of Commons (15-20%). With 20−25 percent Top-up MPs, AV-plus would increase Party Representativity ('proportionality').

In the end, *parties* must nominate parliamentary candidates; while the voter's predicament is paramount. With preferential voting in fairly small Top-up areas, AV-plus essentially places PR on a human scale. Commitment to that principle need not rule out debate on technical improvements (short of multi-member STV) before the Referendum.

References

1. Blake R (1976): The Report of the Hansard Society Commission on Electoral Reform. Hansard Society for Parliamentary Government, London.

2. Bogdanor V (1992): Proportional Representation: WHICH SYSTEM? Electoral Reform Society, London.

3. Gallagher M (1993): 'The Election of the 27th Dáil' in Gallagher M and Laver M eds: How Ireland Voted 1992. PSAI Press, Limerick.

4. Jenkins R (1998): The Report of the Independent Commission on the Voting System. Cm 4090-I. The Stationery Office, London.

5. Jesse E (1988): 'Split-voting in the Federal Republic of Germany: An Analysis of the Federal Elections from 1953 to 1987': Electoral Studies 7, 109-124.

6. Keesing's (1987 and 1996): Record of World Events 33, 35214 and 42, 41340.

7. Kestelman P (1996): Is STV a form of PR?, *Voting matters*, Issue 6, 5-9.

8. Kestelman P (1999): Quantifying Representativity, *Voting matters,* Issue 10, 7-10.

9. Mackie T and Rose R (1997): A Decade of Election Results: Updating the International Almanac. Centre for the Study of Public Policy, Strathclyde University, Glasgow.

Tie-Breaking in STV

Earl Kitchener

It is a fundamental principle of STV that later preferences should not affect the fate of earlier ones; this encourages sincere voting, but means that some arbitrary or random choice must be made to break ties, which can give unreasonable results.

An extreme case can arise where there is one seat and the electors are the same as the candidates; for example, if a partnership is electing a senior partner. Each candidate may put himself first, and all, except candidate A, put A second. Under most present rules, one candidate then has to be excluded at random, and it may be A. There is no way of getting over this unreasonable result without looking at later preferences, and the system of Borda scores is probably as good as any; with N candidates, N-1 points are allotted to a first preference, N-2 to a second, and so on. If it were desired to increase the importance of early preferences, the interval between values could be increased for early preferences. Ties in this system would be very rare, and it could be used to break ties in the normal STV counting.

In the above example no candidate or voter could reasonably object to the result, but in a real election, reported by Hill[1], with four candidates for one place, the voting was:

```
A B C     1
B A D     1
A C D B   1
B C A     1
```

The quota is two, which both A and B have. Under the proposed system A, with nine, beats B's eight. The second voter may complain that his second preference, for A, enabled A to beat his first preference. If the second voter had known in advance how the others were going to vote, he would not have put A second; but it is not unusual in small STV elections for a voter to find that if he had known the other voters' intentions he would have voted differently. He has got his second preference in, so has not much to complain about. In view of the uncertainty of voting intentions it is doubtful whether the proposed rule would lead to insincere voting, and it would avoid the possibility of A being unreasonably excluded in the first example. It has the virtue of satisfying Woodall's “No support” property[2], that no candidate who is not listed by any voter should be elected unless every candidate who receives some support is elected.

Hill has described a Sequential STV system[3] which deals in a more general way with the problem of premature exclusion of a candidate with few first preferences, but many other early ones; Hill does not recommend it, because of the breach of the rule against looking at later preferences. The present proposal, being confined to tie-breaking, might be less likely to lead to insincere voting, which is the main (and perhaps the only) objection to looking at later preferences.

References

1. I D Hill, ERS Document TC 95/13, September 1995.
2. D R Woodall, Properties of Preferential Election Rules. *Voting matters*, Issue 3 (1994), p10
3. l D Hill, Sequential STV. *Voting matters*, Issue 2 (1994), pp5-7

Checking two STV programs

B A Wichmann

Last year, I received a request from the Electoral Reform Ballot Services to ‘validate’ the computer software that they use to perform elections for their customers. Before that work was finished, I had another request from ERS itself to re-certify the program used to perform elections in the Church of England. Since there was a substantial overlap between both of these activities, these are reported together.

The checking undertaken was merely to ensure that the election results reported were as required by the respective rules. Hence many issues which might be of interest were not examined, such as: the user-friendliness, speed and memory requirements, number of satisfied users, maturity of the program, etc. In fact, the two programs which were tested are very different: David Hill's program is a complete system for data entry and edit, counting and presentation of the results and has been available for some years. In contrast, Keith Edkins' program is solely a counting program and is a recent development.

ERBS's requirements were identified as mainly to check a program that implements the ERS rules that were published in 1997[1] (ERS97). However, their requirements are significant in terms of the capacity required, amounting to the ability to handle up to 350 candidates and up to 250,000 votes. In principle, modern computers have no inherent difficulty in handling elections of that size, provide the software is designed appropriately.

If software is to be shown to be reliable, then a large number of test cases need to be run, or an alternative means needs to be devised to show logically that all the relevant functionality is correctly implemented. In performing the first certification of the Church of England rules in 1990, the technique adopted was to ensure that all the code in David Hill's counting engine was executed, and that the election results obtained were correct (checked by Eric Syddique). It was not thought that the same technique could be applied effectively for the ERBS validation, so the use of many tests was used instead.

If high reliability is to be demonstrated then several hundred tests should be run (corresponding to some years of use by ERBS). This immediately gives a difficult problem — how can one be assured that the result produced by the computer is correct? Initially an attempt was made to determine a small number of tests which performs all the relevant functionality which would then make manual checking feasible. However, the individual actions in ERS97 are quite numerous and difficult to identify — for instance, the result sheet does not state many specific actions undertaken during a count. Hence it seemed that the best means for undertaking the checking was to compare two programs for the ERS97 rules which were available.

Comparing two programs to increase reliability is not widely regarded[2], but in this case, the two programs were known to have very different internal workings and were quite independently developed. Hence it was thought that the comparison would be effective.

Unless comparisons can be made automatically by program, the number of tests will be limited to a level which would not give the assurance needed. Hence to facilitate such comparison and to avoid the need for the STV programs to produce elaborate printing, an output format was designed that could be input into a spreadsheet for printing. This format is logically just the conventional Result Sheet, but specified so that mechanical checks, such as those on row and column arithmetic, can be made. I am grateful to both authors that they amended their programs to produce this output since the testing would have been very tedious without that. Two small differences were located between the programs but an analysis showed that neither could change the result. Finally, the comparisons were automated which resulted in a successful validation of Keith Edkins' program.

No formal validation was undertaken of David Hill's program for these rules, but, of course, the same results were obtained. The program is not designed to handle ERBS's very large elections. It currently has 50 as its maximum number of candidates. ERBS would also wish for Colin Rosenstiel's interpretation of the quota reduction rules to be applied, but this has not been implemented, as explained in David Hill's article[3].

A number of issues arose from the validation as follows:

Quota reduction

A logical problem has been noted by David Hill in ERS97 which arises when the quota is reduced before any candidate is elected. This issue is defined and discussed in a separate article in this issue[3]. The consequences for this validation was that no comparison was possible when this situation arose since David Hill's program does not produce a result, due to the uncertainty in the meaning of the rules. The problem can be regarded as serious, since around 25% of those tests which are based upon real elections involved quota reduction. I decided that I could not formally sign my validation report, since, in my opinion, the meaning of the rules was sufficiently uncertain in this respect. Subsequent to undertaking this work, an analysis showed that the problem could only arise when transfers occurred after quota reduction. For instance, this cannot happen when there is only one seat. An analysis of my election data suggests that the quota reduction problem actually arises in about 12% of real elections. Readers can decide for themselves the significance of this problem from the two articles about quota reduction in this issue[3,8].

New data base

The data base of election data described in *Voting matters*[4] has been substantially enhanced as a result of both validations. This data is now available on a CD-ROM. In order to facilitate the collection of data from real elections, a program has been written, available as a MS-DOS/Windows program, which produces an anonymous version of election data by taking a statistical sample. Anybody can therefore add data to the collection without concern for the confidentiality of the source. (The data base contains the results for each election for the two rules being considered here, and also for the Meek rules.)

Capacity tests

In order to check that large elections could be handled, a program was written to generate large test data together with the results in result sheet format. This technique showed that these large tests can indeed be handled by any modern PC.

Tie-breaks

If an election requires the use of a tie-break, then a computer program makes a random choice. When comparing two programs, such a tie-break can result in two valid, but different results. This made the validation awkward, since either that election had to be ignored, or one of the programs had to be re-run with the option taken by the other program enforced. In most such elections the results were not compared, and as a result, a small difference between the two programs was not detected. The proposal to resolve tie-breaks by Borda scores would largely avoid this problem[7].

Church of England validation

Since the objective here was to revalidate David Hill's program, little would be gained in repeating the activity undertaken for the first validation. There were two changes to the Church's specification: a small change to rectify the Lichfield anomaly (which influences the main counting logic, see below), and the much larger change to add the handling of constraints. The logic used to handle constraints is specified in *Voting matters*[5].

The testing of the main counting logic relied upon the previous testing and the clearance of the Lichfield anomaly. Also, all the tests run were checked for the correctness of the row and column arithmetic. Hence the main effort was in checking the constraint handling.

The new Church of England rules (GS1327)[6] merely specify the actions to be taken during the count using the concept of candidates which are *doomed* or *guarded*. A doomed candidate is one that cannot be elected if a conformant result is to be obtained. A guarded candidate is one that must be elected if a conformant result is to be obtained. GS1327 does not specify the forms that the constraint might take, although it is understood that David Hill's program provides direct support for the constraints that are actually used by the Church. The program requires that every candidate is a member of one and only one constraint group. The constraints themselves specify the maximum and minimum number in a set of constraint groups.

A concern was that it might be possible to specify some constraints which would cause the program to compute for an effectively unbounded length of time. This does not seem possible, basically because the constraints are linear. However, a test was devised which produced a very large table of potential solutions which caused the program to produce a message that insufficient computer storage was available. David Hill has subsequently modified his program to use a file for the table within the counting process which now handles even this case.

Although the program provides direct support for only one form of constraint, indirect support is provided for a much larger range of constraints. As an example, suppose that the constraint groups are Scottish, English and Welsh. A constraint that is not directly support would be that the number of English elected is greater than the number of Scottish elected. However, the indirect method was capable of handling this case.

The approach to testing constraints was to take some elections from the data base (which are like real elections) and add constraints and then check for a conformant result. It was thought that 13 tests adequately covered the implementation of the constraint logic. It appears that the released program handles constraints which are very much more complex that would arise with Church of England counts.

Lichfield anomaly

A problem arose with the use of previous rules which resulted in the change to the rules even when constraints are not being used. This is called the Lichfield anomaly after the diocese where it arose. A simple test case (based upon an example from David Hill) would be to elect 2 from 5 with the following voting pattern:

```
20 AC
13 B
12 C
 2 DB
 1 EB
```

Under the old rules, even though exclusions were one at a time, A's surplus redistribution would be deferred, because it could not change who were the bottom two. Under the new rules it is not deferred because it could change who is the bottom one.

Old rules

```
A  20  20  Elected
B  13  +1   14  +2   16  Elected
C  12       12           Elected
D   2        2  −2    0
E   1  −1    0
```

New rules

```
A  20 −4  16  Elected
B  13
C  12 +4  16  Elected
D   2
E   1
```

A large election

The original certification of David Hill's program did not cover (as it really should have done) the data preparation side. Hence this time, an effort was made to use and test the input logic of the program. A large election was input, both by use of a text editor, and by use of David Hill's program with all the checking options enabled. The conclusion from this was that double-entry should be used in almost all circumstances, since several data entry errors would otherwise be undetected. On the other hand, the program behaved perfectly. (A few points were noted on the user interface, which has resulted in some improvements to the released version.)

Conclusions

Suitable techniques can be used to check STV software. The results have revealed some defects in the programs involved, which, of course, have been removed. However, in fairness to the authors, it is unclear if any of these defects would have remained undetected. Hence the main gain is additional confidence in the software and a reduced risk that such a program would fail during an actual count.

Copies of the full report on both validations are available from the author. Electronic copies are available by mailing a request to Brian.Wichmann@freenet.co.uk.

References

1. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote. 3rd Edition, ERS. 1997.

2. J C Knight and N G Leveson. An experimental evaluation of the assumption of independence in multiversion programming. *IEEE Trans. Software Eng*. Volume 12, No 1. 1986.

3. I D Hill. Quota Reduction in hand-counting STV rules. *Voting matters*, Issue 11 pp8-9.

4. B A Wichmann. An STV Database. *Voting matters*, Issue 2. p9. 1994.

5. I D Hill. STV with constraints. *Voting matters*, Issue 9 pp2-4.

6. GS1327: General Synod, Single Transferable Vote regulations 1990 and 1998. (Obtainable from Church House Bookshop, Great Smith Street, London SW1P 3BN.)

7. Earl Kitchener. Tie-breaking in STV. *Voting matters*, Issue 11 p5.

8. C Rosenstiel. The problem of surpluses when the quota is reduced. *Voting matters*. Issue 11 p10.

Quota reduction in hand-counting STV rules

I D Hill

The 1997 ERS rules for STV[1] include a rule for reducing the quota if some votes become non-transferable before anyone has been deemed elected. In general, such a rule is to be welcomed, as the smaller the quota can legitimately be made the better.

However, in attempting to implement this rule in my STV computer program I ran into difficulties of interpretation. It may be that the circumstances that cause such difficulty would rarely arise in practice, but that is irrelevant. Rules, and programs derived from them, have to work in all circumstances. I wished to know whether the difficulties were real, or whether I was being over-fussy in imagining them, so I consulted a number of people, chosen as being knowledgeable in STV, and asked for their views on what the rules required with each of four examples. Their replies were sufficiently varied as to show that there is a real problem.

The rules in question are:

5.3.1 If a surplus arises at the first stage, select for examination all the papers which the candidate has received.

5.3.2 If a surplus arises at a later stage, because of the transfer of another surplus or the exclusion of a candidate or candidates, select only the last received batch of papers, which gave rise to the surplus.

With minor changes of wording those two rules are as in the previous edition, but we now also have:

5.4.8 If any papers have become non-transferable before any candidate has been deemed elected, recalculate the quota as in paragraph 5.1.6, ignoring the non-transferable vote.

The first three examples were as shown below. The fourth was somewhat different as it did not do what was intended and it is better here to show the intended case instead of the unintended one.

```
Election 1  Election 2  Election 3   Election 4
 17 AB..     14 AB..     17 AB..      12 AB..
 11 BC..     11 BC..     11 BC..      11 BC..
 10 CD..     10 CD..     10 CD..      10 CD..
 10 DA..     10 DA..      9 DA..      10 DA..
  6 E(plump)  6 E(plump)  6 E(plump)   6 E(plump)
              3 EAC..     1 EAC..      5 FAD..
```

In each of these there are 2 seats to be filled and 54 votes. In each case the initial quota is 54/3 = 18. In each case 6 votes become non-transferable before any candidate is deemed elected, so the quota is reduced to 48/3 = 16. In each case candidate A now has over a quota of votes. How do the rules require A's surplus to be dealt with?

As a result of the exercise, it seems clear to me that trying to implement these rules would not be sensible until they have been amended for, even in the simplest cases, elections 1 and 2, it is not absolutely clear where A's surplus should go, since it cannot really be said that the papers concerned ‘gave rise to the surplus’. In election 3 there was much disagreement about how much goes to C and how much (if any) to B. If experts disagree, to the extent that was observed, on what the rules mean, what hope is there for an ordinary returning officer?

In election 4 the ‘gave rise to the surplus’ wording is even more far-fetched than in the other cases, and my own view is that this case is not catered for in the rules.

I am grateful for an additional case that was suggested to me later by one of those whom I had consulted:

```
Election 5
14 AB..
11 BC..
10 CD..
10 DA..
 3 E (plump)
 6 EAC..
```

I would probably have got this one wrong, as my first reaction on seeing it was ‘No problem here’, because A has already got more than the original quota by the time it is known that any votes have become non-transferable, so quota reduction would not apply, but not so. Although exceeding the quota, A is not actually deemed elected (para 5.4.9) until after the quota reduction has been made (para 5.4.8).

My own view is that, in principle, the right way to do such quota reduction is to re-start the election after the reduction, with the equivalent of a new Stage 1, treating all excluded candidates as if withdrawn, but the wording of the current rules does not seem to support that. For the moment what is wanted is the publication of a clarifying amendment to the rules, so that users can know how to proceed. This issue can be resolved only by a properly authorised statement from the ERS Council.

Reference

1. Newland R A and Britton F S. How to conduct an election by the Single Transferable Vote. 3rd edition. Electoral Reform Society. 1997.

The problem of surpluses when the quota is reduced

Colin Rosenstiel

Normally a candidate elected with a quota receives ballot papers at the stage at which their votes first exceed the quota. Since the changes to the rules made in 1997 it is now possible for a candidate to be elected with a surplus at a stage where they receive no ballot papers. If the quota is reduced at the same stage from a larger number than the candidate's current vote to a figure below that vote they can be declared elected with a surplus. It has of course been possible for a candidate to be elected without a surplus at a stage where they receive no ballot papers since the introduction of the second edition of the rules in 1976.

The candidate's surplus does not then arise from papers received at that stage, the rule heretofore. However the principle remains that their surplus is derived only from the last-received parcel of papers, their first preferences if no papers have been received since then. The rules in detail say:

5.3 Transfer of a surplus

5.3.1 If a surplus arises at the first stage, select for examination all the papers which the candidate has received.

5.3.2 If a surplus arises at a later stage, because of the transfer of another surplus or the exclusion of a candidate or candidates, select only the last received batch of papers, which gave rise to the surplus.

Any difficulty in interpreting this wording is because of the possibility of different interpretations of the term ‘arises’. The candidate declared elected due to the quota being reduced may not have received any papers at the stage in question. I would therefore maintain that only perversity could lead to the conclusion that the word ‘arises’ could refer to any other stage than the one at which the papers were received and that the most recently received parcel of papers should be the ones used to transfer the surplus as has always been the case.

It is also possible for the papers forming a surplus to be worth less than the value of the surplus. This is again not new, in terms of transferable papers, and is to be treated in the same way — no paper may be transferred at a higher value than it had when received by the candidate with the surplus.

In his article Dr Hill[1] gives a number of examples which he claims there are difficulties over interpretation of the rules quoted above. He doesn't explain what the difficulties are. If the precise wording above is not applicable (which I argue above is not the case anyway) what rules does he imagine are to be followed?

There is also a problem about the importance of this supposed difficulty. The figure of 12% of cases is mentioned by the Dr Wichmann[2], though without supporting evidence. His original claim was for 25% of cases but it turns out that half were AV elections where no surplus can ever be transferred!

Bear in mind that the disputed cases require (a) a reduced quota (b) a surplus arising at a stage where the elected candidate receives no papers (c) that surplus to be transferred. Since the rule came in I have counted many elections. Just three had reduced quotas. In no cases did a surplus arise at a stage where a candidate received no votes, let alone such a surplus requiring to be transferred.

References

1. I D Hill. Quota reduction in hand-counting STV rules. *Voting matters*. Issue 11. p9.

2. B A Wichmann. Checking two STV programs. *Voting matters*. Issue 11. pp6-8.

Brian Wichmann responds

Colin Rosenstiel correctly quotes my article which on reflection might be confusing. The 25% refers to those elections in which, logically, quota reduction takes place. The 12% refers to those elections in which subsequent transfers take place. Nobody knows what fraction of the 12% are truly ‘ambiguous’ in the sense raised by David Hill. I would regard any significant percentage as quite unacceptable, since surely STV should be no less certain that First Past The Post. To avoid any problems, I would suggest that the Council of the ERS formally accepts a small wording change proposed by Colin Rosenstiel in a letter to David Hill dated 8th November 1998.

All correspondence regarding *Voting matters* should be addressed to:

B A Wichmann, The Editor, *Voting matters*, The Electoral Reform Society, 6 Chancel Street, London, SE1 0UU.

or using e-mail to Brian.Wichmann@freenet.co.uk.

Voting matters

for the technical issues of STV

The Electoral Reform Society

Issue 12 November 2000

Editorial

Issue 9 of *Voting matters* contained two articles on the vexed question of ordering candidates when preferential voting is used. In the first article here Joe Otten returns to this question in the light of some problems noted in the previous ‘solutions’.

In the next two articles, David Hill questions the suggestions made in two different articles that appeared in Issue 11. As often happens in this area, a suggestion which seems fine initially, may have subtle difficulties — at least as far as people other than the author are concerned!

My own article for this issue considers the effect of numerical accuracy of STV when using the Meek algorithm. Unlike the hand-counting rules, the algorithm itself does not define the accuracy that should be used, although omitting this information is the convention with numerical algorithms.

Bob Jones questions what one wants from an electoral system and considers the use of Decision Analysis to make sense of the conflicting requirements. Readers are invited to make their own contribution. The editor hopes that, given sufficient response, a further article might be appropriate which should provide a view from the entire readership of *Voting matters*.

A major article is provided by Simon Gazeley in which a new algorithm is proposed for a computer-based STV count. As is to be expected with such an algorithm, it will take a significant effort to validate. No doubt, if a program is produced to implement it, some ambiguities will be noted. Given an implementation, then comparisons should be straightforward. It appears that the algorithm is essentially more complex than, say Meek — but does that matter?

David Hill provides a third article which is surely a warning to all who advocate STV. We have no ‘standard’ for STV and in some Australian elections, the rules do not appear to give the benefits which one would expect.

Recently, an Internet group has been formed on STV. As editor, I will keep a watching brief on this, both to report material in *Voting matters* and also to encourage others to write articles. As is usual with Internet traffic, it is rather informal and not suitable for direct publication.

A combined issue for Volume 1 of *Voting matters* has been prepared. Unfortunately, it is not economic to print it, but it is available from me in the electronic format PDF which can be printed easily on most modern computers.

Brian Wichmann.

Ordered List Selection revisited

J Otten

Joe Otten is the author of a Windows program for the current ERS STV rules.

1. Problems with methods advocated in *Voting Matters* 9

I was struck by a comment by Hill[1] that Rosenstiel's alternative method to the use of constraints violated the principle that later preferences should not be allowed to count against earlier ones (I will refer to this as ‘the principle’ in this paper). This was because the method involved running repeated counts on the same vote profile, and thus a later preference may have its effect in one count when the fate of earlier preferences was still to be decided in a later count.

I realised that the same criticism could be levelled at the method I advocated for selecting an ordered list[2] (in this case of candidates for a party to offer at a European Parliament election conducted using a list system). It could also be levelled at the similar system proposed by Rosenstiel[3]. In each case multiple counts were used, and the result of one count could affect the result of another — by the use of a constraint in my case, or by overriding it in the other.

Example 1:

AC	2
AD	10
BC	10
C	8
DC	6

This gives the following results:

Vacancies	Results
1	C
2	AC
3	ABC

Both methods give the Result: CABD.

Suppose Rosenstiel's method was used, and those voting BC changed their vote to BDC, example 1 gives

Vacancies	Results
1	C
2	AC
3	ABD

Now, C gets last place, and B and D are tied for third. The tie is broken by looking at first preferences, so D is third. Then there is a similar tie for second between A and B, so B is second. Result ABDC. Voters have improved the position of B by changing later preferences.

My method would still give the order CABD with example 2, but would violate the principle given a similar example.

Wichmann[4] suggests using the Meek keep factor for determining the ordering, and this case is not so immediately obvious, since only one count is held. The Meek algorithm does not allow later preferences to influence whether earlier preferences may be elected. However later preferences may affect the size of the keep factor for elected candidates, and so if this is used to order the candidates, the principle is violated. Electing 3, this gives ABCD in example 1 and ABDC in example 2.

2. Using the Orange Book method

The Orange Book simply suggests that if an order is required, the order in which candidates are elected during the count should be used. This seems, on face value, to be inadequate for selecting a long list of candidates, since the contest for the significant top places would be rather similar to a First Past the Post election, with a few candidates above the quota being given positions dependent only on the numbers of first preferences received. Newland himself, author of the early editions of the Orange Book, indicates in his *Comparative Electoral Systems* why he thinks this method is wrong, advocating a top-down method.

The method appears to rest on the assumption that it is the determination of the whole membership of the list that is the primary purpose of the election. That is not the case. The purpose is that however many seats the party wins, the people thus elected are those who were selected by an STV ballot with the appropriate number of vacancies. Thus in the Green Party, where no more than 1 seat was won in any region, the top of the list should be the AV winner (as indeed they were, since the Green Party used a top-down method.) The Liberal Democrats won 2 seats in some regions, so there the appropriate selection would be that of the top two candidates by an STV election with 2 vacancies.

The problem is that the number of seats that a party will win is unknown at the time of selection. However, it may be reasonable to guess at that number. The order of election (orange book method) would give the order of the candidates elected in the selection ballot, and the reverse order of exclusion could determine the order of later candidates. If a party wins 1 more or fewer seats, the distortion might not be that great.

This does not seem entirely satisfactory, but I cannot see how better can be done without abandoning the principle.

3. Abandoning the Principle

A great many articles in *Voting matters* have discussed the principle. Some have suggested that it might be relaxed, for example to allow Borda scores to be used to break ties[5]. Personally I think the only strong argument against Condorcet style election rules is that they violate the principle. Therefore if the principle must be lost, we may as well look at later preferences more freely and use an election rule more in keeping with Condorcet principles, and do better than any of the methods advocated in *Voting matters* 9.

It seems to me that a great many voters would welcome a substantial benefit to a second or third preference at the expense of a small risk to a first preference. STV does seem to rest on the assumption that the strength of a voter's support for their first preference is such that other considerations are overridden. While I don't think this assumption is true for very many voters (except perhaps for die-hard party loyalists), it is right for STV to make it. It is right because it makes the task of voting much easier. The voter does not need to assess how his or her use of later preferences might affect the fate of an earlier one. The principle encourages voters to indicate their true preferences.

Nonetheless, if the price of the principle is reducing a contest to near equivalence to First Past the Post, I believe that price is too high. I suggest the next question is how may we reap the benefits of the information the principle denied us. In the one vacancy election, systems which violate the principle may benefit by being able to guarantee the election of the Condorcet winner if there is one. I seek now to generalise this benefit to the election of an ordered list.

4. Generalising Condorcet principles to multiple vacancies

Hill[6] describes the complexity that can arise when trying to generalise the concept of a Condorcet top-tier to a multiple vacancy election. However, if we are considering a list selection then we are not simply looking for a subset of all the candidates, but adding them one at a time to a list. This simplifies the problem somewhat. Also for the purposes of simplicity I shall refer to Condorcet to mean any single-winner rule satisfying the Condorcet Criterion. The manner in which cycles should be resolved is not a significant concern here; nor is whether Meek or ERS97 rules are used, although computer counting will be necessary.

The method which follows builds an ordered list from the top down. It, like Condorcet, does not use exclusions at all, but considers at every stage, all possible pairs of candidates for the next position to see if one beats all the others. Like STV, votes are retained by elected candidates so they have less or no influence on later positions.

The top position is elected by Condorcet (call this candidate P).

For every pair, X and Y of other candidates, we must determine which is preferred to the other for the second place. We calculate the result of an STV election between P, X and Y for 2 places (other candidates being withdrawn). This calculation determines whether X is preferred to Y or vice versa. We read off the support for X and Y after any surplus for P has been redistributed and this completes one element in the Condorcet result square. (Normally it is only of interest which of X or Y is elected in this election. However the magnitude of the difference in support will be relevant if a cycle-breaking method needs to be employed.) The calculation is repeated for all other pairs of candidates, not including P (or at least for as many pairs as are necessary to determine the winner). Call the candidate thus elected to position 2 Q.

We need to repeat this exercise for position 3, 4, 5, etc, and we now have more than one elected candidate. Each time we perform an STV count including all the elected candidates, PQR..., and a pair of unelected candidates X and Y, and no others, giving one element of the Condorcet result square as before. We then repeat this for every pair of unelected candidates, and add our new Condorcet-style winner to the list.

Applied to Example 1, the result tables look like this:

(+ values imply row candidate beats column candidate)

Condorcet (6 AV counts between 2 candidates)

	A	B	C	D
A		+2	-12	+6
B	-2		-6	-6
C	+12	+6		+4
D	-6	+6	-4	

Position 2: (3 STV counts with 3 candidates, C and two others)

AvB: C has a surplus of 2, which is non-transferable - A 12, B 10

AvD: C has a surplus of 6, which is non-transferable - A 12, D 6

BvD: C has no surplus - B 10, D 16

	A	B	D
A		+2	+6
B	-2		-6
D	-6	+6	

A is elected to position 2

Position 3: (1 STV count with all candidates)

BvD: A has a surplus of 3, which goes 0.5 to C and 2.5 to D - B 10, D 8.5

	B	D
B		+1.5
D	-1.5	

B is elected to position 3

Result: CABD

Changing the 10 votes from BC to BDC as before (example 2) creates a cycle:

Position 1:

	A	B	C	D
A		+2	-12	-4
B	-2		-6	-6
C	+12	+6		-16
D	+4	+6	+16	

D is the Condorcet Winner and is elected to position 1.

Position 2:

AvB: - A 12, B 10

AvC: - A 12, C 8 (D is guarded, so A is not elected)

BvC: D has a surplus of 4 which goes to C (strictly 3.96 with ERS97) - B 10, C 14

	A	B	C
A		+2	+4
B	-2		-4
C	-4	+4	

A is elected to position 2

Position 3:

BvC: - B 10, C 8.5 (C and D are guarded, so B is not elected)

	B	C
B		+1.5
C	-1.5	

A is elected to position 3.

Result: DABC

D and C have swapped places, as is reasonable given the change of votes from BC to BDC.

Instead of using a usual cycle-breaking rule, an alternative would be to combine the election for the position in question with the following one, elect two, and then go back to the first, where there are now only 2 candidates to choose from, so there can be no cycle. (This would be a normal STV election for the top two. Alternatively we could consider every possible triple, but this may lead to further cycles.)

This procedure is a synthesis of STV and Condorcet. At each position a Condorcet-winner is added to the list, once votes cast for already-elected candidates have been discounted (reduced in value) in the manner of STV. It is not vulnerable to the exclusion of potential winners with few first preferences.

It could also form the basis for a synthesis of STV and Condorcet for unordered elections, although this would be a solution looking for a problem as regular STV is available here. Seeking to elect *n* candidates we could apply the STV rule to every subset of *n*+1 of the candidates and see which *n* were able to beat off any individual challenger. As Hill[6] says, the subset of *n* with this property may not exist, or may not be unique. However the generalised Condorcet method above could be adapted in such cases to arbitrate between competing sets of candidates, or to provide a result where there appears to be none.

5. Summary of examples

	Ex 1	Ex 2
Repeated count rules:		
Rosenstiel /Bottom Up Overriding (R):	CABD	ABDC
Otten /Top Down Constrained (O):	CABD	CABD
Top Down Overriding (TDO):	CABD	CABD
Bottom Up Constrained (BUC):	CABD	ABDC
One count rules:		
Wichmann Meek (2 places) (WM2):	CABD	CABD
Wichmann Meek (3 places) (WM3):	ABCD	ABDC
Orange Book (1 place) (OB1):	CABD	CABD
Orange Book (2 places) (OB2):	ACBD	ACBD
Orange Book (3 places) (OB3):	ABCD	ABDC
Generalised Condorcet rule:		
Generalised Condorcet (GC):	CABD	DABC

I have not described the last two repeated count rules — they are hybrids of the Rosenstiel and Otten rules, which might be called Bottom Up Overriding and Top Down Constrained respectively. It is worth noting that BUC, like GC, does not use exclusions, (candidates already allocated to lower positions are withdrawn before the start of the next count) but with different results.

What are the best results? CABD seems to be a clear favourite for example 1. With example 2, the elementary conflict is that if the electorate were to be represented by one person, the best person (from an AV point of view) would be C, and if it were to be three, the best people would be A, B and D. Rules which take greater care over the top end of the list (O, TDO, WM2, OB1) therefore place C highly and those which concentrate on the bottom (WM3, R, BUC, OB3) place C low. Notably WM3 and OB3 place C low even in example 1.

We have, it seems, not entirely escaped from the consideration in point 2, of needing to know what position on the list is the crucial one. If it is believed that a particular position on the list, say 4th, is the key one, an STV count for 4 winners could be followed by BUC to fill the top 3 and O to fill the positions from 5 down (or R and TDO respectively).

As to be expected GC succeeds in finding the Condorcet winner D in Example 2, who is not found by any of the other methods. Obviously this is an example of my choosing, and I have no doubt that other examples may show GC generating inferior results.

6. Conclusions

I have described three broad approaches to ordered list selection, all of which are unsatisfactory. The methods used by the Green Party and Liberal Democrats violate the principle, but fail to take advantage of the information this releases. The Generalised Condorcet method uses this information but also violates the principle. The orange book method, used as described here, may lead to a severely distorted result if the guess is wrong.

While the methods described in 1, appear for the moment to be the most practical solution to the question of ordering, the fact that counts for differing numbers of candidates frequently produce inconsistent results undermines their credibility.

A significant source of these inconsistencies is changes in early exclusions or the order of exclusions and in which parcel of papers elects a candidate, resulting from the higher or lower quota. (Meek should be less vulnerable to two of these effects.) While my generalised Condorcet method conceals any comparable inconsistencies that might be present, the fact that it eliminates exclusions altogether, means that it should be robust against exclusion-related effects.

The disadvantages are greater complexity and probably a more frequent violation of the principle that later preferences should not count against earlier ones. It will also require considerably more computer time than the alternatives, which may be an issue with a very large election, particularly if Meek is used. It would not be desirable to adopt a rule that then had to be abandoned for very large elections.

I do not at this point advocate that a generalised Condorcet method is adopted. However, I think the idea has its merits, and I do believe the question of ordering demands further consideration. While a single rule may not be appropriate for all circumstances, it should be possible to narrow the field somewhat from that in section 5.

References

1. I D Hill. STV with Constraints, *Voting Matters* 9, pp2-4, 1998.
2. J Otten. Ordered List Selection *Voting Matters* 9, pp8-9, 1998.
3. C Rosenstiel. Producing a Party List using STV. *Voting Matters* 9, pp7-8, 1998.
4. B A Wichmann. Editorial Comment on 3, *Voting matters*, 9, p8, 1998.
5. Lord Kitchener, Tie-Breaking in STV, *Voting Matters* 11, pp5-6, 2000.
6. I D Hill, Trying to find a winning set of candidates. *Voting matters*, 4, p3. 1995.

Tie-breaking in STV

I D Hill

Earl Kitchener[1] puts forward a scheme for using Borda scores for tie-breaking within STV. In general Borda scores are not a sensible way of conducting elections, but for this one purpose it will seem preferable to many people, to use something that takes note of the wishes of the voters, rather than a resort to randomness. The question is whether any such scheme would cause more trouble than it is worth.

We need to remember that ties rarely occur except in the case of very small elections, but it is just those very small ones where voters can see what is happening, and where the effect of later preferences upsetting earlier ones may be most troublesome.

In the real case quoted by Kitchener, there were 4 candidates for 1 seat. The 4 candidates were also the voters but not everyone voted for themselves. The votes were

ABC	1
BAD	1
ACDB	1
BCA	1

giving an AB tie for first place whether judged by Alternative Vote or by Condorcet. Using Borda scores as tie-breaker, A is elected, but this is solely because of a third preference for A against a fourth preference for B.

Now Voter 2 has a right to be cross about that. He put A as second choice meaning, according to all the best explanations of STV, "If B is out of the running, then I wish to support A" but B was not out of the running at that point.

Suppose there were the same set-up the following year. Voter 2 is likely to decide to plump because putting in a second preference the previous year was to his disadvantage. But Voter 1 may realise this and decide that he must plump too to counteract Voter 2's plumping — then Voters 3 and 4 will need to think about their strategies.

Whether anyone decides to plump or not is not really the issue. What matters is that tactical considerations have been allowed in, where STV (in its AV version in this case) is supposed to be free of them.

It may seem a pity to decide it at random, but such looking at the votes only decides it on the grounds that Voter 3 preferred D to B whereas Voter 4 preferred A to D. Is that really relevant when D is clearly out of it anyway?

My own conclusion is that to look at later votes in such circumstances, by Borda scores or any other method, is not a good thing to do, but I recognise that it is a matter of judgement, not of a clear right and wrong.

Reference

1. Earl Kitchener, Tie-breaking in STV. *Voting matters*, Issue 11, p5, 2000

Mixing X-voting and preference voting

I.D. Hill

Hugh Warren[1] puts forward a plan to incorporate X-voting into an STV election, so that those who prefer it are not forced into STV against their will. The aim is very sensible but, as he says, the voters must "be assured that it is being done in a fair way". As Hamlet said: "ay, there's the rub". Is it possible to find a way that actually is fair and, equally necessary, will be accepted as fair by those concerned?

The Warren suggestion is to treat Xs as equal first preferences, treating each X as worth 1/*m* where there are *m* places to be filled. Now suppose, as he does, that *m* = 10. If two voters each plump for a single candidate, one using an X and the other using a 1 in marking the paper, would it be regarded as fair for the second of those to be treated as worth 10 times as much as the first? Surely not.

In an editorial footnote, Brian Wichmann suggests an alternative formulation, treating each X as worth 1/*n* where *n* is the number of Xs marked on the paper. That would solve the above difficulty, but only at the expense of introducing a new one.

Suppose two candidates get X-votes only, one getting 20 Xs each of value 0.5, because those voters used two Xs each, the other getting 40 Xs each of value 0.2, because those voters used five Xs each. The first then has a total vote value of 10, the second a total vote value of 8. So if one of the two is elected it will be the one getting 20 Xs, not the one getting 40 Xs. Would X-voters regard that as fair? I am quite sure that they would not. It is just this sort of situation that I presume that the Warren formulation was carefully designed to avoid.

Is there any way of doing it that everyone would think fair in all cases? I doubt it.

Reference

1. C H E Warren. Incorporating X-voting into Preference voting by STV. *Voting matters*, Issue 11, p2. 2000.

The computational accuracy using the Meek algorithm

B A Wichmann

Introduction

The Meek algorithm[1] is specified without regard to the accuracy of the computation (with the exception of the convergence criterion, which is not relevant to this paper). The formulation in Pascal uses the type **real** which is traditionally floating point, but this could have varying accuracy or even be replaced by a rational arithmetic package of unbounded precision. A natural question to ask is what computational accuracy is required to ensure that the 'correct' candidates are elected, ie, the same candidates as if infinite precision was used. We demonstrate by examples, that there are cases in which very high precision is required.

An example

If a candidate A has first preference votes which only just exceed the quota, then those who have given A as their first preference will have only a small fraction of their vote passed on to their subsequent preference. Moreover, if most of A's subsequent preferences are for B (say) and just one for C, then the fraction going to C can be made smaller still.

The above leads to the following example in which 3 seats are to be filled:

333	AX
333	AY
333	AZ
333	BX
333	BY
333	BZ
667	X
667	Y
667	Z
1	ABX
1	ABY
2	BAX

The total number of votes is 4003, which gives an initial quota of 1000.75. Since A and B each have 1001 first preference votes, there is a surplus to transfer after their election of a quarter of a vote. This implies that the weight associated with A and B is roughly (1−1/4000). This further implies that the vote ABX makes a contribution to X of roughly 1/16,000,000th of a vote.

After the election of A and B, one of X, Y or Z must be eliminated. In the cases above, it is clear this should be Z, since that candidate has no contribution from the last three votes, but X and Y do. However, if the implementation of Meek only recorded millionths of a vote, then the last three candidates would be regarded as equal, in which case, a tie-break would occur.

For this test, we are only concerned as to what happens at the third stage. If a tie-break occurs, we know that the implementation does not have the accuracy necessary to compute the same result that would arise from infinite accuracy.

The above example illustrates that the accuracy required to give the same result as with infinite precision is unbounded even with six candidates (since we can just use more votes to increase the accuracy needed). However, the same technique can be employed with more candidates to increase the accuracy without increasing the number of votes. For instance, with 69 candidates and less than 1,000 votes, one can produce an example requiring 127 decimal places! The full details of this are available from the author.

Conclusions

There are somewhat bizarre voting patterns in which the accuracy required by the Meek algorithm is high, if the same result is to be obtained as that which would result from infinite precision.

One cannot expect the accuracy provided by an actual implementation to be high enough to guarantee the same result as that from infinite precision. (The highest available accuracy that is easily provided on a modern computer is 17 decimal places.)

The examples used here involved only the first two stages of a count. However, an important property of the Meek algorithm is that there is no accumulation of rounding error from one stage to the next, since the state is just the (discrete) record of those elected and eliminated. The weights are not really relevant since they only provide a starting point for the next iterative step.

One could gauge the impact of computational accuracy if one knew the rate at which ties arose which are *not* due to an algebraic tie. If such a computational tie arose with my database of around 370 elections, then it should be detected. In work which involved comparing two implementations of Meek(using all these 370 elections), it is likely that one implementation would report a tie-break when the other implementation did not. Such an occurrence did not arise.

Hence the overall conclusion is that the accuracy of the existing implementation of 64-bits is sufficient in practice, but not theoretically if the requirement is to produce the same result as that given by infinite precision.

Reference

1. I D Hill, B A Wichmann and D R Woodall. Algorithm 123 — Single Transferable Vote by Meek's method. *Computer Journal*. 1986.

A Comparison of Electoral Systems using Decision Analysis

H G Jones

Bob Jones is a retired mathematician and former secretary Derbyshire Electoral Reform Group (DERG)

Introduction

Decision Analysis is a method by which comparisons between different courses of action may be evaluated in order to obtain a desired end product. In the field of electoral reform the end product is the best electoral system, and the means of evaluating different systems is by comparing how well they measure up to desirable features of such systems.

The idea of applying Decision Analysis to electoral systems was first suggested by Tony Cooper, chairman of DERG,in the late 1980s and initially the performance of a system

	Feature											
System	PRO-R	PRO-N	CHO-P	ONECM	EASV	EASC	EASBC	EW&E	LOC	PLOC	Total	Ranking
FPTP(SM)	3(9)	4(12)	0(0)	10(20)	10(20)	10(10)	2(2)	2(6)	10(30)	4(8)	55(117)	
AV(SM)	4(12)	5(15)	0(0)	10(20)	9(18)	9(9)	2(2)	2(6)	10(30)	6(12)	57(124)	
PL(MM))	10(30)	10(30)	0(0)	10(20)	10(20)	7(7)	10(10)	5(15)	0(0)	0(0)	62(132)	
PLRC(MM)	10(30)	10(30)	0(0)	10(20)	10(20)	8(8)	10(10)	5(15)	2(6)	3(6)	68(145)	
PLRO(MM)	10(30)	10(30)	5(10)	10(20)	9(18)	7(7)	10(10)	7(21)	2(6)	4(8)	74(160)	3
STV(MM)	8(24)	9(27)	10(20)	10(20)	8(16)	7(7)	10(10)	10(30)	9(27)	10(20)	91(201)	1
AMS(HY)	9(27)	10(30)	0(0)	5(10)	9(18)	9(9)	7(7)	7(21)	8(24)	3(6)	67(152)	4
AV+(HY)	7(21)	8(24)	5(10)	5(10)	8(16)	7(7)	5(5)	5(15)	9(27)	7(14)	66(149)	5
AV+50(HY)	10(30)	10(30)	5(10)	5(10)	8(16)	7(7)	7(7)	7(21)	8(24)	7(14)	74(169)	2

against each feature was evaluated as *excellent*, *good*, *fairly good* and *poor*. More recently the evaluation has been carried out numerically with scores being given up to a maximum of 10.

As well as this scoring procedure, it was realised that certain features were of greater importance than others, and weighting factors (WF) were therefore applied to each feature. For example, proportionality is considered to be very important and is thus given a WF of 3 , the relevant feature score being multiplied by WF. On the other hand ease of counting is not of great importance as the returning officer and his or her staff will have been trained to deal with the relevant system. In this case the weighting factor (WF) is taken as 1.

Notation for systems

1. **Single Member Constituencies**

FPTP(SM): First-past-the-post.

AV(SM): Alternative Vote.

2. **Multi-Member Constituencies**

PL(MM): Party List based on the whole country (as in Israel).

PLRC(MM): Party List based upon regions using a closed list.

PLRO(MM): Party List based on regions with an open list.

STV(MM): Single Transferable Vote.

3. **Hybrid Systems**

AMS(HY): Additional Member System as used in Germany and in differing forms for the Scottish Parliament and Welsh Assembly.

AV+(HY): AV(SM) with a top-up as proposed by Lord Jenkins for Westminster.

AV+50(HY): Similar to AV+(HY) but having equal numbers of local and regional members.

Notation for Features

PRO-R: How proportional is the result within a region? (A region is visualised as, say, ten adjacent single-member constituencies).

PRO-N: How proportional is the total election result?

CHO-P: Is there a choice within a party as well as across party lines?

ONECM: Is there one class of elected members?

EASV: How easy is the system for the voter?

EASC: How easy is it to conduct the count?

EASBC: Does the system ease the task of determining constituency boundaries?

EW&E: Does the system encourage women and persons of ethnic minorities to stand for election?

LOC: How closely is the elected member linked to his or her constituency?

PLOC: How easily can a voter contact an elected member of their own political persuasion?

My Decision Table

1. **Weighting factors**

The weighting factors I have chosen for the features above are:

WF=3 for PRO-R, PRO-N, EW&E, LOC.

WF=2 for ONECM, EASV, PLOC, CHO-P.

WF=1 for EASC, EASBC.

2. **Decision Table**

The figures in parentheses are obtained by multiplying the score (out of 10) by the weighting factor WF, thus obtaining a weighted score. The total (weighted) score is the sum of the weighted scores for each feature of a system. The figures presented in the table gives my own judgement of the features for each system.

Conclusions

On this basis STV appears to be the best system. This, however, is something I have believed for the last 20 years or so. Maybe I have been subconsciously biased!

The scoring and weighting reflects my personal opinions and feelings. Small differences in scoring and, particularly in WFs, can easily change the above conclusions and I would be grateful for other opinions.

STV with Elimination by Electability Scores

Simon Gazeley

1. Introduction

It is widely thought among students of electoral reform that a candidate in a single-seat election who can beat every other in Condorcet pairwise comparisons is the most representative possible of the expressed views of that electorate. This proposition can be disputed, but for present purposes I shall regard it as axiomatic. The Condorcet principle can be extended to cover elections for n seats when $n>1$; one way of achieving this is to conduct mini-elections by STV to select n out of every possible set of $n+1$ candidates, and to elect the set of n candidates that wins the largest number of these mini-elections.

There are two problems with this extended form of Condorcet. One is that, when two or more seats are being contested, it is not practicable for any but the smallest elections: 15 candidates contesting 5 seats would give rise to 5005 contests; 27 candidates standing for the 15 seats on the Council of the Electoral Reform Society would give rise to 13,037,895 contests. Confronted with the result sheet of such an election, the electorate would find it difficult to understand how the winning candidates won and, perhaps more importantly, how the losing candidates lost. The other problem is that there could be more than one set of n candidates (whether $n>1$ or $n=1$) which gain the equal greatest number of victories. We would have to provide some kind of tie-breaker.

I believe that we can achieve the effect of Condorcet for one or more seats without these practical difficulties. Indeed, David Hill[1] has suggested one such scheme which selects sets of n candidates and tests each set against the other candidates one at a time. He admits that his scheme can elect a candidate other than the Condorcet winner in an election for one seat: I believe that the system propounded here will always elect the Condorcet winner, if there is one.

2. A Brief Digression on Proportionality

Woodall[2] has proved that no system can be devised which has all the following properties:

1. Increased support, for a candidate who would otherwise have been elected, should not prevent their election.

2. a. Later preferences should not count against earlier preferences.

   b. Later preferences should not count towards earlier preferences.

3. If no second preferences are expressed, and there is a candidate who has more first-preference votes that any other candidate, then that candidate should be elected.

4. If the number of ballots marked X first, Y second, plus the number marked Y first, X second, is more than half the number of ballots, then at least one of X and Y should be elected.

Given that preferential voting is desirable, few would consider any system which lacks either of properties 3 or 4 to be acceptable. Woodall later[3] extended 4, dubbing it the “Droop Proportionality Criterion” (DPC), which he stated thus:

If, for some whole numbers K and L satisfying 0<K≤L, more than K Droop quotas of voters put the same L candidates (not necessarily in the same order) as the top L candidates in their preference listings, then at least K of those L candidates should be elected.

A voter who puts those L candidates (in any order) as the top candidates in order of preference is said to be “strongly committed” to that set of L candidates. We will refer to a set of candidates to whom one set of voters is strongly committed as a “DPC set”.

Under any of the rules in current use, the elimination of candidates in an STV election makes votes available to other candidates in the DPC sets to which they belong. No candidate who has a quota at the relevant stage is eliminated, and, with insignificant exceptions, eliminations are made one at a time. This ensures that the result of an STV count is consistent with the Droop Proportionality Criterion. STV with Elimination by Electability Scores (STV(EES)) shares this characteristic.

3. The aim of STV(EES)

Conventional STV (whether by Meek's method[4] or one of the manual methods) is directed towards identifying with as little ado as possible the candidates who should get the seats: election takes precedence over elimination. The problem with this approach is that only as many of each voter's preferences are examined as are necessary to award the quota to sufficient candidates within the rules. For example, the second and subsequent preferences of the voters whose first preference was cast for the eventual runner-up are not even examined.

On the other hand, the aim of STV(EES) is to identify those candidates who certainly should not be elected. It does so by taking account of all the preferences of every voter; in some circumstances, this feature will cause the system to fail on Woodall's second property. To identify candidates for elimination, it calculates “electability scores” (see below) for the candidates: as new electability scores are calculated at successive stages, these form the basis for the elimination of candidates one by one until only sufficient are left to fill the available seats. These remaining candidates are elected.

STV(EES) differs in another way from conventional STV. As we are identifying candidates for elimination, not election, we do not have to use the Droop quota, and in fact its use can lead to perverse results. Instead, we calculate the “threshold”, which any of the other candidates must be able to attain in order to survive.

4. How STV(EES) works

STV(EES) is based on Meek's method, the most significant feature of which in this context is that votes are transferred in strict order of the voter's preference, regardless of whether the receiving candidates already have a quota of votes or not. In STV(EES), all candidates start as “contending” candidates. We then calculate the “electability score” (see below) of each contending candidate in turn, and candidates are withdrawn on the basis of those electability scores.

A stage of STV(EES) culminates in the withdrawal, either temporary or permanent, of a candidate. It consists of two substages: the first establishes the threshold of votes which a candidate must be capable of achieving in order to survive; the second is to test whether the candidates who start with less than the threshold can in fact achieve it. At the end of the second sub-stage, one of these candidates is withdrawn from contention. This withdrawal takes one of two forms: the candidate is either “eliminated”, which means that (s)he takes no further part in the count, and is treated from that point on as though (s)he had withdrawn before it started; or is “temporarily excluded”, which means that (s)he is withdrawn for the time being, but comes back in after the next elimination.

Before explaining how to calculate electability scores, we must define the “retention factor”, which Meek calls the “proportion retained”. In a Meek count, a point will be reached when a candidate has more than the quota. Clearly, that candidate should get less of the incoming votes in the next iteration of the count than were credited this time; and in successive iterations, the proportion of each incoming vote that stays with that candidate will diminish. The tendency will be for each new total of votes credited to that candidate to be closer to the quota than the last. To achieve this, an incoming whole vote or fraction of a vote is multiplied by an amount m where $0<m<1$; the result of this multiplication is the fraction of that vote which is credited to that candidate. This amount m is known as the retention factor. Retention factors start with a value of 1.0, and those for the candidates with more than the quota are re-calculated at every iteration; thus retention factors will diminish as the count progresses. The Droop quota is also re-calculated at every iteration on the basis of the votes credited to candidates, ignoring those which have become non-transferable.

In an STV(EES) election, the first sub-stage of each stage is the calculation of the threshold. It does this by calculating the mean of the votes of the n candidates who have the most votes. Surpluses over the mean are transferred, then a new mean is calculated. This process of distributing the votes, calculating the mean, and transferring surpluses is repeated until the first n candidates have the same number of votes. The top n candidates are then known collectively as the “probables”, and their common total of votes is the threshold (T). The value of T remains fixed throughout the second substage, which is the calculation of the contending candidates' electability scores. Let C be the contending candidate whose electability score we are calculating (the “candidate under test”), and let all the contending candidates other than C have a common retention factor of c. C's own retention factor remains at 1.0. In successive iterations, c and the retention factors of the probables are recalculated until the votes credited to all the probables are equal to the threshold and C either has the threshold or has less than the threshold while no other contending candidate has any votes at all. At this point, c is declared to be C's electability score. The electability scores of the remaining contending candidates are calculated in like fashion. The smaller C's electability score, the greater the number of votes that have had to be transferred from contending candidates other than C in order to ensure that C and the probables get their thresholds.

If the votes credited to the candidate under test and the contending candidates have a collective total of less than *T*, this indicates that the probables had a Droop quota of votes each when that candidate's electability score was being calculated. In that case, that contending candidate is eliminated, and all the non-eliminated candidates are re-classified as contending. On the other hand, if all the contending candidates' electability scores are at least 0.0, the one with the highest electability score is temporarily excluded, and only the existing probables are re-classified as contending. The new set of contending candidates proceeds to the next stage.

Stage succeeds stage until there are only *n* candidates who have not been eliminated, and those final candidates are elected. Note that at any stage when there are only *n*+1 "active" candidates (ie, candidates who have not been eliminated or temporarily excluded), one of them is certain to be eliminated. We therefore know that candidates will be eliminated until only *n* active candidates survive; thus an STV(EES) election must come to an end.

STV(EES) aims to identify a set of *n* candidates which can score at least as many victories in Condorcet mini-elections as every other. This means, for every eliminated candidate X, that there must be no set of *n* candidates including X which can score more victories in Condorcet mini-elections than every set of *n* not including X. We know at any given stage that every probable is better supported at that stage than X, and that every temporarily excluded candidate was better supported than X at the time of their temporary exclusion. Any DPC set to which X belongs has more members than can be elected by the number of voters that support it, and every other member of that DPC set is better supported than X. We can therefore be confident, though not certain, that there is no set of *n* candidates including X that can score more victories in Condorcet mini-elections than every set of *n* not including X. We can, however, state with certainty that in a count for one seat, the Condorcet winner (if there is one) will win. This is because, by definition, the Condorcet winner will win a contest with any one other candidate: and since no candidate is eliminated unless *n* other candidates have a Droop quota of votes each, the Condorcet winner cannot be eliminated.

5. An Illustration

Six candidates are contesting two seats, and votes are:

```
ABCDEF 3670
CBAEFD 3436
DEFABC 1936
EFDBCA 1039
FDECAB 1919
       =====
       12000
```

After sub-stage 1.1, A and C are probables, and the threshold (the number of votes held by both A and C when transfers are complete) is 3436. At sub-stage 1.2, electability scores are:

```
B 0.1319
D 0.4125
E 0.2860
F 0.3478
```

This means that if D, E, and F had a common retention factor of 0.1319, A, B, and C would have 3436 votes each when surpluses have been transferred; if B, E, and F had a common retention factor of 0.4125, A, C, and D would have 3436 votes each when surpluses have been transferred, and so on. As D has the largest electability score at this stage, we act on the presumption that D has a better chance of being elected than B, E, or F, and so we ensure by temporary exclusion that D does not run the risk of being eliminated at substage 1.2. Note that this presumption is like the presumption of innocence in a criminal trial: the process tests it and may very well overturn it.

At substage 2.1, effective votes are:

```
ABCEF 3670
CBAEF 3436
EFABC 1936
EFBCA 1039
FECAB 1919
      =====
      12000
```

Again, A and C are probables and the threshold is 3436. At 0.7608, E's electability score is higher than B's or F's, so E is temporarily excluded at substage 2.2. Effective votes are now:

```
ABCF 3670
CBAF 3436
FABC 1936
FBCA 1039
FCAB 1919
     =====
     12000
```

At substage 3.1, A and F are probables, and the threshold is 4016.9493, more than the current Droop quota. As neither B nor C can get that many votes if the other is temporarily withdrawn, we can eliminate both. D and E are now reclassified as contending, making effective votes:

```
ADEF 3670
AEFD 3436
DEFA 1936
EFDA 1039
FDEA 1919
     =====
     12000
```

At substage 4.1, A and D are probables, and the threshold is 3696.7554. At substage 4.2, E's electability score is 0.4741 and F's is 0.3385, so E is temporarily excluded. Active votes are now:

```
ADF 3670
AFD 3436
DFA 1936
FDA 1039
FDA 1919
    =====
    12000
```

At substage 5.1, A and F are probables, and the threshold is 4309.9757, more than the current Droop quota. D cannot get that many votes and therefore is eliminated. E now comes back in, and active votes are:

```
AEF 3670
AEF 3436
EFA 1936
EFA 1039
FEA 1919
    =====
    12000
```

At substage 6.1, the threshold is 5040.5, more than the current Droop quota, and A and E are probables. There is no prospect that F can attain the threshold, so we eliminate F. A and E are the only active candidates left, so they are elected.

6. Discussion

The example above is unusual in that there are two discrete DPC sets, ABC and DEF, supported respectively by 7106 and 4894 voters. The result is consistent with the Droop Proportionality Criterion in that each set contributes one winning candidate. In fact, an exhaustive Condorcet count produces a three-way tie for first place between AD, AE, and AF. This results from a paradox whereby AD wins the ADE contest, AE wins the AEF contest, and AF wins the ADF contest. Any of these outcomes is as valid as either of the others. It is noteworthy that STV(EES) does not “hang up” on a Condorcet paradox.

If there are too few DPC sets with sufficient support to “soak up” all *n* seats being contested, can the system still produce a reasonable outcome? Let there be 4 candidates contesting 2 seats with votes:

```
ABCD 41
BCDA 30
CDAB 25
DABC 24
     ===
     120
```

The results of an exhaustive Condorcet count are:

Contest Winners

```
ABC AB
ABD AD
ACD AC
BCD BC
```

We have a paradox in that AB wins the ABC contest, but AD wins the ABD contest and AC wins the ACD contest; there is also a four-way tie. As A starts with a quota of first preferences, A must be one of the winning candidates, but which of the other three should take the second seat?

Under STV(EES), A and B are probables, and the initial threshold is 35.5. At stage 1, the electability scores of C and D are respectively 0.5625 and 0.54, so C is temporarily excluded. At stage 2, A and D win the ABD contest, and B is eliminated. At stage 3, A, C, and D remain in the contest, so A and C are elected.

How can the elimination of B and D be justified? Part of the answer is that D was in only one winning set in the exhaustive Condorcet count, whereas the other candidates were in at least two. But is there any objective reason why B rather than C should be eliminated? Here we must confess that the system may be said to be perverse: 95 voters prefer B to C, but only 25 prefer C to B. In defence of this outcome, we can say that set AC is one of the joint Condorcet winners, so it meets the aim of STV(EES); and that when a tie is the result of a paradox, it will be arbitrary to some extent. But I would still have preferred AB to be the winning set in this case.

I submit that STV(EES) will in most cases (perhaps all) give a result that is compatible with an exhaustive Condorcet count: and that even if it does not, the result will still be defensible.

Acknowledgement

I am grateful to David Hill, Hugh Warren, and Brian Wichmann, whose queries, comments, suggestions, and advice have made this paper very much better than it would otherwise have been. My deep gratitude is due also to the independent referee, who not only made many suggestions but also pointed out a fatal flaw in an earlier version of the system.

References

1. I D Hill, Sequential STV, *Voting matters*, Issue 2 (1994), 5-7.
2. D R Woodall, An impossibility theorem for electoral systems, *Discrete Mathematics,* **66,** (1987) 209-211.

3. D R Woodall, Properties of Preferential Election Rules, *Voting matters*, Issue 3 (1994), 8-15.

4. B L Meek, A New Approach to the Single Transferable Vote, *Voting matters*, Issue 1 (1994), 1-6. (Paper 1), 6-11 (Paper 2).

Annex - An Algorithm for STV(EES)

All candidates start as contending candidates with a retention factor(RF) of 1.0. They should be in random order.

The suggested procedure is as follows:

Substage 1

1. Set every active candidate's retention factor to 1.0.

2. Repeat the following procedure until n candidates have T votes each.

   a. Distribute the votes in accordance with "Distributing the Votes" below.

   b. Calculate T, the mean of the votes of the n candidates with most votes.

   c. For every candidate who has more than T votes, calculate a new retention factor by multiplying their present RF by T and dividing the result by the number of votes credited to that candidate.

3. If n candidates have T votes each, classify those n candidates as probables. If more than n candidates have T votes each, classify the first n in ranking order as probables.

Substage 2

1. Select each contending candidate in turn to be the "candidate under test" and calculate their electability scores as follows:

a. If $T>V/(n+1)$, where V is the total of votes credited to all the candidates, mark the candidate under test for elimination. Otherwise, set the retention factor of the contending candidates, the candidate under test, and the probables, to 1.0, then repeat the following procedure until the probables and the candidate under test have T votes each, or until $T>V/(n+1)$:

i. Distribute the votes in accordance with "Distributing the Votes" below.

ii. Recalculate the retention factor (RF) of any probable who has more than T votes by multiplying it by T and dividing the result by the number of votes credited to that candidate. Recalculate the common RF of the contending candidates by multiplying it by $(V-(n+1)T)/C$, where C is the total of votes credited to the contending candidates other than the candidate under test.

b. If $T=V/(n+1)$ and there are only $n+1$ active candidates. or if $T>V/(n+1)$ mark the candidate under test for elimination. Otherwise, set the electability score of the candidate under test to the common RF of the other contending candidates.

2. Award the probables a notional electability score of 1.0, then rank the active candidates in their present order within descending order of electability score.

3. If any contending candidate is marked for elimination, eliminate all the marked candidates, reclassify all the non-eliminated candidates as contending, and rank them in random order. Otherwise, temporarily exclude the highest-ranked contending candidate, set that candidate's RF to 0.0, and reclassify only the probables as contending candidates.

Distributing the Votes

Examine each vote in turn and:

1. Multiply the value of the vote by the retention factor of the voter's first preference. Award that amount of the vote to that candidate.

2. If any of the vote is unallocated, multiply it by the retention factor of the candidate of the voter's next preference. Award that amount of the vote to that candidate. Repeat until none of the vote is left, or until the voter's preferences are exhausted.

3. If any of the vote is left when all the candidates have had their shares, put it to non-transferable.

How to ruin STV

I D Hill

To ruin STV by turning it, in effect, into merely a party list system, the following steps may be taken:

1. Make voting compulsory so that even the laziest have to turn out;

2. Insist that votes, as given by voter-defined preferences, are not valid unless every candidate (from a long list) is given a preference number, without gaps or repetition;

3. Allow the voter the alternative option of merely ticking a party box, and take that to indicate an STV vote as specified by the chosen party;

4. Use traditional STV counting rules, so that it can be guaranteed that, if you choose your own order, either your first choice will not be elected, or if elected but not on the first count, then all your hard work entering later preferences will be totally ignored;

5. Insist that, as the party box method is optional, this is not taking anything away from the voters.

Since many voters are lazy, most can then be expected (save in very exceptional circumstances) to use the party box method, as to do anything else is a lot of work and almost certainly for no benefit. Is it unimaginable that party politicians would try to pervert STV in this way? Unfortunately not; all these things now happen in Australia, and nearly all the virtues of STV have consequently been lost.

To see the dire effects of this, consider the election of 6 Senators for New South Wales at the 1998 Federal Election, for which there were 69 candidates. In some Australian STV elections not all the candidates have to be given preference numbers, though they usually require a substantial number. In this one all 69 had to be put in strict preference order. Just imagine doing that when the alternative of merely ticking a party box was available.

Probably many voters would not be aware of the effect mentioned in item 4 above, so that may not have much effect on what happens, but it would certainly add to the frustration for anyone who did know about it.

The remarkable thing in the circumstances is not that practically everyone used the party option but that 19012 voters, or 0.51%, did not.

The whole output table is much too vast for reproduction here, but the sense of it can be derived by looking at just the party that did best, with candidates A1, A2, A3 and A4 in that order on the party ticket. The first four stages for those candidates were:

A1	1446231	−909698	536533		536533		536533
A2	2914	+908567	911481		911481	−374948	536533
A3	864	+196	1060	+11	1071	+374505	375576
A4	2551	+130	2681	+3	2684	+199	2883

Eventually A3 also was elected. It can be seen, just from this small part of the information, how the party listing is totally dominant, and crushes all individualism. In particular, note how the party's preference for A3 over A4 overwhelms the fact that A4 got three times as many first preferences as A3. In fact, after transfers, all the votes ended up pointing at the three candidates highest on the list of the above party that took three seats, the two candidates highest on the list of another party that took two seats, the candidate first on the list of a further party that took one seat, and the candidate first on the list of the runner-up party. For the candidates, it is clear that getting a high place on the party list, rather than being liked by the voters, is what matters, as with party list systems in general.

Is it wise to tell politicians that STV can be perverted like this? Given that it has already happened in Australia, it can hardly be hidden from them anyway. The important thing is to bring the facts to the attention of STV supporters, so that they know that it is something to be ready to fight against.

Editorial Note

Unfortunately, there was a very misleading typographical error in Issue11. This was on the table marked **Old rules** on page 8. The entry against candidate C should have the word 'Elected' deleted. I am sorry if this caused any confusion. A corrected version is available electronically from me.

All correspondence regarding *Voting matters* should be addressed to:

B A Wichmann, The Editor, *Voting matters*, The Electoral Reform Society, 6 Chancel Street, London, SE1 0UU.

or using e-mail to Brian.Wichmann@freenet.co.uk.

To aid production, the Editor would welcome contributions on IBM-PC discs (with a printed copy as well) or to Brian.Wichmann@freenet.co.uk.

Voting matters

for the technical issues of STV

The Electoral Reform Society

Issue 13 April 2001

Editorial

As the political debate intensifies prior to a General Election, the search for a better technical understanding continues here.

Hugh Warren responds to remarks made about his suggestion for merging X voting with STV.

The issue of undertaking recounts with STV is very unclear. Even with computer counts, ensuring that there are no errors whatsoever in the data input is unrealistic. My own paper provides details of a computer system designed to detect if an STV election is close enough to justify a recount.

Joe Otten provides details of an algorithm for handling STV elections with complex constraints. Even though such constraints override the voters' intentions, it seems that several elections are of this type and hence there is a demand for such an implementation.

David Hill provides an analysis of STV when equality of preference is permitted. It seems that there are problems in this area, so the fact that conventional STV does not provide equality is not necessarily a disadvantage.

Lastly, I provide a paper concerned with the *transparency* of STV. The conclusion is to call for the partial disclosure of the votes so that anybody can perform an effective check on the counting process. Comments on this and all the other papers are welcome!

CD-ROM Publication

With the support of the McDougall Trust, I am collecting electoral material with the aim of publishing it in CD-ROM format. It is intended, for example, that the publication will include all of *Voting matters*. (As a separate exercise, back issues of the journal *Representation* may be made available on CD ROM as well.) The main emphasis will be on the collation of election data, especially that involving STV or preferential voting. If you have or know of material which you think could be suitable, please contact me. A key advatange of the CD ROM media is that well over 5,000 pages can be placed on one disc.

Brian Wichmann

The principal objects of the McDougall Trust (The Arthur McDougall Fund) are to advance knowledge of and research into representative democracy, its forms, functions and development and associated institutions. The Trust is governed by a High Court Scheme issued in 1959 which states its charitable purposes as being ‘to advance knowledge of and encourage the study of and research in: political or economic science and functions of government and the services provided to the community by public and voluntary organisations: and methods of election of and the selection of and government of representative organisations whether national, civic, commercial, industrial or social.’

Mixing X-Voting and Preference Voting

C H E Warren

In my paper on incorporating X-voting into preference voting by STV[1], without saying so I had treated it as axiomatic that a method of mixing X-voting and preference voting should reduce to either X-voting or preference voting by STV should all the voters be of one sort.

In a comment at the end of my paper, the Editor suggested an alternative formulation which, sadly, would not reduce to X-voting as it is always practised should all the voters be X-voters. The Editor's formulation would not therefore satisfy the axiom mentioned above.

The answer to the question at the end of David Hill's paper[2] "Is there a way of doing it that everyone would think fair in all cases?" is surely "No".

There are the hardliners on both sides — those who think that anything other than X-voting is not fair, and those who think that anything other than preference voting by STV, which I imagine includes David Hill, is not fair.

The most that one can hope for, then, is not a way of doing it that everyone would think fair, but a way that a majority of considered opinion would think fair.

The major response that I have had to my paper[1] so far is that "it is a good idea".

References

1. C H E Warren. Incorporating X-voting into Preference voting by STV. *Voting matters*, Issue 11, p2, 2000.

2. I D Hill. Mixing X-voting and preference voting. *Voting matters*, Issue 12, p6 , 2000.

Recounts with STV

B A Wichmann

Introduction

With Westminster elections, if a result is sufficiently close, a recount is undertaken to reduce the risk of an incorrect result being declared. Of course, with First Past The Post, a simple measure of the closeness of the result is possible, so that the criteria for a recount can be easily given. (A virtually identical problem has arisen with the US elections in Florida in which obsolete technology is employed!)

With STV, recounts are very rarely undertaken due to the problems that this would give. In Newland and Britton rules[1], both first and second edition, there was an instruction, at the end of each stage "Ascertain that candidates and/or their agents are content" and a recount of the stage could be called for if not. The difficulty with this is that it may not become evident that an early stage needs checking until a later one has occurred, and the only sure strategy for candidates was always to ask for a recount after every stage. In the latest edition of the rules, those words have, in any case, been omitted.

However, when the count is conducted by computer, the computer itself can be used to assess the need for a 'recount'. The article is not concerned with the actual process of undertaking a recount (merely running the counting program again would be pointless), but with providing a tool to assess the risks of an incorrect result being obtained due to a typing error when the papers are entered manually.

This article describes a set of computer programs, developed for Electoral Reform Ballot Services, which assesses the need for a recount.

The concept

At first, I thought that the problem was too difficult to undertake, since if a change is made to even one ballot paper, it is hard (in general) to predict any change of result. However, given a computer program that can undertake a count in a matter of minutes (if not seconds) then an alternative method is available which does not require any analysis of the result of changes in specific papers.

The stages are as follows:

1. A simple model is produced of the manual data entry process, together with the likely data entry errors.

2. From the data entry error analysis, a computer program is produced which simulates such errors.

3. The above computer program is used to construct a hundred (or more) copies of the original election data with simulated errors.

4. The simulated elections are counted by program and the results compared with the original results to see if an incorrect result is likely.

This process can be made effective since the speed of modern computers allows a hundred of more copies of an election to be counted in a reasonable time. (It is surely sufficient for an overnight batch computer run to produce the result — although for smaller elections, a result should be obtained in a few minutes. Examples so far have only taken about an hour to run.)

The system

The system consists of two programs: one which produces copies of the original election with data errors added, and another which analyses the results from all the elections. Provision is made to handle the Meek rules[2] or the ERS97 rules[1]. In addition, a batch execution run is produced to call the relevant election counting program on all the simulated elections.

The data entry model is essentially one of key depressions using ballot papers in which the voter adds preference numbers. Since typing errors have known patterns, a reasonable guess can be made of the potential errors in terms of those errors. However, it is difficult to accurately calibrate the rate of errors. Such errors are naturally rare, say 1 in 5,000 characters, but at this rate one would need to double-check many thousands of characters to obtain a good estimate of the error rate. In addition, the computer entry programs used for ballot entry already include some checks and hence the error simulation program ensures that these checks will be passed. Also, the staff of ERBS are naturally familiar with the requirements and appear to take special care with the first preference (not actually allowed for in the current program). There is some evidence that the staff at ERBS may realise at the end of the ballot paper that they are 'out-of-step' and hence go back to correct an error. In view of the above, there is clearly some doubt as to the accuracy of the model of data errors, but the statistical nature of the problem makes some doubt inevitable.

After some experimentation, the data error rate was set at one key depression per 6,000 characters. However, if the error would then be detected by the STV program, such as arising from a repeated preference, the corresponding change is not made.

Results

This can be illustrated by an example taken from a real election (which has been made anonymous).

```
Data error analysis program, version 1.01
Basic data of original election:
 Title: R048: STV Selection Example 1
 To elect 10 from 29 candidates.
 Number of valid votes: 944
 Count according to Meek rules

Data used to simulate input errors to count:
 Key errors taken as 1 in 6000 key depressions.
 Duplication and removal of papers taken
                    as 1 in 6000 papers.
 Number of simulated elections produced: 100
 Seeds were initially:  16215,  15062 and   7213
          and finally:  17693,  15003 and  25920
```

```
Some statistics from the generated election data:
 Average number of commas added for each election: 1
 Average number of commas deleted for each
                                         election: 1
 Average number of interchanges for each election: 2
 Average number of papers deleted for each
                                         election: 0
 Average number of papers duplicated for
                             each election: 0
 Average number of papers changed for each
                                  election: 4
 Average number of papers changed at
                            preference: 1 is 1

Candidates elected in the original election and all
simulated ones:
 Jane BENNETT
 Robert BROWNING
 Joan CRAWFORD
 Francis DRAKE
 Mary-Ann EVANS
 Kate GREENAWAY
 John MASEFIELD
 Alfred TENNYSON
 Sybil THORNDIKE

Candidates not elected in the original election or
any of the simulated ones:
 James BOSWELL
 Emily BRONTE
 George BYRON
 Eric COATES
 Ella FITZGERALD
 Stella GIBBONS
 Graham GREENE
 Sherlock HOLMES
 Samuel JOHNSON
 John KEATS
 Alice LIDDELL
 Harold PINTER
 Walter RALEIGH
 Margaret RUTHERFORD
 Will SHAKESPEARE
 Percy SHELLEY
 John WESLEY
 Virginia WOOLF
```

The program records the known details of the election which includes the type of count used: Meek in this case. Then the statistics are recorded on the simulated elections. Firstly, there is the key depression error rate used, then the seeds used for the pseudo-random generator so that the process can be re-run if required. Then a summary is produced of the changes made to the papers. Note that one of the changes is that of repeating *and* duplicating a paper (both changes are needed to reflect the checks made on the total number of papers). The commas indicate moving onto the next preference. Note that of nearly 1,000 papers, typically one change is made to the first preference position.

Of course, the changes that will be of most interest are those relating to the election of the candidates. The first two lists are the candidates which are always elected or always excluded — there should be no doubt about the status of these.

```
Number of other candidates: 2
 Original Result  Simulated Result(95% conf. limits) Name
     Elected          Elected  98% ( 93% to 100%)     Clara BOW
  Not  Elected    Not Elected  98% ( 93% to 100%)     Benjamin FRANKLIN

End of report
```

The last table indicates the position with those candidates whose status varied in the 101 elections performed (1 original and 100 simulated).

The number of such candidates is two. In the case of Clara Bow, she was elected in the original election and also in 98% of the simulated ones, ie in two cases she was not elected. The case with Benjamin Franklin is exactly the opposite. However, merely knowing that percentage is not what is required. We need an estimate of the probability of an incorrect result, which is the likely value of the percentage in the long run, that is if infinitely many simulated elections were used. This long-term value is estimated to lie between 93% and 100% (to a 95% probability).

In this particular case the result is not seriously in doubt. However if the percentage range included the 50% figure, then it is proposed that this would be sufficient to require a recount.

Conclusions

The method proposed here appears to be an effective means of determining if a recount should be undertaken for an STV election. However, the technique does depend upon a statistical model of the nature of the data preparation errors which is always going to be hard to produce.

The method can be applied to assess the impact of data errors arising from mechanically produced data, assuming the data error rate is high enough to warrant its use.

I am grateful to David Hill who provided some Pascal code which gives the 95% probability ranges — a vital part of the system.

References

1. R A Newland and F S Britton. How to conduct an election by Single Transferable Vote. ERS, 1973, 1976 and 1997.

2. I D Hill, B A Wichmann and D R Woodall. Algorithm 123 — Single Transferable Vote by Meek's method. *Computer Journal*. 1987.

STV with multiple constraints

J Otten

Joe Otten is the author of an STV program for Windows which is being extended to handle constraints

The problem

David Hill writes in *Voting matters*[1] that the handling of constraints should be undertaken by marking as *doomed* candidates who cannot be elected if a conformant result is to be obtained, and marking as *guarded* those candidates who must be elected for a conformant result. A doomed candidate is eliminated immediately so that the next preference can be taken into account, while guarded candidates await attaining a quota (if that is possible). However, where multiple constraints are to be applied, then Hill states we should list all the possible ways that the constraints might be met, so that we can tell when it is necessary to guard or doom continuing candidates. If you are unfamiliar with these details, I recommend reading Hill's article first.

In this paper we consider the situation with two independent sets of constraints, such as nationality and gender. A group of candidates are those sharing the same constraining characteristics. While I agree that Hill's method works, and that simpler methods do not, there is a problem when the numbers of candidates and groups of candidates become large. For instance, suppose there are 20 candidates to be elected from 30 groups, with 2 candidates in each group, there would be astronomic number of cases ($\approx 3^{30}$), of which maybe only half can be ruled out by the constraints. Such a list of possibilities would take far too long to calculate on a fast computer with efficient code, and occupy an excessive amount of storage. This is clearly not feasible. It might appear that such complexity of constraints should not arise in practice — unfortunately it has arisen which has prompted the approach given here.

A worked Example

We re-work Hill's example which is that of 14 to be elected, where must be 7 English, 6 Scottish and 1 Welsh, and additionally 7 Men and 7 Women. We refer to each of these by the initial letter with the nationality first. In this example, there are 8 possibilities listed:

EM	EW	SM	SW	WM	WW
0	7	6	0	1	0
1	6	5	1	1	0
1	6	6	0	0	1
2	5	4	2	1	0
2	5	5	1	0	1
3	4	3	3	1	0
3	4	4	2	0	1
4	3	3	3	0	1

Each time an election or exclusion causes one or more of these results to become impossible, we cross it out. We can then see when it is necessary to guard or doom candidates.

This problem requires a solution that does not involve listing every combination since the size of the list rises exponentially with the number of groups. I believe this is possible if we deduce and keep track of every constraint as it applies to every group. In Hill's example this is possible. At the crucial point he argues that “...only 2 Scottish women remain, we have to elect 6 Scottish altogether and have elected none as yet. Therefore we must elect at least 4 Scottish men. But we are restricted to 7 men in total and we have already elected 3. It follows that we must elect exactly 4 Scottish men, and that means that the remaining 2 Scottish women must be guarded, and that the 2 English men must be excluded as soon as possible,...”

This argument is sound, and does not itself rely on an exhaustive listing of all the possible combinations. I propose a procedure which implements this sort of logic in a way that can be automated and performed at the start of the count and after every election and exclusion.

The way I propose to represent this is as in the following grid.

		English	Scottish	Welsh	Total
Men	*Elected*	0	0	0	0
	Min	0	0	0	7
	Max	7	6	1	7
	Cands	4	11	2	17
Women	*Elected*	0	0	0	0
	Min	0	0	0	7
	Max	7	6	1	7
	Cands	7	3	1	11
Total	*Elected*	0	0	0	0
	Min	7	6	1	14
	Max	7	6	1	14
	Cands	11	14	3	28

A row (of 4 lines) corresponds to each gender constraint and a column to each nationality constraint. A cell, with 4 entries, *Elected, Min, Max, Cands*, corresponds to a candidate group or to a row or column total or to the grand total. The grid has been initialized with the numbers of candidates in each group, and the various totals required by the constraints (as from Hill's example). Of course, we have none elected in this initial table, the constraints are as given before, and the new information is that concerning the candidates.

The basic method is to repeatedly apply five rules to a table until a stable condition is produced which essentially provides a bounding box which must enclose any conformant solution. We need to apply these rules initially (to confirm that a solution is possible) and at each election and elimination. Each rule is triggered by a condition which should be satisfied by a conformant solution.

1. In each group we require: *Elected* ≤ *Min* ≤ *Max* ≤ *Cands*. **Rule** — increase *Min* or decrease *Max*. If as a result of applying the rules *Min* > *Max* then no conformant result is possible (there is no bounding box) and we do not regard this as a settled state.

2. In each group, the *Min* must be possible — i.e. it must be possible for this few to be elected, even if the current minimum is elected from the row/column, and the maxima elected from each other group in that row/ column. **Rule** — increase *Min*.

3. Like 2, for maxima — in each group, it must be possible for this many to be elected, even if the current maximum is elected from the row/column, and the minimum elected from each other group in the row/ column. **Rule** — decrease *Max*.

4. The row/column minimum must be at least the sum of the minima of the items in the row/column. **Rule** — increase *Min*.

5. The row/column maximum must be no more than the sum of the maxima of the items in the row/column. **Rule** — decrease *Max*.

Hence if any of the conditions required is violated, we apply the associated rule until a settled state is reached.

Once the grid is in a settled state, and if in any cell *Elected* = *Max* then continuing candidates in that cell are doomed. If in any cell *Min* = *Cands* then all continuing candidates in that cell are guarded.

I hope it is clear that each of these rules is a logical necessity, as is its **Rule** when it applies. What is not so clear is that following these rules is sufficient to ensure that candidates are always doomed or guarded as necessary.

To see what is going on, let us apply the above now before we start counting the votes, as we need to in order to ensure that there is a conformant result and to identify any candidates which may be initially guarded or doomed.

		English	Scottish	Welsh	Total
Men	*Elected*	0	0	0	0
	Min	0	3	0	7
	Max	4	6	1	7
	Cands	4	11	2	17
Women	*Elected*	0	0	0	0
	Min	3	0	0	7
	Max	7	3	1	7
	Cands	7	3	1	11
Total	*Elected*	0	0	0	0
	Min	7	6	1	14
	Max	7	6	1	14
	Cands	11	14	3	28

a) By 1, *Max* English Men must be reduced from 7 to 4 because there are not enough candidates. Similarly, *Max* Scottish Women must be reduced from 6 to 3.

b) By 2, *Min* Scottish Men = 2. There are at most 5 non-Scottish men, and we need 7 men altogether.

c) Similarly by 2, *Min* English Women = 3. Since *Min* English + *Max* English Men = 7.

d) By 2, *Min* Scottish Men = 3. Since *Min* Scottish Men + *Max* Scottish Women = 6.

This is a settled state, so we conclude that a conformant result is possible, and we can start counting the votes. The first event is the election of a Welsh man, which we mark as a 1 in the space referring to the number of Welsh men elected. This requires the following alterations:

a) By 1, *Min* Welsh Men = 1.

b) By 3, *Max* Welsh Women = 0.

c) By 2, *Min* English Women = 4.

d) By 3, *Max* English Men = 3.

This is a settled state. We now have 2 cells where *Elected* = *Max*, so the continuing candidates in those cells, a Welsh Man and the Welsh Woman are doomed. The doomed candidates are removed from the grid by reducing the *Cands* entry.

		English	Scottish	Welsh	Total
Men	*Elected*	0	0	1	0
	Min	0	3	1	7
	Max	3	6	1	7
	Cands	4	11	2	17
Women	*Elected*	0	0	0	0
	Min	4	0	0	7
	Max	7	3	0	7
	Cands	7	3	1	11
Total	*Elected*	0	0	1	1
	Min	7	6	1	14
	Max	7	6	1	14
	Cands	11	14	3	28

The next events are — the election of 2 English Men and 2 English Women, and the exclusion of a Scottish Woman. We would in practice update the grid after each of these 5 events, but for the purpose of this example, we will do it in one go.

a) By 1, *Min* English Men = 2, due to the election.

b) By 1, *Max* Scottish Women = 2.

		English	Scottish	Welsh	Total
Men	*Elected*	2	0	1	3
	Min	2	3	1	7
	Max	3	6	1	7
	Cands	4	11	1	17
Women	*Elected*	2	0	0	2
	Min	4	0	0	7
	Max	7	2	0	7
	Cands	7	2	0	9
Total	*Elected*	4	0	1	5
	Min	7	6	1	14
	Max	7	6	1	14
	Cands	11	13	1	25

This completes the actions directly as a result of the elections, but now we must continue to give a settled state

c) By 2, *Max* Scottish Men = 4.

d) By 2, *Min* English Women = 5.

e) By 3, *Max* English Men = 2.

f) By 2, *Min* Scottish Men = 4.

g) By 2, *Min* Scottish Women = 2.

h) By 3, *Max* English Women = 5.

		English	Scottish	Welsh	Total
Men	*Elected*	2	0	1	3
	Min	2	4	1	7
	Max	2	4	1	7
	Cands	2	11	1	16
Women	*Elected*	2	0	0	2
	Min	5	2	0	7
	Max	5	2	0	7
	Cands	7	2	0	9
Total	*Elected*	4	0	1	5
	Min	7	6	1	14
	Max	7	6	1	14
	Cands	11	13	1	25

At this point, the grid is in a settled state, and we know precisely how many are in each group, so the constraints problem has been solved. *Elected* = *Max* for English Men, so the 2 continuing English Men must be doomed, and *Min* = *Cands* for the Scottish Women, so these must both be guarded. The count will continue to determine which of the English Women and which of the Scottish Men are elected.

All I have demonstrated here is that this method achieves the same result in this case as Hill's method. However, I hope that it is clear how it works and why it should therefore work for all 2-dimensional constraints problems.

Rules 4 and 5 were not needed as none of the Row total or Column total *Min* and *Max* could be altered. This was because the constraints were of the rigid “must equal 7” variety rather than the more flexible “must be between 5 and 9” variety.

Constraints and the STV rules

Given the logic above for handling constraints, then this must be integrated into an STV system which would use a specific rule set in the unconstrained case. We consider this with three sets of rules: The Church of England rules[2] (a hand-counting system which makes provision for constraints), the current ERS rules[3] (hand-counting with no provision for constraints) and Meek[4] (computer-counting with no provision for constraints).

The logic above, using *guarded* and *doomed*, depends upon electing and excluding candidates one at a time. None of the three sets satisfy this, and in consequence, the integration of these STV rules with the constraint logic is non-trivial. The addition is naturally simplest with the Church rules, since they have been written with that intent. However, the rules themselves are without constraints and a separate section gives a series of amendments to the rules which are to be applied in the case of constraints. The wording of the special section is reasonably straightforward since elections and exclusions take place one at a time.

Consider the following situations:

i) Suppose A is excluded, and this causes C and D to be doomed. The Church rules just exclude A at this stage, and then exclude C and D at the next stage. It seems possible to exclude all three together, but this surely makes no difference.

ii) Suppose A and B are to be excluded (with A having fewer votes than B), and the exclusion of A causes B to be guarded, and C and D to be doomed. This then is essentially the same case as above.

iii) Suppose A and B are to be excluded (with A having fewer votes than B), and the exclusion of A causes C and D to be doomed, but does not affect the status of B. It is clear that C and D should be excluded before B, since transfers from C and D could spare B from exclusion.

This last case shows the importance of exclusions being undertaken one at a time. This implies that the rules in ERS for multiple exclusions should be changed to handle constraints. Indeed, whatever method is used to handle constraints, the serialization of elections and exclusions is needed.

With Meek, the published algorithm only allows single exclusions, but the version implemented by I D Hill allows for a single exclusion and multiple elections at one stage. Both the elections and the exclusion need to be serialized to apply the constraints logic.

With all the rules, if two candidates achieve the quota at the same stage, then the election of one could cause the other to be doomed. Hence, if this is a tie, the tie-breaking logic would need to be applied to produce a result, even though this was not necessary without constraints.

Conclusions

The logic for handling constraints which was first specified by David Hill can be implemented in a manner that does not involve the use of large lists. This can be combined with the conventional STV rules, provided changes are made to elect and exclude candidates one by one.

Our illustration here was with an example having two independent types of constraint and therefore requiring two-dimensional tables. However, the same logic can be applied with higher dimensions if required.

With larger problems, the size and number of dimensions, and hence the computational requirements, will increase in proportion, not suffering the combinatorial explosion that the listing of all possible combinations does.

Software has been written to implement this procedure and successfully tested on a 4×16×9×3 hypercube.

References

1. I D Hill. STV with constraints. *Voting matters*, Issue 9 pp2-4. 1998.

2. GS1327: General Synod, Single Transferable Vote regulations 1990 and 1998. (Obtainable from Church House Bookshop, Great Smith Street, London SW1P 3BN.)

3. R A Newland and F S Britton. How to conduct an election by Single Transferable Vote. ERS, 1997.

4. I D Hill, B A Wichmann and D R Woodall. Algorithm 123 — Single Transferable Vote by Meek's method. *Computer Journal*. 1986.

Difficulties with equality of preference

I D Hill

One of the things that some people do not like about STV is the fact that voters have to give a strict order of preference of those candidates whom they mention, where they would sometimes prefer to be allowed to express equality. Even where they are clear about the ordering of their first few preferences, and their last few, they may well wish to separate out their middle candidates from their high ones and their low ones without ordering those middle ones.

Instructions to voters

Difficulties arise in deciding how such equality is to be specified. Suppose candidate A is first preference, then B and C equal, followed by D. Some voters will wish to mark those first four candidates as 1, 2, 2, 3. Others will insist that logic requires 1, 2, 2, 4, while still others may want to use 1, 2½, 2½, 4. What is allowed has to be specified and made not too difficult to follow.

One way out of such difficulties is to say that any numbers the user may wish can be used, but only their order will be taken into account. But if such freedom is to be allowed to those who use equality, it must in fairness also be allowed to those not using equality. This disables some useful tests that can be made for correctness of data input to a computer file. Furthermore suppose someone uses 0; is this to be regarded as better than 1? Then suppose that there are 17 candidates in total and that one voter marks four candidates as 1, 2, 3, 4 while another marks four candidates as 1, 2, 3, 17. Did they both really mean the same thing? I doubt it.

Such difficulties are not fatal, but they need careful thought, and they may complicate the instructions to voters. If they lead to less secure input of data to the computer because of the checks that can no longer be made, that also matters.

Counting the votes

There are other difficulties though in how to count such votes. The basic idea is as set out by Brian Meek[1], that a vote for A(BC)D, where the brackets indicate equality of preference for B and C, should be treated as half a vote reading ABCD and half a vote reading ACBD, and similarly with equalities of more than two candidates. This needs careful handling to avoid a “combinatorial explosion” if equality of large numbers of candidates is allowed.

However there is a difficulty of principle, rather than merely of the mechanics of the operation, that arises if voters choose to mention all candidates and to put two or more of them in equal last place. Meek's paper mentioned this possibility with approval, as allowing voters the option of indicating all remaining candidates as equal, as an alternative to not mentioning them at all. It is the one point in Meek's STV papers where I have to disagree with him, for allowing that option would mean having to explain to voters how to choose which method to use and what their different effects could be; not a task that I would wish on anyone. Or alternatively, just not to mention it, leaving voters uninformed about what they are doing.

The trouble is that there are two important principles in counting votes that are here in conflict:

1. that a vote should be interpreted in accordance with what is actually written on it, and in no other way;

2. that votes of identical meaning should be treated identically.

Now, with five candidates, for example, if one voter marks ABC as the first three preferences and stops there, while another voter marks ABC(DE), the strict interpretation of how to handle the two votes, once the fate of A, B and C has been settled, is different, but their meaning, in terms of preferences, is identical. If voters had been asked to express *degrees* of preference in some way, perhaps those two things might not be thought identical, but all that they have been asked for is an *order* of preference, and I cannot see how those two orders could possibly be thought different. This difficulty does not arise where equality is not allowed, since it so happens that two votes ABCD and ABCDE are treated identically by STV in any case, if those five are the only candidates.

There are three options: (1) to treat them differently even though their meanings are identical; (2) to treat both votes as if they had been ABC(DE); (3) to treat both votes as if they had been ABC. Of these I believe the third option to be the most satisfactory, in that there are cases where an abstention gives a better result than an equality of all remaining candidates, but I know of no case where the opposite can be claimed. (See Woodall's discussion of “symmetrical completion”[2]). I have therefore adopted this approach in my STV computer program.

The difference comes out very clearly in the results of an actual election, that used my program and allowed equality. Some voters, believe it or not, put all the candidates (not merely enough to fill all seats) as equal first choice. The program did not blink an eyelid but put those votes at once into non-transferable, treating them merely as a new way of abstaining. Surely this is right, rather than the alternative of diluting the meaningful votes with this useless information.

Having decided on option (3) then, there arises yet another problem. One of the two fundamental principles on which the Meek system is based is “If a candidate is eliminated, all ballots are treated as if that candidate had never stood”.

Suppose then that we have 5 candidates and someone has voted AB(CD). The (CD) equality has to be included as these are not last places; it is important to the voter's wishes that C and D, though not differentiated from each other, are both preferred to the unmentioned E.

If E is now excluded, we must behave as if E had never been a candidate. With E gone, all four remaining candidates are mentioned and, in accordance with the option adopted above, the AB(CD) vote must now be treated as AB. Any part of the vote that was previously awarded equally to C and D now becomes non-transferable instead. This still treats them equally, of course, but it can have the odd effect that somebody's vote may go down in the course of the count, whereas normally votes can only go up until the candidate is elected or excluded. This is certainly an extra complication that one has to be ready to explain if it occurs.

Overall, my conclusion is that, although allowing equality has some advantages, and it can be implemented, the complications may be too many to be worth it. On the other hand, those bodies that have actually used it report no difficulties, and say that the facility is strongly valued by a significant number of electors.

References

1. B L Meek, A new approach to the single transferable vote, paper II. *Voting matters*, issue 1, 7-11.

2. D R Woodall, Properties of preferential election rules. *Voting matters*, issue 3, 8-15. (Elections 11 and 12 on page 14).

Is STV transparent?

B A Wichmann

'When I use a word' Humpty Dumpty said in a rather scornful tone, 'it means just what I choose it to mean, neither more nor less'. *Through the Looking-Glass, Lewis Carroll, 1832-1898.*

Introduction

The problem with the issue of transparency is to decide what it means. Even then, to be useful, we need something which can be measured, at the very least in an informal sense. Is transparency just a matter of assurance? In which case this can be assisted by auditing, such as is used in the ISO 9000 quality management standard. I think not, since we surely accept that we need to trust those performing the election count. Even with a witnessed count, such as in public elections, we still need to trust those handling the ballot as any conjurer can testify. Even given that trust, we expect *evidence* that the count has been conducted according to the relevant rules.

Use of computers

Even if the election rules are such as to permit a manual count, it is quite likely that an STV count will be conducted using a computer. Hence we now have to question the validity and evidence for such a computer-based operation.

The public perception of computers is mixed. Few check the arithmetic in their bank statements — so surely we should accept such arithmetic when it can be checked by hand. On the other hand, the very complex calculations in weather forecasting cannot be checked, and we all know that the results are far from perfect. Fortunately, an STV count is nearer to a bank statement than to weather forecasting and hence public trust is not unreasonable.

An interesting analogy to trusting a computer-based count is that of safety-critical software which *must* be trusted. The recent problems in the railway industry, specifically passing a signal at red, is being tackled by the automatic train protection system which uses computers to stop the train. Indeed, on the Docklands Light Railway, the problem has been solved by having no drivers! In other words, we trust computers to be *more* reliable than people, at least when the situations are well-defined.

Nevertheless, there is something comforting about seeing piles of ballot papers building up against each candidate which is lost when machine counting takes place. For those witnessing a manual count, it is comforting because it is easier to place trust in people you can see. The experience in Florida is a warning that machine counting can be flawed unless sufficient controls are exercised.

Complexity

It cannot be denied that the counting process of First Past The Post (FPTP) is simple. This, in itself, is a substantial aid to transparency. Hence the simpler the rules, the easier it is to demonstrate beyond reasonable doubt, that the rules have been applied. Indeed, *transparency* might be a euphemism for *to understand* rather than anything associated with verification and auditing.

All the different STV rules must be regarded as complex. The nature of the complexity is different in the hand-counting variety compared with the machine-based versions like Meek. If rules designed for manual counting are used, but implemented using a computer, then the issue of transparency is different — since one must be concerned with the correctness of the software.

Proponents of a specific rule are likely to claim it is simple — not unreasonable if they know it well. The fact is that we have no widely accepted measure of complexity and hence we cannot use complexity as a means of quantifying transparency.

Criteria for Transparency

The main approach is to demonstrate traceability from the ballot papers through to the election result. The actual papers themselves are of no concern (in this article) and hence it is assumed that they can be (or have been) transcribed without error. There is no doubt that FPTP is 100% transparent.

We now consider three examples of STV from the point of view of transparency.

ERS97, by hand

We are assuming that ERS97 is followed to the letter[2]. Hence we have a defined result sheet. Hence the question arises as to whether this information provides complete traceability. It does not since the following information is missing:

1. The transfers at substages are merged and just the total transferred listed.
2. The quota is listed only once, and hence if quota reduction takes place, one assumes that only the final quota is listed. Hence it will not be clear that quota reduction has taken place.
3. When a tie-break is required, there is no indication as to how this should be recorded (if at all).
4. In ERS97, a tie can be broken on the basis of (the first difference of) a substage result, but these results are not recorded on the result sheet.

Church of England, by computer

The Church of England regulations[3] do not specify in detail the form of the result sheet, but a pro-forma result sheet is provided by Church House. This is similar to the ERS97 result sheet and therefore does not list substages as above.

Items 2 and 4 of the previous case do not apply to the Church of England rules, and therefore the remaining issue is the manner for recording tie-breaks.

However, all the computer programs that conduct STV counts provide substantial detail on the actions performed — much greater than the typical result sheet. This includes the resolution of any tie-break. Hence one has a reasonable degree of transparency if the fullest form of computer output is available.

On the question of checking the computer software, the Church of England rules are relatively easy to program and the corresponding checking of the software is also manageable (at least without the facility of constraints which is not considered here).

Meek, by computer

The issues here are quite different from those with the two previous cases with hand-counting rules. The algorithm is defined[1], and hence the correctness of the software is relatively simple to address.

The problem is that at each stage, a computation is required which needs at least a Spreadsheet to handle with ease. Moreover, without any other information than the votes and keep values for each candidate at each stage, it is not possible (in general) to determine the preferences which gave the observed result. In other words, we have lost traceability to the actual ballot papers. (A similar situation arises with multiple exclusions with ERS97[2], but it is not so common.)

Other issues

Two questions a voter could reasonably ask need consideration:

What happened to my vote? In the case of hand-counting rules, a detailed knowledge of the rules is required as well as the result sheet. The rules are devised so that relatively few of the preferences given are used — this is deliberate to minimise the actual work involved in a count. Hence, in most cases, it is simple to trace the position of the paper amongst the piles of papers within the count. For the Church of England rules which does not allow multiple exclusions, it is more straightforward to trace your vote. It is even simpler with Meek[5], since at each stage, all the papers are re-considered. The formula using the keep values for each candidate gives the fraction of the paper going to each candidate. If issues of security could be resolved, a voter could interrogate the voting system to validate and trace his/her vote.

What if I changed my vote? This is similar to the last question except that if the change was sufficient to alter the decisions on election and elimination, then the subsequent stages would be in doubt. The uncertainty arises because preferences may then be inspected which were never examined before — and hence cannot be determined from the result sheet.

The Data Protection Acts of 1984 and 1999 imply that the candidates have some rights of access to the information about them contained in the preferential ballots. The 1984 Act is reasonably straightforward to follow and my view was that the candidate should be told, if a request is made, of the number of votes he/she attained in each preference position, assuming the data was held on a computer.

The 1999 Act is much more complex and very hard for anybody other than a trained lawyer to interpret. It does cover manual as well as computer counts. As I understand it, ERBS has never been asked for information under the Act nor has concluded what information should be disclosed.

Conclusions

The transparency of STV is nowhere near that of FPTP, regardless of the voting rules in use. Currently, it is not really possible for a voter to obtain the same level of understanding for an STV ballot as for FPTP. This is a serious loss, since in many cases, the impact of a single vote with STV could change the result and the voter should be aware of this. (Of course, this loss is more than compensated by the additional information STV uses.)

I conclude that the above should be rectified by two changes to current practice:

Preferences should be published if they contribute to the count.

This is not complete publication of the ballot papers. My own experience suggests that complete publication might allow some individual papers to be identified which would be contrary to the overriding need for a secret ballot. For the last remaining candidate, say, only the initial preference is inspected, and hence all that would be stated would be the total number of first preferences attained. Similarly, many papers differing in some preferences would be grouped together, since the differences were not used in the ballot.

It has been suggested to me that full publication would be possible for large elections in which the identification of a single paper would be more difficult. I have rejected this since it would imply an arbitrary decision as to when an election is *large*. Moreover, for large ballots, the published summary of the papers would be small compared with the total, and hence would not be an excessive requirement.

Full publication would also allow candidates to try other STV rules which would not necessarily encourage acceptance of the declared results. For some (small) elections, the summary proposed here for publication would be the complete data from the ballot papers. However, in this situation, it may well be possible to derive that information directly from the result sheet anyway, so formal publication could not be regarded as sacrificing ballot secrecy.

In the case of the Meek rules, the removal of the unseen preferences is undertaken as follows (where KV is the Keep Value of a candidate):

1. Remove all preferences for withdrawn candidates
2. For each eliminated candidate A, compute at the point of elimination, the set X of candidates having KV=1.0 (must be continuing or elected candidates). Remove all preferences for A that appear after any candidate within X (in each paper).
3. For those candidates B for which KV=1.0 at the end of the count, eliminate all preferences after B. (Hence a first preference for B will have only a first preference.)

Similar logic can be produced to remove unseen preferences for the hand-counting rules.

Joe Otten made an interesting comment about a witnessed count. If you could not go in person, could you provide your own copy of a vote-count program to *observe* the count? I think not, since it would provide terrible problems if the results did not agree, and the returning officer could not be expected to ensure that the provided program only undertook appropriate actions. (David Hill[4] made a similar point that the data should be available for people to run their own program.)

Internet facilities should be available for voters to determine what happened to their vote.

This would be simple to provide and can be made secure by means of a Java applet that runs on the voter's computer.

Assuming that the used preferences are available in an electronic format, then anybody would be able to re-run the election count with suitable software. This is surely as transparent as possible. The Internet facility would allow voters to understand the impact of their vote without having to be an expert in the particular STV rules in use.

Acknowledgements

The author received very detailed comments on an earlier draft of this paper from: David Hill, Michael Hodge, Joe Otten, Steve Todd and Joe Wadsworth. I have tried to take these comments into account, while not necessarily accepting all of them.

References

1. I D Hill, B A Wichmann and D R Woodall. Algorithm 123 — Single Transferable Vote by Meek's method. *Computer Journal*. Vol 30, pp277, 1987.
2. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote. ERS 3rd Edition. 1997.
3. Church of England, General Synod. Regulations for the conduct of elections by the method of the single transferable vote. GS1327. 1990 and 1998.

4. I D Hill. Are better STV rules worthwhile? *Voting matters* 3, Page 8. December 1994.

5. I D Hill. How was my vote used? *Voting matters* 7, Page 7. September 1996.

All correspondence regarding *Voting matters* should be addressed to:

B A Wichmann, The Editor, *Voting matters*, The Electoral Reform Society, 6 Chancel Street, London, SE1 0UU.

or using e-mail to Brian.Wichmann@freenet.co.uk.

Voting matters

for the technical issues of STV

The Electoral Reform Society

Issue 14 December 2001

Editorial

Readers will no doubt be pleased to know that New Zealand has passed legislation to use STV in area Health Board elections and also in some local elections. Some may be surprised that the legislation specifies the use of the Meek algorithm and hence means that a computer count will be undertaken. Although these elections will not be until 2004, work is in progress to ensure that appropriate software is available and fully meets the requirements. I hope that developments can reported via *Voting matters*.

In a separate move, the Republic of Ireland is considering the use of computers to undertake its counts, although in this case, the rules are those in the Irish constitution which were designed for a manual count.

In the first article in this issue Simon Gazeley reports a means of undertaking a manual count which avoids the need to elect candidates with less that the quota of votes. Comments on the logic of this proposal or its feasibility would be welcome.

In the second article, I report on the observed differences in those elected with the current ERS rules compared with the Meek algorithm — somewhat topical in view of the New Zealand decision (although it was motivated by preparing an election data-based for publication on the McDougall Trust CD-ROM).

In the third article, David Chapman makes a proposal for electing one candidate which is described as *preferential approval voting*. The counting method seems straightforward to undertake manually and yet claims some of the benefits of the more complex algorithms.

In the last article Bob Jones reports on the questionnaire which was circulated with Issue 12. Unfortunately, the number of responses was rather small and hence it is difficult to deduce much from the replies. The Decision Analysis table that Bob produced can be recommended as a means of encouraging people to think more deeply about the issues involved.

McDougall Trust CD-ROM

The CD-ROM, mentioned in the last editorial, should be available early in 2002. Hence if you have material that would be suitable, or know the source of such material, please let me know. Election data from the UK, Ireland and Malta will be included.

The CD will contain an acknowledgement to the many referees would have aided in this publication and especially to Dr David Hill who has proof-read all 14 issues.

Brian Wichmann.

STV with Symmetric Completion

Simon Gazeley

Meek's[1] formula for STV differs from manual systems in significant ways which have been explained by Hill[2]. These differences make Meek more acceptable to many than manual STV, but it means that a computer is necessary for any but the very simplest Meek counts. I believe it is possible to improve manual STV without either losing the ability to do it manually, or introducing some unintended unacceptable effect. The current ERS rules[3] are taken as a starting point in formulating the changes proposed, and will be referred to as N-B.

When a candidate has a surplus, N-B transfers the "parcel" of votes which gave rise to that surplus — ie, the votes which that candidate received most recently. Note that the ballot-papers will all be of the same value, which can be 1.0 or less. The papers in the parcel are sub-divided into transferable votes (those on which a subsequent preference has been expressed for a candidate who is not yet elected or eliminated), and non-transferable (those on which all the candidates for whom a preference has been expressed are either elected or eliminated). If the total of transferable votes at their present value is less than or equal to the surplus, they are all transferred at that value to the voters' next preferences, and sufficient of the non-transferable votes are left with the elected candidate to preserve that candidate's quota with no surplus; any non-transferable votes over and above the quota are put to the non-transferable pile. If the total of transferable votes is greater than the surplus, a new value is calculated for each transferable vote such that when all of them are transferred at that value, their total value is equal to the surplus, and the elected candidate is left with the quota.

This procedure in effect shares out the non-transferable votes among the continuing candidates in the proportions of the transferable votes, and can give a result which I consider perverse. Consider the following count for two seats, adapted from one devised by David Hill:

```
Case 1
A        60
AB       60
CD       51
DC        9
```

The quota is 60, so A gets the first seat. N-B ignores the 60 voters who expressed no preference after A. It transfers the 60 AB votes at full value to B, who now gets the other seat. On the other hand, Meek transfers all the votes credited to A, in this case at a value of 0.5. Thus B gets 30 of the AB votes, while 30 of the A votes go to non-transferable. The new total of effective votes is now 150, making the new quota 50. C, with 51 votes, has attained this new quota and gets the second seat.

Now suppose that the 60 A voters had in fact expressed second preferences, three for C, the rest for B. Votes would be:

```
Case 2
AB      117
AC        3
CD       51
DC        9
```

In Case 2, the N-B count is identical to the Meek count. A gets the first seat, but this time all the votes credited to A are transferred at a value of 0.5, leaving A with 60. B gets 58.5 of the transferred votes and C gets 1.5, increasing C's total to 52.5. Now, nobody other than A has the quota, so we eliminate D. C's total of votes now goes up to 61.5, more than the quota, so C gets the second seat. Comparing Cases 1 and 2, we see that the additional 57 votes on which the second preference is for B are counteracted under N-B by just three voters whose second preference is for C.

Owing to the habit of many voters of not casting preferences for all candidates, the total number of votes credited to candidates tends to decline as the count proceeds. This is countered in some rules by requiring the voters to cast preferences for all candidates, forcing them to register preferences they do not feel and perhaps cannot justify. This means that in N-B counts, the final candidates to be elected often have less than a quota. As the quota is higher in these cases than it needs to be, the opportunity is lost to transfer as many surplus votes as could have been transferred if the quota had been lower from the beginning but still attainable by only as many candidates as there are seats. In a Meek count, the quota is recalculated at every stage to take account of the votes which become non-transferable and all surpluses over each successive value of the quota are transferred. Thus, the only criterion for election in a Meek count is attainment of the quota.

It is reasonable to presume that a voter who does not rank all the candidates is indifferent to the fates of the candidates left unranked, and therefore does not wish the vote to favour any of the unranked candidates over the others. As the example above clearly shows, N-B can give second and subsequent preferences more votes than the voters are presumed to have intended them to receive. Note that the A voters have no right to feel aggrieved; if they had wanted to cast further preferences, they were perfectly entitled to do so. However, the CD voters are certainly entitled to protest that the 60 A votes were treated by N-B in effect as AB votes, thus denying the second seat to C.

In a manual count, the option of reducing the quota as in Meek is not available, as the count would have to be restarted at every change of the quota. The other option is to share among the continuing candidates the votes which would otherwise have been non-transferable, treating them as if they had in fact been cast as equal lowest preferences for the candidates concerned. Following Woodall[4], I shall call this "symmetric completion". To those who are against symmetric completion on the grounds that it is never justified to award any part of a vote to a candidate for whom no preference has been expressed, my response is that symmetric completion treats all short votes alike and does not give too much weight to surplus votes on transfer. In both these respects, it is superior in my view to N-B.

With symmetric completion, the numbers of votes credited to the continuing candidates will usually be greater than they would have been under N-B, especially at the later stages. This means that there will be a tendency for more surpluses to be available for transfer, and therefore for more voters' preferences to be taken into account. Applying symmetric completion to Case 1 above, we get at the first stage

A	120
C	51
D	9

The quota is 60, and A is elected. A's votes are all transferred at a value of 0.5 to next preferences: the 60 AB votes go to B, who now has (60 × 0.5) = 30 votes, and the 60 A votes go equally to B, C, and D, who each get (20 × 0.5) = 10 votes. Votes are now:

A	60
B	40
C	61
D	19

and C gets the second seat.

Implementing STV with symmetric completion (STV-SC) would entail some changes to the N-B procedure. This is best illustrated by an example. Six candidates are contesting three seats, with votes:

A	59
AEFB	66
B	172
BCAE	12
C	112
CABD	86
D	11
DFEA	195
E	33
EDCF	148
F	21
FBDC	85
	====
	1000

The quota is 250. As no candidate has the quota, F, with fewest votes, is eliminated. As in N-B, the 85 FBDC votes are transferred to B. Although STV-SC puts the 21 F ballot-papers to the non-transferable pile, it does not put the 21 F votes to non-transferable, as all votes in STV-SC are transferred. Instead, we call these 21 votes on which no further preferences are expressed "dividend votes", because they are divided equally among the continuing candidates, in this case 21/5=4.2 to each. The number of dividend votes is calculated as the difference between the total of votes currently credited to candidates and the original total of valid ballot-papers; a running total is kept against each candidate's name of the number of dividend votes (s)he has received, and the stage at which they were gained. Effective votes at stage 2 are:

A	129.20
B	273.20
C	202.20
D	210.20
E	185.20

Now, the sum of A's votes and B's surplus is less than the votes credited to E, the candidate in last-but-one place. Under N-B rules, and therefore under STV-SC rules, the transfer of B's surplus is deferred, and we eliminate A at once. The 66 AEFB votes go to E, the 59 A papers to non-transferable. The total of votes credited to the candidates is now 936.80; the 63.2 dividend votes are awarded equally to C, D, and E, 21.06 to each. Votes are now:

B	273.20
C	223.26
D	231.26
E	272.26

We now transfer B's surplus, as that is the larger. The most recent parcel received by B contains the 85 transferred FBDC votes, plus A's share of the 21 dividend F votes, making 89.2 in all. We now transfer the 85 FBDC votes to D and the 4.2 F votes to C and D @ 23.2/89.2=0.26. As this boosts D's total above the quota, we end the count.

The only criterion for election in STV-SC, as in Meek, is attainment of the quota. To cater for rounding errors in transferred votes, the number of dividend votes is recalculated at each stage as the difference between the original total (in this case, 1000) and the total of the votes credited to candidates after all transferable votes have been transferred; the number of dividend votes awarded to each continuing candidate is truncated if necessary to two decimal places. As the total of the votes credited to the candidates is the same after each stage as it was after the previous one (except perhaps for rounding error), surpluses can arise at any point, giving the voters concerned a greater opportunity than under conventional N-B to influence the subsequent course of the election.

Should symmetric completion be imported into Meek? The answer is emphatically ***no***. Woodall[4], using an example provided by David Hill, has shown that quota reduction in Meek is preferable to symmetric completion, even though Meek himself was equivocal on the point. The purpose of this paper has been to show that, given the practical constraints of a manual count, symmetric completion can deal with a problem that may arise in N-B without in general substituting one that is as bad or worse.

References

1. B. L. Meek, A New Approach to the Single Transferable Vote, *Voting matters*. Issue 1 (1994), 1-6 (Paper 1), 6-11 (Paper 2).

2. I. D. Hill, Meek style STV — a simple introduction, *Voting matters*. Issue 7 (1996), 5-6.

3. Robert A. Newland & Frank S. Britton, How to Conduct an Election by the Single Transferable Vote, Electoral Reform Society, 3rd Edition, 1997.

4. D. R. Woodall, Properties of Preferential Electoral Systems, *Voting matters*. Issue 3 (1994), 8-15.

Do the differences matter?

Brian Wichmann

Introduction

In preparing material for a CD-ROM which contains ballot data[1], I have revised and extended the data which makes it feasible to undertake meaningful comparisons between the different STV counting rules.

It is naturally regrettable that the counting rules do indeed produce different results, that is, elect different candidates. This is to be expected, especially when comparing the Meek algorithm with the hand counting rules. Approximations must be made to provide a feasible manual process, so if it is required that a witnessed count be undertaken (and hence the moving of ballot papers between piles for each candidate) then a manual counting rule is required.

Unfortunately, real election data is hard to collect due to the confidentiality that usually applies to such data. However, a computer program has been written to produce such data anonymously by a random process which would not invalidate statistical tests on the anonymous data. This has resulted in a few more data sets from which a comparison can be made.

The two counting algorithms being compared here are Meek[2] and ERS 97[3].

Data selection and comparison

The total election data contains many examples used to test counting software which is not representative of real ballot data. However, 188 ballot sets have been identified as appropriate in three classes, as follows:

R001-R060. Data from real elections. This includes a few in which a random selection has been made from the total in the real election.

M001-M091. This data has been constructed from result sheets in such a way as to reflect real ballot data. In particular, the ones constructed from elections in the Irish Republic has been adjusted to reflect the observed transfers between the parties.

S001-S019, S021-S038. This set is constructed from data such as the Eurovision Song Contest, in which preferential voting could have been applied.

When a count is conducted, if a random choice has to be made, it is hard to conclude that a real difference has occurred. In fact, 29 of the above elections produced a different result, but in 10 of these a random choice was made and hence we ignore these.

We are therefore left with 19 differences out of 188 elections, ie 10.1% different. (I could have omitted those for electing one person, but I did not. These are mainly the third class above in which no difference was observed.)

Case	Votes	Candidates	Seats	Non-transferables (Meek-ERS97)/Votes	Difference
M005	27,757	7	3	-0.23%	1 (0.57%)
M010	38,410	9	4	1.83%	1 (0.71%)
M019	29,193	13	4	-1.06%	1 (*)
M028	44,454	13	5	-0.76%	1 (0.05%)
M051	39,991	10	4	0.13%	1 (0.45%)
M059	35,038	11	4	-0.13%	1 (0.58%)
M060	25,553	9	3	-0.96%	1 (0.33%)
M066	24,825	9	3	-1.79%	1 (*)
M070	44,914	13	5	-0.54%	1 (0.03%)
M073	36,407	8	4	0.16%	1 (1.01%)
M078	27,881	8	3	-0.07%	1 (0.38%)
R004	42	10	5	0.12%	1 (2.50%)
R005	58	8	7	3.79%	1 (0.40%)
R033	211	14	7	-2.61%	2
R040	257	20	15	0.07%	1 (*)
R045	2,908	12	5	5.95%	1 (0.83%)
R046	853	10	9	13.69%	1 (0.09%)
R048	944	29	10	0.04%	1 (0.15%)
R059	1,147	10	6	-0.40%	1 (0.03%)

In the table, the last entry records the number of seats whose occupancy changed and, in brackets, the number of votes less than the quota which the Meek algorithm recorded against the candidate which ERS97 elected (expressed as a percentage of the total number of votes). Hence for M005, the last remaining candidate which the Meek algorithm did not eliminate was the one elected by ERS97 and had 6358.85 votes against a quota of 6517.76 (6517.76-6358.85=158.91 votes = 0.57% of 27,757). The star indicates that the remaining candidate in the Meek count was not the one elected by ERS97 and hence the two counts diverged at an earlier point — not just the last stage. Of course, in the one case in which two seats differed, it is not possible to provide a simple numerical difference.

It can be seen from the table that the differences are significant and large in some cases. In five cases (M070, R004, R005, R046 and R048) the differences are small and perhaps could be regarded as acceptable. The total number of seats in these 19 elections is 106 with 20 differences and hence a discrepancy in those elected of 18.8%, or 2.1% difference if all the elections are considered.

The difference in the handling of non-transferables between the two algorithms is a matter of controversy. To indicate whether the number of non-transferables is a factor, the difference that the two algorithms give in the number of non-transferables is expressed as a percentage of the total votes. In the case of R046, ERS97 has a very much lower number of non-transferables which surely has a key effect on the result. However, in general, the pattern is not so clear.

It could be that the method of constructing the Mddd data (first class above) produces results which would not be typical of real elections. However, the table clearly shows that the Rddd (real elections, second class) examples show similar differences.

Conclusions

I conclude that unless it is essential to have a manual, witnessed count, the Meek rules should be used for STV counting. The approximations introduced to enable a manual count produces too many differences for the hand counting rules to be used otherwise.

Any of the data upon which this paper is based can be provided to interested parties.

References

1. See Editorial, *Voting matters*, Issue 13. April 2001.

2 I D Hill, B A Wichmann and D R Woodall. Algorithm 123 — Single Transferable Vote by Meek's method. *Computer Journal*. Vol 30, pp277-281, 1987.

3. R A Newland and F S Britton. How to conduct an election by the Single Transferable Vote. ERS 3rd Edition. 1997.

Preferential Approval Voting

D E Chapman

Dr David Chapman, who has carried our research on the electoral system at the Universities of Lancaster and Virginia, is currently Director of the Democracy Design Forum, a consultancy on electoral systems.

Introduction

This paper puts forward a new method for electing, by use of preferential voting, a candidate to fill a single seat. It is proposed as an improvement on the normally used single-seat electoral systems such as Plurality (as used for the Westminster Parliament), Second Ballot (previously used in France) or Alternative Vote (used in Australia). The new system is similar in its working to Approval Voting (the system proposed in 1982 by Brams and Fishburn[1]). However, it achieves this effect by means of *preferential* voting instead of the simple X voting of the latter system. It is therefore called *Preferential Approval Voting*, or PAV for short.

The advantage claimed for PAV is one of equity, that as compared with other systems, it gives candidates and parties a stronger incentive to be *equally responsive* to the different sections of the electorate. Also, PAV appears to be a highly practicable method of election. It is not complicated to count, having about the same level of complication as the Alternative Vote, and it could easily be counted by hand, not needing to be counted by computer, however large is the number of candidates.

PAV can best be explained by means of its relation to Approval Voting. The procedure of Approval Voting is simply this: the electors vote (non-preferentially) for as many candidates as they like, for one or for more than one, and the candidate who gets most votes is elected. PAV simulates this procedure by use of preferential voting (that is, voting where the elector votes by marking the candidates in order of preference, 1 for a first preference, 2 for a second preference, and so on, for as many candidates as he wishes).

Now under Approval Voting, the voter will always vote for the candidate whom he most prefers. But under what circumstances will he vote further down his preference ordering, voting in addition for his next-preferred candidate, or for several of the next-preferred candidates? It seems likely that he will do so if he expects that a candidate whom he very much less prefers has some chance of being elected, and if he thinks that voting for the next-preferred candidate or candidates will reduce this chance. For example, a voter whose first preference is Labour, second is Liberal Democrat, and third is Conservative, will always vote for the Labour candidate, and might vote for the Liberal Democrat in addition, if he thinks that the Conservative has a significant chance of winning.

PAV approximately simulates this voting behaviour, by use of the preference orderings provided by the voters. Thus PAV always counts the voter as voting for his first-preferred candidate. PAV counts him as voting for his *next*-preferred candidate *when the latter is preferred to the leading candidate, that one who so far in the counting has obtained most votes*. In other words, this leading candidate is treated as one who has a significant chance of being elected, and therefore voters are assumed to vote for the candidates they prefer to him.

The rules of PAV

Here are the full rules of PAV. The electors vote by putting the candidates in order of preference. “Points” are assigned to candidates, according to the preferences for them, and the candidate with most points is elected. For this purpose, the counting of the votes proceeds in stages, as follows.

The first stage. In respect of each ballot paper, a point is given to the candidate marked as first preference on that paper. The points of each candidate are counted, and the *leading candidate* is found (that is, the candidate who has most points). If there is a tie between two or more candidates, one of them is selected by lot to be the leading candidate.

Any further stage. Those ballot papers are considered, in respect of which a point has *not* so far been given to the leading candidate of the *previous* stage. In respect of each such ballot paper, a point is given to the candidate next-preferred to the last candidate to receive a point, *provided this next-preferred candidate is preferred to the leading candidate of the previous stage*. The leading candidate (who will possibly be a new one) is then found, that is, the candidate who has obtained most points up to and including the current stage.

These further stages are repeated, each one giving more points to the candidates, until the final stage is reached, at which none of the electors' next preferred candidates is preferred to the leading candidate, so that no candidate is entitled to receive any further point. At this final stage, the candidate who has most points is elected.

It will be seen that the method of counting the votes for PAV, is somewhat similar to that for the Alternative Vote. Under both PAV and AV, the first stage is to count the first preferences on all ballot papers. In each later stage, the next preferences are counted on a limited number of the ballot papers, until the winning candidate is found.

A preferential system which bears some resemblance to PAV is that of Descending Acquiescing Coalitions (DAC). DAC is a new preferential election method for filling a single seat, which was recently proposed by Woodall[2,3], as an improvement on the Alternative Vote (which is discussed more fully below). DAC resembles PAV in that both can be regarded as a preferential simulation of Approval Voting. However, Woodall[2] admits DAC is “much more complicated than [the Alternative Vote]”, and would be likely to require a computer to carry out the counting. Thus it is clear that PAV will be much simpler than DAC (see below).

The effects of PAV

In order to illustrate the working of PAV, and to demonstrate the properties of the system, let us consider some numerical examples. We first consider Election 1, where the electors' preferences are single-peaked, that is, preferences are based on some dimension (such as that of left-to-right positions in policy), on which each voter has his own most-preferred point, and on which he prefers any other point less, the further it is from his most-preferred point.

(The notation used to describe the election is explained as follows. The first lines show the voters' preference listings of the candidates. Thus in the top line, 35 voters rank L first, C second, and R third. The subscripts against some of the candidates in a preference listing, show in what stage points are given to the candidate. Thus in the third line, 16 points are given to C in the first stage, and 16 points are given to R in the second stage. After the preference listings, each column shows the total points which have been obtained by each candidate by the specified stage. Thus by stage 2, L has obtained 35 points, C 65, and R 49. The greatest total of points, that of C, is shown in underlined, C being the leading candidate at stage 2.)

Election 1

35 L_1 C R
16 C_1 L R
16 $C_1 R_2$ L
33 $R_1 C_2$ L

	Stage 1	Stage 2	Stage 3
L	35	35	35
C	32	65	65
R	33	49	49

In stage 1, each candidate gets one point for each first preference. L is the leading candidate, getting most points. In stage 2, candidate C (who is the next preference of the 33 first-preference supporters of R, and who is preferred by them to L, the leading candidate of the previous stage) therefore gets 33 more points. Similarly, R gets 16 more points, by being preferred to L by 16 first-preference supporters of C. C, now having most points, becomes the new leading candidate. In stage 3, none of the next-preferred candidates is preferred to C, the leading candidate of the previous stage, and so no candidate gets any more points. Thus C, having most points in the final stage, is elected.

We can use the results of Election 1 to illustrate how PAV deals with *incomplete preference listings*, that is, ballot papers which do not express a preference for all the candidates. It makes no difference whether or not a last preference is expressed by the voter. For example, if the 33 voters voting RCL voted RC instead, this would not alter the result, since we would still know, for stage 2, that they preferred C to L, the leading candidate, so that C would still get 33 extra points. However, it *does* make a difference if a *non-last* preference is not expressed. For example, if the 33 voters voted just R, that is, first preference for R, with no preference given for any other candidate, then C would get no extra points in stage 2, since no preference for C over L would have been expressed.

But let us return to the original results of Election 1 as shown above. In this situation of single-peaked preferences, PAV has elected the centre candidate in the left-to-right dimension. This candidate elected by PAV is also the so-called Condorcet winner, that is, the candidate who beats each other candidate, always being preferred to the other candidate by a majority of voters. (C is preferred over L by 65 voters to 35 and over R by 67 to 33.) Note that PAV achieves this result (that is, of electing the centre candidate or Condorcet winner) despite the fact that C has fewest first preferences, which would prevent C from being elected under the Alternative Vote, that form of preferential system which is most commonly used for electing to one seat.

However, if PAV is actually in use for a series of elections, then it is unlikely that the electors' preferences between the candidates will remain single-peaked. For candidates L and R will surely come to realise that under PAV, their respective extremist positions are going to lose them election after election, and so they will adjust their appeals to give themselves a better chance of winning. Thus L will appeal to the supporters of R, to persuade more of them to change their preference listing to RLC instead of RCL, and R will appeal to supporters of L to get them to change to LRC. The pattern of the electors' preferences will then no longer be single-peaked, but will tend towards what might be called a symmetrical pattern, where there is about the same number of voters with each possible preference listing (that is, in this case, one-sixth LCR, one-sixth LRC, and so on). Thus a typical election might be something like Election 2.

Election 2

18	L_1C_4R
17	L_1R_3C
17	C_1L_4R
15	C_1R_2L
17	R_1C_2L
16	R_1L_3C

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
L	$\underline{35}$	35	51	$\underline{68}$	$\underline{68}$
C	32	$\underline{49}$	49	67	67
R	33	48	$\underline{65}$	65	65

Thus by broadening their appeal, L and R have got more points, and L has succeeded in getting elected. L now gets second preferences, not only from first-preference supporters of C as before, but also from the first-preference supporters of R, and similarly R now gets second preferences from the first-preference supporters of L. This illustrates how PAV gives a candidate or party the incentive to appeal to, and to be responsive to, all sections of electors.

Election 2 can be used to illustrate the general strategy by which a candidate will seek to win under PAV. A candidate wins by getting a point from the most voters. A candidate C gets a point from any one voter V either if C gets V's first preference, or otherwise *if C is preferred by V to that one of the leading candidates who is least preferred by V*. Thus in Election 2, L gets a point not only from the 18 LCRs and 17 LRCs, but also from the 17 CLRs and the 16 RLCs.

This has implications for a candidate's general strategy. He will be primarily concerned to persuade voters to prefer him over their least preferred leading candidate. Once they do this, he will not seek to persuade them to give him a still higher preference (that is, a first preference in Election 2), since this will tend to be difficult to achieve, and in any case *it will not bring him any more points*. Thus when there are three leading candidates, as in Election 2, each one will direct his appeal primarily at those electors who have tended to give him last preference, and in general, each candidate will be seeking to get second preferences rather than first preferences.

Further properties of PAV

PAV has the same property as does the Alternative Vote, and also DAC, that a candidate who gets an absolute majority of first preferences is necessarily elected. This can be simply shown as follows. Suppose A has the first preferences of more than half the voters. Thus A is the leading candidate at the first stage, with a point from more than half the voters. At the second stage, the best that any other candidate can do is to get a point from every voter who did not vote first preference for A, that is, he must get points from less than half the voters. Thus A, with a point from more than half the voters, must be the leading candidate at the second stage. By a similar argument, A must be the leading candidate at the next stage, and at any stage after that. Thus A must be elected.

However, PAV is *unlike* the Alternative Vote in that the candidate with fewest first preferences can be elected, as was the case in the single-peaked example of Election 1 above. Indeed, PAV can enable a candidate to get elected who has very few or even no first preferences. A *non*-single-peaked example of this, which might well occur occasionally in practice, is Election 3. Here a candidate C, who has few first preferences, gets more points than either A or B (each of whom have close to half the first preferences) by persuading many of As and Bs first-preference supporters to give C their *second* preferences.

Election 3

32	A_1C_3B
16	A_1B_4C
15	B_1A_4C
32	B_1C_2A
2	C_1B_2A
3	C_1A_3B

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
A	<u>48</u>	48	51	66	66
B	47	<u>49</u>	49	65	65
C	5	37	<u>69</u>	<u>69</u>	<u>69</u>

This lack of the need for first preferences under PAV, can be expected to reduce the entry barrier against new candidates. For it is likely to be easier to gain second preferences than first preferences, thus making it easier under PAV for a new candidate to compete successfully with already established candidates, than it would be under the Alternative Vote, or in particular under Plurality. Thus under PAV, at least when it has been in use for some time, it is likely that few candidates will obtain a majority of first preferences, and that the most usual situation in each constituency will be for there to be three strong candidates (or perhaps sometimes more than three) in not very unequal competition. In other words, it is likely that under PAV, there will be a tendency towards a symmetrical situation like that shown in Election 2.

In all the examples given above, Elections 1, 2 and 3, there were only three candidates competing. How then will PAV operate, if there is a larger number of candidates? The same procedure will be followed, that of sorting and counting the next preferences stage by stage, until that stage is reached, where no next-preferred candidate is preferred to the leading candidate, and thus no candidate is entitled to receive any further points. Because there are more candidates, there will of course be more next preferences to sort and to count. But the extra counting need not be in proportion to the number of extra candidates. The reason for this is that on any one ballot paper, only the top preferences need to be counted, down to the preference for the candidate who is one preference step above that one of the "leading candidates" whom the voter least prefers. It is likely that the extra candidates will be given a very low preference (or no preference) by most of the voters, and that because of this their preferences for them will not need to be counted.

Election 4 is given below, as an example of a four-candidate election. Election 4 is assumed to be a re-run of Election 2, in which one party, the party which previously ran L as its candidate, now runs two candidates L and M, one a woman and one a man, in order to give the electors a wider choice. Electors are assumed to put L and M in the same position in their preference listings as they put L in Election 2.

Election 4

10	L_1 M_2 C R
8	M_1 L_2 C R
9	L_1 M_2 R C
8	M_1 L_2 R C
9	C_1 L_2 M R
8	C_1 M_2 L R
8	C_1 R_3 L_4 M
7	C_1 R_3 M_5 L
9	R_1 C_3 L_4 M
8	R_1 C_3 M_5 L
9	R_1 L_4 M C
7	R_1 M_3 L C

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
L	19	<u>44</u>	44	<u>70</u>	<u>70</u>	<u>70</u>
M	16	43	<u>50</u>	50	65	65
C	32	32	49	49	49	49
R	<u>33</u>	33	48	48	48	48

PAV and Condorcet

Another question of some interest is how PAV differs from Condorcet, the well-known method of electing to a single seat by means of preferential voting. Under Condorcet, A beats B if there are more voters who prefer A to B than those who prefer B to A. But under PAV, A beats B if there are more voters who give A a first preference, or otherwise prefer A to a "leading candidate", than those who give B a first preference, or prefer B to a leading candidate. Thus an important difference between the two systems, is that under Condorcet, a voter supports either A or B, but cannot support both; whereas under PAV, it will often be the case that the same voter supports both A and B, preferring A to a leading candidate, and also preferring B to a leading candidate. Not surprisingly, PAV is in this respect similar to normal Approval Voting, where any one voter can vote (in this case with an "X") for both A and B.

But how far does PAV tend to elect the Condorcet winner (CW)? The CW was elected in Election 1, where preferences were single-peaked, and also in the more likely preference situation of Election 2 (L, the PAV winner, being preferred over C by 51 voters to 49 and over R by 52 to 48). However, in Election 3, where C, the PAV winner, got most of his votes from second preferences, the CW was *not* elected, the CW being candidate A (who was preferred over B by 51 voters to 49, and over C by 63 to 37). It thus appears that in practice, in the preference situations most likely to occur, PAV has a very high probability of electing the CW, but that it might not elect the CW in some unusual situations, where the PAV winner obtains an especially high proportion of his points from lower preferences.

PAV and Descending Acquiescing Coalitions (DAC)

It is of especial interest to compare PAV with DAC, which is another new single-seat preferential system which can be regarded as a preferential simulation of Approval Voting. The rules of DAC can be explained as follows.

A voter is said to *acquiesce* to a set of candidates if there is no candidate outside the set whom he prefers to any candidate in the set. (In other words, in respect of any pair of candidates, one in the set and one outside the set, he always *either* prefers the candidate in the set, *or* expresses no preference between them.) The set of all those voters who acquiesce to the candidates A and B is referred to as the *coalition* acquiescing to A and to B, or as {A, B}. For example, if there are only three candidates A, B and C, then {A, B}will be all those voters voting as follows: ABC, AB, BAC, BA, A or B.

That candidate is elected who obtains the acquiescence of a greater number of voters than any other candidate. This is determined as follows. A candidate A is said to *beat* a candidate B if the greatest coalition acquiescing to A and *not* acquiescing to B, is greater than the greatest coalition acquiescing to B and not acquiescing to A. That candidate is elected who beats each other candidate.

This can be illustrated by the following two examples, taken from Woodall[2] .

Election 5 (Election 3 of Woodall)

11	AB
7	B
12	C

This produces acquiescing coalitions as follows, in descending order of size.

{A, B, C}	30
{B, C}	19
{A, B}	18
{A, C}	12
{C}	12
{A}	11
{B}	7

B beats A, because {B, C} > {A, C}. B beats C, because {A, B}> {A, C}. Thus B is elected.

Election 6 (Election 4 of Woodall)

5	ADCB
5	BCAD
8	CADB
4	DABC
8	DBCA

This produces a set of the greatest acquiescing coalitions as follows.

{A, B, C, D}	30
{A, B, C}	13
{D}	12
{A, D}	9
{A, C}	8
{B, C, D}	8
{B, D}	8
{C}	8

A beats B, because{A, D} > {B, C, D}. A beats C, because {A, D}> {B, C, D}. A beats D, because{A, B, C}> {D}. Thus A is elected.

Let us now compare DAC with PAV. Under DAC, A beats B if more voters are in the greatest coalition acquiescing to A and *not* acquiescing to B, than are in the greatest coalition acquiescing to B and not acquiescing to A. Under PAV, A beats B if there are more voters who give A a first preference, or otherwise prefer A to a “leading candidate”, than those who give B a first preference, or prefer B to a leading candidate.

DAC is like PAV, and *un*like the Alternative Vote, in that it does not require a candidate to get first-preference votes in order to get elected, and so it can elect the candidate with fewest first preferences (as it does in Election 5). The two systems DAC and PAV are similar to each other, and to Approval Voting, in that each of them can give value to one or more of the highest non-first preferences of an elector, and in that if it does, the value of a non-first preference is the same as that of a first. DAC can thus be regarded as a preferential simulation of Approval Voting, as can PAV.

PAV and lack of monotonicity

A system is non-monotonic if it is possible under it for a candidate who gets more voting support, to lose the election as a consequence. The ten monotonicity properties, that is, ways in which a system can be monotonic or not, are analysed in Woodall[2,3]. Elections 7 to 9 below, show PAV to be non-monotonic in at least two of these ways.

Election 7

10	A_1 B_2
9	B_1
2	C_1 B
9	C_1
8	D_1 A_2

	Stage 1	Stage 2	Stage 3
A	10	18	18
B	9	19	19
C	11	11	11
D	8	8	8

Thus B is elected.

Now suppose that in Election 8, the two voters who voted CB in Election 7, change to voting BC instead. The stages of the count will then be as shown below, and A will be elected. Thus by moving up the preference listing of these two voters, B will have lost the election.

Election 8

10	A_1 B
9	B_1
2	B_1 C_3
9	C_1
8	D_1 A_2

	Stage 1	Stage 2	Stage 3	Stage 4
A	10	18	18	18
B	11	11	11	11
C	9	9	11	11
D	8	8	8	8

Alternatively, suppose that in Election 9, the profile is as in Election 7, except that three new voters enter the election, and vote first preference for B, so that the second line in the election profile is 12 B instead of 9 B.

Election 9

10	A_1 B
12	B_1
2	C_1 B_3
9	C_1
8	D_1 A_2

	Stage 1	Stage 2	Stage 3	Stage 4
A	10	18	18	18
B	12	12	14	14
C	11	11	11	11
D	8	8	8	8

Thus A is elected. Again, B has lost the election, this time by getting more voters to vote for him.

It should be pointed out that the Alternative Vote is also non-monotonic, whether more or less so than PAV I am unable to determine. DAC, on the other hand, was designed to satisfy as many monotonicity properties as possible, and in fact satisfies eight out of ten of them.

How far, then, would this lack of monotonicity in PAV be a problem not just in theory, but in actual practice in real elections? The main objective of PAV is to give each candidate the incentive to be responsive to each section of electors. Thus the important question is, how far will lack of monotonicity interfere with this incentive? Will a candidate (such as B in Elections 7 to 9 above) ever have the incentive to *displease* the electors, so that they give him a lower preference, or so that fewer of them vote for him?

This seems unlikely, for two reasons. First, a non-monotonic profile of votes such as those of Elections 7 to 9 seems itself unlikely when candidates are competing strongly, not only for first preferences, but for second and third preferences as well. Then the profile tends towards a more symmetrical pattern such as that shown in Election 2 above, which would be monotonic. Second, in order for the candidate to be provided with this negative incentive, he must be able to predict that the overall profile of votes at the next election will be such as to produce this non-monotonicity, and furthermore that his own votes will be in that presumably narrow range where he will benefit from losing votes. In the absence of this prescience, the candidate will have the incentive to respond positively to the electors, in the expectation that nearly always it will be beneficial for him to get more votes rather than fewer of them. Thus it seems unlikely that this lack of monotonicity will affect the candidates' incentives, or will be of practical importance.

Strategic voting

It is well known that any non-probabilistic method of election provides the opportunity, in some situation or other, for electors to engage in strategic voting. What form then will this strategic voting take, under PAV? It appears that the most likely strategy will be for the voter to give a *truncated preference listing*. For example, if it is expected that either A or B will get most points, and that both will get considerably more than C, then some of the ABCs (that is, electors whose preferences are A first, B second, C third) might adopt the strategy of voting only a first preference for A, and giving no preference for the other candidates (and similarly some BACs might vote only a first preference for B). Thus by not giving any votes to B, the ABCs make it more likely that A, their first preference, will be elected.

The other systems similar to PAV are liable to strategy in a similar way. Thus under normal (non-preferential) Approval Voting, a similar strategy is very likely to be used—ABCs voting only for A and BACs voting only for B, when the election is expected to be a two-horse race between A and B. Under DAC, a preferential system with some similarity to PAV, this same strategy of the truncated preference listing is likely to be used (according to Woodall[4]).

How far, then, is it a problem, that there is this opportunity for strategy under PAV? The strategy will be used in constituencies where two of the candidates are clearly stronger than the others, and it is expected that the winner will be one or other of them. But in constituencies where there are *three* or more strong candidates, and it is unclear which of them is going to get most points, the electors will tend *not* to vote strategically, but to express fully their preferences between these candidates.

However, there are reasons to expect that any constituency will tend to move from the former situation towards the latter, that is, from one with two strong candidates to one with three or more. Firstly, as it was shown above, PAV does not require a candidate to get many first preferences in order to win, and so it presents relatively little entry barrier to an effective new candidate. Secondly, when there are two strong candidates, let us say A and B, and a weaker candidate C, the strategic voting which this situation encourages actually benefits C. For some ABCs will vote only first preference for A, which will reduce Bs votes, and some BACs will vote only first preference for B, thus reducing As votes. This reduces the number of first or second preferences which C needs to get, to approach about the same number of votes as A or B, making it easier for C to become a third strong candidate. It will then be uncertain which of the three candidates is going to get most votes, and strategic voting will become unlikely.

Thus in conclusion, it seems that the tendency in any constituency is towards a situation where there are three (or perhaps more than three) strong candidates, each with some chance of winning. To the extent that this situation occurs, the truncation strategy will tend not to be used, and voters will express fully their preferences for the candidates.

An evaluation of PAV

In the view of this paper, the main objective of an electoral system is to provide the elected candidates, and the parties to which they belong, with the incentive *to respond to the needs of the electors*; and to respond not just to a part of the electorate, even a majority part, but to respond equitably to *each section* of electors, each possible minority. How far then does PAV provide the incentive to this equitable all-round responsiveness?

To answer this question, let us consider the examples of Elections 1 and 2 above. In Election 1, candidates L and R fail to respond to all sections of electors, L not responding to the right-wing electors, and so getting a last preference from them, and R not responding to the left-wing electors. Consequently, they lose points, and neither of them has any prospect of getting elected.

However, in Election 2, each of them has broadened his appeal to include the whole electorate, L responding to right-wing electors, and R to left-wingers. L now gets second preferences, not only from centre electors as before, but also from right-wing electors, and similarly R gets second preferences from left-wingers. Thus by broadening their appeal, L and R get more points, and L succeeds in getting elected. This illustrates how PAV gives each candidate the incentive to respond to each section of electors.

Note that in Election 2, the situation between all three candidates is *symmetrical* in the sense that any two candidates compete with each other for the second preferences of the third candidate's first-preference supporters. Thus L and R compete for the second preferences of centre electors (just as they did in Election 1). But now L competes with C for right-wingers' second preferences, and similarly R competes with C for left-wingers' second preferences. Any one candidate thus needs to be responsive to the first-preference supporters of any other candidate, in order to compete with the third candidate for their second preferences. For example, L needs to be responsive to centre electors to compete with R, and to right-wing electors to compete with C. Thus PAV gives each candidate the incentive to be responsive to each section of the electorate.

Another way of understanding the incentives provided by PAV is as follows. In the likely situation where there are three candidates competing, and each becomes a leading candidate at some stage in the counting, a candidate receives one point for each first preference and one point for each second preference. Thus (assuming all voters express their second preferences), a candidate needs to get either a first-preference or a second-preference vote from *at least two-thirds of the voters* in order to get elected. He is not likely to achieve this, in competition with two other candidates also trying to do the same thing, unless he appeals to each section of electors. Thus the candidate has the incentive to respond to each section of the electorate.

Furthermore, a first preference is worth no more than a second preference--both are worth only one point. Thus there will be no need for a candidate to appeal to a given section of electors any more strongly than is necessary to get second preferences from it, and no reason to give the section any specially favourable treatment, in order to obtain from it a higher proportion of first preferences. This is clearly a factor making for the candidates' *more equal* responsiveness to each section.

It is interesting to compare the situation under PAV as described above, with that under the Alternative Vote. Here, in order to get elected, a candidate needs to obtain the support not of two-thirds of the voters, but of only one-half. Thus he is likely to appeal less widely. Further, each candidate must strive for first preferences, since the candidate with fewest first preferences will be excluded. This seems likely to create an incentive for a candidate to favour some sections of electors over others, in order to get first preferences from them.

To illustrate this, let us consider an example with three candidates, A, B and C, where it is expected that C will be excluded, and that it will be a close finish between A and B. Each of A and B will have his core supporters, to whom he is strongly responsive, in order to obtain first preferences from them. Also, each of A and B will be strongly responsive to those voters giving first preference to C, in order to compete with the other candidate for these voters' second preferences. But A will tend to be unresponsive to the core supporters of B, because of the difficulty of persuading them to switch from first preference for B to first preference for A. Similarly, B will tend to be unresponsive to the core supporters of A. Thus the Alternative Vote, by forcing candidates to strive for first preferences, makes for their unequal responsiveness to the different sections of electors. In comparison, PAV, which makes no requirement for first preferences, will give candidates the incentive to respond more equally to each different section of electors.

PAV in the UK

If PAV were introduced in the UK for the Westminster Parliament, the present single-member constituencies would be retained. The only difference for the electors would be that they would vote by putting candidates in order of preference, instead of X-voting for only one candidate.

What then would be the effect on the parties' shares of seats? The present Plurality system, which essentially gives a seat to the candidate with most first preferences, discriminates strongly against the Liberal Democrats, who have third most first preferences. However, under PAV, they would be likely to get many more seats than now, since there seems no reason why they should not get about as many second preferences as either of the other two major parties. Thus it seems likely that the three major parties would be more equal in their seats than they are now, and that no one party would get a majority; so that a coalition government would need to be formed, by some two of them.

The point of most interest, and the main advantage claimed for the new system, is that it would give parties the incentive to change their policies to be more inclusive, more equitably responsive to the different sections of the electorate. For example, the Conservative Party currently tends to be unresponsive to strong Labour supporters, since under the present Plurality system few of them could be persuaded to switch to voting for the Conservatives. But under PAV, the Conservative Party would become more responsive to them, in order to compete with the Liberal Democrats for their second preferences. Similarly, the Conservatives would become more responsive to strong Liberal Democrat supporters, in order to compete for their second preferences with Labour. Thus the three major parties would tend to converge in policy, towards a policy more equally responsive to each section of electors; and as a result of this convergence, a coalition government formed by any two of them would be likely to be stable, and acceptable to all sections of the electorate.

Other uses of PAV

PAV could be used with advantage, instead of the Two-Ballot System, for the election by popular vote of individual office-holders, such as the president of France, the president of Russia, or the prime minister of Israel. The advantage of PAV for this purpose, can be explained as follows. Under the Two-Ballot System, the usual rule is that if there are more than two candidates on the first ballot, and no-one gets a majority of the votes, then the two strongest candidates go forward to the second ballot, where one of them must get a majority. Thus a moderate or centre candidate, who is widely acceptable to the electorate, and who could win in the second ballot if he got there, may well fail to get elected, because he gets too few votes on the first ballot. But as was explained above, under PAV there is no requirement to get first preferences (corresponding to first-ballot votes in the Two-Ballot System), and a candidate can be elected just as well by second as by first preferences. Thus this moderate or centre candidate, with few first preferences but many second preferences, is likely to get elected under PAV, where he would not be elected under the Two-Ballot System.

For similar reasons, it might be desirable to use PAV for purposes such as the following: the election of a president, or of a chairman, by the members of a legislature; the election of the party leader by the party membership, or by the party's MPs.

PAV could also be used for a multi-option referendum, to enable the electorate to choose one option out of three or more. This can be justified as follows.

In the usual type of referendum, electors choose between two options, these options being some proposed action, let us say A, and the status quo S. Proposers will be concerned to find an A which will get a majority over S, and in doing so they may come up with an A which is very harmful to the minority, while perhaps only marginally beneficial to many people in the majority. Thus the two-option referendum might lead to very unequal treatment of different sections of the electorate, and to division and conflict.

However, if a PAV-using multi-option referendum is introduced, a compromise option C is likely to be proposed, one which is better than A for S preferrers and some A preferrers, and better than S for other A preferrers. Thus there will be three options on the ballot paper, A, S and C, for the electors to place in order of preference. Since C will have many second preferences, it is likely that C will be adopted. This illustrates how a PAV-using multi-option referendum tends to improve the outcome, reducing the risk that any section of the electors will be severely harmed.

References

1. S J Brams and P C Fishburn, *Approval Voting*, 1982.

2. D R Woodall, Monotonicity and Single-Seat Election Rules, *Voting Matters*, Issue 6 (May 1996), 9 - 14.

3. D R Woodall, Monotonicity of Single-Seat Preferential Election Rules, *Discrete Applied Mathematics*, 77 (1997), 81 - 98.

4. D R Woodall, private communication, 3 July 1996.

Decision Analysis — Responses to a Questionnaire

H G Jones

Bob Jones is a retired mathematician and former secretary of Derbyshire Electoral Reform Group (DERG).

Introduction

An article describing the application of Decision Analysis to choice of "best" electoral system was given in Issue 12 of *Voting matters*. Readers were invited to complete their own version of the Analysis Table supplied. The present article gives an analysis of the responses received.

Not surprisingly, in view of the readership of *Voting matters* nearly all favoured STV. It was therefore decided to invite a wider population to respond. This was just before the General Election on June 7th and candidates from the local "Jenkins AV+" area were contacted. The area consists of the present constituencies of Cheltenham, Gloucester, Tewkesbury, Stroud, Cotswold, and Forest of Dean. Responses from some 20 candidates was sparse so other political and non-political people were contacted.

A total of 14 responses was received.

Method of averaging

For each FEATURE (of a voting system) the average value from respondents was evaluated. These features are plotted in Figure 1 in the order giving the most liked feature first. In that order, the features are:

PRO-N: How proportional is the national result?

EASYV: How easy is the system for the voter to use?

PRO-R: How proportional is the result within a region? (A region is visualised as, say, 10 of the present neighbouring constituencies.)

LOC: Local link — How closely are MPs linked to an area?

EW&E: Does the system encourage women and people from ethnic minorities to stand for election?

CHO-MP: Is there a choice within a party as well as across party lines?

PLOC: How easily can constituents contact an MP of their preferred political persuasion?

ONECMP: Is there one class of MP? (Some systems have regional as well as local MPs)

EASYC: How easy is the process of counting?

EASYBC: How easy is the task of the Boundary Commission?

STAB: Stability of government. STAB really asks the question "Is the government likely to complete its normal period of office?" Critics of PR sometimes say it results in "weak" coalition government. This has some validity with Party Lists, particularly when based on the whole country as in Israel. Experience in Germany since 1945 with AMS, and in Eire since 1922 with STV are to the contrary.

It should be noted that the *Voting matters* article used a range of weighing factors from 0 to 3, whereas from March 2001 a range from 0 to 10 was in use. Furthermore the additional FEATURE of STAB was not considered as it did not appear in the original *Voting matters* article.

For each voting system, a similar plot is produced in Figure 2. Here the systems in reducing order of preference are:

STV: Single Transferable vote.

PLRO: Party List based upon a region and using open lists.

PLRC: Party List based upon a region and using closed lists.

AV50: Similar to AV+, but having a 50% top-up element.

PL: Party List.

AV+: The proposal made by Lord Jenkins.

AMS: Additional Member System as used in Germany since 1945 and in differing forms for the Scottish Parliament and the Welsh Assembly.

AV: Alternative vote.

FPTP: First Past the Post (as used in Westminster).

Readers who would like to fill in their own questionnaire can obtain a copy from the Editor by writing to ERS or electronically by e-mailing Brian.Wichmann@freenet.co.uk.

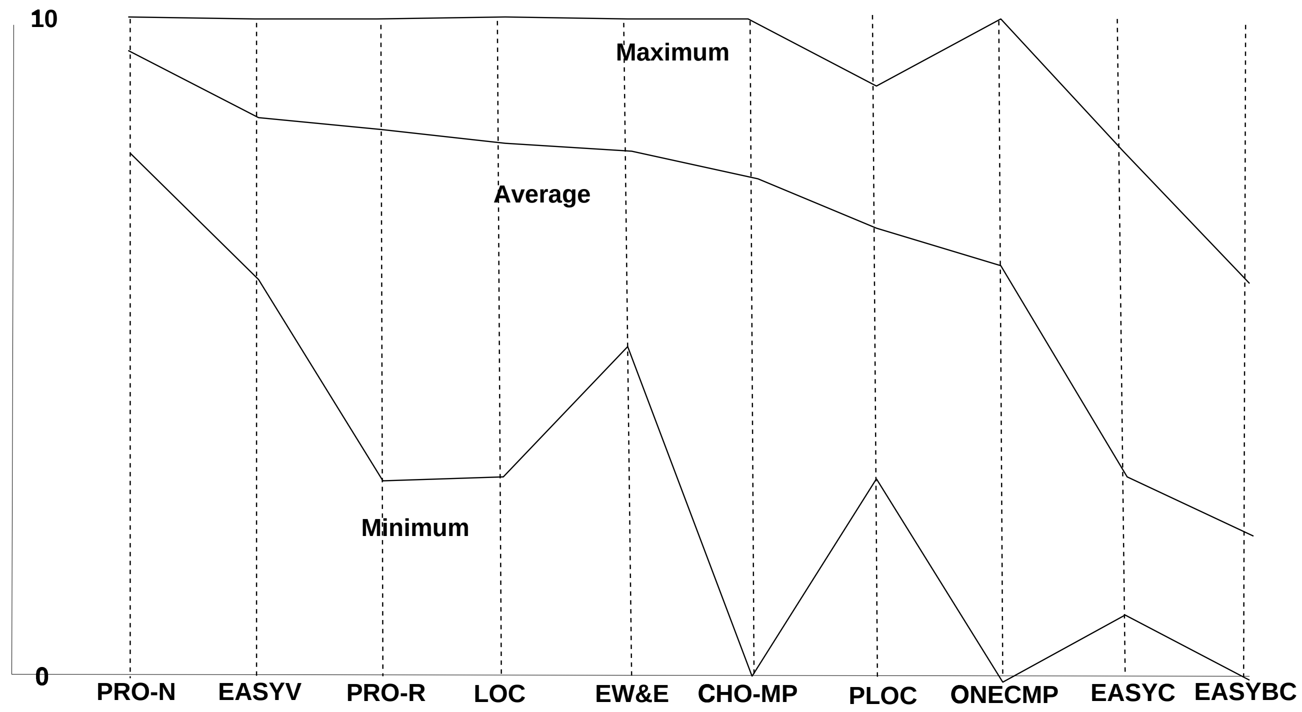


Figure 1: Features

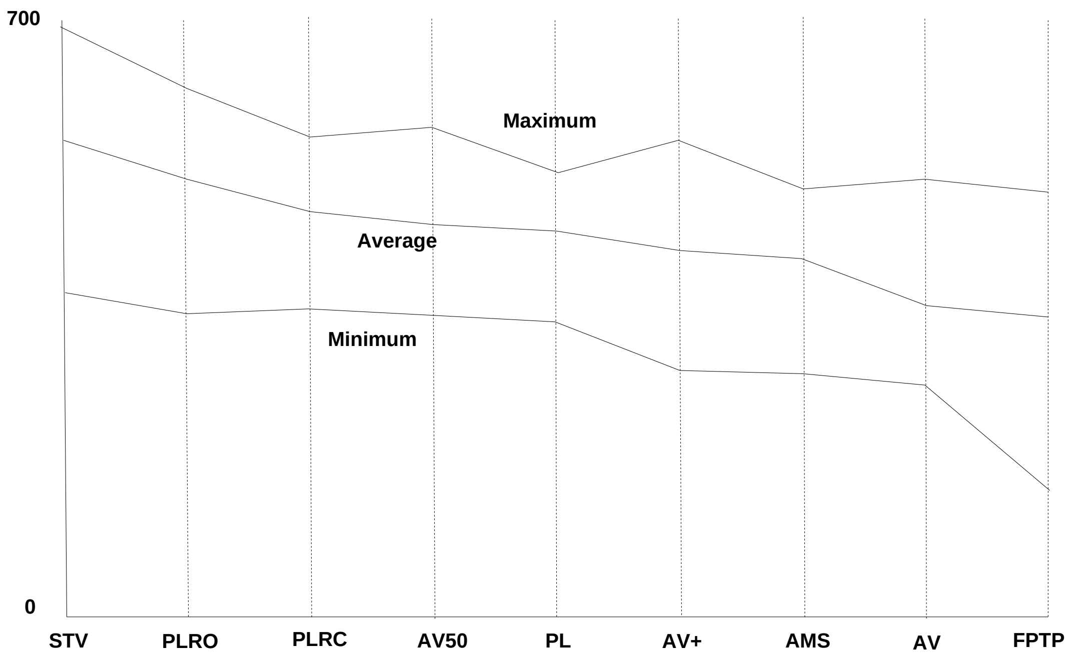


Figure 2: Systems

All correspondence regarding *Voting matters* should be addressed to:

B A Wichmann, The Editor, *Voting matters*, The Electoral Reform Society, 6 Chancel Street, London, SE1 0UU.

or using e-mail to Brian.Wichmann@bcs.org.uk.

To aid production, the Editor would welcome contributions on IBM-PC discs (with a printed copy as well) or to Brian.Wichmann@bcs.org.uk.

Voting matters

for the technical issues of STV

The Electoral Reform Society

Issue 15 June 2002

Editorial

On the 17th May 2002, the Dáil constituencies of Meath, Dublin North and Dublin West used an experimental system for electronic voting. It is expected that this system will be used exclusively for local and national elections in the Irish Republic in the near future.

Of course, the three constituencies used the same electoral rules as in the other 40 — essentially a hand-counting system which has at least one ‘problem’ in that the result can depend upon the order in which papers are transferred on a surplus. Examination of such issues has traditionally always been hampered by the lack of complete information of all the preferences expressed by the voters.

We now have a significant step forward for electoral studies since the Irish electronic voting results includes the complete data input to the electronic counting software. One can reasonably expect future issues of *Voting matters* to analyse this data.

The first paper in this issue is indeed an analysis of Irish election data, but only uses the result sheets. Philip Kestelman shows statistically significant bias according to the alphabetic position (on the ballot paper). I might add that even a casual inspection of the full data mentioned above shows a tendency for the final few preferences to be in strictly ascending or descending order.

In the second paper, Eivind Stensholt considers the problem when additional support for a candidate results in that otherwise elected candidate not being elected. This property of *non-monotonicity* applies even to the case of electing a single candidate, as shown in this paper. On the other hand, the paper indicates that it is relatively rare.

In the third paper, Markus Schulze considers an algorithm for electing candidates with preference voting proposed by Professor Sir Michael Dummett. Sir Michael has chosen not to respond to the criticisms made.

In the last paper, David Hill and Simon Gazeley produce a new STV-like algorithm which merges the ideas of Condorcet and STV. The advantage of this algorithm is to avoid the property of all conventional STV algorithms of premature exclusion, such as for a universal second-choice candidate. On the other hand, this method has the disadvantage of later preferences could possibly upset earlier ones in rare cases.

McDougall Trust: STV Resources CD

A proof copy was prepared in February, but the publication date has not yet been agreed.

Brian Wichmann.

Positional Voting Bias Revisited

Philip Kestelman

Introduction

It is widely supposed that candidates appearing high on ballot-forms enjoy a considerable electoral advantage. In a highly influential paper on the 1973 General Election to the Irish Dáil, by multi-member Single Transferable Voting (STV), Robson and Walsh (1974) observed that Deputies (TDs) over-represented candidates with A-C surnames. Compared to randomly sampled Irish electors, “The under-representation of M-O names among politicians is very striking”.

Proportionality conventionally measures the relationship between numbers of Party *votes* and seats (regardless of candidates). Despite a probable age bias, we are hardly concerned that seats considerably over-represent first preferences for *incumbent* candidates; let alone that incumbents are far more likely to be elected than ‘excumbents’ (non-incumbents).

On the other hand, we are concerned not only that seats should proportionally represent votes for *women* candidates, but also that seats should be proportional to women *candidates*, in the interests of Parliament representing society. In respect of ballot-form *position*, we are primarily concerned with the relationship between numbers of *candidates* and seats (regardless of votes), by surname initial, when candidates are listed surname-alphabetically on ballot-forms.

Electability

This article mainly evaluates positional voting bias in the last 12 general elections in the Irish Republic (1961-97). Electability is quantified in terms of an Electability Index (S%/C%): the ratio of a seat-fraction (S%) to a candidate-fraction (C%); and of a Relative Electability Ratio: the ratio between specified Electability Indices.

Aggregating all 12 elections (Total S/C = 1,875/4,594), Upper/Lower half surname A-J/K-Z Electability Indices were 1.11/0.88, with a statistically highly significant Relative Electability Ratio of 1.26 (P<0.001). By comparison, alphabetically Upper/Lower half forename A–L/M–Z Electability Indices were 1.01/0.99, with an insignificant Relative Electability Ratio of 1.01 (P > 0.05).

Cumbency

In 1961-97, most incumbent candidates (S/C = 1,404/1,687 = 83 percent) were re-elected; whereas few excumbents (471/2,907 = 16 percent) were elected, rendering them more susceptible to alphabetic disproportionality. Surname A-J/K-Z Electability Indices (S%/C%) were 1.01/0.98 for incumbents, and 1.15/0.86 for excumbents, with Relative Electability Ratios of 1.03 (P>0.05) and 1.34 (P<0.05), respectively.

The last 12 Irish general elections have consistently over-represented excumbent candidates with A-C surnames; under-representing those with K-M surnames (overall S%/C%, 1.27 and 0.81: Table A). Even combining the 12 elections into three quartets leaves considerable variability in both forename and surname Electability Indices.

Table A: *Excumbent* Electability Index, by Elections and Forename/Surname initial letter: Irish Republic, 1961-97 (12 general elections: Dáil Éireann, 1962-98).

Name		Elections: Electability Index (S% / C%)			
Initial letter		1961-97	1961-73	1977-82	1987-97
Forename	A-F	1.04	1.34	1.11	0.88
	G-L	1.08	1.08	0.98	1.16
	M-P	1.02	0.82	1.14	1.08
	Q-Z	0.82	0.84	0.72	0.85
Ratio (A-L/M-Z)		1.14	1.42*	1.10	1.01
Surname	A-C	1.27	1.38	1.19	1.24
	D-J	1.04	1.04	1.06	1.04
	K-M	0.81	0.73	0.84	0.83
	N-Z	0.92	0.91	0.91	0.94
Ratio (A-J/K-Z)		1.34*	1.47*	1.29	1.27

* P < 0.05

In 1961-73, excumbent forename and surname alphabetic biases were equally convincing (P<0.05); but insignificant subsequently. Ironically in 1973, the Relative Electability Ratio for A-L / M-Z forenames (2.76) exceeded that for A-J / K-Z surnames (1.57)! The pitfalls of generalising from a single election are manifest.

District Magnitude

Surname disproportionality was virtually confined to four- and five-member STV constituencies: only three-member constituencies returned TDs more-or-less faithfully reflecting excumbent surnames (Table B). Magnitude-specific surname A-J/K-Z Relative Electability Ratios proved statistically insignificant, but much closer to unity in three-member constituencies (1.25, 0.89 and 0.72) than in four-member constituencies (1.62, 1.36 and 1.51), or in five-member constituencies (2.05, 1.77 and 1.42), in 1961–73, 1977–82 and 1987–97, respectively.

Table B: *Excumbent* Electability Index, by District Magnitude (seats per constituency) and Surname initial letter: Irish Republic, 1961-97 (12 general elections: Dáil Éireann, 1962-98).

Surname Initial letter	Magnitude: Electability Index (S% / C%) Total+	3	4	5
A-C	1.27	1.10	1.51	1.21
D-J	1.04	0.87	0.98	1.28
K-M	0.81	1.05	0.66	0.73
N-Z	0.92	0.98	0.96	0.79
Ratio (A-J/K-Z)	1.34*	0.96	1.51	1.64

*P < 0.05 +Including a few two-member constituencies.

District Canditude and Position

Interestingly, the 1961-97 aggregate, excumbent Relative Electability Ratio by surname (A-J/K-Z) proved identical with that by ballot-form position (Upper/Lower = 1.34: P<0.05). Like the surname A-J/K-Z Relative Electability Ratio with district magnitude (the number of seats per constituency), the positional Upper/Lower Relative Electability Ratio increased with district 'canditude' (the number of candidates per constituency: Table C).

Table C: *Excumbent* Electability Index, by District Canditude (candidates per constituency) and Ballot-form Position: Irish Republic, 1961-97 (12 general elections: Dáil Éireann, 1962-98).

Ballot-form Position+	Canditude: Electability Index (S% / C%) Total	4-8	9-11	12-21
Top	1.30	1.22	1.26	1.41
Upper-middle	0.98	0.86	1.06	1.01
Lower-middle	0.90	1.15	0.91	0.76
Bottom	0.83	0.83	0.81	0.86
Ratio (Upper/Lower)	1.34*	1.13	1.37	1.49

* P < 0.05 +Excluding odd-Canditude mid-candidates.

Party Policy

Both main political parties in the Irish Republic (Fianna Fáil and Fine Gael) have staunchly denied over-nominating candidates appearing high on ballot-forms[9]. Table D analyses the surname-alphabetic distribution of FF and FG excumbent candidates, compared with other (non-FF + FG) excumbents, in terms of a Relative Nomination Index, over time.

Table D: Two Main Party *Excumbent* Relative Nomination Index, by Elections and Surname initial letter: Irish Republic, 1961-97 (12 general elections: Dáil Éireann, 1962-98).

Surname Initial letter	Elections: Relative Nomination Index (Fianna Fáil +Fine Gael C%/ Other C%) 1961-97	1961-73	1977-82	1987-97
A-C	1.46	1.07	1.63	1.57
D-J	0.96	1.02	1.13	0.84
K-M	0.92	1.09	0.79	0.87
N-Z	0.81	0.85	0.69	0.92
Ratio (A-J/K-Z)	1.35***	1.08	1.82***	1.25

*** P < 0.001

Evidently since 1977, both main parties have greatly over-nominated A-C surname candidates (and/or other parties have under-nominated them); with the honours evenly divided between Fianna Fáil and Fine Gael. Relative to the publication of Robson and Walsh (1974), the timing may not have been entirely coincidental!

Electorate

Robson and Walsh (1974) observed that the alphabetic distribution of surname initial letters differed insignificantly between randomly sampled Irish electors and excumbent candidates at the 1973 Irish General Election. Presumably nowadays, the surname initials of *electors* are rather better represented by excumbent, non-FF + FG candidates; and Table E compares overall seat-fractions, by surname initial, with excumbent, non-FF + FG candidate-fractions.

This Surname Concentration Index (Total S%/Excumbent, non-FF + FG C%) highlights Dáil Éireann over-representing A-C surnames in the Irish electorate; while under-representing K-M surnames. Despite the lower surname A-J/K-Z Concentration Ratio since 1987 (1.35), A-C surname electors remain over twice as likely as K-M surname electors to become TDs.

Table E: Surname Concentration Index, by Elections and Surname initial letter: Irish Republic, 1961-97

(12 general elections: Dáil Éireann, 1962-98).

Surname Initial letter	Elections: Surname Concentration Index (Total S% /Excumbent, non-Fianna Fáil +Fine Gael C%) 1961-97	1961-73	1977-82	1987-97
A-C	1.65	1.56	1.66	1.58
D-J	1.01	1.13	1.13	0.91
K-M	0.70	0.76	0.63	0.73
N-Z	0.85	0.68	0.85	0.99
Ratio (A-J/K-Z)	1.65***	1.85***	1.86***	1.35**

** P < 0.01 *** P < 0.001

Other STV Elections

Compared to the last 12 Irish general elections, with a total surname A-J/K-Z Relative Electability Ratio of 1.26 (P<0.001), the last five European elections in the Irish Republic (1979-99: Total S/C = 75/234) have yielded a higher but statistically insignificant surname A-J/K-Z Relative Electability Ratio of 1.37 (P > 0.05) [5] .

On the other hand, the last five Irish Local Elections (1979–99: Total S/C = 4,918/10,250) disclosed a lower surname A-J/K-Z Relative Electability Ratio of 1.12 (P<0.001)[3]. Perhaps better acquainted with local government candidates, voters discriminate more individually; numbering their preferences regardless of alphabetical order.

At the 1973 Assembly Election in Northern Ireland, Robson and Walsh (1974) attributed eight out of 78 Seats to positional voting bias. Yet at the 1998 Northern Ireland Assembly Election (Total S/C = 108/296), the surname A-J/K-Z Relative Electability Ratio fell below unity (0.87: P>0.05)[4]. Certainly, parties are more sharply differentiated in Northern Ireland than in the Irish Republic.

Discussion

Using forenames as controls, surname-alphabetic electability valuably measures voters' lack of discrimination between candidates within parties. Neither voters nor the Irish electoral system (STV) can be reproached for any positional voting bias.

However, Dáil Éireann remains surname-alphabetocratic, over-representing candidates with A-C surnames, while under-representing excumbents (non-incumbents) with K-M surnames (Table A: compare Table E): especially in constituencies with over three seats (Table B), and/or over eight candidates (Table C).

Perhaps aware of Robson and Walsh (1974), Ireland's two main parties (Fianna Fáil and Fine Gael) have apparently over-nominated A-C excumbents (notably since 1977: Table D). However, thus acting on the belief of increased electability may itself increase A-C surname over-representation: aggregating the last 12 Irish general elections (1961-97), excumbent S%/C% for FF + FG (1.67) was considerably higher than for other parties (0.43).

Reassuringly, aggregating all 12 general elections (1961-1997), the excumbent Surname Relative Electability Ratio (S/C ratio: A-J/K-Z=1.34 overall) proved significantly higher for FF+FG (1.28: P<0.05) than for the other candidates (1.03: P>0.05). however, it remains unclear whether the two main parties have benefited from A-C over-nomination.

Darcy and McAllister (1990) found "no evidence for position advantage for political parties in any election". Their review concluded that positional voting bias may be eliminated by removing its causes: notably, compulsory voting; completion of all preferences; and ballot-forms not indicating candidates' Party affiliation (as in Ireland before 1965[7]).

On the strength of the 1973 Irish General Election, Robson and Walsh (1974) advocated randomising the order of candidates on ballot-forms. Citing Robson and Walsh (1974), Sinnott[8] suggested that the problem could "easily be eliminated by arranging the names in a number of different randomised orders on different sets of ballot papers".

At the Dublin High Court in 1986, Mr Justice Murphy accepted that candidates with surname initials high in the alphabet were over-represented but, noting that alphabetic order helped voters to find candidates, he found it constitutional[9]. Indeed, the voter's predicament is paramount; and to avoid the palpable frustrations of randomised ballot-forms in locating preferred candidates, a reasonable compromise might be to print half the ballot-forms in surname-alphabetic order, with the other half in the reverse order — if positional voting bias really matters.

Acknowledgement

I am indebted to David Hill for statistical advice. Statistical significance was calculated by combining election-specific, one-tailed exact two-by-two table probabilities[6].

References

1. Dáil Éireann (1962-98): *Election Results and Transfer of Votes in General Election*: October 1961 — June 1997. Stationery Office, Dublin.

2. Darcy R and McAllister I (1990): Ballot Position Effects: *Electoral Studies* 9, 5-17.

3. Department of the Environment and Local Government (1980-2000): *Local Elections*, 1979-99. Stationery Office, Dublin.

4. Chief Electoral Officer for Northern Ireland (1998): *The New Northern Ireland Assembly Election 25 June 1998*. Belfast.

5. European Parliament Office in Ireland/Department of the Environment and Local Government (1980-2000): *European Parliament Election Results*, Ireland, 1979-99.

6. Fisher R A (1970): *Statistical Methods for Research Workers*, 14th Edition. Hafner Publishing Company, New York (pages 99-101).

7. Robson C and Walsh B (1974): The Importance of Positional Voting Bias in the Irish General Election of 1973: *Political Studies* 10, 191-203.

8. Sinnott R (1999): The Electoral System: in Coakley J and Gallagher M eds: *Politics in the Republic of Ireland* (Third Edition). Routledge, London (page 124).

9. Trench B *et al* (1987): The Alphabet Advantage: in *Magill Book of Irish Politics: Election February 87* (page 23).

Nonmonotonicity in AV

E Stensholt

Eivind Stensholt is from the Norwegian School of Economics and Business Administration

Introduction

Nonmonotonicity arises with STV when apparent additional support for a candidate, A, at the expense of another candidate, C, causes a third candidate, B, to be elected. Without the additional support, A would be elected. Thus the additional support actually costs A the election. This unfortunate property in the standard variations of STV is linked to the elimination of candidates in the counting process[6], and it is unavoidable unless some compromise is made with the principle that a voter's later preferences cannot influence the fate of the voter's earlier preferences.

How frequently will it happen that a candidate is not elected, but might have been elected if some of his or her support had gone to another candidate instead? That depends on the voters' behaviour. Based on standard assumptions on the distribution of voter preference, modified by empirical evidence of voter behaviour, the frequency is estimated for elections in which 1 candidate is elected from 3. This is the Alternative Vote (AV), a single-seat version of STV.

It is also shown how the nonmonotonicity is related to the Condorcet paradox in which one majority prefers B to A, another majority prefers A to C, and a third majority prefers C to B. In all elections considered, each voter is assumed to give a complete preference list.

For example, consider an election (from a simulation with 10000 voters) with

475	ABC
3719	ACB
390	CAB
2110	CBA
41	BCA
3265	BAC

No candidate has 50% of the first preference votes. C, with only 2500 first preference votes is eliminated, and finally B defeats A with 5416 votes to 4584. However, if x of the ACB-voters vote "strategically" CAB instead, the election may turn out differently. Then the profile is

475	ABC
3719−x	ACB
390+x	CAB
2110	CBA
41	BCA
3265	BAC

If $x > 806$, C with 2500+x overtakes B with 3306, and if $x<888$, A is still ahead of B with 4194−x to 3306. Thus, with $806 < x < 888$, B gets eliminated, and finally A defeats C with 7459−x votes to 2541+x.

The example also shows the Condorcet paradox of cyclic majorities. In pair-wise encounters A defeats C with 7459−x to 2541+x, C defeats B with 6219 votes to 3781, and B defeats A with 5416 votes to 4584. However, in real elections with 3 candidates cyclic majorities become very rare as the number of voters increases. One indicator of unrealism is that the cyclic order ABCA receives only 475+390+41+x = 906+x votes while ACBA receives 3719+2110+3265−x = 9094−x votes. In real elections the votes are distributed in the 6 categories in a more harmonious way.

If nonmonotonicity occurs in a real election, the scenario is most likely that there is a plurality winner, A (with the largest number of first preference votes), another Condorcet winner, B (who defeats each other candidate in pair-wise encounters), and a third candidate, C (who is last in first preference votes). Such an example, from the same simulation, is

2996	ABC
1122	ACB
875	CAB
2046	CBA
1431	BCA
1530	BAC

Here C is eliminated and B wins the AV-election. If x voters switch from ACB to CAB, and $40 < x < 648$, then B is eliminated and A wins. It turns out that if AV is modified and A declared winner in the few cases like this, nonmonotonicity is eliminated. Instead, however, another principle will be violated: B may win by a suitable vote transfer from BAC to BCA.

3-candidate elections may be classified according to how well the "electoral cake model" in Stensholt[5] may be fitted; the figure on page 7 shows a good fit. The model may be fitted quite well to most real elections. When simulated elections are classified, election P is considered more "realistic" than election Q if the model fits P better than Q. When better fit, i.e. more "realism", is demanded, the frequency of the

Condorcet paradox will approach 0. Nonmonotonicity, however, occurs in about 0.90% of all simulated "realistic" elections. Two real elections (37 candidates, 63 voters) and (14 candidates, 115 voters) have been checked, with nonmonotonicity in, respectively, 0.66% and 1.10% of the candidate triples.

A description of nonmonotonicity by means of inequalities

A possible preference distribution P in an election with 3 candidates, A, B, and C (a *profile* in the social choice vernacular), consists of a sequence of 6 non-negative numbers.

$P = (p\ \ q\ \ r\ \ s\ \ t\ \ u)$,

These are the numbers (absolute or relative) of voters with preference ranking respectively: ABC, ACB, CAB, CBA, BCA, BAC.

If x of the ABC-voters and y of the ACB-voters change to vote CAB, there is a new profile Q:

$Q = (p-x\ \ q-y\ \ r+x+y\ \ s\ \ t\ \ u)$.

Nonmonotonicity occurs if B is AV-winner in P and A in Q despite the natural expectation that the candidate A is weaker in Q than in P. The story is told in 9 inequalities.

$$r+s+t+u > p+q \qquad (1)$$
$$p+q+r+s > t+u \qquad (2)$$
$$p+q > r+s \qquad (3)$$
$$t+u > r+s \qquad (4)$$
$$s+t+u > p+q+r \qquad (5)$$
$$p+q+t+u-(x+y) > r+s+(x+y) \qquad (6)$$
$$r+s + (x+y) > t+u \qquad (7)$$
$$p+q-(x+y) > t+u \qquad (8)$$
$$u+p+q-(x+y) > r+s+t + (x+y) \qquad (9)$$

A translation to non-mathematical language links the inequalities to the AV rules. (1, 2): In P, neither A nor B have 50% of the first preference votes. (3, 4): In P, C has the lowest number of first preference votes. (5): In P, B wins over A (after elimination of C). (6): In Q, C does not reach 50% first preference votes. (7): In Q, C passes B in first preference votes. (8): In Q, A keeps more first preference votes than B. (9): In Q, A wins over C (after elimination of B).

However, the mathematical version (1-9) is easier to analyse. Write (7, 8, 9) equivalently as

$$\min[p+q-t-u, (u+p+q-r-s-t)/2] > x+y > t+u-r-s \qquad (10)$$

Thus numbers x and y satisfying (7, 8, 9) exist if and only if (11) and (12) hold:

$$p+q+r+s > 2t+2u \qquad (11)$$
$$p+q+r+s > 3t+u \qquad (12)$$

Moreover, (1), (2), (3), and (6) are redundant because of (5), (11), (4 and 8), and (9), respectively. Therefore the $p+q$ supporters of candidate A can turn defeat in P to victory in Q if and only if (4, 5, 11, 12) all hold. Then $x+y$ of them vote "strategically" CAB, with $x+y$ as in (10).

A profile where a candidate may be helped by being ranked lower in some ballots without any other change in any ballot will be called a nonmonotonic profile for the election method considered. In discussing various election rules, it is also useful to have an "absolute" definition: A profile is then nonmonotonic if it is so for AV. A monotonic election method is one without nonmonotonic profiles. AV, and the usual STV-variations are nonmonotonic because of the elimination rules. By the criteria (4, 5, 11, 12)

$$p+q > t+u > r+s \qquad (13)$$

Thus, in P, A is plurality winner (first past the post), while B beats A and A beats C in pair-wise comparisons by (5) and (9). This we will call nonmonotonicity of type ABC. There are six types of nonmonotonic profiles: ABC, ACB, CAB, CBA, BCA, and BAC.

Connection to the Condorcet paradox; a geometric description

The Condorcet paradox occurs together with ABC-type nonmonotonicity when also C beats B in pair-wise comparison, i.e.

$$q+r+s > t+u+p \qquad (14)$$

Otherwise B is the Condorcet winner, i.e. B defeats each opponent in a pair-wise contest. The strategic voting of the $x+y$ voters who honestly support A is then designed to take the AV victory away from Condorcet winner B to plurality winner A. Define E, F, G, H, K as functions of the profile:

$$E=-r-s+t+u$$
$$F=-p-q-r+s+t+u$$
$$G= p+q+r+s-2t-2u \qquad (15)$$
$$H= p+q+r+s-3t-u$$
$$K=-p+q+r+s-t-u$$

When all possible profiles are standardized, e.g. to $p+q+r+s+t+u$=12, as in the table below, they form a 5-dimensional simplex with 6 corners — a higher dimensional analogue of the familiar 3-dimensional simplex (tetrahedron) with 4 corners and 4 triangular sides.

Profile	*p*	*q*	*r*	*s*	*t*	*u*	E	F	G	H	K	100ε
P01	4	0	2	2	2	2	0	0	0	0	–4	5.61
P02	4	0	2	2	0	4	0	0	0	4	–4	0.00
P03	4	0	0	4	2	2	0	4	0	0	–4	0.00
P04	4	0	0	4	0	4	0	4	0	4	–4	0.00
P05	6	0	0	3	3	0	0	0	3	0	–6	0.00
P06	6	0	0	3	0	3	0	0	3	6	–6	0.00
P07	6	0	0	2	2	2	2	0	0	0	–8	0.00
P08	6	0	0	2	0	4	2	0	0	4	–8	0.00
P09	0	4	2	2	2	2	0	0	0	0	4	5.61
P10	0	4	2	2	0	4	0	0	0	4	4	20.69
P11	0	4	0	4	2	2	0	4	0	0	4	20.69
P12	0	4	0	4	0	4	0	4	0	4	4	41.35
P13	0	6	0	3	3	0	0	0	3	0	6	32.54
P14	0	6	0	3	0	3	0	0	3	6	6	39.77
P15	0	6	0	2	2	2	2	0	0	0	4	17.27
P16	0	6	0	2	0	4	2	0	0	4	4	38.72

By (4, 5, 11, 12) the nonmonotonic profiles of ABC-type form a convex subset S of this simplex, given by

$$E > 0,\ F > 0,\ G > 0,\ H > 0 \qquad (16)$$

The Condorcet paradox occurs if K > 0 too. The profiles in the table are the corners of the closure of S and have non-negative E, F, G, H.

In the right hand column, $\varepsilon = \varepsilon(P)$ is a continuous function of the profile P, defined in Stensholt[5]. By its definition, $0 < \varepsilon < 3\sqrt{3}/4\pi \approx 0.4135$. Generally ε is well below 0.01 in profiles from real elections with many voters. Any profile P satisfying (16) may be written as

$$P = k0l{\cdot}P01 + k02{\cdot}P02 + k03{\cdot}P03 + ... + kl6{\cdot}Pl6 \qquad (17)$$

with non-negative k_j and k01 + k02 + k03 +... + kl6 = 1.

To a profile P = (*p q r s t u*) we may assign a twin profile P* = (*q p r s t u*). Thus P**=P and $P_i^* = P_{i+8}$, i=1, 2,.., 8. If P is a nonmonotonicity profile of type ABC, so is P*. With P as in (17), then

$$P^*=k09{\cdot}P01+k10{\cdot}P02+..+k16{\cdot}P08+k01{\cdot}P09+k02{\cdot}P10+..+k08{\cdot}P16, \quad (18)$$

$$K(0.5{\cdot}[P + P^*]) = 0.5{\cdot}[K(P) + K(P^*)] = -2{\cdot}(k07 + k08 + kl5 + kl6) \le 0 \quad (19)$$

Thus the profile 0.5 [P + P*], midway between P and P*, will never give the Condorcet paradox, but it is on the borderline if and only if k07 = k08 = kl5 = kl6 = 0 Somewhere between 1/3 and 2/3 along the line segment from P to P*, K = 0. From the K-column in Table 1 it is clear that, with many voters, somewhere between 33% and 50% of all nonmonotonicity profiles also have a Condorcet cycle. However, they are not all equally likely to occur in real elections.

Simulation and reality

One million random 3-candidate profiles were generated with uniform probability in the simplex. The distribution is known as the *Impartial Anonymous Culture* (IAC). The IAC also depends on the number of voters, but the simulation corresponds to the limit case of infinitely many voters. Actually about 100 voters would give quite similar results.

In 3621 of the IAC-generated profiles were E>0, F>0, G>0, H>0. As there are six nonmonotonicity types, about 6×0.3621% ≈ 2.17% of the profiles are nonmonotonic. Among these 3621, 1602, i.e. ≈ 44.24% also had K>0, indicating a Condorcet cycle in the profile: A beats C beats B beats A. For comparison, 6.25% of all IAC-profiles have a Condorcet cycle [2,5].

In real elections the cycle frequency is much lower. That is due to a structure in the profiles, which may come from the voters having some common perception of the "political landscape" although they have placed themselves in different positions and rank the candidates accordingly [5].

Imagine that the voters are distributed with uniform density in a circular disc, that candidates A, B and C are among them, and that a voter ranks the candidates according to their distance from the voter's position. In a pair-wise comparison between A and B, B wins if and only if B is closer than A to the circle centre. A and B divide the voters between them with the mid-normal to the line segment AB as dividing line. Similarly the mid-normals for BC and AC divide the disc. The three candidates split the "voter cake" in six pieces by three straight cuts through one common point, each piece getting an area proportional to the number of votes with the corresponding ranking of the candidates. In a model like this, the Condorcet paradox can never occur except in a degenerate form with all cuts through the circle center, and *p*=*s*, *q*=*t*, *r*=*u*.

Empirically, the electoral cake model fits reasonably well for 3-candidate profiles from real elections with a large number of voters. That is why the Condorcet paradox is rare. The function ε(P) measures the deviation of P from the model. For the examples in the introduction,

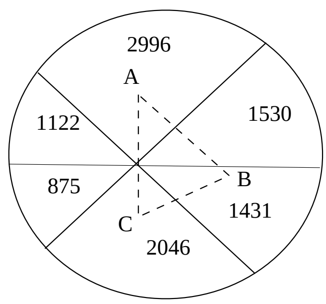


Figure giving the profile (2996 1122 0875 2046 1431 1530) which fits well with the "electoral cake" model.

ε(0475, 3719, 0390, 2110, 0041, 3265) = .174035768

ε(2996, 1122, 0875, 2046, 1431, 1530) = .000000108
(see figure).

Among the simulation profiles with small $\varepsilon(P)$, about 0.15% were nonmonotonic of ABC-type. This suggests an estimate of 0.90% for the probability for nonmonotonicity in a candidate triple in real elections with many voters.

In an election with 63 voters and 37 candidates at the author's institution, 51 of the 37×36×35/6 = 7770 triples were nonmonotonic, a fraction of 0.66%. In these 51 triples, the Condorcet paradox occurred only 7 times, i.e. much less than the 44.24% in the full IAC-simulation. In another election in the same place, with 115 voters and 14 candidates there were 4 nonmonotonicity triples out of 14×13×12/6 = 364, i.e. 1.10% and the Condorcet paradox occurred in none of them. Comparison with the simulation requires some caution since the triple profiles in an election with many candidates cannot be assumed stochastically independent.

Conclusion

In an election with 3 candidates, A, B and C, let A be plurality winner. In the vast majority of elections, there will also be a unique Condorcet winner. If A also happens to be Condorcet winner, A wins the AV-election. That cannot be very controversial.

So assume B is Condorcet winner, which means that B wins if A or C is eliminated. B may win with very few first preference votes in the ballots, but electing B means that there are no "wasted votes". The "plurality ideology" may also be modified to avoid wasting votes by eliminating B; then the supporters of B are allowed to influence the choice between A and C. An election method that *always* eliminates a Condorcet winner who is not also a plurality winner, may seem strange. However, it would, arguably, be a democratic improvement of the plurality method that is in wide use today. It preserves the "plurality ideology" as well as possible, preferring to let centre voters decide between "right" and "left" rather than filling an assembly with centre politicians.

AV can be seen as a compromise between the "plurality ideology" and the "Condorcet ideology". There are two possibilities.

(I) If B has the smallest support in terms of first preference votes, i.e. $p+q > r+s > t+u$, then B is eliminated.

(II) If B is number 2 in terms of first preference votes, i.e. $p+q > t+u > r+s$, then B is the AV-winner.

Nonmonotonicity occurs in (II) if A has a number of surplus first preference votes that could be transferred to C in a way that benefits A. Such transfer is not a part of AV, but this can be remedied in the spirit of STV if the transfer rule is extended. When (16) holds, let the necessary number of surplus votes be transferred from voter categories ABC and ACB to CAB if this lets C become number 2 in terms of first preference votes, and still lets A win against C after elimination of B. This transfer of first preference votes from A to C involves only voters who prefer A to B (categories ABC, ACB, CAB), and it may be implemented in the counting process when it helps A to win instead of B.

An obvious argument against such a procedure is that it occasionally may violate the cherished principle that my second preference should never hurt my first preference. To see this, consider first standard AV. Then C is eliminated after examination of first preferences only. The second preferences of C's supporters become available, and either A (plurality winner) wins or B (Condorcet winner if one exists) wins. Among the conditions in (16) for an extra transfer of votes from A to C, the three first only involve first preference votes: $p+q$, $r+s$, $t+u$. The inequality $H > 0$ requires information about t and u, i.e. about the second preferences of B's supporters. This allows for strategic voting on behalf of B. Let z voters move from BAC to BCA. Then according to (15) the requirement $H > 0$ is sharpened to

$$p+q+r+s-3t-u-2z>0.$$

The strategy is to break this condition, which is achieved if and only if

$$p+q+r+s-3t-u \le 2z \le 2u$$

Such strategy is possible if and only if

$$p+q+r+s+t+u \le 4(t+u),$$

i.e. if and only if B has at least 25% of the first preference votes. This will, however, always be the case when the extra transfer rule is invoked, because by (1) A has less than 50% of the first preference votes and by (4) B has more first preference votes than C. AV with extra transfer violates the principle exactly when standard AV violates monotonicity.

In 3-candidate elections, voters may be offered one of two guarantees:

1) You can never hurt a candidate by an upwards move;

2) You can never hurt a candidate by a change in the subsequent ranking.

In about 99% of the elections, the profile is monotonic. Then AV and AV with extra transfer satisfy both 1) and 2), as no extra transfer is done. In the remaining cases, standard

AV picks the Condorcet winner and violates 1) but not 2), while AV with extra transfer picks the plurality winner and violates 2) but not 1). Which of the two guarantees is then most important?

With more candidates, it becomes more complicated to study nonmonotonicity in AV. With 5 candidates, A, B, C, D, and E, there are 10 triples, and each candidate takes part in 6 triples:

{A,B,C}, {A,B,D}, {A,B,E}, {A,C,D}, {A,C,E}, {A,D,E}, {B,C,D}, {B,C,E}, {B,D,E}, {C,D,E}.

After all but 3 candidates are eliminated, there is a final triple, say {A,B,C}. If AV is adopted in more than 600 constituencies, as in a Westminster election, there will generally be some with nonmonotonicity in {A,B,C}. How bad will criticism from frustrated supporters of a non-elected plurality winner in such cases be for people's trust in standard AV?

If A, B and C are much stronger than all other candidates, it may be enough to implement the extra vote transfer in {A,B,C} in order to cope with most nonmonotonic profiles. Nonmonotonicity is reduced, at a price: How bad will criticism from frustrated supporters of a non-elected Condorcet winner in such cases be for people's trust in AV with extra transfer?

The purpose of elimination is to find the opponent for A in the final pair, so B or C must be eliminated. The extended transfer rule only adjusts the border between elimination of B and elimination of C. Is an election of B due to honest first priority from A's supporters more tolerable than election of A due to honest subsequent ranking from B's supporters?

Can we achieve monotonicity with more than 3 candidates, at a reasonable price? Perhaps a recursive idea may work. Assume that the set of profiles S with *n* candidates has been subdivided into *n* subsets $S = S_1 \cup S_2 \cup ... \cup S_n$, so that candidate *i* wins with profile in S_i and that this election method is monotonic. With *n*+1 candidates left, eliminate Z with the lowest number of first preference votes. If that leads to a profile in S_Y and $X \neq Y$, then allow an extra transfer of first votes from X to Z or even to more candidates in order to eliminate another candidate and obtain an *n*-candidate profile *i* S_x. The possibility of saving more candidates than Z from elimination by an extra transfer raises the question of whether X is uniquely defined.

A more radical measure is to count in each triple separately, implementing the extra transfer. "Triple-AV" then gives a candidate one point for a triple victory, and achieves monotonicity. It is similar to Copeland's method[1,3,4], which gives one point for each victory in a pair-wise comparison and avoids Condorcet cycles. On the other hand, the price for monotonicity with triple-AV may well be too high in terms of violations of the principle.

An axiomatic study of election theory reveals some basic impossibilities. Certain combinations of nice properties cannot be realized simultaneously in one election method. To achieve monotonicity, one must sacrifice the principle. On the other hand, only in the few cases where (16) holds, will triple-AV find another triple winner than standard AV.

Three papers in *Voting matters*[6,7,8] deal with nonmonotonicity and related problems. One theme is the axiomatic understanding of election methods: which combinations of desirable properties are theoretically incompatible? That kind of knowledge is important for everyone concerned with "how to choose how to choose". An axiomatic approach, however, needs a clearly formulated and manageable conceptual frame. As part of this frame, it must be clearly stated what kind of preference relations the voters are allowed to express. One may restrict ballots to be complete, or to conform to a linear listing of the alternatives (single-peak condition), etc. Within this frame the axiomatic investigator must take into account all possible profiles without any extra screening against unrealistic profiles. Even a highly concocted profile may be a counter-example that kills a hypothesis; lack of realism is no objection if the profile formally is within the axiomatic frame. According to Stensholt[5] a bound on the function $\varepsilon(P)$ of the 3-candidate profile P is useful to screen off most of the unrealistic profiles generated in a simulation. However, a criterion like $\varepsilon(P) < 0.01$ does not seem suitable for axiomatic treatment. Axiomatics must be followed up by other approaches, e.g. comparisons of election methods on simulated and empirical data.

Acknowledgements

It is a pleasure to recognize the many exchanges with the editor and the comments of an anonymous referee. The latter led to a complete rewriting of the conclusion section, where however, there still are viewpoints for which the author dares not presume full agreement.

References

1. A.H. Copeland, A "reasonable" social welfare function, Seminar on Mathematics in Social Sciences, University of Michigan, 1951.

2. W. Gehrlein and P.C. Fishburn, Probabilities for election outcomes for large electorates, *J. Econom. Theory,* 19, (1978) pp. 38-49

3. V.R. Merlin, and D.G. Saari, "Copeland Method. II. Manipulation, Monotonicity, and Paradoxes"; *Journal of Economic Theory;* Vol. 72, No. 1; January, 1997; 148-172.

4. D.G. Saari. and V.R. Merlin, "The Copeland Method. I. Relationships and the Dictionary"; *Economic Theory;* Vol. 8, No. l; June, 1996; 51-76.

5. E. Stensholt(1996) Circle pictograms for vote vectors, SIAM Review Vol 38 No 1, pp.96-119.

6. D.R. Woodall (1994) "Properties of Preferential Election Rules", *Voting matters*, Issue 3, 8-15;

7. D.R. Woodall (1995) "Monotonicity — An In-Depth Study of One Example", *Voting matters*, Issue 4, 5-7;

8. D.R. Woodall (1996) "Monotonicity and Single-Seat Election Rules", *Voting matters*, Issue 6, 9-14.

On Dummett's "Quota Borda System"

M Schulze

Markus Schulze is a physicist and mathematician from Germany

In two books[1,2], in his submission to the Jenkins Commission[3], and at a number of conferences, Michael Dummett has promoted a preferential voting method where one successively searches for solid coalitions of increasing numbers of candidates and where, when one has found such a solid coalition, one declares the candidates with the best Borda scores elected. Dummett calls his method "Quota Preference Score" (QPS) or "Quota Borda System" (QBS). He writes that his method "has never been in use, but was voted the best at a conference on electoral reform held in Belfast with representatives of all parties" [3]. In his book *Voting Procedures*, he describes this method as follows (where *v* is the number of voters, *S* is the number of seats, **C** is the number of candidates, and the "preference score" is the Borda score) [1, pp. 284-286]:

The assessment will proceed by stages, all but the last of which may be called "qualifying stages": it will of course terminate as soon as all *S* seats have been filled. We may first describe the assessment process for the case when *S* is 2 or 3. At stage 1, the tellers will determine whether there are any candidates listed first by more than 1/(*S*+1) of the total number *v* of voters: if so, they immediately qualify for election. If seats remain to be filled, the preference scores of all candidates not qualifying at stage 1 will then be calculated. At stage 2, the ballot papers will be scrutinized to see if there is any pair of candidates, neither of whom qualified at stage 1, to whom more than *v*/(*S*+1) voters are solidly committed: if so, that member of the pair with the higher preference score now qualifies for election. If seats remain to be filled, the tellers will proceed to stage 3, at which they will consider sets of three candidates, none of whom has already qualified. If more than *v*/(*S*+1) voters are solidly committed to any such trio, that one with the highest preference score qualifies for election. In general, at the qualifying stage *i*, the tellers determine whether, for any set of *i* candidates none of whom has so far qualified, there are more than *v*/(*S*+1) voters solidly committed to those candidates; if so, the member of the set with the highest preference score qualifies for election at stage *i*. If there still remain seats to be filled after all the qualifying stages have been completed, they will be filled at the final stage by those candidates having the highest preference scores out of those who have not yet qualified. (. . .)

When *S* = 4, however, it may be thought that a body of voters, amounting to more than two-fifths of the electorate and solidly committed to two or more candidates, is entitled to 2 of the 4 seats. To achieve this, the assessment process must be made a little more complex. Stage 1 will proceed as before, and, at stage 2, the same operation must be carried out as described above. Before proceeding to stage 3, however, the tellers must also consider every pair of candidates of whom one qualified at stage 1 and the other did not: if more than 2·*v*/(*S*+1) voters are solidly committed to such a pair, that one who did not qualify at stage 1 qualifies at stage 2. (Note that, if more than 2·*v*/(*S*+1) voters are solidly committed to two candidates, one of them must qualify at stage 1.) Likewise, at each qualifying stage *i*, the tellers must ask, of every set of *i* candidates of whom at most one has already qualified, whether more than 2·*v*/(*S*+1) voters are solidly committed to those candidates. If so, and none of them has previously qualified, the two with the highest preference scores will now qualify; if one of them qualified at an earlier stage, that one, of the rest, who has the highest preference score will qualify at stage *i*. (. . .)

In general, at stage *i*, the tellers must ask, of each set of voters solidly committed to i candidates, what multiple of *v*/(*S*+1) members it contains, up to *i*·*v*/(*S*+1). If it contains more than *v*/(*S*+1) voters, at least one of the *i* candidates will qualify for election; if it contains more than 2·*v*/(*S*+1), at least two will qualify; if 3 ≤ *i* and it contains more than 3·*v*/(*S*+1), at least three will; and so on, up to the case in which it contains more than *i*·*v*/(*S*+1) voters, when all *i* candidates will qualify.

This description of QBS seems unnecessarily long. Usually, Dummett offers a significantly shorter description. For example, in his submission to the Jenkins Commission he writes[3]:

The scruntineers can first mark as elected any candidate ranked highest by a sufficiently large minority (one-sixth of the voters in a five-member constituency, etc.). Then, having calculated the Borda counts of all remaining candidates, they can discover whether any set of from two to five candidates receives solid support from a sufficiently large minority: if so, that candidate in the set with the highest Borda count is marked as to be elected. The remaining seats will be filled by the candidates most generally acceptable to the electorate as a whole, i.e. those with the highest Borda counts.

In my opinion, a problem of the shorter description is that readers could mistakenly believe that the order in which the solid coalitions are considered at each stage and the question at which stages the different candidates have qualified were unimportant. However, example 1 demonstrates that they are decisive.

Example 1 (v = 100; S = 2; $\mathcal{C}$ = 5):

```
29 DBCEA.
17 ABDCE.
17 BADCE.
17 CADBE.
13 ACDBE.
 7 CABDE.
```

The Borda scores are 243 for candidate A, 250 for candidate B, 227 for candidate C, 251 for candidate D, and 29 for candidate E. Table 1 lists all solid coalitions. At stage 1, no candidate qualifies for election. At stage 2, it is observed that more than $v/(S+1)$ voters are solidly committed to the candidates A and B and that more than $v/(S+1)$ voters are solidly committed to the candidates A and C. When one uses only the short description of QBS, then one could mistakenly believe that there are two different possibilities how to proceed resulting in two different sets of winners. **First**: When one starts with the set A and B, candidate B qualifies for election because he has a better Borda score than candidate A. Then one has to consider the set A and C; candidate A qualifies for election because he has a better Borda score than candidate C. As no seats remain to be filled, QBS terminates and the candidates A and B are the winners. **Second**: When one starts with the set A and C, candidate A qualifies for election because he has a better Borda score than candidate C. Then one has to consider the set A and B; however, as this set has already won one seat no additional candidate qualifies at stage 2. At stage 3, one observes that

Table 1: Solid Coalitions in Example 1

one candidate		two candidates		three candidates		four candidates	
candidate	No.	candidates	No.	candidates	No.	candidates	No.
A	30	A,B	34	A,B,C	7	A,B,C,D	71
B	17	A,C	37	A,B,D	34	A,B,C,E	
C	24	A,D		A,B,E		A,B,D,E	
D	29	A,E		A,C,D	30	A,C,D,E	
E		B,C		A,C,E		B,C,D,E	29
		B,D	29	A,D,E			
		B,E		B,C,D	29		
		C,D		B,C,E			
		C,E		B,D,E			
		D,E		C,D,E			
	100		100		100		100

more than $v/(S+1)$ voters are solidly committed to the candidates A, B and D; however, as this set has already won one seat no additional candidate qualifies at stage 3. At stage 4, one observes that more than $2 \cdot v/(S+1)$ voters are solidly committed to the candidates A, B, C and D; as candidate D has the best Borda score candidate D qualifies for election. As no seats remain to be filled, QBS terminates and the candidates A and D are the winners.

However, the long description in "Voting Procedures" states clearly that when one has to decide how many additional seats a given solid coalition gets at a given stage then one has to consider as already qualified only those candidates who have already qualified at strictly earlier stages. In example 1, when one starts with the set A and C, candidate A qualifies for election because he has a better Borda score than candidate C. Then one has to consider the set A and B; as none of these candidates has already qualified at a strictly earlier stage, candidate B qualifies for election because he has a better Borda score than candidate A.

In short, to guarantee that the result doesn't depend on the order in which the solid coalitions are considered at a given stage, it is important that one looks only at those candidates who have qualified at *strictly* earlier stages. For example, suppose, at stage 10, one finds a set of 10 candidates such that more than $5 \cdot v/(S+1)$ voters, but not more than $6 \cdot v/(S+1)$ voters, are solidly committed to these 10 candidates. Suppose that already 4 of these 10 candidates have qualified at stages 1-9. Then that candidate of this set who has the best Borda score of all those candidates of this set who did not qualify at stages 1-9 qualifies at stage 10 *even if this set has already won additional seats at stage 10*.

At first sight, it isn't clear whether the QBS winners can be calculated in a polynomial runtime since there are $2^{\mathcal{C}}$ possible sets of candidates. However, a set of candidates has to be taken into consideration only when at least one voter is committed to this set. In so far as at each of the $\mathcal{C}$ stages there cannot be more than v sets of candidates such that at least one voter is committed to this set, one has to take not more than $v \cdot \mathcal{C}$ sets of candidates into consideration to calculate the QBS winners. Therefore, a polynomial runtime is guaranteed.

When not each voter ranks all candidates, then Dummett's intention is met best when in each stage i those voters who don't strictly prefer all the candidates of some set of i candidates to every other candidate are allocated to no solid coalition.

Nicolaus Tideman writes about QBS [4]:

To avoid sequential eliminations, Michael Dummett suggested a procedure in which a search would be made for solid coalitions of a size that deserved representation, and when such a coalition was found, an option (or options) that the coalition supported would be selected. If the solid coalition supported more than one option, the option (or options) with the greatest "preference score" (Borda count) would be selected. Preference scores would also be used to determine which options would fill any positions not filled by options supported by solid coalitions. I find Dummett's suggestion unsatisfying. Suppose there are voters who would be members of a solid coalition except that they included an "extraneous" option, which is quickly eliminated, among their top choices. These voters' nearly solid support for the coalition counts for nothing, which seems to me inappropriate.

At first sight, it isn't clear whether Tideman's criticism is feasible. It is imaginable that whenever there are "voters who would be members of a solid coalition except that they included an 'extraneous' option" there is also an STV method (i.e. a method where surpluses of elected candidates are transferred according to certain criteria to the next available preference and where, when seats remain to be filled, candidates are eliminated according to certain criteria and their votes are transferred to the next available preference) where this "nearly solid support for the coalition counts for nothing". If this is the case, then it is not appropriate to criticize QBS for ignoring this "nearly solid support". However, example 2 demonstrates that there are really situations where the QBS winners differ from the STV winners independently of the STV method used.

Example 2 (v = 100; S = 3; $\mathbf{C}$ = 5):

```
40 ACDBE.
39 BCDAE.
11 DABEC.
10 DBAEC.
```

The Borda scores are 252 for candidate A, 248 for candidate B, 237 for candidate C, 242 for candidate D, and 21 for candidate E. Table 2 lists all solid coalitions. At stage 1, the candidates A and B qualify for election because both candidates are preferred to every other candidate by more than $v/(S+1)$ voters each. At stage 2, it is observed that more than $v/(S+1)$ voters are solidly committed to the candidates A and C and that more than $v/(S+1)$ voters are solidly committed to the candidates B and C; but as both sets of candidates have already won one seat each, no additional candidate qualifies for election at stage 2. At stage 3, it is observed that more than $v/(S+1)$ voters are solidly committed to the candidates A, C, and D and that more than $v/(S+1)$ voters are solidly committed to the candidates B, C, and D; but as both sets of candidates have already won one seat each, no additional candidate qualifies for election at stage 3. At stage 4, it is observed that more than $3 \cdot v/(S+1)$ voters are solidly committed to the candidates A, B, C, and D; as this set has already won 2 seats, candidate D, the candidate with the best Borda score of all those candidates who haven't yet qualified, qualifies for election. As no seats remain to be filled, QBS terminates and the candidates A, B, and D are the winners. However, STV methods necessarily choose the candidates A, B, and C because, independently of how surpluses are transferred, candidate C always reaches the quota. In my opinion, example 2 questions whether compliance with proportionality for solid coalitions is sufficient for being a proportional preferential voting method.

Table 2: Solid Coalitions in Example 2

one candidate		two candidates		three candidates		four candidates	
candidate	No.	candidates	No.	candidates	No.	candidates	No.
A	40	A,B		A,B,C		A,B,C,D	79
B	39	A,C	40	A,B,D	21	A,B,C,E	
C		A,D	11	A,B,E		A,B,D,E	21
D	21	A,E		A,C,D	40	A,C,D,E	
E		B,C	39	A,C,E		B,C,D,E	
		B,D	10	A,D,E			
		B,E		B,C,D	39		
		C,D		B,C,E			
		C,E		B,D,E			
		D,E		C,D,E			
	100		100		100		100

Dummett's justification for his method is his claim that, unlike traditional STV methods, QBS is less "quasi-chaotic". He writes [3]:

The defect of STV is that it is quasi-chaotic, in the sense that a small change in the preferences of just a few voters can have a great effect on the final outcome. This is because it may affect which candidate is eliminated at an early stage, and thus which votes are redistributed, this then affecting all subsequent stages of the assessment process.

However, in my opinion, example 3 demonstrates that also QBS is "quasi-chaotic". This is because a small change in the preferences can affect which candidate qualifies at an early stage, this then affecting all subsequent stages of the assessment process.

Example 3 (v = 100; S = 2; $\mathbf{C}$ = 5):

```
26 BCAED.
24 DCEBA.
10 EADBC.
 8 ABCED.
 7 EABDC.
 7 EDBCA.
 6 CDEBA.
 6 DEBCA.
 3 DCEAB.
 2 EBADC.
 1 DCBEA.
```

The Borda scores are 142 for candidate A, 216 for candidate B, 215 for candidate C, 204 for candidate D, and 223 for candidate E. Table 3 lists all solid coalitions. At stage 1, candidate D qualifies for election because more than $v/(S+1)$ voters strictly prefer candidate D to every other candidate. At stage 2, it is observed that more than $v/(S+1)$ voters are solidly committed to the candidates C and D; but as this set of candidates has already won one seat, no additional candidate of this set qualifies for election at stage 2. At stage 3, it is observed that more than $v/(S+1)$ voters are solidly committed to the candidates A, B, and C; as none of these candidates has already qualified, candidate B, the candidate with the best Borda score, qualifies for

Table 3: Solid Coalitions in the original Example 3

one candidate		two candidates		three candidates		four candidates	
candidate	No.	candidates	No.	candidates	No.	candidates	No.
A	8	A,B	8	A,B,C	34	A,B,C,D	
B	26	A,C		A,B,D		A,B,C,E	34
C	6	A,D		A,B,E	9	A,B,D,E	19
D	34	A,E	17	A,C,D		A,C,D,E	3
E	26	B,C	26	A,C,E		B,C,D,E	44
		B,D		A,D,E	10		
		B,E	2	B,C,D	1		
		C,D	34	B,C,E			
		C,E		B,D,E	13		
		D,E	13	C,D,E	33		
	100		100		100		100

election. As no seats remain to be filled, QBS terminates and the candidates B and D are the winners.

When a single DEBCA ballot is changed to BDECA, the Borda scores are 142 for candidate A, 218 for candidate B, 215 for candidate C, 203 for candidate D, and 222 for candidate E. Table 4 lists all solid coalitions for this modified example. At stage 1, no candidate qualifies for election. At stage 2, it is observed that more than $v/(S+1)$ voters are solidly committed to the candidates C and D; as candidate C has a better Borda score, candidate C qualifies for election. At stage 3, it is observed that more than $v/(S+1)$ voters are solidly committed to the candidates A, B, and C; but as this set of candidates has already won one seat, no additional candidate of this set qualifies for election at stage 3. At stage 4, it is

Table 4: Solid Coalitions in the modified Example 3							
one candidate		two candidates		three candidates		four candidates	
candidate	No.	candidates	No.	candidates	No.	candidates	No.
A	8	A,B	8	A,B,C	34	A,B,C,D	
B	27	A,C		A,B,D		A,B,C,E	34
C	6	A,D		A,B,E	9	A,B,D,E	19
D	33	A,E	17	A,C,D		A,C,D,E	3
E	26	B,C	26	A,C,E		B,C,D,E	44
		B,D	1	A,D,E	10		
		B,E	2	B,C,D	1		
		C,D	34	B,C,E			
		C,E		B,D,E	13		
		D,E	12	C,D,E	33		
	100		100		100		100

observed that more than $v/(S+1)$ voters are solidly committed to the candidates A, B, C, and E and that more than $v/(S+1)$ voters are solidly committed to the candidates B, C, D, and E; but as both sets have already won one seat each, no additional candidates qualify for election at stage 4. At stage 5, candidate E qualifies for election because he has the best Borda score of all candidates who have not already qualified. Thus, by ranking candidate B higher candidate B is changed from a winner to a loser. By changing a single ballot the QBS winners are changed from the candidates B and D to the candidates C and E.

References

1. Michael Dummett, *Voting Procedures*, Clarendon Press, Oxford, 1984

2. Michael Dummett, *Principles of Electoral Reform*, Oxford University Press, 1997

3. Michael Dummett, Submission to the Independent Commission on Electoral Reform, 3rd July 1998

4. T. Nicolaus Tideman, *Collective Decisions and Voting* (draft), 1993

Sequential STV - a new version

I.D. Hill and Simon Gazeley

In Issue 2 of *Voting matters*, a system was reported called Sequential STV [1], designed to overcome, at least to some extent, the problem of premature exclusion of a candidate, which occurs when the one who has the fewest votes at the time is excluded, though due to receive many transfers later if only that exclusion had not taken place. That system has now been improved and we report here on the new version. One particular result of the improvement is that, in the case of a single seat, it is now certain to find the Condorcet winner if there is one.

The aim is to find a set of candidates of size n, where n is the number of seats to be filled, such that any set of n+1 candidates consisting of those n and 1 more, will result in the election of those n when an STV election is performed. When n=1 this reduces, of course, to the Condorcet rule. In a small election, or when n=1, it would be relatively easy and quick to do a complete analysis to find if there is such a set. The challenge is to find a way of doing so that will work in a reasonable time in large elections, where such a complete analysis would be impracticable. We recognise that the meanings of 'a reasonable time' and 'impracticable' are open to dispute, and that what is practicable will change as computers continue to get faster.

In the old version of Sequential STV, an initial STV count divided the candidates into probables and others, but the others were regarded as 'in a heap' and all of equal status. Consequently, if a challenger was successful, it would have been contrary to the axioms of anonymity and neutrality[2] to make a change of probables until all the others had been tested too, and that could lead to more than one challenger in the next main stage. In the new version the others are not put in a heap but in a queue, where the order depends upon the voting pattern. It is then fair to implement any change of probables at once, and the division of the method into main stages and sub-stages is no longer necessary.

How it works − the easy part

An initial STV count is made but instead of dividing into those elected and not elected, it classifies those who would have been elected as probables, and puts the others into a queue, in the reverse order of their exclusion in that initial count, except that the runner-up is moved to last place as it is already known that an initial challenge by that candidate will not succeed. Having found the probables and the order of the queue, further rounds each consist of n+1 candidates, the n

probables plus the head of the queue as challenger, for the *n* seats.

It should be noted that, apart from the initial count, which is only to get things started, all counts are of *n*+1 candidates for *n* seats, so the ‘exclude the lowest’ rule, which is the least satisfactory feature of STV, is not used.

If the challenger is not successful, the probables are unchanged for the next round and the challenger moves to the end of the queue, but a successful challenger at once becomes a probable, while the beaten candidate is put to the end of the queue. The queue therefore changes its order as time goes on but its order always depends upon the votes.

The reordering of the queue during the count, by putting any losing candidate to the end of the queue, is to make sure that it cannot ever get into a state where, say, a set X are probables, A, B and C are all near the top of the queue and X+A beats X+B beats X+C beats X+A, while D is further down and X+D has not been tested. Putting losing candidates to the end means that D must head the queue at some point before A, B and C come round again.

This continues until either we get a complete run through the queue without any challenger succeeding, in which case we have a solution of the type that we are seeking, or we fall into a Condorcet-style loop. In the latter case, we have to enter the more difficult part, set out below, but it should be emphasised that in real elections, as distinct from specially devised test cases, that rarely happens.

How it works − the more difficult part

To decide that a loop has been found, a set that has been seen before must recur as the probables. If the queue is in the same order as before then a loop is certain and action must be taken at once. If, however, a set recurs but the queue is in a different order, it is conceivable though unlikely that something different, that breaks the loop, could happen. So, in that case, a second chance is given and the counting continues but, if the same set recurs yet again, a loop is assumed and action taken.

In either event the action is the same, to exclude all candidates who have never been a probable since the last restart (which means the start where no actual restart has occurred) and then restart from the beginning except that the existing probables and queue are retained instead of the initial STV count.

If there is no candidate who can be excluded, then a special procedure is used, in which any candidate who has always been a probable since the last restart is classified as a certainty and any other remaining candidate as a contender. From each possible set of *n*+1 candidates that includes all the certainties, an election for *n* seats is conducted. Since, at this point, most of the original candidates will be either excluded or certainties, there is no need to fear an astronomical number of tests needing to be made.

At the end of each test, the one candidate who has not reached the quota is assigned a fractional value calculated by dividing that candidate's votes by the quota. When all the tests have been done, the average of these fractions is calculated for each candidate. Additionally candidates are awarded one point for each contest in which they did reach the quota. It is these complete points that mainly decide, the average fraction being really only a tie-breaker.

The contender with the highest score is then reclassified as a certainty and, if the number of certainties is less than the number of seats, the special procedure is repeated with one contender fewer and one seat fewer to fill.

While this process may look complicated, it should be remembered that, on most occasions, only the part called ‘the easy part’ above is used, while the complications are used to sort out a Condorcet paradox if it occurs.

Programming

Where loops occur it will often be found that a particular set of candidates is being tested more than once. Storing results and accessing them as necessary would obviously be much quicker than repeating the same STV count many times. However, since most voting patterns do not have such loops, such storing of results would usually be unproductive extra work. For the present, the system has been programmed with repetition rather than storing.

The name ‘Sequential STV’

From now on the name Sequential STV will be used to mean this new version.

A random version

The initial STV count, to choose the initial probables and to determine the initial order of the queue, turns out to be not very important, in that an alternative version that selects the initial probables at random, and orders the initial queue at random, nearly always reaches the same eventual answer. It is fun to watch it getting from an initial nonsense selection to end up at the correct solution, but this version should not be used in practice because of rare cases where it can get a different result from that given by starting with an STV count and, where this is so, we suspect that it would usually be a less good result.

An example of such a rare case has been given previously[3] with a fictitious set of votes, having 4 candidates for 2 places, in which testing ABC elects AB and testing ABD elects AB, yet testing ACD elects CD and testing BCD elects CD. In that example, Sequential STV elects AB (which is, in fact, the better choice) whereas the random version has a 50-50 chance of finding either AB or CD. Such an example seems unlikely ever to occur in reality but the fact that it is possible means that it is better to guard against it by not using the random version.

Examples

With 5 candidates for 2 seats, consider the voting pattern

```
104 ABCD
103 BCDA
102 CDBA
101 DBCA
  3 EABCD
  3 EBCDA
  3 ECDBA
  3 EDCBA
```

Plain STV elects BC. Sequential STV chooses BC as probables, then tests BCD, BCE and BCA in that order. BC win each time and are elected.

Suppose, however, that the voters for A, B, C and D had all put in E as second preference to give (the example used in reference 1).

```
104 AEBCD
103 BECDA
102 CEDBA
101 DEBCA
  3 EABCD
  3 EBCDA
  3 ECDBA
  3 EDCBA
```

This evidently makes E a very much stronger candidate, for if any one of A, B, C or D had not stood, E would have been the first elected, but plain STV takes no notice, electing BC just as before. Sequential STV chooses BC as probables but then tests BCD, where BC stay as probables and D goes to the end of the queue, followed by BCE where BE become the new probables and C goes to the end of the queue. It then tests BEA and BED, BE winning each time. There is no need to test BEC again as that result is already known, so BE are elected.

Real voting patterns

In 43 real elections held on file, the sequential method merely confirmed the original result in 38 of them, and replaced just 1 candidate in 3 more of them. In only 2 cases were loops found, making it necessary to do more than the easy part of the method.

Timings

Some timings were made on an 11-year old PC with a 386 chip. In a real election with 10 candidates for 6 seats and 841 voters, simple STV took 11 seconds. Sequential STV made no change in those elected and took 23 seconds.

In a much more difficult case with 30 candidates for 15 seats and 563 voters, simple STV took 1 minute 6 seconds. Sequential STV found 1 candidate to be definitely replaced and 3 others who were in a loop for the final seat. It took a total of 18 minutes 30 seconds.

Should it be used?

With this new version, should it be recommended for practical use? That depends upon whether the user is willing to abandon the principle that it should be impossible for a voter to upset earlier preferences by using later preferences. Many people regard that principle as very important, but reducing the frequency of premature exclusions is important too. We know that it is impossible to devise a perfect scheme, and it is all a question of which faults are the most important to avoid.

In considering this, we need to take into account, among other things, that the true aim of an election should not be solely to match seats as well as possible to votes, but to match seats to the voters' wishes. Since we do not know the wishes we must use the votes as a substitute, but that makes it essential that the votes should match the wishes as far as possible. That, in turn, makes it desirable that the voters should not be tempted to vote tactically.

They would not be so tempted if they felt confident that later preferences were as likely to help earlier ones as to harm them, and if they could not predict the effect one way or the other. At present, we see no reason to doubt that these requirements are met.

All things considered, we believe that Sequential STV is worthy of serious consideration.

Comparison with STV(EES) and with CPO-STV?

STV(EES) [4] was designed to meet much the same aims as Sequential STV, and also has the same disadvantage that later preferences can upset earlier ones. A comparison of the two would be interesting. As at present defined, however, STV(EES) is so slow that a comparison is not easy. For an electoral method to be slow should not be considered too much of a disadvantage for real elections if it can be shown to get better results, but it is certainly a disadvantage for research purposes where a large number of counts of different data may be required within a reasonable time.

Using the examples above, STV(EES) elects BC from the first but BE from the second, just as Sequential STV does.

In the example given in section 6 of reference 4, AC were elected by STV(EES), which was not wrong as there was a paradox in the votes, but the paper admitted that 'I would still have preferred AB to be the winning set in this case', so it may be worth noting that Sequential STV does indeed elect AB.

CPO-STV [5, 6] was designed to search for an outcome that is globally optimum rather than merely locally stable. Again a comparison would be interesting.

Acknowledgements

We are grateful to Douglas Woodall and Nic Tideman for helpful comments on earlier versions of this paper.

References

1. I.D. Hill, Sequential STV. *Voting matters*, issue 2, 5-7. 1994.

2. D.R. Woodall, Properties of preferential election rules. *Voting matters*, issue 3, 8-15. 1994.

3. I.D. Hill, Trying to find a winning set of candidates. *Voting matters*, issue 4, 3. 1995.

4. Simon Gazeley, STV with elimination by electability scores. *Voting matters*, issue 12, 9-13. 2000.

5. T. Nicolaus Tideman, The single transferable vote. *Journal of Economic Perspectives*, 9, 27-38. 1995.

6. T. Nicolaus Tideman, Better voting methods through technology: the refinement-manageability trade-off in the single transferable vote. *Public Choice*, 103, 13-34. 2000.

All correspondence regarding *Voting matters* should be addressed to:

B A Wichmann, The Editor, *Voting matters*, The Electoral Reform Society, 6 Chancel Street, London, SE1 0UU.

or using e-mail to Brian.Wichmann@bcs.org.uk.

To aid production, the Editor would welcome contributions on IBM-PC discs (with a printed copy as well) or to Brian.Wichmann@bcs.org.uk.

Voting matters

for the technical issues of STV

The Electoral Reform Society

Issue 16 February 2003

Editorial

The year 2002 has seen significant advances with the technology of STV, from opposite sides of the world.

In the Republic of Ireland, plans for the introduction of electronic voting (in the polling booth, not at this stage, via the Internet) have advanced to a key stage. Suitable technology has been developed for the polling stations and software has been written to undertake the count. ERBS was contracted to test the counting software to ensure it adhered to the rules which are identical to the hand-counting ones. On the 17th May, the Dáil elections were held in which three constituencies were handled electronically as an experiment, while the others were handled by the traditional manual means. The software validation was completed in time under the direction of Joe Wadsworth using a program for the Irish rules written by Joe Otten and with the editor running over 400 tests, some specially written for the occasion. I am glad to report that the counting went smoothly on the day.

The Irish election data for the three constituencies (Meath, Dublin North and Dublin West) was placed on the Internet with the full results of the count. To my knowledge, this is the first time over 2,000 STV votes (ie, the full set of preferences given by each voter) has been made publicly available. It is now possible to analyse this data. It is immediately clear, even by a manual inspection that many final preferences are in ballot paper order.

The developments with STV in New Zealand have been continuing throughout 2002 and are reported in the final article in this issue by Stephen Todd.

Other articles in this issue includes a note by Peter Dean showing how the actual administration of STV has changed over the years in Tasmania (even without the impact of computers). David Hill also considers a disturbing example of changes to the preferences on ballot papers which are not visible to the traditional rules.

Eivind Stensholt presents a rather technical article about the implementation of Meek STV rules when equality of preference is permitted. (Does the observed ballot-paper ordering with the Irish election indicate that equality of preference should be allowed?)

The remaining article is a short one by myself about the vexed question of proportionality.

Welcome to the McDougall Trust

This issue is the last one under the ERS banner. Following discussions between ERS and the Trust, *Voting matters* is being transferred to the Trust for publication for the time being. At this point, no significant changes are envisaged.

Brian Wichmann.

STV in Tasmania

P Dean

Peter Dean has been involved with ERS for many years.

In his article in *Voting matters*[1], Philip Kestelman raises the issue of positional voting bias. In Tasmania, there has been a continuous process of changing some details of the STV voting system to make it fairer. The problem of positional voting bias was addressed in 1979 and first used in 1980.

A summary of STV in Tasmania from Newman[2] is as follows:

1897 First experimental use of STV.

1903 Women given the vote.

1909 First state-wide election by STV.

1917 By-elections and vacancies filled by a recount of the original ballots. First used in 1922.

1921 Women allowed to stand as candidates.

1922 Deposit lost if less than 20% of the quota if excluded or at the end of the count.

1930 Compulsory vote, previously 63-67% turnout, up to 82% in 1928.

1941 Grouping by party labels.

1954 Parliamentary term reduced from 5 to 3 years.

1955 Speaker to be chosen from party with the lower statewide vote.

1957 Assembly of 35 instead of 30 to overcome potential deadlock.

1972 Term changed to 5 years, and 4 years thereafter.

1973 Voters required to make 7 choices instead of 3. Previously 90% of electors restricted their choice to a single party. Franchise reduced to 18.

1976 Draw for ballot position, and position within party list.

1980 first use of rotated ballot. The printer must issue equal numbers of papers showing different names in the favoured position, starting with the first name alphabetically. Thus with a columnar ballot paper 2, 8, 3 and 7 members in the 4 columns, 16 different printings are made.

A 1957 Select committee reported that it provided the Tasmanian elector with a wider freedom of choice, and a more effective vote than any other method of Parliamentary election in the world.

References

1. P Kestelman. Positional Voting Bias Revised. *Voting matters*, Issue 15, pp 2-5. June 2002.

2. Terry Newman. *Hare-Clark in Tasmania*. Joint Library Committee of the Parliament of Tasmania, Parliament House. Hobart. 1992.

Implementing a suggestion of Meek's

E Stensholt

Eivind Stensholt is from the Norwegian School of Economics and Business Administration

Introduction

In preferential elections voters are often assumed to have linear rankings, i.e. they rank all candidates without ties. Here the topic is STV elections where only a “complete order” is required, which means that a voter must give each candidate a rank, but may declare equal preference. Hence in a 10-candidate election a voter V may rank

PQ(ABCDE)RST

which, in Hill's notation[1], means that A, B, C, D, E share third to seventh rank.

At an iterative step in an algorithm for Meek's method a candidate P has a certain current retention factor: $1-p$, which is a positive number less than or equal to 1. Voter V starts on top of his list, offers P his full vote, for which $1-p$ is retained and offers Q p votes, has $p(1-q)$ retained and has $w = pq$ votes when coming to the set of equal preferences {A, B, C, D, E}.

Meek[2] suggested to count as if there were 5! = 120 “minivoters”, each with a weight of $w/120$ votes, with one minivoter for each possible way to split up the {A, B, C, D, E} into 5 singleton classes. With n candidates ranked equal, there are $n!$ possible linear rankings, and the work soon becomes too much even for computers if each minivoter is considered separately. However, the counting can be systematized, so that the necessary work grows as n^2. Thus there need not be a “combinatorial explosion”, but the algorithm does not otherwise relate to Hill's discussion of how to cope with equality of preference.

A count with five candidates equal

One minivoter ranks ABCDE, and contributes

$(1-a)w/120$, $a(1-b)w/120$, $ab(1-c)w/120$, $abc(1-d)w/120$, $abcd(1-e)w/120$

to A, B, C, D, and E, respectively. Each minivoter keeps weight $abcdew/120$, and hence voter V keeps $abcdew$ to influence the ranking of R, S, and T.

What is the total contribution from the 120 minivoters to candidate E? The contribution has 5 parts:

24 minivoters have E as number 1: $24(1-e)w/120$
24 minivoters have E as number 2: $6(a+b+c+d)(1-e)w/120$
24 minivoters have E as number 3:
$4(ab+ac+ad+bc+bd+cd)(1-e)w/120$
24 minivoters have E as number 4:
$6(bcd+acd+abd+abc)(1-e)w/120$
24 minivoters have E as number 5: $24(abcd)(1-e)w/120$.

The total contribution from V to E is therefore

$$[1/5 + (a+b+c+d)/20 + (ab+ac+ad+bc+bd+cd)/30 + (bcd+acd+abd+abc)/20 + (abcd)/5](1-e)w$$

An efficient algorithm is possible because the factors that depend on a, b, c, and d are easily calculated as the coefficients in a polynomial:

$$Q(E, x) = (x+a)(x+b)(x+c)(x+d) = x^4 + (a+b+c+d)x^3 + (ab+ac+ad+bc+bd+cd)x^2 + (bcd+acd+abd+abc)x + abcd.$$

How much computational effort is involved in calculating $Q(E, x)$? Writing

$$Q(E, x) = [x^3 + (a+b+c)x^2 + (bc+ca+ab)x + abc](x+d)$$
$$= [x^4 + (a+b+c)x^3 + (bc+ca+ab)x^2 + (abc)x] + [dx^3 + (a+b+c)dx^2 + (bc+ca+ab)dx + (abc)d],$$

we see that the factor $(x+d)$ involves first 3 multiplications of two real numbers with d as a factor and then 3 additions of two real numbers to get the coefficients of x^3, x^2, and x. Multiplying $(x+a)(x+b)$ needs one multiplication and one addition, and $(x+a)(x+b)(x+c)$ is calculated with two more of each. Hence $Q(E,x)$ requires 1+2+3 = 6 multiplications and 1+2+3 = 6 additions. Moreover, the contribution formula contains 6 multiplications, 4 additions, and 1 subtraction.

The general case

In general, consider n candidates, C1, ..., Cn, with retention factors $1-p(1)$, ... , $1-p(n)$. Consider the polynomials

$$Q(Ci, x) = [x+p(1)][x+p(2)] \ldots\ldots [x+p(n)]/[x+p(i)]$$
$$= B(0)x^{n-1} + B(1)x^{n-2} + B(2)x^{n-3} + \ldots + B(n-1)$$

for i from 1 to n. Clearly $B(0) = 1$ while the other $B(k)$ depend on i. They are the elementary symmetric polynomials in the $p(j)$ where $j \neq i$. The multiplication of $n - 1$ factors of type $[x + p(j)]$ involves $1 + 2 + 3 + \ldots + (n-2) = (n-1)(n-2)/2$ multiplications of two real numbers and equally many additions.

Suppose the candidates C1, ..., Cn form an equal preference set for voter V, who has weight w left after contributing to the higher ranked candidates. The contribution from V to candidate Ci, i.e. the votes to Ci from $n!$ minivoters, is given by the contribution formula Rev(i) =

$$[K(n-1,0)B(0) + K(n-1,1)B(1) + \ldots + K(n-1,t)B(t) + \ldots + K(n-1,n-1)B(n-1)][1-p(i)]w$$

where the $K(n-1,t)$ are determined as follows: There are $n!$ minivoters, with weight $w/(n!)$ each. Among them, $(n-1)!$ have candidate Ci as number $t+1$. The t candidates ranked ahead of Ci can be permuted in $t!$ ways. The $n-t-1$ candidates ranked after Ci can be permuted in $(n-t-1)!$ ways. Thus $t!(n-t-1)!$ of the $(n-1)!$ minivoters have the same t candidates ahead of Ci and they offer the same support to candidate Ci. The total revenue collected by Ci from these $(n-1)!$ minivoters is $t!(n-t-1)!\ B(t)\ [1-p(i)]w/(n!)$. Thus $K(n-1,t) = t!(n-t-1)!/(n!)$, i.e.

$$K(n,t) = t!(n-t)!/((n+1)!).$$

For the use of the contribution formula, it is practical to tabulate the coefficients $K(n-1,t)$.

If each $Q(Ci,x)$ is calculated as a product with $n-1$ factors, i from 1 to n, the total requirement is $n(n-1)(n-2)/2$ multiplications of two real numbers and $n(n-1)(n-2)/2$ additions. Thus the work grows with the third power of n. Here we leave out the $n+1$ multiplications and $n-1$ additions and 1 subtraction that must be performed each time the contribution formula is used.

However, with $n>5$ one may reduce the work by first calculating $Q(x)$ =

$$[x+p(1)][x+p(2)] \ldots [x+p(n)] = A(0)x^n + A(1)x^{n-1} + A(2)x^{n-2} + \ldots + A(n)$$

by means of $n(n-1)/2$ multiplications and $n(n-1)/2$ additions, and then for each i perform the division with $[x+p(i)]$:

$$A(0)x^n+A(1)x^{n-1}+A(2)x^{n-2} + \ldots + A(n) = [B(0)x^{n-1}+B(1)x^{n-2}+B(2)x^{n-3} + \ldots + B(n-1)][p(i)+x]$$

leads to $A(0) = B(0) = 1$ and

$A(1) = B(0)p(i) + B(1)$,
$A(2) = B(1)p(i) + B(2), \ldots$,
$A(n\text{-}1) = B(n\text{-}2)p(i) + B(n\text{-}1)$.

Hence $Q(Ci,x)$ is calculated as follows:

$B(1) = A(1) - B(0)p(i)$,
$B(2) = A(2) - B(1)p(i) , \ldots$,
$B(n\text{-}1) = A(n\text{-}1) - B(n\text{-}2)p(i)$.

The division with $[x+p(i)]$ requires n-1 multiplications with $p(i)$ as a factor and n-1 subtractions. All the divisions for i from 1 to n require $n(n\text{-}1)$ multiplications and $n(n\text{-}1)$ subtractions. Thus it is enough to perform $3n(n\text{-}1)/2$ multiplications and $3n(n\text{-}1)/2$ additions/subtractions instead of $n(n\text{-}1)(n\text{-}2)/2$ of each.

There are of course also $n(n+1)$ multiplications and n^2 additions/subtractions associated with the use of the contribution formula for n candidates, and so we arrive at $n(5n\text{-}1)/2$ multiplications and $n(5n\text{-}3)/2$ additions/ subtractions.

Further small savings are obviously possible, e.g. by keeping $Q(Cn,x)$ as an intermediate result from the calculation of $Q(x)$ instead of dividing $Q(x)$ by $[x+p(n)]$, but they do perhaps not justify the extra programming.

A program for calculating the contributions

Here is a Maple routine for calculating the contribution from a voter with weight 1 to each candidate in an equal preference set of n candidates 1, 2, ... , n, with given retention factors. The total number of candidates is denoted by C.

Set n = *number of candidates ranked equally by the voter*:

```
> n:=9;

                 n := 9
```

Set p(i) *for candidates* 1, 2, ... , n, *so that* 1-p(i) *is the current retention factor for candidate* i.

```
> for i from 1 to n do p(i):=0.5+0.04*i; od;

                p(1) := 0.54
                p(2) := 0.58
                p(3) := 0.62
                p(4) := 0.66
                p(5) := 0.70
                p(6) := 0.74
                p(7) := 0.78
                p(8) := 0.82
                p(9) := 0.86
```

As an example we use these equidistant values for the p(i). *The routine consists of a "preparation" and two instructions. The preparation is used only once per run of the election program. It sets the coefficients* K(i,j) = j!(i-j)!/(i+1)! *by first calculating the binomial coefficients* " i - choose - j " = i!/(j!(i-j)!).

Preparation. *Set the table of constants. Let C be the total number of candidates:*

```
> C:=20: for i from 0 to C-1 do K(i,0):=1.0; od:
for j from 1 to C-1 do K(0,j):=0.0; od:
for i from 1 to C-1 do for j from 1 to C-1 do
K(i,j):=K(i-1,j-1)+K(i-1,j); od: od:
for i from 1 to C-1 do for j from 0 to i do
K(i,j):=1.0/((i+1)*K(i,j)); od: od:
```

Instruction 1. *Calculate the polynomial of degree* n:

```
> A(0):=1.0: B(0):=1.0: for j from 1 to n do A(j):=0.0; od:
for j from 1 to n do for i from 0 to j-1 do
A(j-i):= A(j-i-1)*p(j) + A(j-i); od; od;
```

Instruction 2. *Calculate the polynomial of degree* n-1 *for candidate s and simultaneously set* Rev(s) = the revenue for candidate s, s=1, 2, ..., n:

```
> for s from 1 to n do Pr:=K(n-1,0): q:=p(s):
for j from 1 to n-1 do B(j) := A(j)-B(j-1)*q;
Pr:=Pr+B(j)*K(n-1,j); od: Rev(s):=Pr*(1-q); od:
```

Another instruction shows the revenue Rev(s) *collected by candidate s from all* n! *"minivoters"* :

```
> for s from 1 to n do Rev(s):=Rev(s); od;

        Rev(1) := .171708815169
        Rev(2) := .154311932284
        Rev(3) := .137512907077
        Rev(4) := .121258700936
        Rev(5) := .105503965732
        Rev(6) := .0902095328389
        Rev(7) := .0753412681397
        Rev(8) := .0608691895186
        Rev(9) := .0467667763417
```

These contributions sum to 0.963483088037.

The voter keeps p(1) p(2) ... p(n) = 0.036516911963.

What happens in the example above?

Consider 9 candidates sharing ranks 1 to 9 in a vote, and assume the retention factors are as above. The preparation has calculated a table including (K(8,0), ..., K(8,8)) =

(0.1111111111, 0.01388888889, 0.003968253968, 0.001984126984, 0.001587301587, 0.001984126984, 0.003968253968, 0.01388888889, 0.1111111111)

With the p(i) above, instruction 1 gets the polynomial of degree n = 9,

Q(x)= [x+0.54] [x+0.58][x+0.62] [x+0.66] [x+0.70] [x+0.77] [x+0.78] [x+0.82] [x+0.86] =

$1\,x^9 + 6.30\,x^8 + 17.5920\,x^7 + 28.576800\,x^6 + 29.75937888\,x^5 + 20.60302608\,x^4 + 9.482569153\,x^3 + 2.797730344\,x^2 + 0.4801360978\,x + 0.03651691196$.

Then for s=9, instruction 2 gets Q(C9,x) = Q(x)/[x+0.86] =

$1\,x^8 + 5.44\,x^7 + 12.9136\,x^6 + 17.471104\,x^5 + 14.73422944\,x^4 + 7.93158876\,x^3 + 2.661402819\,x^2 + 0.508923920\,x + 0.0424615266$,

and at the same time it calculates the contribution from the voter with weight 1 to candidate 9:

$[1 \times 0.1111111111 + 5.44 \times 0.01388888889 + 12.9136 \times 0.003968253968 + 17.471104 \times 0.001984126984 + 14.73422944 \times 0.001587301587 + 7.93158876 \times 0.001984126984 + 2.661402819 \times 0.003968253968 + 0.508923920 \times 0.01388888889 + 0.0424615266 \times 0.1111111111] \times (1- 0.86)$

$= 0.3340484026 \times (1- 0.86) = 0.04676677636$.

Acknowledgement

The author is grateful to two referees for suggestions that have made the presentation clearer.

References

1. I D Hill, Difficulties with equality of preference, *Voting matters*, Issue 13, April 2001, pp8-9.

2. B L Meek, Une nouvelle approche du scrutin transferable (fin), *Mathématiques et sciences humaines*, 9 no.29 pp 33-39. 1970

3. B L Meek, A new approach to the Single Transferable Vote, *Voting matters*, Issue 1, March 1994, pp1-11.

What would a different method have done?

I D Hill

Following an election, the question is often raised of what the result would have been had a different electoral method been used. In general, no reply can be given to this question not only because sufficiently detailed information is not available on the votes, but also because voters can be expected to behave differently if a different system is used.

In comparing one STV system with another, however, rather than totally different systems, it seems unlikely that there would be very much difference in how voters behave, and a reasonable reply is possible provided that the full voting pattern is divulged. It is very welcome that it has been divulged for the three constituencies counted by computer in the recent general election in Eire. Such openness is to be commended. Too often, though, the full voting pattern is regarded as confidential, and the only information is a result sheet, which is quite insufficient for the purpose.

As an example, the question might be whether the result of the 2002 ERS Council election would have been different had the Meek system been used. Working solely from the result sheet (the only information available) I have constructed a voting pattern in which some votes have the character # inserted within their preferences. Before running such data on a computer the # characters have to be replaced, either by a number representing a candidate, or by a space which is then ignored by the STV program.

If the # characters are all replaced by a space, and ERS97 rules used, the actual result sheet is reproduced. If Meek rules are used the same candidates are elected, following a similar order of events.

However, if the # characters are all replaced by the number that represents any one of the defeated candidates, and ERS97 rules used, the same result sheet appears, identical in every particular, but if Meek rules are used, that defeated candidate is elected, at the expense, of course, of one of those who was actually successful.

There is no suggestion that this artificial voting pattern is anything like the true one. I am absolutely sure that it is not, but it is somewhat remarkable that it is possible to devise such a voting pattern with no effect at all on the ERS97 result sheet. The fact that it is possible shows the extent to which the information available is totally inadequate to answer the question. I believe it to be impossible to do the reverse, leaving the Meek result unchanged while varying the ERS97 result.

The artificial voting pattern can be supplied on request.

What sort of proportionality?

I D Hill

In pure mathematics proportionality is a well-defined concept, but that is because we can always go into fractions whenever necessary. For proportionality within voting systems we are restricted to whole numbers in those elected for each party (using "party" in the general sense of any relevant grouping of the candidates, not only in the sense of a formal political party). Under such circumstances it is in many cases not at all easy to say whether one result is more nearly proportional than another. This is particularly so where some parties, quite correctly, get zero seats, while none get zero votes.

I agree with Philip Kestelman[1] that none of the measures that he discusses is perfect. I agree also that the comparative answers that they produce are so similar that, if using any, we might as well settle on one of them. But as I have said before[2] they are all fundamentally flawed in basing their calculations on first-preference votes only, and this can be very misleading, particularly where there is a substantial amount of cross-party voting for successive preferences.

However there is an additional point to be considered, even where first preferences do give full information on party popularity, there being no cross-party voting at all. Under such circumstances it could be the rule that if n is the minimum value, across parties, of votes per seat, then any party with at least n votes must get at least 1 seat, any party with at least $2n$ votes must get at least 2 seats, any party with at least $3n$ votes must get at least 3 seats, and so on. Given the restriction to whole numbers, and that some parties may get zero seats, what could be more proportional than that? Yet none of the measures that Kestelman considers meets that rule.

For simplicity, consider the case of only 2 parties and only 2 seats to be filled. Suppose the votes are 70 for party A and 30 for party B. We can at once rule out the option of giving both seats to party B, but is it better to give both to A or one to each?

Suppose we allot them as 1 to each. Then $n = 30 / 1$ so party A with more than $2n$ votes must get at least 2 seats and the rule is violated. Suppose we allot them as both to party A. Then $n = 70 / 2$ and the rule is satisfied for party B does not reach 35 to be worth a seat. Yet every one of the measures that Kestelman considers says that 1 to each is a better answer than both to party A. To my mind that shows all those measures to be unsatisfactory. I regret that I do not know of a better alternative, but to do without a measure is preferable to using a defective one.

If anyone doubts that both to party A is the better answer, let them assume that there had been only 3 candidates and votes 36 A1 A2, 34 A2 A1, 30 B. The measures all say that to elect A1 and B, or even A2 and B, is preferable to A1 and A2, which is surely nonsense.

However, I am grateful to Philip Kestelman for the suggestion that we might, perhaps, say that to elect A1 and B is more party-representative, while to elect A1 and A2 is more candidate-representative. There might be something in that.

References

1. P Kestelman, Quantifying representativity. *Voting matters*, Issue 10, 7-10. 1999.

2. I D Hill, Measuring proportionality. *Voting matters*, Issue 8, 7-8. 1997.

Proportionality Revisited

B A Wichmann

Introduction

The issue of proportionality in the last article[1], raised two problems in my mind which are addressed here.

A flaw

Consider the hypothetical case of an STV election in the UK, in which there is a United Kingdom Independence Party (UKIP) candidate together with a Tory candidate. A Tory voter who is on the Europhobic wing of the party could well decide to give his/her first preference to the UKIP candidate. On the other hand, if the Tory candidate was also Europhobic, then the voter would surely place his/her first preference with the Tory. In other words, the first preference votes for the Tory and UKIP cannot reasonably be analysed in isolation.

Of course, this issue is not specific to the Tory party — the same problem could arise with a Socialist Party candidate standing against a New or Old Labour candidate.

I conclude from this that an analysis of party support based upon first preferences alone is doomed to failure.

Granularity

In this section, we set aside the flaw noted above, and analyse the issue of proportionality from just one point of view: the granularity imposed by the size of the constituencies. If a constituency elects 4 members, then it is clear that strict proportionality could only be obtained if each party had a multiple of 25% of the first preference votes. Obviously, there will always be a mismatch between the first preference votes and the proportion of candidates elected.

As an example, we consider the 1997 Irish General election[2]. The 166 seats for the Dáil are from 41 constituencies having 3, 4 or 5 seats each. In this analysis, we consider three categories for the first preference votes: those of Fianna Fáil (FF), those for Fine Gael (FG) and the others. It can reasonably be said that the 'others' does not represent a party, but if strict proportionality is obtained for FF and FG, then the others as a single group will also be represented proportionally. We return to this problem later.

Kestelman[3] considers several measures of proportionality. Here, we consider some of those measures as applied to each individual constituency and compare this with the actual result. The measures used here are the Loosemore-Hanby Index, Gallagher Index of Disproportionality, Sainte-Laguë Index and the Farina Index (all taken from the above paper).

Given a specific index, then one can determine the number of seats for each party which would give the closest fit with respect to that index. In fact, all the indices give the same result with one exception: the Sainte-Laguë Index gives a different result for the Dublin Central constituency. Ignoring this isolated value we have the table as follows:

Constituency	*Actual*	*Best*	*Fit (%)*	*Comparison*
Carlow-Kilkenny	(2,2,1)	(2,2,1)	13.998	=
Cavan-Monaghan	(2,2,1)	(2,2,1)	8.850	=
Clare	(3,1,0)	(2,1,1)	7.452	FF to Other
Cork East	(2,2,0)	(2,1,1)	16.773	FG to Other
Cork North-Central	(3,2,0)	(2,1,2)	12.473	**Two changes**
Cork North-West	(2,1,0)	(2,1,0)	24.912	=
Cork South-Central	(3,2,0)	(2,2,1)	11.923	FF to Other
Cork South-West	(1,2,0)	(1,1,1)	20.608	FG to Other
Donegal North-East	(2,0,1)	(1,1,1)	17.801	FF to FG
Donegal South-West	(1,1,1)	(1,1,1)	12.710	=
Dublin Central	(2,1,1)	(2,0,2)	17.771	FG to Other
Dublin North	(2,1,1)	(1,1,2)	16.756	FF to Other
Dublin North-Central	(2,1,1)	(2,1,1)	4.487	=
Dublin North-East	(2,1,1)	(2,1,1)	19.113	=
Dublin North-West	(2,0,2)	(2,1,1)	15.808	FG to Other
Dublin South	(2,2,1)	(2,1,2)	11.999	FG to Other
Dublin South-Central	(2,1,1)	(1,1,2)	13.301	FF to Other
Dublin South-East	(1,1,2)	(1,1,2)	4.042	=
Dublin South-West	(2,1,2)	(1,1,3)	12.192	FF to Other
Dublin West	(2,1,1)	(1,1,2)	11.492	FF to Other
Dun Laoghaire	(2,2,1)	(1,2,2)	11.226	FF to Other
Galway East	(2,2,0)	(2,1,1)	7.923	FG to Other
Galway West	(2,1,2)	(2,1,2)	10.324	=
Kerry North	(1,1,1)	(1,1,1)	19.729	=
Kerry South	(1,0,2)	(1,0,2)	18.479	=
Kildare North	(1,1,1)	(1,1,1)	9.214	=
Kildare South	(1,1,1)	(1,1,1)	8.464	=
Laoighis-Offaly	(3,2,0)	(3,1,1)	13.281	FG to Other
Limerick East	(2,1,2)	(2,1,2)	9.015	=
Limerick West	(1,2,0)	(1,1,1)	4.945	FG to Other
Longford-Roscommon	(2,2,0)	(2,1,1)	15.181	FG to Other
Louth	(2,1,1)	(2,1,1)	12.575	=
Mayo	(2,3,0)	(2,3,0)	14.288	=
Meath	(3,2,0)	(2,2,1)	3.803	FF to Other
Sligo-Leitrim	(2,2,0)	(2,1,1)	15.211	FG to Other
Tipperary North	(2,0,1)	(1,0,2)	24.890	FF to Other
Tipperary South	(1,1,1)	(1,1,1)	11.361	=
Waterford	(2,1,1)	(1,1,2)	14.951	FF to Other
Westmeath	(1,1,1)	(1,1,1)	15.218	=
Wexford	(2,2,1)	(2,2,1)	3.036	=
Wicklow	(2,1,2)	(1,1,3)	13.758	FF to Other

The content of the table is best explained by taking an entry: say Waterford, with 4 seats. The *Actual* and *Best* entries give the seats in the order (FF, FG, Other). The *Best* entry is computed according to all the indices apart from the isolated result already noted. The Fit% figures are calculated from the formula:

Fit% $=\sqrt{(\sum(S\%-V\%)^2)}$, which is related to the Gallagher index.

The last column gives the comparison between the actual and best entries in seats. For Waterford, a single change in the actual result by a FF seat becoming an Other seat would produce the 'best' result.

One can see from this result that 18 constituencies would remain unchanged if they gave the best fit to first preference proportionality. The major difference is that the two major parties have gained over the others — the best fit giving 56 seats in the Dáil for 'others' against the actual number of 35.

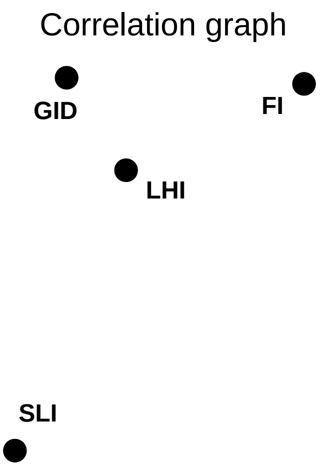
Two constituencies are different from the others. In the case of Cork North-Central, a two seat change is needed from the actual result to get the best fit. The reason for this is a high level of transfers from the other candidates to the two major parties. The case of Donegal North-East is special because the difference in the actual and best does not involve an increase in the 'other' seats. The reason for this was a significant transfer from FG to FF in the actual election when an FG candidate was still available for transfers.

As would be expected, there is a wide variation in the *Fit* entries. Also, the *Fit* values decrease with increased constituency seats: an average of 15.7% for 3-seats, 12.8% for 4-seats and 10.7% for 5-seats.

The under-representation of the Other group is to be expected as many of those candidates are excluded early in the count with many transfers to the major parties (as well as to non-transferables). This effect clearly indicates the dubious nature of grouping all the parties other than the major two into one.

The conclusion from this analysis seems to be that there is little loss in proportionality due to the natural granularity of the STV system. The lack of proportionality compared to the first preferences is caused by the vote transfers. There is a capital T in STV.

In addition to the above analysis of granularity, the same data reveals a very close correlation between the indices used. This is gratifying, since they are clearly supposed to be measuring the same property. However, the correlations can be represented approximately in a graph as follows in which the indices are indicated by their initials and the distance between them increases with a lack of correlation. From this it appears that the Loosemore-Hanby Index is centrally placed which reinforces Kestelman' s support for that index.



References

1. I D Hill, What sort of proportionality? *Voting matters*, Issue 16, pp5-6. 2003

2. 28th Dáil General Election, June 1997, Election Results and Transfer of Votes. *The Stationery Office*, Dublin 1998.

3. P Kestelman, Quantifying Representativity. *Voting matters*. Issue 10. pp7-10. March 1999.

STV in New Zealand

S W Todd

Stephen Todd is a member of the Electoral Reform Coalition and has advocated STV in New Zealand for many years

In May 2001, the New Zealand Parliament enacted the Local Electoral Act 2001. At section 3 of the Act, it is stated that its purpose "is to modernise the law governing the conduct of local elections and polls ..." including, to "allow diversity (through local decision-making) in relation to ... the particular electoral system to be used for local elections and polls[.]"

Section 5 of the Act defines "electoral system" as "... any of the following electoral systems that are prescribed for use at an election or poll:

(a) the system commonly known as First Past The Post:

(b) the system commonly known as Single Transferable Voting (STV) using Meek's method of counting votes[.]"

As a result of this legislation, New Zealand becomes the first country in the world to adopt STV by Meek's method for use in public elections. Indeed, although local authorities have the choice of switching to STV if they or their electors want it, the Act, at section 150, amends the New Zealand Public Health and Disability Act 2000, to make it mandatory for the seven elected members of the country's twenty-one district health boards to be elected by STV.

It will come as no surprise to learn that the road to STV becoming a reality in New Zealand was not an easy one. In 1994, on behalf of the Electoral Reform Coalition, I prepared a draft bill for the Deputy Leader of the Opposition (Labour Party), the Hon David Caygill, MP. After consulting the Electoral Reform Society in the UK, I incorporated the Northern Ireland rules in the relevant Schedule of the bill. Mr Caygill took the bill to a subsequent meeting of the Labour caucus, which agreed that it should be accepted as a private member's bill.

At that point it became the responsibility of the opposition spokesperson on Local Government, Richard Northey, MP. He placed it in the fortnightly ballot of members' bills in October 1994, and it was drawn from that ballot the following April. Mr Northey introduced the bill (Local Elections (Single Transferable Vote Option) Bill) into the House of Representatives on 19 July 1995.

Ten of 78 submissions on the bill were heard by the Electoral Law select committee, in November 1995. On 31 July 1996, the committee established a subcommittee, comprising Richard Northey (Chairperson) and Hon. David Caygill, to consider the bill. Advice was received from officials in the Department of Internal Affairs, and the subcommittee reported its findings to the committee on 21 August 1996. The bill was reported back to Parliament in early September, just as Parliament was dissolved so that New Zealand's first MMP election could be held (on 12 October). The bill was held over for consideration by the new Parliament.

Part of the "advice [...] received from officials" was to abandon the Northern Ireland rules on the ground that they did not treat all votes equally, particularly with regard to those votes given for successful candidates that were not in the actual parcel of votes that put a candidate up over the quota. Such inequality in the treatment of votes was seen as unfair.

Furthermore, knowing that computer technology was increasingly being used in local elections, the committee wanted counting rules that were more compatible with the use of such technology.

Unfortunately, the rules written to replace the Northern Ireland rules in the Report copy of the bill were logically unsound. The main problem was that the word "votes", as used in the rewritten rules, did not always mean the same thing. Sometimes it referred to transferable papers and other times to the value of those papers. In undertaking the rewrite, the authors overlooked the fact that, regardless of whether hand-counting rules are carried out by hand or by computer, it is voting papers that are being transferred, sometimes at full value, sometimes at a reduced transfer value, rather than votes. A number of consequential errors arising from this and other misunderstandings, rendered the rules inoperable.

The rule pertaining to the calculation of the transfer value was a case in point. In the case of the transfer of a surplus resulting from a previous transfer of votes, the transfer value of the votes transferred [was to] be "the result of dividing the surplus by the total number of votes transferred in that previous transfer to the candidate from whom the surplus is transferred."

A transfer value is calculated by dividing the surplus by the number of transferable papers, not by the sum of the value of those papers and non-transferable papers, i.e. total votes. Under normal hand-counting rules, for example, an elected candidate may obtain the quota upon receiving a batch of 280 voting papers, each having a transfer value of 0.35 — a total of 98 votes. If this candidate now has a surplus of 60 votes and only 240 of the 280 papers last received are transferable, then they would be transferred at a transfer value of 0.25.

The above-mentioned rule, however, states that the transfer value shall be calculated by dividing the surplus of 60 votes by the 98 votes transferred at the previous transfer, which comes to 0.612244... If this transfer value (0.61?) were then applied to the 240 transferable papers (although there was nothing to say it should be), a total of 146.40 votes would be transferred instead of 60, and the total number of non-transferable votes would be increased unnecessarily by 24.40!

Not only was there no direction as to how many decimal places the transfer value was to be taken to, but it was very obvious that the votes would not sum to the correct totals. Something had to be done.

The Electoral Law Committee of the new Parliament called for submissions on the Report copy of the bill, to be received by 30 October 1997.

During the course of my efforts to make sense of the re-written counting rules, I realised quite suddenly that what officials had been attempting to do, was to replace the Northern Ireland rules with Meek-equivalent rules, unaware that Meek's method of counting votes had already been invented, and subsequently perfected.

Consequently, in the Electoral Reform Coalition's submission, we recommended to the committee that the counting rules be replaced by Meek's method. Our efforts were all to no avail, however, with the bill being lost following a tied vote (4-4) in committee in May 1998.

That month, I set to work drafting a completely new bill, this time for opposition Green Party MP, Rod Donald, in which I incorporated Meek's method of counting votes. The Explanatory Note to the bill explained that Meek's method was a significant improvement over the various hand-counting rules, and why; that it treated all votes equally; and that a Meek count had to be carried out by computer.

The draft was completed in December 1998 and sent out to interested parties for comment. Reaction from the local government sector was generally unsupportive, but two prominent political scientists with a particular interest in local government agreed that Meek's method was an improvement over hand-counting rules.

The local government sector was resisting the STV option because local returning officers (now called electoral officers) were terrified at the thought of having to learn how to conduct a complicated hand-count of votes. They imagined dozens of people constantly shuffling thousands of pieces of paper from one pile to another over several days. In these cost-conscious times, when the public demands instant results, they simply didn't want to know about it.

Although sector representatives indicated continued resistance, this new bill happened to coincide with a push by the sector to have the local electoral legislation completely re-written and up-dated.

In June 1999, I was invited to attend a workshop on matters pertaining to the administration and conduct of local elections to give a presentation on Meek's method. Soon after, perhaps realising that their main objection to STV (fear of hand-counts) need not be a relevant consideration, and that the issue of STV was not going to go away, sector representatives decided to include provision for an STV option in their list of proposed improvements to the legislation.

A year later, in July 2000, Rod Donald's bill was drawn out in the fortnightly ballot of members' bills and given its first reading. At this time, the newly-elected Labour-led government decided that seven of the 11 members of the 21 district health boards (DHBs) that it intended to set up to replace the structure put in place by the previous government, would be elected by STV.

A significant reason for this decision was to ensure that the Maori population would have the means to ensure they were represented on these boards by people they helped to elect, if that was what they wanted. The legislation stipulates that at least two of the 11 positions must be filled by Maori, so enabling Maori to elect Maori members would enable, in most cases, the four appointed positions to be filled having regard to criteria other than ethnicity.

The government, which generally relied on the Green Party for its majority, and needing the support of the Greens to ensure the Local Electoral Bill would be enacted during the first half of 2001, agreed to include provisions for local authorities to adopt STV in that bill. In turn, Rod Donald allowed his bill to lapse in select committee.

At this point, late-July 2000, a decision needed to be made as to which of the several forms of STV would be included in the Local Electoral Bill. Relevant officials in the Department of Internal Affairs consulted well-known political scientists, and with myself, and reduced the choice to four — Tasmania's Hare-Clark rules, Northern Ireland's "senatorial" rules, the "original" STV rules, as used in the Republic of Ireland, and Meek's method.

In September 2000, a paper was submitted to Cabinet recommending that Meek's method be accepted as the form of STV best suited for New Zealand. Meek was "preferred to the hand counting forms of STV because it best contributes to effective and fair representation, and public confidence and understanding of local elections."

Two factors which contributed to this recommendation being made were that writing a computer program to implement Meek's method would be far more straightforward than if one of the forms of hand-counting rules were adopted, and because Meek's method reduces the number of "wasted" votes to an absolute minimum, and ensures all successful candidates achieve the required quota for election.

Furthermore, officials "noted that in 1996, the Electoral Law Committee proposed that Richard Northey's STV Option Bill be amended from the senatorial rules to a form that reflected the intent of the Meek rules, in order to remove the necessary arbitrariness generated by hand counting."

As alluded to in the first paragraph above, the Local Electoral Act provides for local authorities to resolve to change to STV, or to hold a poll on the electoral system, and also for electors to demand a poll be held on the electoral system.

In August and September 2002, eight (out of a total of 86) local authorities resolved to adopt STV to elect their councils and community boards (if any) in October 2004. A further two councils (Wellington and Whangarei) resolved to hold a poll of electors, on 30 November and 5 December, respectively. Wellington voted narrowly to adopt STV; Whangarei voted by a margin of almost 2 to 1 to retain the first-past-the-post (FPTP) system.

Since then, the Opotiki District Council, which was one of the eight local authorities to resolve to change to STV, and the Masterton District Council, which resolved to stay with FPTP, have further resolved to hold a poll of electors.

At the time of writing (January 2003), there have been 10 successful poll demands, with possibly a handful more by the end of February. All polls must be held no later than 21 May 2003, the results of which are binding on the councils concerned for the next two triennial general elections of the country's local authorities (9 October 2004 and 13 October 2007).

Voting matters

for the technical issues of STV

published by

The McDougall Trust

Issue 17 October 2003

About the McDougall Trust (reg. charity no. 212151)

The McDougall Trust is a charitable trust formed in 1948. The charity's purposes as stated in its governing scheme of 1959 are to advance knowledge of and encourage the study of and research into:

- political or economic science and functions of government and the services provided to the community by public and voluntary organisations; and
- methods of election of and the selection and government of representative organisations whether national, civic, commercial, industrial or social.

The Trust's work includes the maintenance and development of the Lakeman Library for Electoral Studies, a unique research resource, the production and publication of Representation: The Journal of Representative Democracy, and, of course, this publication **Voting matters**, that examines the technical issues of the single transferable vote and related electoral systems.
For further information on the Trust, please contact:

The Secretary,
McDougall Trust,
6 Chancel Street,
London SE1 0UX, UK.
Telephone: +44 (0)20 7620 1080
Facsimile: +44 (0)20 7928 1528
Email: admin@mcdougall.org.uk
Web: www.mcdougall.org.uk

For futher information on this publication, please contact B A Wichmann, the Editor at the above address or by email at: `Brian.Wichmann@bcs.org.uk`

Editorial

This is the first issue under the auspices of the McDougall Trust. The Editor has taken the opportunity of this change to make a number of stylistic changes. These are mainly as follows:

- Use of the LaTeX typesetting system so that, if they wish, authors can submit material in a format that can be directly typeset.
- Starting papers on a new page so that individual papers can be handled more easily.

This issue also has a slight departure in having two papers which are more mathematical in nature than is usual. It has been decided that the Editor should ensure that the main points of such papers are intelligible to non-mathematical readers by placing an appropriate summary here.

There are four papers in this issue:

- D R Woodall: *QPQ, a quota-preferential STV-like election rule*,
- J Otten: *Fuller Disclosure than Intended*,
- M Schulze: *A New Monotonic and Clone-Independent Single-Winner Election Method* and
- J Gilmour: *Calculation of Transfer Values — Proposal for STV-PR Rules for Local Government Elections in Scotland*.

In Douglas Woodall's paper he defines a new way of counting preferential votes which is analogous to conventional STV. To understand the counting process, it is probably best to work through the examples in the paper with the general definition in mind. It is clear that undertaking this form of counting without a computer is viable. Hence the interest here would be to see if QPQ has any appeal to those who think it inappropriate to use computers to count an election. The main mathematics in Woodall's paper is to show that QPQ has several desirable properties — hence this part can be skipped and the results taken on trust.

The paper of Joe Otten arose from a resolution put to the ERS AGM requesting that the full election data of the preferences specified should be available for STV elections. (Such disclosure was available for the three Irish constituencies for which electronic voting was employed in the June 2002 elections.) The paper explains a potential danger from full disclosure with a proposed resolution.

Markus Schulze in his paper considers the question of electing just one person, which would be the Alternative Vote (AV) with STV. Many would consider that AV is inappropriate since it does not necessarily elect the Condorcet winner (if there is one). The paper starts from the position of electing the Condorcet winner but with the objective of ensuring as many desirable properties are satisfied as possible. The proof that certain properties are satisfied involves some logical analysis which I hope most readers can follow.

James Gilmour's paper has arisen as a result of the recent consultation process for the introduction of STV in Scottish local elections. Here, he shows by analysis and example that the calculation of the transfer values can be improved by using more precision in the calculation than is often the case.

QPQ, a quota-preferential STV-like election rule

Douglas R. Woodall
Email: douglas.woodall@nottingham.ac.uk

1 Introduction

Olli Salmi, in a posting to an Election Methods list [6], has suggested a new quota-preferential election rule, which is developed slightly further in this article, and which is remarkably similar to the Single Transferable Vote (STV) in its effects. I shall call it QPQ, for Quota-Preferential by Quotient. Both in its properties and in the results it gives, it seems to be more like Meek's version of STV [2] than the traditional version [3]. This is surprising since: (i) in marked contrast with STV, the quota in QPQ is used only as a criterion for election, and not in the transfer of surplus votes; (ii) QPQ, unlike Meek's method, involves no iterative processes, and so the votes can be counted by hand; and (iii) QPQ derives from the European continental tradition of party list systems (specifically, d'Hondt's rule), which is usually regarded as quite different from STV. I do not imagine that anyone who is already using STV will see any reason to switch to QPQ; but people who are already using d'Hondt's rule may feel that QPQ is a natural progression of it, and so more acceptable than STV.

D'Hondt's rule for allocating seats to parties was proposed by the Belgian lawyer Victor d'Hondt [1] in 1882. The seats are allocated to the parties one by one. At each stage, a party with v votes and (currently) s seats is assigned the quotient $v/(1+s)$, and the next seat is allocated to the party with the largest quotient. This continues until all seats have been filled.

Many variations of this rule were subsequently proposed, in which the divisor $1+s$ is replaced by some other function of s. However, the next contribution of relevance to us is an adaptation of d'Hondt's rule to work with STV-type preferential ballots. This adaptation has been part of Sweden's Elections Act for many years; we will call it the *d'Hondt–Phragmén method*, since it is based on a method proposed by the Swedish mathematician Lars Edvard Phragmén [4, 5] in 1895. The seats are again allocated one by one, only this time to candidates rather than parties; at each stage, the next seat is allocated to the candidate with the largest quotient (calculated as explained below). In the event that the voters effectively vote for disjoint party lists (e.g., if every ballot is marked for *abcd*, *efg* or *hijkl*), then the d'Hondt–Phragmén method gives exactly the same result as d'Hondt's rule. However, it was introduced in the Swedish Elections Act as a means of allocating seats *within* a party, at a time when voters were allowed to express a choice of candidates within the party. It does not guarantee to represent minorities proportionally.

Salmi's contribution has been to introduce a quota into Phragmén's method. In this version, which he calls the *d'Hondt–Phragmén method with quota*, the candidate with the largest quotient will get the next seat if, and only if, this quotient is larger than the quota; otherwise, the candidate with the smallest quotient is excluded, and the quotients are recalculated. In this respect it is like STV. However, unlike in STV, this is the only way in which the quota is used; it is not used in transferring votes. QPQ, as described here, differs from Salmi's original version only in that the quota is defined slightly differently, and the count is preferably restarted after every exclusion.

Both the d'Hondt–Phragmén method (with or without quota), and QPQ, can be described in terms of groups of voters rather than individuals, and this is naturally how one thinks when processing piles of ballots by hand. But it seems to me that they are easier to understand when rewritten in terms of individual ballots rather than groups, and they are described here in this form. From now on, s denotes the total number of seats to be filled.

2 The details of QPQ

2.1. The count is divided into a sequence of stages. At the start of each stage, each candidate is in one of three states, designated as *elected*, *excluded* and *hopeful*.

At the start of the first stage, every candidate is hopeful. In each stage, either one hopeful candidate is reclassified as elected, or one hopeful candidate is reclassified as excluded.

2.2. At the start of each stage, each ballot is deemed to have elected some fractional number of candidates, in such a way that the sum of these fractional numbers over all ballots is equal to the number of candidates who are currently classed as elected. At the start of the first stage, every ballot has elected 0 candidates.

2.3. At the start of each stage, the quotients of all the hopeful candidates are calculated, as follows. The ballots contributing to a particular hopeful candidate c are those ballots on which c is the topmost hopeful candidate. The quotient assigned to c is defined to be $q_c = v_c/(1+t_c)$, where v_c is the number of ballots contributing to c, and t_c is the sum of all the fractional numbers of candidates that those ballots have so far elected.

2.4. A ballot is *active* if it includes the name of a hopeful candidate (and is a valid ballot), and *inactive* otherwise. The *quota* is defined to be $v_{\mathrm{a}}/(1 + s - t_{\mathrm{x}})$, where v_{a} is the number of active ballots, s is the total number of seats to be filled, and t_{x} is the sum of the fractional numbers of candidates that are deemed to have been elected by all the *inactive* ballots.

2.5a. If c is the candidate with the highest quotient, and that quotient is greater than the quota, then c is declared elected. In this case each of the v_c ballots contributing to c is now deemed to have elected $1/q_c$ candidates *in total* (regardless of how many candidates it had elected before c's election); no change is made to the number of candidates elected by other ballots. (Since these v_c ballots collectively had previously elected t_c candidates, and they have now elected $v_c/q_c = 1 + t_c$ candidates, the sum of the fractional numbers of candidates elected by all voters has increased by 1.) If all s seats have now been filled, then the count ends; otherwise it proceeds to the next stage, from paragraph 2.3.

2.5b. If no candidate has a quotient greater than the quota, then the candidate with the smallest quotient is declared excluded. No change is made to the number of candidates elected by any ballot. If all but s candidates are now excluded, then all remaining hopeful candidates are declared elected and the count ends; otherwise the count proceeds to the next stage, from paragraph 2.3.

The details of the calculations of the quotients and quota may become clearer from a study of Election 2 in the next section.

The specification above contains two stopping conditions, in paragraphs 2.5a and 2.5b. These are included for convenience, to shorten the count. However, they are not necessary; they could be replaced by a single rule to the effect that the count ends when there are no hopeful candidates left. We shall see below (in Propositions 5 and 6) that, left to its own devices in this way, QPQ will elect exactly s candidates. It shares this property with Meek-STV but not with conventional STV, in which the stopping condition of paragaph 2.5b is needed in order to ensure that enough candidates are elected.

The most important proportionality property possessed by STV is what I call the *Droop proportionality criterion*: if more than k Droop quotas of voters are solidly committed to the same set of $l \geqslant k$ candidates, then at least k of those l candidates should be elected. (Here the *Droop quota* is the total number of valid ballots divided by one more than the number of seats to be filled, and a voter is *solidly committed* to a set of l candidates if the voter lists those candidates, in some order, as the top l candidates on their ballot.) We shall see in Proposition 7 that QPQ also satisfies the Droop proportionality criterion.

We shall see in Proposition 4 that if two candidates a and b are elected in successive stages, first a and then b, with no exclusion taking place between them, then b's quotient at the time of b's election is no greater than a's quotient at the time of a's election. (Thus with the d'Hondt–Phragmén method, which is essentially the same as QPQ but with no quota and no exclusions, each candidate elected has a quotient that is no greater than that of the previous candidate elected.)

This is not necessarily true, however, if an exclusion occurs between the elections of a and b. Consider the following election.

Election 1 (3 seats)

16 ab, 12 b, 12 c, 12 d, 8 eb.

There are 60 votes, and so the quota is $60/4 = 15$. The initial quotients are the numbers of first-preference votes; a, with a quotient of 16, exceeds the quota and is elected. Now b's quotient becomes $(16 + 12)/2 = 14$, and this is the only quotient to change, so that no other candidate reaches the quota. Thus e is excluded. Now b's quotient becomes $(16+12+8)/2=18$, and so b is elected with a quotient that is larger than a's was at the time of a's election. This means that each of the ab ballots was deemed to have elected $\frac{1}{16}$ of a candidate after a's election, but only $\frac{1}{18}$ of a candidate after b's election. This conveys the impression that these ballots have elected a negative proportion of b, or else (perhaps worse) that the b and eb ballots are being treated as having elected part of a.

To avoid this, it is proposed here that the count should be restarted from scratch after each exclusion. We shall see below, in Proposition 8, that if c is the first candidate to be excluded, and the count is then restarted with c's name deleted from all ballots, then all the candidates who were elected before c's exclusion will be elected again (although not necessarily first or in the same order). With this variant of the method, the count is divided into rounds, each of which apart from the last ends with an exclusion; the last round involves the election of s candidates in s successive stages, with no intervening exclusions. Now no ballot can ever be regarded as contributing a negative amount to any candidate, or a positive amount to a candidate not explicitly mentioned on it.

With Meek's method, a voter can tell from the result sheet exactly how their vote has been divided between the candidates mentioned on their ballot, and therefore how much they have contributed to the election of each candidate. QPQ does not explicitly divide votes between candidates; but with the multi-round version just described, as with the d'Hondt–Phragmén method itself, a voter can tell from the result sheet what proportion of each candidate they have elected; and multiplying these proportions by the final quota could be regarded as indicating how much of their vote has gone to each candidate, implicitly if not explicitly. For example, suppose candidates a and b are elected with quotients (at the time of election) $q_a > q_b$, candidate c is hopeful to the end, and the final quota is Q. Then a voter whose ballot (after the deletion of any excluded candidates) reads abc has elected $1/q_a$ of a, $1/q_b - 1/q_a$ of b, and was able to contribute $1/Q - 1/q_b$ towards the election of c (which, however, was insufficient to get c elected). And a voter whose ballot reads bac or bca has elected $1/q_b$ of b, nothing of a, and was again able to contribute $1/Q - 1/q_b$ towards the election of c. The fact that the abc and bac voters make the same contribution to c is a property that is shared with Meek-STV but not with conventional STV.

3 Examples

The first of these examples is intended to clarify the method of calculation of the quotients and quota.

Election 2 (3 seats)

5 a, 15 abc, 15 ac, 10 b, 15 bc,
20 c, 15 d, 5 e.

There are 100 votes, and so the initial quota is $100/4 = 25$. The initial quotients are the numbers of first-preference votes; a's quotient of 35 is the largest, and exceeds the quota, and so a is elected. Each of the 35 ballots that has a in first place is deemed to have elected $\frac{1}{35}$ of a; 5 of these plump for a and now become inactive, 15 have b in second place, and 15 have c in second place. So the quota now becomes $(100-5)/(4-\frac{5}{35}) \approx 24.62$, b's quotient becomes $(25+15)/(1+\frac{15}{35}) = 28.0$, and c's quotient becomes $(20+15)/(1+\frac{15}{35}) = 24.5$. Now b's quotient exceeds the quota, and so b is elected. Each of the 40 ballots that contributed to b's election is deemed to have elected $\frac{1}{28}$ of a candidate *in total*; 10 of these plump for b and now become inactive, and the remaining 30 have c in the place after b. So the quota now becomes $(100-5-10)/(4-\frac{5}{35}-\frac{10}{28}) \approx 24.29$, and c's quotient becomes $(20+15+30)/(1+\frac{15}{35}+\frac{30}{28}) = 26.0$. Now c is elected. We can set out the count as follows.

Election 2

	quotients					quota	result
	a	b	c	d	e		
Stage 1	35	25	20	15	5	25.00	a elected
Stage 2	–	28	$24\frac{1}{2}$	15	5	24.62	b elected
Stage 3	–	–	26	15	5	24.29	c elected

We have already mentioned that QPQ satisfies the Droop proportionality criterion, which is one important test of proportionality. The next two elections provide another test of proportionality. In both of these there are two parties, one with candidates a, b, c and the other with candidates d, e, f. The voters vote strictly along party lines. However, the abc-party voters all put a first, b second and c third, whereas the def-party voters are evenly divided among the three candidates. In Election 3, the abc party has just over half the votes, and so we expect it to gain 3 of the 5 seats, whereas in Election 4 it has just under half the votes, and so we expect it to gain only 2 seats. We shall see that this is what happens.

Election 3 (5 seats)

306	abc
99	def
98	efd
97	fde

Election 4 (5 seats)

294	abc
103	def
102	efd
101	fde

In each case there are 600 votes, and so the quota is $600/6 = 100$. In Election 3, after the election of a, b and c the abc ballots become inactive, and, since these ballots are electing 3 seats, the quota reduces to $294/(6-3) = 98$. The counts proceed as follows.

Election 3

	quotients						quota	result
	a	b	c	d	e	f		
Stage 1	306	0	0	99	98	97	100	a elected
Stage 2	–	153	0	99	98	97	100	b elected
Stage 3	–	–	102	99	98	97	100	c elected
Stage 4	–	–	–	99	98	97	98	d elected
Stage 5	–	–	–	–	$98\frac{1}{2}$	97	98	e elected

Election 4

	quotients						quota	result
	a	b	c	d	e	f		
Stage 1	294	0	0	103	102	101	100	a elected
Stage 2	–	147	0	103	102	101	100	b elected
Stage 3	–	–	98	103	102	101	100	d elected
Stage 4	–	–	98	–	$102\frac{1}{2}$	101	100	e elected
Stage 5	–	–	98	–	–	102	100	f elected

We see that in each case the result is the one expected by proportionality. This is the same result as is obtained using STV (using the Droop quota—but not if the Hare quota is used).

In a single-seat election, QPQ and STV both reduce to the Alternative Vote. It is not clear how many seats and candidates are needed for QPQ to give a different result from Meek-STV, but here is an example with three seats and five candidates.

Election 5 (3 seats)

12 $acde$, 11 b, 7 cde, 8 dec, 9 ecd.

There are 47 votes, and so the quota (in STV or QPQ) is $47/4 = 11\frac{3}{4}$. STV elects a with a surplus of $\frac{1}{4}$ of a vote, which goes to c. No other candidate exceeds the quota, and so c, having the smallest vote, is excluded. Now d is elected with a surplus of $3\frac{1}{2}$ votes, which all goes to e, causing e to be elected. In QPQ, each candidate's initial quotient is their number of first-preference votes. So a is elected, and c's quotient then becomes $(12+7)/2 = 9\frac{1}{2}$. The candidate with the smallest quotient is now d, and so d is excluded. If the election is not restarted at this point, e now has a quotient of 17 and is elected, and this gives c a quotient of $(12+7+8+9)/3 = 12$ so that c is elected. If the election is restarted after d's exclusion, then e is elected first, and then there is a tie between a and c for the second place; whichever gets it, the other will get the third place. So in all cases the results are: STV: a, d, e; QPQ: a, c, e.

4 Proofs

In this section we will use the term *single-round QPQ* to refer to the version where one does not restart the count after an exclusion, and *multi-round QPQ* to refer to the version where one does. In the event that no exclusion occurs, both methods proceed identically, being then equivalent to the d'Hondt–Phragmén method. 'A count in which no exclusions occur' could refer to this possibility, in which exclusions are absent by chance, but it covers also the final round of a multi-round QPQ count, which is guaranteed to be free of exclusions; this final round is again equivalent to d'Hondt–Phragmén, although applied to ballots from which some candidates may already have been deleted.

It will be helpful to start by recalling some simple inequalities.

Proposition 1. *If m, n, x, y are positive real numbers such that $m/n \leqslant x/y$, then*

$$\frac{m}{n} \leqslant \frac{m+x}{n+y} \leqslant \frac{x}{y}\,. \tag{1.1}$$

If, in addition, $y < n$, then

$$\frac{m-x}{n-y} \leqslant \frac{m}{n}\,. \tag{1.2}$$

Proof. Since the denominators are all positive, the conclusions are equivalent to the inequalities $m(n+y) \leqslant (m+x)n$, $(m+x)y \leqslant x(n+y)$, and $(m-x)n \leqslant m(n-y)$. These all follow from the hypothesis, which is that $my \leqslant xn$. ☐

Proposition 2. *During a multi-round QPQ count, the quota never increases.*

Proof. To obtain a contradiction, suppose that the quota does increase at some stage, and consider the first stage at which this happens. Let the quota at the start of this stage be $Q = v_{\mathrm{a}}/(1+s-t_{\mathrm{x}})$, where v_{a} is the number of active ballots at the start of this stage, and t_{x} is the sum of the fractional numbers of candidates that are deemed to have been elected by all the *inactive* ballots at the start of this stage. For each active ballot that becomes inactive in this stage, the effect is to subtract 1 from v_{a} and add t to t_{x}, where t is the fractional number of candidates that that ballot has elected. This number t is either 0 or $1/q$, where q is the quotient possessed by some already-elected candidate at the time of their election. In order for this candidate to have been elected, necessarily q was greater than the quota at that time, which we are supposing was at least Q. Thus in all cases $t < 1/Q$. It follows that if x ballots become inactive in the current stage, then the effect is to subtract x from v_{a} and add a number $y < x/Q$ to t_{x}. Let Q' be the quota at the end of the current stage. If $y = 0$ then clearly $Q' < Q$. If $y \neq 0$ then $Q < x/y$, so that (1.2) gives

$$Q' = \frac{v_{\text{a}} - x}{1 + s - t_{\text{x}} - y} \leqslant \frac{v_{\text{a}}}{1 + s - t_{\text{x}}} = Q.$$

This contradicts the supposition that the quota increases in the current stage, and this contradiction proves the result. □

Proposition 3. *In any QPQ count, if a is elected with quotient q_a, and b is a hopeful candidate whose quotients at the start and end of the stage in which a is elected are q_b and q'_b respectively, then $q_b \leqslant q'_b \leqslant q_a$.*

Proof. Clearly $q_b \leqslant q_a$, since otherwise a would not have been elected in this stage. Suppose there are x ballots that contribute to a at the start of this stage and to b at the end of this stage, and let $y = x/q_a$, so that $x/y = q_a \geqslant q_b$. Then, after a's election, each of these x candidates is deemed to have elected $1/q_a$ candidates, so that collectively they have elected y candidates. If at the start of the current stage there were v_b ballots contributing to b, which collectively had already elected t_b candidates, then

$$q_b = \frac{v_b}{1 + t_b} \leqslant q'_b = \frac{v_b + x}{1 + t_b + y} \leqslant \frac{x}{y} = q_a$$

by (1.1). □

Proposition 4. *In a QPQ count in which no exclusions occur, each candidate to be elected has a quotient (at the time of election) that is no larger than the quotient (at the time of election) of the previous candidate to be elected.*

Proof. If candidates a and b are elected in successive stages, with quotients q_a and q'_b respectively, and if b's quotient at the start of the stage in which a is elected is q_b, then $q_b \leqslant q'_b \leqslant q_a$ by Proposition 3. In particular, $q'_b \leqslant q_a$, which is all we have to prove. □

Proposition 5. *Even if the stopping condition in paragraph 2.5a is deleted, it is not possible for more than s candidates to be elected by any form of QPQ (single-round or multi-round).*

Proof. Suppose it is. Consider the stage in which the $(s+1)$th candidate, c, is elected. At the start of this stage, let the quota be Q; let there be v_c ballots contributing to c, and suppose these v_c ballots collectively are currently electing t_c candidates; let there be v_{o} ballots contributing to other hopeful candidates, which are currently electing t_{o} candidates; let the number of active ballots be $v_{\text{a}} = v_c + v_{\text{o}}$; and let the number of candidates being elected by the inactive ballots be $t_{\text{x}} = s - t_c - t_{\text{o}}$. As in the proof of Proposition 2, every ballot has elected at most $1/Q$ candidates, and so $t_{\text{o}} \leqslant v_{\text{o}}/Q$. Thus

$$\frac{v_{\text{o}}}{t_{\text{o}}} \geqslant Q = \frac{v_{\text{a}}}{1 + s - t_{\text{x}}} = \frac{v_c + v_{\text{o}}}{1 + t_c + t_{\text{o}}},$$

and, by (1.2), c's quotient q_c satisfies

$$q_c = \frac{v_c}{1 + t_c} = \frac{(v_c + v_{\text{o}}) - v_{\text{o}}}{(1 + t_c + t_{\text{o}}) - t_{\text{o}}} \leqslant \frac{v_c + v_{\text{o}}}{1 + t_c + t_{\text{o}}} = Q.$$

This shows that c cannot be elected in the current stage, and this contradiction shows that at most s candidates are elected in total. □

Proposition 6. *Even if the stopping condition in paragraph 2.5b is deleted, at least s candidates must be elected by any form of QPQ (single-round or multi-round).*

Proof. Suppose this is not true, and consider the stage in which the number of nonexcluded candidates first falls below s. Suppose that at the start of this stage there are e elected candidates and (therefore) $s - e$ hopeful candidates. Since no hopeful candidate has a quotient greater than the quota,

$$v_c \leqslant Q(1 + t_c) \tag{1.3}$$

for every hopeful candidate c, where Q is the quota, v_c is the number of ballots contributing to c, and t_c is the number of candidates that these ballots collectively have elected, all measured at the start of the current stage. Now, the sum of the $s - e$ numbers v_c is v_{a}, the number of active ballots, and the sum of the $s - e$ numbers t_c is the number of candidates elected by all the active ballots, which is $e - t_{\text{x}}$, where t_{x} is the number of candidates elected by the inactive ballots. So summing (1.3) over all $s-e$ hopeful candidates gives $v_{\mathrm{a}} \leqslant Q(s-e+e-t_{\mathrm{x}}) = Q(s-t_{\mathrm{x}})$. Thus

$$Q = \frac{v_{\mathrm{a}}}{1+s-t_{\mathrm{x}}} < \frac{v_{\mathrm{a}}}{s-t_{\mathrm{x}}} \leqslant Q.$$

This contradiction shows that at least one of the $s-e$ hopeful candidates must have a quotient greater than the quota Q, and so be elected in the current stage. This contradicts the supposition that the number of nonexcluded candidates falls in the current stage, and this contradiction proves the result. □

Propositions 5 and 6 together show that, left to its own devices, QPQ will always elect the right number of candidates; the only stopping condition required is that the election must terminate when there are no hopeful candidates left.

Proposition 7. *Every form of QPQ satisfies the Droop proportionality criterion: if more than k Droop quotas of voters are solidly committed to the same set of $l \geqslant k$ candidates, then at least k of those l candidates must be elected.*

Proof. The argument is rather similar to the proof of the previous proposition. Let L be the set of l candidates in question. In view of Proposition 5, we may assume that the stopping condition in paragraph 2.5a is deleted, so that the count cannot end because we have elected too many candidates outside L. Thus if Proposition 7 is not true then there must come a point in some election at which the number of nonexcluded candidates in L falls below k. Consider the stage in which this happens. Suppose that at the start of this stage there are e elected candidates and (therefore) $k-e$ hopeful candidates in L. Since no hopeful candidate has a quotient greater than the quota Q, (1.3) holds for all these $k-e$ hopeful candidates. Since the quota at the start of the count was equal to the Droop quota, and, by Proposition 2, the quota never increases, the number of ballots solidly committed to L is greater than kQ, and so the sum of the $k-e$ numbers v_c is greater than kQ. Moreover, none of these ballots can have contributed to electing any candidate outside L, and so the sum of the $k-e$ numbers t_c is at most e. So summing (1.3) over all $k-e$ hopeful candidates in L gives

$$Qk < \sum_c v_c \leqslant Q\Big(\sum_c (1+t_c)\Big) \leqslant Q(k-e+e) = Qk.$$

This contradiction shows that at least one of the hopeful candidates in L must have a quotient that is greater than Q, and so the number of nonexcluded candidates in L cannot fall in the current stage. This contradiction in turn proves the result. □

Proposition 8. *Suppose that in the first k stages of a QPQ count candidates $a_1, \ldots, a_k$ are elected (in that order) with quotients $q_1, \ldots, q_k$ respectively, and in the $(k+1)$th stage candidate b is excluded. Suppose that the count is restarted with b's name deleted from every ballot. Then, in the new count, candidates $a_1, \ldots, a_k$ will all be elected before any exclusions take place, and each candidate a_i will have quotient at least q_i at the time of their election.*

Proof. Suppose that in the first count the quota at the time of a_i's election is Q_i, so that $q_i > Q_i$, for each i. The deletion of b cannot decrease any candidate's initial quotient, nor increase the quota, and so at the start of the new count a_1 has quotient at least q_1 and the quota is at most Q_1. Since, by Propositions 2 and 3, the election of other candidates cannot increase the quota nor decrease a_1's quotient, a_1 will have a quotient greater than the quota as long as a_1 remains hopeful. Thus a_1 will eventually be elected, before any exclusions take place, with a quotient that is at least q_1.

In order to obtain a contradiction, suppose that the conclusion of the Proposition does not hold for all these values of i, and consider the smallest value of i for which it fails to hold. Then $i \geqslant 2$, since we have just seen that the conclusion holds for a_1. Consider the first point at which $a_1, a_2, \ldots, a_{i-1}$ are all elected, and let a_j be the last of these candidates to be elected; a_j may, but need not, be a_{i-1}. Since the conclusion holds for all of $a_1, a_2, \ldots, a_{i-1}$, we know that a_j had quotient at least q_j at the time of election. By Proposition 4 applied to the first count and then to the new count, $q_j \geqslant q_i$, and every candidate elected so far in the new count has been elected with a quotient that is at least q_j and hence at least q_i. So if a_i has already been elected in the new count then the conclusion of the Proposition holds for a_i. Since we are supposing that this is not the case, it must be that a_i has not yet been elected. We will consider a_i's quotient and the quota at the start of the next stage, immediately following the election of a_j.

In the first count, a_i was elected with quotient $q_i = v_i/(1+t_i)$, where v_i is the number of ballots that contributed to a_i after a_{i-1}'s election, and t_i is the fractional number of candidates that these ballots had so far elected. These v_i ballots are the ones on which no candidate other than $a_1, \ldots, a_{i-1}$ is preferred to a_i, and so they again contribute to a_i at this point in the new count. So in the new count, a_i now has quotient $\hat{q}_i = (v_i + v'_i)/(1 + \hat{t}_i + \hat{t}'_i)$, where v'_i is the number of ballots contributing to a_i at this point that did not contribute to a_i at the time of a_i's election in the first count, and $\hat{t}_i$ and $\hat{t}'_i$ are the fractional numbers of candidates elected by the original v_i contributors and the new v'_i contributors at this point in the new count. Each of these $v_i + v'_i$ ballots is deemed to have elected either 0 candidates or a number of candidates of the form $1/\hat{q}$, where $\hat{q}$ is the smallest quotient of any elected candidate listed above a_i on that ballot. For all the ballots of this second type, $\hat{q} \geqslant q_j \geqslant q_i$; thus $\hat{t}'_i \leqslant v'_i/q_i$ and $v'_i/\hat{t}'_i \geqslant q_i$. Moreover, for each of the original v_i ballots that is of this second type, the number $\hat{q}$ for that ballot is the smallest of a new set of quotients, each of which is at least as large as the corresponding quotient in the original count, so that if the ballot was electing $1/q$ candidates at the time of a_i's election in the original count then $\hat{q} \geqslant q$ and $1/\hat{q} \leqslant 1/q$; thus $\hat{t}_i \leqslant t_i$. It follows from (1.1) that

$$\hat{q}_i = \frac{v_i + v'_i}{1 + \hat{t}_i + \hat{t}'_i} \geqslant \frac{v_i + v'_i}{1 + t_i + \hat{t}'_i} \geqslant \frac{v_i}{1 + t_i} = q_i. \quad (1.4)$$

Now let us consider the quota. Let v be the number of valid ballots. In the first count, the quota at the time of a_i's election was $Q_i = (v - v_x)/(1 + s - t_x)$, where v_x is the number of *inactive* ballots at the time of a_i's election, and t_x is the fractional number of candidates that these ballots have elected. These v_x inactive ballots are the ones that contain the name of no candidates other than $a_1, \ldots, a_{i-1}$, and so they are again inactive at this point in the new count. So in the new count, the quota at this point is $\hat{Q}_i = (v - v_x - v'_x)/(1 + s - \hat{t}_x - \hat{t}'_x)$, where v'_x is the number of ballots that were active at the time of a_i's election in the first count but are inactive at this point in the new count, and $\hat{t}_x$ and $\hat{t}'_x$ are the fractional numbers of candidates elected by the original and the new inactive ballots at this point in the new count. By the same argument we used in the previous paragraph to prove that $\hat{t}_i \leqslant t_i$, we can now deduce that $\hat{t}_x \leqslant t_x$. Moreover, by Propositions 2 and 4 and the criterion for election in paragraph 2.5a, every candidate elected so far has been elected with a quotient that is greater than the current quota $\hat{Q}_i$, so that $\hat{t}'_x \leqslant v'_x/\hat{Q}_i$ and $v'_x/t'_x \geqslant \hat{Q}_i$. It follows from (1.1) that

$$\hat{Q}_i = \frac{v - v_x - v'_x}{1 + s - \hat{t}_x - \hat{t}'_x} \leqslant \frac{v - v_x}{1 + s - \hat{t}_x} \leqslant \frac{v - v_x}{1 + s - t_x} = Q_i. \quad (1.5)$$

It follows from (1.4) and (1.5) that $\hat{q}_i \geqslant q_i > Q_i \geqslant \hat{Q}_i$, so that a_i's current quotient is greater than the current quota. Since, by Propositions 2 and 3, the election of other candidates cannot increase the quota nor decrease a_i's quotient, a_i will have a quotient greater than the quota as long as a_i remains hopeful. Thus a_i will eventually be elected, before any exclusions take place, with a quotient that is at least q_i. This contradicts the supposition that the conclusion of the Proposition failed to hold for a_i, and this contradiction completes the proof of the Proposition. □

5 References

[1] V. d'Hondt, *Système Pratique et Raisonné de Représentation Proportionelle [A Practical and Rational System for Proportional Representation]*, Librairie C. Muquardt, Brussels, 1882.

[2] I. D. Hill, B. A. Wichmann and D. R. Woodall, Algorithm 123, Single transferable vote by Meek's method, *The Computer Journal* **30** (1987) 277–281.

[3] R. A. Newland and F. S. Britton, *How to Conduct an Election by the Single Transferable Vote*, Electoral Reform Society of Great Britain and Ireland, London, 1972; 2nd ed. 1976.

[4] E. Phragmén, Proportionella val [Proportional elections], *Svenska Spörsmål* **25**, Stockholm, 1895.

[5] E. Phragmén, Till frågan om en proportionell valmetod [On the question of a proportional election method], *Statsvetenskaplig Tidskrift* **2** (1899), 87–95.

[6] O. Salmi, RE: [EM] D'Hondt without lists, http://groups.yahoo.com/group/election-methods-list/message/10068.

Fuller Disclosure than Intended

Joe Otten
Email: joe@datator.co.uk

1 Introduction

The full disclosure of preferences in the case of an STV election carries one danger of abuse. That is the potential for a unique preference list to identify a particular voter. Suppose there are 10 candidates in an election. Then there are $10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 = 3\ 628\ 800$ possible complete preference lists as well as a number of incomplete lists. In an electorate of a few tens or hundreds of thousands, it is obvious that the vast majority of the possible preference lists will not be used.

Of the preference lists that are used, they will generally follow some sort of pattern, such as the candidates of one party, followed by the candidates of another party, etc. It will therefore be fairly easy to create a large number of different preference lists that favour a particular candidate (with first preferences), and are most unlikely to be used by any voter.

2 The problem

The full disclosure of preference data facilitates the following fraud: The fraudster bribes or coerces a large number of voters to vote according to an exact preference list that is provided, and is different for each voter. The preference lists provided will be different unlikely sequences, such as the preferred candidate followed by alternate liberals and fascists or conservatives and communists.

Disclosure of the full preference data will then disclose, with a high probability, the voting behaviour of the bribed voters. There may be some false positives, but there will be no false negatives — i.e. if a preference list is missing then it is certain that a bribed voter welched.

3 The solution

One solution has been proposed — that of anonymising the preference data in a similar way to how census data is anonymised. Changes are made to the individual records in such a way as to minimise changes that result to any statistical aggregates an analyst might be interested in. The problem with this is that the statistical analysis of preference data is in such infancy that it is not clear what aggregates should be preserved, or how they might be preserved.

My preferred solution is that prior to disclosure, preference lists should be aggregated by censoring lower preferences until there are at least, say, 3 instances of every preference list to be published. So for example, if there are 10 votes of ABCDEFG then that fact can be published. If there is 1 vote of BCDEFGA, 1 of BCDEFAG and 1 of BCDEGAF then the fact that there were 3 votes of BCDExxx would be published. This would mean that no single individual's vote would be identifiably disclosed.

A New Monotonic and Clone-Independent Single-Winner Election Method

M Schulze
Email: markus.schulze@alumni.tu-berlin.de

Markus Schulze has studied mathematics and physics at the Technische Universität Berlin.

1 Introduction

In 1997, I proposed to a large number of people who are interested in mathematical aspects of election methods a new method that satisfies Pareto, monotonicity, resolvability, independence of clones, reversal symmetry, Smith-IIA, and Schwartz. This method immediately attracted a lot of attention and very many enthusiastic supporters. Today, this method is promoted e.g. by Diana Galletly [1], Mathew Goldstein [2], Jobst Heitzig [3], Raul Miller, Mike Ossipoff [4], Russ Paielli, Norman Petry, Manoj Srivastava, and Anthony Towns and it is analyzed e.g. in the websites of Blake Cretney [5], Steve Eppley [6], Eric Gorr [7], and Rob LeGrand [8]. Today, this method is taught e.g. by James E. Falk of George Washington University and Thomas K. Yan of Cornell University [9]. In January 2003, the board of Software in the Public Interest (SPI) adopted this method unanimously [10]. In June 2003, the DEBIAN Project adopted this method with 144 against 16 votes [11, 12]. Therefore, a more detailed motivation and explanation of the method is overdue.

There has been some debate about an appropriate name for the method. Some people suggested names like "Beatpath Method", "Beatpath Winner", "Path Voting", "Schwartz Sequential Dropping" (SSD) or "Cloneproof Schwartz Sequential Dropping" (CSSD or CpSSD). However, I prefer the name "Schulze method", not because of academic arrogance, but because the other names do not refer to the method itself but to specific heuristics for implementing it, and so may mislead readers into believing that no other method for implementing it is possible. In my opinion, although it is advantageous to have an intuitive and convincing heuristic, in the end only the properties of the method are relevant.

I have already found some implementations of my method in the internet. Unfortunately, most implementations that I have seen were inefficient because the programmers have not understood the Floyd algorithm so that the implementations had a runtime of $O(N^5)$ although the winners of this method can be calculated in a runtime of $O(N^3)$, where N is the number of candidates.

It is presumed that each voter casts at least a *partial ranking* of all candidates. That means: It is presumed that for each voter V the relation "voter V strictly prefers candidate A to candidate B" is irreflexive, asymmetric, and transitive on the set of candidates. But it is not presumed that each voter casts a *complete ranking*. That means: It is not presumed that this relation is also linear.

Suppose that d[X,Y] is the number of voters who strictly prefer candidate X to candidate Y. Then the *Smith set* is the smallest non-empty set of candidates with d[A,B] $>$ d[B,A] for each candidate A of this set and each candidate B outside this set. *Smith-IIA* (where IIA means Independence from Irrelevant Alternatives) says that adding a candidate who is not in the new Smith set should not change the probability that a given and already running candidate is elected. Smith-IIA implies the majority criterion for solid coalitions and the Condorcet criterion. Unfortunately, compliance with the Condorcet criterion implies violation of other desired criteria like participation [13], later-no-harm, and later-no-help [14].

A *chain from candidate A to candidate B* is an ordered set of candidates C(1),...,C(n) with the following three properties:

1. C(1) is identical to A.
2. C(n) is identical to B.
3. d[C(i),C(i+1)] − d[C(i+1),C(i)] > 0 for each i = 1,...,(n−1).

A *Schwartz winner* is a candidate A who has chains at least to every other candidate B who has a chain to candidate A. The *Schwartz set* is the set of all Schwartz winners. *Schwartz* says that the winner must be a Schwartz winner.

In section 2, the Schulze method is defined. In section 3, well-definedness of this method is proven. In section 4, I present an implementation with a runtime of $O(N^3)$. In section 5, I prove that this method satisfies Pareto, monotonicity, resolvability, independence of clones, and reversal symmetry. From the definition of the Schulze method, it is clear that this method meets Smith-IIA and Schwartz.

Another election method that satisfies Pareto, monotonicity, resolvability, independence of clones, reversal symmetry, Schwartz, and Smith-IIA is Tideman's Ranked Pairs method [15, 16]. However, appendix A demonstrates that the proposed method is not identical with the Ranked Pairs method. Appendix B demonstrates that the proposed method can violate the participation criterion in a very drastic manner. A special provision of the implementation used by SPI and DEBIAN is described in appendix C. Appendix D explains how the proposed method can be interpreted as a method where successively the weakest pairwise defeats are "eliminated." Appendix E presents a concrete example where the proposed method does not find a unique winner.

2 Definition of the Schulze Method

Stage 1: Suppose that d[A,B] is the number of voters who strictly prefer candidate A to candidate B.

A *path from candidate A to candidate B* is an ordered set of candidates C(1),...,C(n) with the following two properties:

1. C(1) is identical to A.
2. C(n) is identical to B.

The *strength* of the path C(1),...,C(n) is min { d[C(i),C(i+1)] − d[C(i+1),C(i)] | i = 1,...,(n−1) }.

Thus a chain from candidate A to candidate B, as defined in the introduction, is simply a path with positive strength.

p[A,B] := max { min { d[C(i),C(i+1)] − d[C(i+1),C(i)] | i = 1,...,(n−1) } | C(1),...,C(n) is a path from candidate A to candidate B }.

In other words: p[A,B] is the strength of the strongest path from candidate A to candidate B.

Candidate A is a *potential winner* if and only if p[A,B] ≥ p[B,A] for every other candidate B.

When p[A,B] > p[B,A], then we say: "Candidate A disqualifies candidate B".

Stage 2: If there is only one potential winner, then this potential winner is the *unique winner*. If there is more than one potential winner, then a *Tie-Breaking Ranking of the Candidates* (TBRC) is calculated as follows:

1. Pick a random ballot and use its rankings; consider ties as unsorted with regard to each other.
2. Continue picking ballots randomly from those that have not yet been picked. When you find one that orders previously unsorted candidates, use the ballot to sort them. Do not change the order of the already sorted.
3. If you go through all ballots, and some candidates are still not sorted, order them randomly.

The winner is that potential winner who is ranked highest in this TBRC.

3 Well-Definedness

On first view, it is not clear whether the Schulze method is well defined. It seems to be possible that candidates disqualify each other in such a manner that there is no candidate A with p[A,B] ≥ p[B,A] for every other candidate B. However, the following proof demonstrates that path defeats are transitive. That means: When candidate A disqualifies candidate B and when candidate B disqualifies candidate C, then also candidate A disqualifies candidate C.

Claim: (p[A,B] > p[B,A] and p[B,C] > p[C,B]) ⇒ p[A,C] > p[C,A].

Proof: Suppose
(1) p[A,B] > p[B,A] and
(2) p[B,C] > p[C,B].

The following statements are valid:
(3) min { p[A,B]; p[B,C] } $\leq$ p[A,C].
(4) min { p[A,C]; p[C,B] } $\leq$ p[A,B].
(5) min { p[B,A]; p[A,C] } $\leq$ p[B,C].
(6) min { p[B,C]; p[C,A] } $\leq$ p[B,A].
(7) min { p[C,A]; p[A,B] } $\leq$ p[C,B].
(8) min { p[C,B]; p[B,A] } $\leq$ p[C,A].

For example: If min { p[A,B]; p[B,C] } was strictly larger than p[A,C], then this would be a contradiction to the definition of p[A,C] since there would be a route from candidate A to candidate C via candidate B with a strength of more than p[A,C]; and if this route was not itself a path (because it passed through some candidates more than once) then some subset of its links would form a path from candidate A to candidate C with a strength of more than p[A,C].

Case 1: Suppose
(9a) p[A,B] $\geq$ p[B,C].
Combining (2) and (9a) gives:
(10a) p[A,B] $>$ p[C,B].
Combining (7) and (10a) gives:
(11a) p[C,A] $\leq$ p[C,B].
Combining (3) and (9a) gives:
(12a) p[B,C] $\leq$ p[A,C].
Combining (11a), (2), and (12a) gives:
(13a) p[C,A] $\leq$ p[C,B] $<$ p[B,C] $\leq$ p[A,C].

Case 2: Suppose
(9b) p[A,B] $<$ p[B,C].
Combining (1) and (9b) gives:
(10b) p[B,C] $>$ p[B,A].
Combining (6) and (10b) gives:
(11b) p[C,A] $\leq$ p[B,A].
Combining (3) and (9b) gives:
(12b) p[A,B] $\leq$ p[A,C].
Combining (11b), (1), and (12b) gives:
(13b) p[C,A] $\leq$ p[B,A] $<$ p[A,B] $\leq$ p[A,C].

Therefore, the relation defined by p[A,B] $>$ p[B,A] is transitive.

4 Implementation

The strength of the strongest path p[i,j] from candidate i to candidate j can be calculated with the Floyd algorithm [17]. The runtime to calculate the strengths of all paths is $O(N^3)$. It cannot be said frequently enough that the order of the indices in the triple-loop of the Floyd algorithm is **not** irrelevant.

Input: d[i,j] with i $\neq$ j is the number of voters who strictly prefer candidate i to candidate j.

Output: “w[i] = true” means that candidate i is a potential winner. “w[i] = false” means that candidate i is not a potential winner.

```
for i := 1 to N do
for j := 1 to N do
  if ( i ≠ j ) then
    p[i,j] := d[i,j] − d[j,i] ;

for i := 1 to N do
for j := 1 to N do
  if ( i ≠ j ) then
    for k := 1 to N do
      if ( i ≠ k ) then
        if ( j ≠ k ) then
          {
          s := min { p[j,i], p[i,k] } ;
          if ( p[j,k] < s ) then
            p[j,k] := s ;
          }

for i := 1 to N do
  {
  w[i] := true ;
  for j := 1 to N do
    if ( i ≠ j ) then
      if ( p[j,i] > p[i,j] ) then
        w[i] := false ;
  }
```

5 Properties

5.1 Pareto

Pareto says that when no voter strictly prefers candidate B to candidate A and at least one voter strictly prefers candidate A to candidate B then candidate B must not be elected.

The Schulze method meets Pareto.

Proof: Suppose no voter strictly prefers candidate B to candidate A and at least one voter strictly prefers candidate A to candidate B. Then d[A,B] $>$ 0 and d[B,A] = 0.

Case 1: If BA is already the strongest path from candidate B to candidate A, then p[B,A] = d[B,A] − d[A,B] < 0. Therefore, candidate A disqualifies candidate B because p[A,B] ≥ d[A,B] − d[B,A] > 0, so that p[A,B] > p[B,A].

Case 2: Suppose that B,C(1),…,C(n),A is the strongest path from candidate B to candidate A. As every voter who strictly prefers candidate B to candidate C(1) also necessarily strictly prefers candidate A to candidate C(1), we get d[A,C(1)] ≥ d[B,C(1)]. As every voter who strictly prefers candidate C(1) to candidate A also necessarily strictly prefers candidate C(1) to candidate B, we get d[C(1),B] ≥ d[C(1),A]. Therefore, d[A,C(1)] − d[C(1),A] ≥ d[B,C(1)] − d[C(1),B]. For the same reason, we get d[C(n),B] − d[B,C(n)] ≥ d[C(n),A] − d[A,C(n)]. Therefore, the path A,C(1),…,C(n),B is at least as strong as the path B,C(1),…,C(n),A. In so far as B,C(1),…,C(n),A is the strongest path from candidate B to candidate A by presumption, we get p[A,B] ≥ p[B,A].

Suppose that candidate B is a potential winner. Then also candidate A is a potential winner.

Proof: Suppose that B,C(1),…,C(n),X is the strongest path from candidate B to candidate X. Then, A,C(1),...,C(n),X is a path, but not necessarily the strongest path, from candidate A to candidate X with at least the same strength because d[A,C(1)] − d[C(1),A] ≥ d[B,C(1)] − d[C(1),B]. Therefore, p[A,X] ≥ p[B,X] for every candidate X other than candidate A or candidate B. Suppose that X,C(1),…,C(n),A is the strongest path from candidate X to candidate A. Then, X,C(1),...,C(n),B is a path, but not necessarily the strongest path, from candidate X to candidate B with at least the same strength because d[C(n),B] − d[B,C(n)] ≥ d[C(n),A] − d[A,C(n)]. Therefore, p[X,B] ≥ p[X,A] for every candidate X other than candidate A or candidate B.

Since candidate B is a potential winner, p[B,X] ≥ p[X,B] for every other candidate X. With p[A,X] ≥ p[B,X], p[B,X] ≥ p[X,B], and p[X,B] ≥ p[X,A], we get p[A,X] ≥ p[X,A] for every other candidate X. Therefore, also candidate A is a potential winner.

Therefore, when no voter strictly prefers candidate B to candidate A and at least one voter strictly prefers candidate A to candidate B then when candidate B is a potential winner also candidate A is a potential winner. Therefore, candidate B cannot be elected at stage 1 of the Schulze method. Candidate B cannot be elected at stage 2, either, since candidate A is necessarily ranked above candidate B in the TBRC.

5.2 Monotonicity

Monotonicity says that when some voters rank candidate A higher without changing the order in which they rank the other candidates relatively to each other then the probability that candidate A is elected must not decrease.

The Schulze method meets monotonicity.

Proof: Suppose candidate A was a potential winner. Then p_{old}[A,B] ≥ p_{old}[B,A] for every other candidate B.

Part 1: Suppose some voters rank candidate A higher without changing the order in which they rank the other candidates. Then d_{new}[A,X] ≥ d_{old}[A,X] and d_{new}[X,A] ≤ d_{old}[X,A] for every other candidate X. d_{new}[X,Y] = d_{old}[X,Y] when neither candidate X nor candidate Y is identical to candidate A. Therefore d_{new}[A,X] − d_{new}[X,A] ≥ d_{old}[A,X] − d_{old}[X,A] for every other candidate X. And d_{new}[X,Y] − d_{new}[Y,X] = d_{old}[X,Y] − d_{old}[Y,X] when neither candidate X nor candidate Y is identical to candidate A. For every candidate B other than candidate A the value p[A,B] can only increase but not decrease with d[A,X] − d[X,A] since only AX but not XA can be in the strongest path from candidate A to candidate B and the value p[B,A] can only decrease but not increase with d[A,X] − d[X,A] since only XA but not AX can be in the strongest path from candidate B to candidate A. Therefore p_{new}[A,B] ≥ p_{old}[A,B] and p_{new}[B,A] ≤ p_{old}[B,A]. Therefore p_{new}[A,B] ≥ p_{new}[B,A] so that candidate A is still a potential winner.

Part 2: Suppose that candidate E is not identical to candidate A. It remains to be proven that when candidate E was not a potential winner before then he is still not a potential winner. Suppose that candidate E was not a potential winner. Then there must have been a candidate F other than candidate E with

(1) p_{old}[F,E] > p_{old}[E,F].

Then, of course, also p_{new}[F,E] > p_{new}[E,F] is valid unless XA was a weakest link in the strongest path from candidate F to candidate E and/or AY was the weakest link in the strongest path from candidate E to candidate F. Without loss of generality, we can presume that candidate F is not identical to candidate A and that

(2) p_{old}[A,E] = p_{old}[E,A]

because otherwise with p_{old}[A,E] $>$ p_{old}[E,A] we would immediately get p_{new}[A,E] $>$ p_{new}[E,A] (because of the considerations in Part 1) so that we would immediately get that candidate E is still not a potential winner. Since candidate A was a potential winner, we get

(3) p_{old}[A,F] $\geq$ p_{old}[F,A].

The following statements are valid for the same reason as in section 3:

(4) min { p_{old}[A,E]; p_{old}[E,F] } $\leq$ p_{old}[A,F].
(5) min { p_{old}[A,F]; p_{old}[F,E] } $\leq$ p_{old}[A,E].
(6) min { p_{old}[E,A]; p_{old}[A,F] } $\leq$ p_{old}[E,F].
(7) min { p_{old}[E,F]; p_{old}[F,A] } $\leq$ p_{old}[E,A].
(8) min { p_{old}[F,A]; p_{old}[A,E] } $\leq$ p_{old}[F,E].
(9) min { p_{old}[F,E]; p_{old}[E,A] } $\leq$ p_{old}[F,A].

Case 1: Suppose XA was a weakest link in the strongest path from candidate F to candidate E. Then

(10a) p_{old}[F,E] = p_{old}[F,A] and
(11a) p_{old}[A,E] $\geq$ p_{old}[F,E].
Now (3), (10a), and (1) give
(12a) p_{old}[A,F] $\geq$ p_{old}[F,A] = p_{old}[F,E] $>$ p_{old}[E,F],
while (2), (11a), and (1) give
(13a) p_{old}[E,A] = p_{old}[A,E] $\geq$ p_{old}[F,E] $>$ p_{old}[E,F].
But (12a) and (13a) together contradict (6).

Case 2: Suppose AY was the weakest link in the strongest path from candidate E to candidate F. Then

(10b) p_{old}[E,F] = p_{old}[A,F] and
(11b) p_{old}[E,A] $>$ p_{old}[E,F].
Now (11b), (10b), and (3) give
(12b) p_{old}[E,A] $>$ p_{old}[E,F] = p_{old}[A,F] $\geq$ p_{old}[F,A],
while (1), (10b), and (3) give
(13b) p_{old}[F,E] $>$ p_{old}[E,F] = p_{old}[A,F] $\geq$ p_{old}[F,A].
But (12b) and (13b) together contradict (9).

Conclusion: When some voters rank candidate A higher without changing the order in which they rank the other candidates relatively to each other, then (a) when candidate A was a potential winner candidate A is still a potential winner and (b) every other candidate E who was not a potential winner is still not a potential winner and (c) candidate A can only increase in the TBRC while the positions of the other candidates are not changed relatively to each other. Therefore, the probability that candidate A is elected cannot decrease.

5.3 Resolvability

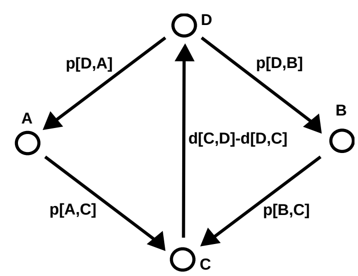
Resolvability says that at least in those cases in which there are no pairwise ties and there are no pairwise defeats of equal strength the winner must be unique.

The Schulze method meets resolvability.

Proof: Suppose that there is no unique winner. Suppose that candidate A and candidate B are potential winners. Then:

(1) p[A,B] = p[B,A].

Suppose that there are no pairwise ties and that there are no pairwise defeats of equal strength. Then p[A,B] = p[B,A] means that the weakest link in the strongest path from candidate A to candidate B and the weakest link in the strongest path from candidate B to candidate A must be the same link, say CD. Then this situation looks as follows:



As the weakest link of the strongest path from candidate B to candidate A is CD, we get:
(2) p[D,A] $>$ p[B,A].
As the weakest link of the strongest path from candidate A to candidate B is CD, we get:
(3) p[A,D] = p[A,B].
With (2), (1), and (3) we get:
(4) p[D,A] $>$ p[B,A] = p[A,B] = p[A,D] which contradicts the presumption that candidate A is a potential winner.

5.4 Independence of Clones

An election method is *independent of clones* if the following holds:

Suppose that candidate D and candidate E are two different candidates.

1. Suppose (a) that there is at least one voter who either strictly prefers candidate D to candidate E or strictly prefers candidate E to candidate D or (b) that candidate D is elected with zero probability.
2. Suppose that candidate D is replaced by a set of candidates D(1),...,D(m) in such a manner that for every candidate D(i) in this set, for every candidate F outside this set, and for every voter V the following two statements are valid:
   a) V strictly preferred D to F ⇔ V strictly prefers D(i) to F.
   b) V strictly preferred F to D ⇔ V strictly prefers F to D(i).

Then the probability that candidate E is elected must not change.

The Schulze method is independent of clones.

Proof: Suppose that candidate D is replaced by a set of candidates D(1),…,D(m) in the manner described above. Then d_{new}[A,D(i)] = d_{old}[A,D] for every candidate A outside the set D(1),…,D(m) and for every i = 1,…,m. And d_{new}[D(i),B] = d_{old}[D,B] for every candidate B outside the set D(1),…,D(m) and for every i = 1,…,m.

(1) Case 1: Suppose that the strongest path C(1),…,C(n) from candidate A to candidate B did not contain candidate D. Then C(1),…,C(n) is still a path from candidate A to candidate B with the same strength. Therefore: p_{new}[A,B] ≥ p_{old}[A,B].

Case 2: Suppose that the strongest path C(1),…,C(n) from candidate A to candidate B contained candidate D. Then C(1),…,C(n) with D replaced by an arbitrarily chosen candidate D(i) is still a path from candidate A to candidate B with the same strength. Therefore: p_{new}[A,B] ≥ p_{old}[A,B].

(2) Case 1: Suppose that the strongest path C(1),…,C(n) from candidate A to candidate B does not contain candidates of the set D(1),…,D(m). Then C(1),…,C(n) was a path from candidate A to candidate B with the same strength. Therefore: p_{old}[A,B] ≥ p_{new}[A,B].

Case 2: Suppose that the strongest path C(1),…,C(n) from candidate A to candidate B contains some candidates of the set D(1),…,D(m). Then C(1),…,C(n) where the part of this path from the first occurrence of a candidate of the set D(1),…,D(m) to the last occurrence of a candidate of the set D(1),…,D(m) is replaced by candidate D was a path from candidate A to candidate B with at least the same strength. Therefore: p_{old}[A,B] ≥ p_{new}[A,B].

With (1) and (2), we get: p_{new}[A,B] = p_{old}[A,B].

When we set A ≡ D in (1) and (2), we get: p_{new}[D(i),B] = p_{old}[D,B] for every candidate B outside the set D(1),…,D(m) and for every i = 1,…,m.

When we set B ≡ D in (1) and (2), we get: p_{new}[A,D(i)] = p_{old}[A,D] for every candidate A outside the set D(1),…,D(m) and for every i = 1,…,m.

Suppose candidate A, who is not identical to candidate D, was a potential winner, then p_{old}[A,B] ≥ p_{old}[B,A] for every other candidate B; because of the above considerations we get p_{new}[A,B] ≥ p_{new}[B,A] for every other candidate B; therefore, candidate A is still a potential winner. Suppose candidate B, who is not identical to candidate D, was not a potential winner, then p_{old}[B,A] < p_{old}[A,B] for at least one other candidate A; because of the above considerations we get p_{new}[B,A] < p_{new}[A,B] for at least this other candidate A; therefore, candidate B is still not a potential winner.

Presumption 1 in the definition of independence of clones guarantees that at least in those situations in which the TBRC has to be used to choose from the candidates D(1),…,D(m),E (a) candidate E is ranked above each of the candidates D(1),...,D(m) when he was originally ranked above candidate D. (b) candidate E is ranked below each of the candidates D(1),…,D(m) when he was originally ranked below candidate D. Therefore, replacing candidate D by a set of candidates D(1),…,D(m) can neither change whether candidate E is a potential winner nor, when the TBRC has to be used, where this candidate is ranked in the TBRC.

5.5 Reversal Symmetry

Reversal symmetry says that when candidate A is the unique winner then when the individual preferences of each voter are inverted then candidate A must not be elected.

The Schulze method meets reversal symmetry.

Proof: Suppose candidate A was the unique winner. Then there must have been at least one other candidate B with p_{old}[A,B] > p_{old}[B,A]. (Since the relation defined by p[X,Y] > p[Y,X] is transitive there must have been at least one candidate B other than candidate A with p[B,E] ≥ p[E,B] for every candidate E other than candidate A or candidate B. Since candidate A was the unique winner and since no candidate other than candidate A has disqualified candidate B, candidate A must have disqualified candidate B, i.e. p_{old}[A,B] > p_{old}[B,A].)

When the individual preferences of each voter are inverted then d_{new}[Y,X] = d_{old}[X,Y] for each pair XY of candidates. When C(1),…,C(n) was a path from candidate X to candidate Y of strength Z then C(n),…,C(1) is a path from candidate Y to candidate X of strength Z. Therefore, p_{new}[Y,X] = p_{old}[X,Y] for each pair XY of candidates. Therefore, p_{new}[B,A] > p_{new}[A,B] so that candidate B disqualifies candidate A.

6 References

[1] http://www.chiark.greenend.org.uk/∼galletly/university/speech2.pdf

[2] http://condorcet-dd.sourceforge.net

[3] http://www.groucho.info

[4] http://www.electionmethods.org

[5] http://www.condorcet.org

[6] http://www.alumni.caltech.edu/∼seppley

[7] http://www.ericgorr.net/condorcet

[8] http://userfs.cec.wustl.edu/∼rhl1/rbvote/desc.html

[9] http://www.cs.cornell.edu/Courses/cs100m/2001SP/projects.htm

[10] http://www.spi-inc.org/corporate/resolutions/resolution-2003-01-06.wta.1

[11] http://www.debian.org/devel/ constitution

[12] http://www.debian.org/vote/2003/ vote_0002

[13] Hervé Moulin, Condorcet's Principle Implies the No Show Paradox, *Journal of Economic Theory*, vol. 45, p. 53-64, 1988

[14] Douglas R. Woodall, Monotonicity of single-seat preferential election rules, *Discrete Applied Mathematics*, vol. 77, p. 81-98, 1997

[15] T. Nicolaus Tideman, Independence of Clones as a Criterion for Voting Rules, *Social Choice and Welfare*, vol. 4, p. 185-206, 1987

[16] Thomas M. Zavist, T. Nicolaus Tideman, Complete Independence of Clones in the Ranked Pairs Rule, *Social Choice and Welfare*, vol. 6, p. 167-173, 1989

[17] Robert W. Floyd, Algorithm 97 (Shortest Path), *Communications of the ACM*, vol. 5, p. 345-350, 1962

[18] Joaquín Pérez, The Strong No Show Paradoxes are a common flaw in Condorcet voting correspondences, *Social Choice and Welfare*, vol. 18, p. 601-616, 2001

[19] M.J.A.N. de Condorcet, *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Imprimerie Royale, Paris, 1785

[20] Peyton Young, Condorcet's Theory of Voting, *American Political Science Review*, vol. 82, p. 1231-1244, 1988

[21] Clarence G. Hoag, George H. Hallett, *Proportional Representation*, MacMillan Company, New York, 1926

A Tideman's Ranked Pairs Method

Tideman's Ranked Pairs method [15, 16] is very similar to my method in so far as both methods meet Pareto, monotonicity, resolvability, independence of clones, reversal symmetry, Schwartz and Smith-IIA. However, the following example demonstrates that these methods are not identical.

Example:

3	ACDB
5	ADBC
4	BACD
5	BCDA
2	CADB
5	CDAB
2	DABC
4	DBAC

The matrix d[i,j] of pairwise defeats looks as follows:

	A	B	C	D
A	—	17	18	14
B	13	—	20	9
C	12	10	—	19
D	16	21	11	—

The matrix p[i,j] of the path strengths looks as follows:

	A	B	C	D
A	—	6	6	6
B	2	—	10	8
C	2	8	—	8
D	2	12	10	—

Candidate A is the unique Schulze winner because candidate A is the unique candidate with p[A,X] ≥ p[X,A] for every other candidate X.

Tideman suggests to take successively the strongest pairwise defeat and to lock it if it does not create a directed cycle with already locked pairwise defeats or to skip it if it would create a directed cycle with already locked pairwise defeats. The winner of the Ranked Pairs method is that candidate X who wins each pairwise comparison which is locked and in which candidate X is involved.

Tideman's Ranked Pairs method locks D > B. Then it locks B > C. Then it skips C > D since it would create a directed cycle with the already locked defeats D > B and B > C. Then it locks A > C. Then it locks A > B. Then it locks D > A. Thus, the Ranked Pairs winner is candidate D.

B The Participation Criterion

The *participation* criterion says that adding a set of identical ballots on which candidate A is strictly preferred to candidate B should not change the winner from candidate A to candidate B. Moulin [13] proved that the Condorcet criterion and the participation criterion are incompatible. Pérez [18] demonstrated that most Condorcet methods can violate the participation criterion in a very drastic manner. That means: It can happen that adding a set of identical ballots on which candidate A is strictly preferred to every other candidate changes the winner from candidate A to another candidate or that adding a set of identical ballots on which every other candidate is strictly preferred to candidate B changes the winner from another candidate to candidate B. The following example demonstrates that also the Schulze method can violate the participation criterion in a very drastic manner. (The basic idea for this example came from Blake Cretney.)

Example:

4	ABCDEF
2	ABFDEC
4	AEBFCD
2	AEFBCD
2	BFACDE
2	CDBEFA
4	CDBFEA
12	DECABF
8	ECDBFA
10	FABCDE
6	FABDEC
4	FEDBCA

The matrix d[i,j] of pairwise defeats looks as follows:

	A	B	C	D	E	F
A	—	40	30	30	30	24
B	20	—	34	30	30	38
C	30	26	—	36	22	30
D	30	30	24	—	42	30
E	30	30	38	18	—	32
F	36	22	30	30	28	—

The matrix p[i,j] of the path strengths looks as follows:

	A	B	C	D	E	F
A	—	20	8	8	8	16
B	12	—	8	8	8	16
C	4	4	—	12	12	4
D	4	4	16	—	24	4
E	4	4	16	12	—	4
F	12	12	8	8	8	—

Candidate A is the unique winner since he is the only candidate with p[A,X] ≥ p[X,A] for every other candidate X. However, when 3 AEFCBD ballots are added then the matrix d[i,j] of pairwise defeats looks as follows:

	A	B	C	D	E	F
A	—	43	33	33	33	27
B	20	—	34	33	30	38
C	30	29	—	39	22	30
D	30	30	24	—	42	30
E	30	33	41	21	—	35
F	36	25	33	33	28	—

The matrix p[i,j] of the path strengths looks as follows:

	A	B	C	D	E	F
A	—	23	5	5	5	13
B	9	—	5	5	5	13
C	7	7	—	15	15	7
D	7	7	19	—	21	7
E	7	7	19	15	—	7
F	9	9	5	5	5	—

Now, candidate D is the unique winner since he is the only candidate with p[D,X] ≥ p[X,D] for every other candidate X. Thus the 3 AEFCBD voters change the winner from candidate A to candidate D.

C A Special Provision of the Implementation used by SPI and DEBIAN

There has been some debate about how to measure the strength of a pairwise defeat when it is presumed that on the one side each voter has a sincere complete ranking of all candidates, but on the other side some voters vote only a partial ranking because of strategical considerations. I suggest that then the strength of a pairwise defeat should be measured primarily by the absolute number of votes for the winner of this pairwise defeat and secondarily by the margin of this pairwise defeat. The purpose of this provision is to give an additional incentive to the voters to give different preferences to candidates to which the voters would have given the same preference because of strategical considerations otherwise.

The resulting version of this method is used by SPI and DEBIAN because (a) here the number of candidates is usually very small and the voters are usually well-informed about the different candidates so that it can be presumed that each voter has a sincere complete ranking of all candidates and (b) here the number of voters is usually very small and the voters are usually well-informed about the opinions of the other voters so that the incentive to cast only a partial ranking because of strategical considerations is large.

The resulting version still satisfies Pareto, monotonicity, resolvability, independence of clones, reversal symmetry, Smith-IIA, and Schwartz. When each voter casts a complete ranking then this version is identical to the version defined in section 2. I suggest that in the general case the version as defined in section 2 should be used. Only in situations similar to the above described situation in SPI and DEBIAN, the version as defined in this appendix should be used.

When the strength of a pairwise defeat is measured primarily by p1 (= the absolute number of votes for the winner of this pairwise defeat) and secondarily by p2 (= the margin of this pairwise defeat), then a possible implementation looks as follows:

Input: d[i,j] with $i \neq j$ is the number of voters who strictly prefer candidate i to candidate j.

Output: “w[i] = true” means that candidate i is a potential winner. “w[i] = false” means that candidate i is not a potential winner.

```
for i := 1 to N do
for j := 1 to N do
  if ( i ≠ j ) then
    {
    p2[i,j] := d[i,j] − d[j,i] ;
    if ( d[i,j] > d[j,i] ) then
      p1[i,j] := d[i,j] ;
    if ( d[i,j] ≤ d[j,i] ) then
      p1[i,j] := −1 ;
    }

for i := 1 to N do
for j := 1 to N do
  if ( i ≠ j ) then
    for k := 1 to N do
      if ( i ≠ k ) then
        if ( j ≠ k ) then
          {
          s := min { p1[j,i], p1[i,k] } ;
          t := min { p2[j,i], p2[i,k] } ;
          if ( ( p1[j,k] < s ) or ( ( p1[j,k] = s ) and
                ( p2[j,k] < t ) ) ) then
            {
            p1[j,k] := s ;
            p2[j,k] := t ;
            }
          }

for i := 1 to N do
  {
  w[i] := true ;
  for j := 1 to N do
    if ( i ≠ j ) then
    if ( ( p1[j,i] > p1[i,j] ) or ( ( p1[j,i] = p1[i,j] ) and
          ( p2[j,i] > p2[i,j] ) ) ) then
      w[i] := false ;
  }
```

D The Schwartz Set Heuristic

Another way of looking at the proposed method is to interpret it as a method where successively the weakest pairwise defeats are “eliminated”. The formulation of this method then becomes very similar to Condorcet’s original wordings.

Condorcet writes [19] p. 126: “Create an opinion of those $N(N\text{-}1)/2$ propositions that win most of the votes. If this opinion is one of the $N!$ possible then consider as elected that subject to which this opinion agrees with its preference. If this opinion is one of the $(2^{(N(N-1)/2)}) - (N!)$ impossible opinions then eliminate of this impossible opinion successively those propositions that have a smaller plurality and accept the resulting opinion of the remaining propositions."

In short, Condorcet suggests that the weakest pairwise defeats should be eliminated successively until the remaining pairwise defeats form a ranking of the candidates. The problem with Condorcet's proposal is that it is not quite clear what it means to "eliminate" a pairwise defeat (especially in so far as when one successively eliminates the weakest pairwise defeat that is in a directed cycle of not yet eliminated pairwise defeats until there are no directed cycles of non-eliminated pairwise defeats any more then the remaining pairwise defeats usually do not complete to a unique ranking [20]). It is clear what it means when a candidate is "eliminated"; this candidate is treated as if he has never stood. But what does it mean when the pairwise defeat A $>$ B is "eliminated" although candidate A and candidate B are still potential winners?

A possible interpretation would be to say that the "elimination" of a pairwise defeat is its replacing by a pairwise tie. However, when this interpretation is being used then the Smith set, as defined in the Introduction, can only grow but not shrink at each stage. But when the Schwartz set, as defined in the Introduction, is being used, then the number of candidates decreases continuously. With the concept of the Schwartz set the Schulze method can be described in a very concise manner:

Step 1: Calculate the Schwartz set and eliminate all those candidates who are not in the Schwartz set. Eliminated candidates stay eliminated.

If there is still more than one candidate and there are still pairwise comparisons between non-eliminated candidates that are not pairwise ties: Go to Step 2.

If there is still more than one candidate, but all pairwise comparisons between non-eliminated candidates are pairwise ties, then all remaining candidates are *potential winners*: Go to Step 3.

If there is only one candidate, then this candidate is the *unique winner*.

Step 2: The weakest pairwise defeat between two non-eliminated candidates is replaced by a pairwise tie. Pairwise comparisons that have been replaced by pairwise ties stay replaced by pairwise ties.

In the version in section 4, the *weakest pairwise defeat* is that defeat where |d[i,j] − d[j,i]| is minimal.

In the version in appendix C, the *weakest pairwise defeat* is that defeat where the number of votes for the winner of this pairwise defeat is minimal or —if there is more than one pairwise defeat where the number of votes for the winner is minimal— of all those pairwise defeats where the number of votes for the winner is minimal that pairwise defeat where the number of votes for the loser of this pairwise defeat is maximal.

If the weakest pairwise defeat between non-eliminated candidates is not unique, then all weakest pairwise defeats between non-eliminated candidates are replaced by pairwise ties simultaneously. Go to Step 1.

Step 3: The TBRC is calculated as described in section 2. The winner is that potential winner who is ranked highest in this TBRC.

E An Example without a Unique Winner

Example [21], p. 502:

3	ABCD
2	DABC
2	DBCA
2	CBDA

The matrix d[i,j] of pairwise defeats looks as follows:

	A	B	C	D
A	—	5	5	3
B	4	—	7	5
C	4	2	—	5
D	6	4	4	—

The matrix p[i,j] of the path strengths looks as follows:

	A	B	C	D
A	—	1	1	1
B	1	—	5	1
C	1	1	—	1
D	3	1	1	—

Candidate X is a potential winner if and only if p[X,Y] $\geq$ p[Y,X] for every other candidate Y. Therefore, candidate B and candidate D are potential winners.

When the Schwartz set heuristic is being used then at the first stage the Schwartz set is calculated. The pairwise defeats are A $>$ B, A $>$ C, B $>$ C, B $>$ D, C $>$ D, and D $>$ A. Hence, the Schwartz set is: A, B, C, and D. At the second stage, the weakest pairwise defeat that is not a pairwise tie between candidates who have not yet been eliminated is replaced by a pairwise tie. The weakest pairwise defeats are A > B, A > C, B > D, and C > D each with a strength of 5:4. All these pairwise defeats are replaced by pairwise ties simultaneously. The remaining pairwise defeats are B > C and D > A. Hence, the new Schwartz set is: B and D. Since there are now no pairwise defeats between candidates who have not yet been eliminated, the algorithm stops and candidate B and candidate D are the winners.

Since 5 voters strictly prefer candidate B to candidate D and 4 voters strictly prefer candidate D to candidate B, candidate B is ranked higher than candidate D in the TBRC with a probability of 5/9 and candidate D is ranked higher than candidate B in the TBRC with a probability of 4/9. Therefore, the winner of the Schulze method is candidate B with a probability of 5/9 and candidate D with a probability of 4/9.

Calculation of Transfer Values — Proposal for STV-PR Rules for Local Government Elections in Scotland

James Gilmour
Email: jgilmour@globalnet.co.uk

1 Introduction

The Local Governance (Scotland) Bill [1] will make provision for future local government elections in Scotland to be by the Single Transferable Vote. Those responsible for drafting the legislation have indicated that they do not intend simply to copy the legislation used for the comparable STV elections in Northern Ireland. They believe they can express some points in the counting procedure more clearly. Thus we have a "painless" opportunity to consider some other changes that might usefully be incorporated at the same time. I suggest one of these should be the calculation of transfer values.

2 Precision of calculation

Some discussion in the Election Methods web group [2] prompted me to look in some depth at the calculation of transfer values in STV-PR. The discussion was started by a reference to Wichmann's review [3] of the ERS97 Rules [4]. Wichmann made a number of points about transfer values, starting with what I would call "apparent precision", but going into the arithmetical realities of the truncated calculations adopted in ERS97 and other sets of rules based on Newland and Britton 1972 [5], including those currently used in Northern Ireland. Wichmann's proposal to give results with an actual accuracy of 0.01 votes was to compute transfer values to [(number of digits in total votes) + 1].

Another member of the EM web group drew attention to the procedures of the Australian Electoral Commission [6]. The AEC calculates transfer values to eight decimal places and then truncates as shown in the example on their website. This requirement to calculate to eight decimal places is not specified in any Australian legislation, but only in the AEC's internal working documents [7]. The relevant law [8] makes no reference to the accuracy or precision for any of the STV calculations. The AEC adopted eight decimal places because that was the limit of the desktop calculators available at the time they framed that working rule [7].

The AEC example shows that while they calculate the transfer value of a ballot paper to eight decimal places (8dp) and then use that 8dp result to calculate the transfer values of the votes being transferred, they truncate the candidates' transferred votes to integer values. They do not show decimal parts of a vote anywhere on their result sheets. This truncation to integer values might seem perverse, but does not result in the loss of significant numbers of votes.

In the AEC example there is a surplus of 992,137 votes carried on 1,518,178 papers, of which one candidate receives 1,513,870 papers. The AEC calculation shows an 8dp truncated transfer value of 0.65350505 for each paper. This results in a candidate integer truncated transfer vote of 989,321. The "full" calculation with the 8dp transfer value would have been 989321.69, so they have lost only 0.69 of a vote by integer truncation. This amounts to only 0.000131% of the quota. Had the transfer value been calculated to 15dp (limit of numerical precision for Microsoft Excel 2002), the loss by integer truncation of the votes transferred would have been only 0.700215653, amounting to 0.000133% of the quota.

In contrast, using the ERS/NI rules and calculating the same example to only two decimal places and then truncating, gives a transfer value of 0.65, and a candidate transfer vote of 984,015.50. In this case there would be a loss of 5,306.20 votes from the "true" transfer value, amounting to 1.01% of the quota.

3 Examples from elections

For practical examples I have looked at the immediately available results from the Australian Federal Senate elections in 1998 [9] and the Northern Ireland Assembly elections of 1998 [10]. To make sure there were no complications in the calculations, I looked only at separate transfers arising from the surpluses of candidates whose first preference votes exceeded the quota, i.e. who were elected at stage 1. The relevant figures are in the Tables 1 and 2. In the Australian results they show "non-transferable votes" separately for "exhausted ballots" and for "lost by fraction", ie due to truncation.

The losses arising from truncation are expressed as percentages of the quotas for the relevant elections because this offers the most valid basis for comparisons among the different elections. The results are sorted in ascending order by the size of these percentages. The losses in the Australian transfers range from 0.0043% to 0.032%. In only six of those 14 transfers did the loss exceed 0.01% of the quota. The losses in the Northern Ireland transfers range from 0.10% to 1.36%. In five of those 23 transfers the loss exceeded 1.0% of the quota.

The size of the loss in any individual transfer will depend on just how the calculation tumbles out as that will determine the size of the fraction truncated. For example, in the Newry and Armagh election the transfer value was 0.43 (excluding 222 exhausted papers), leading to a loss of 0.0077245 votes on every one of the 13,360 papers actually transferred. In the Australian elections the losses are increased by the large numbers of candidates who stand and to whom transfers are made.

4 Proposal for change

It now seems clear to me that when the STV rules were formalised for Newland and Britton and the Northern Ireland STV regulations in 1972, there was a confusion of two objectives. It is illogical to calculate transfer values to only two decimal places if candidates' votes are to be recorded to of 0.01 of a vote. This approach was probably taken because the 'Senatorial Rules' [11], devised to remove the element of chance when selecting full value ballot papers for the transfer of surpluses, had given each valid ballot paper a value of one hundred before any calculations were done.

For public elections, with large numbers of electors, there is no intrinsic merit in recording candidates' votes with a precision greater than one vote, provided that does not result in the loss of significant numbers of votes. For elections with small numbers of electors (quota less than 100), there may be a benefit in recording candidates' votes with greater precision, perhaps to 0.01 of a vote. Whatever level of precision is required in the recorded vote, calculating transfer values of ballot papers to only two places of decimals is not consistent with that reported precision. There may be a theoretical case for varying the numbers of decimal places in the calculation according the magnitude of the numbers of votes, but the practical approach of the AEC has been shown to give very satisfactory results.

The AEC adopted eight decimal places for the calculation of transfer values because that was the capacity of the desktop calculators available at the time. Most currently available electronic calculators (hand-held and desktop models) display eight decimal digits, i.e. it is possible to enter '12345678' but not '123456789'. However, when a division to obtain a transfer value is made on such a calculator, the result does not contain eight decimal places, but only seven. Thus, to use the example from the AEC website, (surplus = 992137; transferable papers = 1518178), an 8-digit electronic calculator would display a result of 0.6535050 and not the 0.65350505 quoted. It would be possible to obtain eight significant figures on such a calculator by scaling the calculation, eg 992137 / 151817.8 or 9921370 / 1518178. The transfer value would then be displayed as '6.5350505'. However, there would an additional risk of mistakes being made if calculations were scaled in this way and the increase in precision would be very small.

Taking a practical approach, I would recommend that transfer values should be calculated to 7 decimal places, reflecting the capacity of the commonly available electronic calculators. If the calculation loss is minimised in this way, there is then no need to record decimal fractions of votes for each candidate on the result sheet. The loss that would be incurred in discarding the fractional values when summing the votes for each candidate is very small compared to the calculation loss. This would greatly simplify the presentation of STV-PR result sheets for public elections.

Table 1 Australian Federal Senate Elections 1998
Non-transferable Votes arising on Transfer of Surpluses from First Preferences of Candidates elected at Stage 1

State	Total Vote	Quota	Candidate	Candidate's F P Vote	Surplus	Candidates receiving votes	Exhausted Ballots	Lost by Fraction	LbF as Percentage of Quota
NSW 2	3,755,725	536,533	Heffernan	1,371,578	835,045	35	12	23	0.0043%
NSW 1	3,755,725	536,533	Hutchins	1,446,231	909,698	39	18	25	0.0047%
QLD 3	2,003,710	286,245	Hill	295,903	9,658	15	1	14	0.0049%
VIC 2	2,843,218	406,175	Troeth	1,073,551	667,376	27	9	22	0.0054%
VIC 1	2,843,218	406,175	Conroy	1,148,985	742,810	28	10	24	0.0059%
QLD 1	2,003,710	286,245	McLucas	653,183	366,938	31	15	23	0.0080%
QLD 2	2,003,710	286,245	Parer	568,406	282,161	26	8	24	0.0084%
SA 2	946,816	135,260	Bolkus	301,618	166,358	23	6	13	0.0096%
WA 1	1,063,811	151,974	Ellison	405,617	253,643	26	10	16	0.0105%
WA 2	1,063,811	151,974	Cook	366,874	214,900	33	11	16	0.0105%
SA 1	946,816	135,260	Vanstone	381,361	246,101	27	8	17	0.0126%
ACT	197,035	65,679	Lundy	83,090	17,411	15	4	10	0.0152%
TAS 2	308,377	44,054	Abetz	98,178	54,124	18	18	12	0.0272%
TAS 1	308,377	44,054	O'Brien	121,931	77,877	22	30	14	0.0318%

Table 2 Northern Ireland Assembly Elections 1998
Non-transferable Votes arising on Transfer of Surpluses from First Preferences of Candidates elected at Stage 1

State	Total Vote	Quota	Candidate	Candidate's F P Vote	Surplus	Candidates receiving votes	Non-transferable votes	NTV as Percentage of Quota
East Antrim 2	35,610	5,088	Neeson	5,247	159	11	4.89	0.10%
Belfast East 1	39,593	5,657	Robinson	11,219	5,562	15	6.00	0.11%
South Antrim	43,991	6,285	Wilson	6,691	406	9	10.96	0.17%
Belfast North 2	41,125	5,876	Maginness	6,196	320	15	12.25	0.21%
Upper Bann 1	50,399	7,200	Trimble	12,338	5,138	16	20.30	0.28%
Belfast West 1	41,794	5,971	Adams	9,078	3,107	13	22.10	0.37%
North Antrim	49,697	7,100	Paisley	10,590	3,490	15	28.30	0.40%
East Londonderry	39,564	5,653	Campbell	6,099	446	10	25.44	0.45%
West Tyrone	45,951	6,565	Gibson	8,015	1,450	12	32.29	0.49%
Mid-Ulster 2	49,798	7,115	McGuinness	8,703	1,588	7	45.40	0.64%
Fermanagh & South Tyrone	51,043	7,292	Gallagher	8,135	843	11	50.80	0.70%
Mid-Ulster 1	49,798	7,115	McCrea	10,339	3,224	10	49.60	0.70%
Upper Bann 2	50,399	7,200	Rodgers	9,260	2,060	14	55.36	0.77%
Belfast North 1	41,125	5,876	Dodds	7,476	1,600	15	45.79	0.78%
Belfast West 2	41,794	5,971	Hendron	6,140	169	10	50.80	0.85%
North Down	37,313	5,331	McCartney	8,188	2,857	18	47.55	0.89%
Strangford 1	42,922	6,132	Robinson	9,479	3,347	18	59.80	0.98%
East Antrim 1	35,610	5,088	Beggs	5,764	676	14	49.99	0.98%
Foyle	48,794	6,971	Hume	12,581	5,610	14	69.60	1.00%
Belfast East 2	39,593	5,657	Alderdice	6,144	487	18	58.81	1.04%
Strangford 2	42,922	6,132	Taylor	9,203	3,071	20	73.61	1.20%
South Down	51,353	7,337	McGrady	10,373	3,036	16	90.76	1.24%
Newry & Armagh	54,136	7,734	Mallon	13,582	5,848	13	104.92	1.36%

5 Benefits in local government elections in Scotland

The numbers of electors in the constituencies in both the Australian Federal Senate elections and the Northern Ireland Assembly elections are considerably larger than those likely in the multi-member wards for local government elections in Scotland. It is, therefore, useful to make an assessment of the potential effects of changing the precision of calculation of transfer values from 2dp to 7dp using local data.

For this example I have used Glasgow City Council which has an electorate of 453,552 and 79 councillors. I have examined two possible implementations of STV-PR: nine 8-member wards plus one 7-member ward; and nineteen 4-member wards plus one 3-member ward (Table 3). I have assumed there would be equal numbers of electors per councillor in all wards and a turnout of 50%. I have also assumed that the Labour Party would get 47.58% of the first preference votes (= city-wide average in the 2003 FPTP council elections), that 75% of those first preference votes would be for the party's leading candidate in the ward and that all those papers would be transferable. For the calculation with 7dp I have also truncated the transferred votes to integer values as I recommend above. The results in Table 3 show that the effect of truncating the calculation of transfer values at 2dp could be considerable even in the smaller 4-member wards. The losses when the calculation is truncated at 7dp are negligible.

Table 3 Comparison of Effects of Calculating Transfer Values to 2dp and 7dp

Implementation	*8-member ward*	*4-member ward*
Electorate	45,929	22,964
Valid votes	22,964	11,482
Quota	2,552	2,297
Party FP votes	10,926	5,463
Leading candidate's FP votes	8,194	4,097
Surplus for transfer	5,642	1,800
Transfer value 2dp	0.68	0.43
Transferred votes 2dp	5,571.92	1,761.71
Votes lost by truncation at 2dp	70.08	38.29
Votes lost as percentage of quota	2.75%	1.67%
Transfer value 7dp	0.6885525	0.4393458
Transferred votes 7dp	5641	1,799
Votes lost by truncation at 7dp	1	1
Votes lost as percentage of quota	0.039%	0.044%

The actual loss in transfer value due to truncating the calculation at 2dp compared to truncating at 7dp can vary from 0.0000000 to 0.0099999. The general effect can be assessed by considering only the loss that occurs in the third decimal place. The results in Table 4 have been calculated using the same two example wards as above. The ten potential losses all have equal probabilities of occurrence. The loss due to truncation at 2dp in the 8-member ward will exceed 1% of the quota in six cases out of ten and will exceed 2% in three cases out of ten. Even in the smaller ward, the loss due to this truncation will exceed 1% of the quota in four cases out of ten. These losses are substantial and could be avoided by a simple change to the rules for STV-PR elections.

Table 4 Loss of Votes due to Truncation of Transfer Value before 3dp

Implementation	*8-member ward*		*4-member ward*	
Transferable papers	8,194		4,097	
Loss in transfer value	Votes lost	% of quota	Votes lost	% of quota
0.000	0	0.00%	0	0.00%
0.001	8	0.31%	4	0.17%
0.002	16	0.63%	8	0.35%
0.003	24	0.94%	12	0.52%
0.004	32	1.25%	16	0.70%
0.005	40	1.57%	20	0.87%
0.006	49	1.92%	24	1.04%
0.007	57	2.23%	28	1.22%
0.008	65	2.55%	32	1.39%
0.009	73	2.86%	36	1.57%

6 References

[1] Draft Local Governance (Scotland) Bill: Consultation. October 2003. http://www.scotland.gov.uk/consultations/localgov/dlgsc-00.asp

[2] Election Methods web group. http://electorama.com/modules.php?op=modload&name=News&file=article&sid=3&mode=thread&order=0&thold=0

[3] B A Wichmann (1999). A review of the ERS97 rules. *Voting Matters*, Issue 10, 5-7

[4] R A Newland and F S Britton (1997). *How to conduct an election by the single transferable vote*. Electoral Reform Society, London. 1997 ISBN 0 903291 18 5.

[5] R A Newland and F S Britton (1972). *How to conduct an election by the single transferable vote*. Electoral Reform Society, London. 1973 ISBN 903291 02 9.

[6] Australian Electoral Commission Info Centre. Counting the votes: Senate. http://www.aec.gov.au/_content/what/voting/count_senate.htm

[7] G Newman (2003) Director, Research Section, Australian Electoral Commission. Personal communication.

[8] Commonwealth Electoral Act 1918 Section 273. Scrutiny of votes in Senate elections. http://scaleplus.law.gov.au/html/pasteact/0/57/0/PA003460.htm

[9] Australian Election Archive: Federal Election of 3 October 1998. http://psephos.adam-carr.net/countries/a/australia/index1998.shtml

[10] UK Election Results: Northern Ireland Assembly 1998 — Full STV results. http://www.election.demon.co.uk/niassembly.html

[11] British Statutory Rules and Orders No. 727. 1921.

Voting matters

for the technical issues of STV

published by

The McDougall Trust

Issue 18 June 2004

About the McDougall Trust (reg. charity no. 212151)

The McDougall Trust is a charitable trust formed in 1948. The charity's purposes as stated in its governing scheme of 1959 are to advance knowledge of and encourage the study of and research into:

- political or economic science and functions of government and the services provided to the community by public and voluntary organisations; and
- methods of election of and the selection and government of representative organisations whether national, civic, commercial, industrial or social.

The Trust's work includes the maintenance and development of the Lakeman Library for Electoral Studies, a unique research resource, the production and publication of Representation: The Journal of Representative Democracy, and, of course, this publication **Voting matters**, that examines the technical issues of the single transferable vote and related electoral systems.
For further information on the Trust, please contact:

The Secretary,
McDougall Trust,
6 Chancel Street,
London SE1 0UX, UK.
Telephone: +44 (0)20 7620 1080
Facsimile: +44 (0)20 7928 1528
Email: admin@mcdougall.org.uk
Web: www.mcdougall.org.uk

For further information on this publication, please contact B A Wichmann, the Editor at the above address or by email at: `Brian.Wichmann@bcs.org.uk`

Editorial

I would strongly recommend to all readers that the Interim Report of the Commission on Electronic Voting, issued by the Irish Government, is studied closely. This Commission, was formed on the 1st March and required to report by 1st May, on the suitability of the system chosen for use in elections in Ireland. They recommended that the chosen system should *not* be used for the local/European elections to be held on 11th June. The Commission's Report can be downloaded at: http://www.cev.ie/htm/report/download_report.htm

To avoid any confusion, I need to declare an interest in this report, since I worked with Joe Wadsworth of Electoral Reform Services in testing the counting engine of the official software. Our work was not finished until the end of March, which was only 5 weeks before the Commission reported.

Some aspects of their report are of particular interest here:

- The desirability of removing random selection in the counting process;
- Problems associated with full disclosure of the ballot data (discussed further in this issue);
- Some shortcomings with regard to secrecy;
- The need for a Voter Verifiable Audit Trial.

A Voter Verifiable Audit Trial might work by having a printer attached to the electronic voting machine which printed out the filled-in paper after it had been recorded electronically. The voter would then check this, and place the paper in a conventional ballot box. Hence the ballot box papers can be used as a (manual) check against the computer count.

Technically, such a scheme has a number of problems. Firstly, printers are less reliable than a purely electronic device; should the printer jam, the election officials might inadvertently see a ballot paper. Secondly, the conventional record would presumably be used for a recount; however, a manual recount is likely to be less reliable than the initial electronic count. The process whereby the printed papers are used needs to be very carefully considered.

There is no doubt that the undertaking of a manual count is one that the public feels gives confidence in the democratic process. What, therefore, needs to be done to gain the same confidence in an electronic count? The Irish report gives some insight into this important issue. Is it necessary to have a Voter Verifiable Audit Trial, it spite of the problems noted above? Since the Irish Government is still planning to use electronic voting, we will soon be able to see how these issues are being addressed.

Returning to *Voting matters*, there are 6 papers in this issue:

- I. D. Hill: What is meant by 'monotonic'? What is meant by 'AV'?
- M. Schulze: Free riding.
- I. D. Hill: An odd feature in a real STV election.
- I. D. Hill: Full disclosure of data.
- B. A. Wichmann: A note on the use of preferences.
- J. C. O'Neill: Tie-Breaking with the Single Transferable Vote.

David Hill highlights the problem of the meaning of terms and even abbreviations. As Editor, I am always concerned about this, since the terminology in common use varies substantially, especially now that papers are authored from outside the UK/Ireland.

Markus Schulze raises the interesting and important question of the extent to which strategic voting is used in STV elections. Two forms of strategic voting are analysed, which in one case, can be identified from US ballot data in which voters can *write-in* a candidate. Fortunately for STV, the analysis gives no evidence of strategic voting in the analysable case.

The next three article are all about the use of preferences. David Hill first provides an example in which a single paper with a large number of preferences has a crucial effect. His subsequent papers respond to an earlier *Voting matters* paper on full disclosure. In my own article, I consider the actual use made of the preferences specified by the voter, and how this information could be altered to avoid any undesirable consequences of full disclosure.

In the final article, Jeff O'Neill analyses the various ways in which ties are broken which results in a proposal to change the tie-breaking logic in the current Electoral Reform Society rules.

What is meant by 'monotonic'? What is meant by 'AV'?

I. D. Hill
No email available.

It is said that, during the 1939-1945 war, Winston Churchill and President Roosevelt had a disagreement when Churchill wished to table a document and Roosevelt did not wish it to be tabled. It turned out that they both wanted the same thing: that to the British, tabling a document means putting it on the table for discussion; whereas to the Americans, it means putting it in a drawer and forgetting it. Such confusion, caused by language difficulties, can be serious.

1 Monotonic

Schulze [1] explains a method for single seat elections that finds the Condorcet winner if there is one, and has a strategy for choosing a winner where there is a Condorcet paradox. He claims that the method is "monotonic and clone-independent".

The main purpose of this note is to warn others who may have been misled, as I was myself at first, by that claim. The trouble lies in definitions, because I am told that his usage of 'monotonic' is as normally used in the social choice literature, but it is a much narrower definition than is often taken as the meaning in electoral reform literature.

He gives an example where his method certainly violates the condition that Woodall [2] calls mono-add-top: "A candidate x should not be harmed if further ballots are added that have x top (and are otherwise arbitrary)", but Schulze is only claiming to meet mono-raise: "A candidate x should not be harmed if x is raised on some ballots without changing the orders of the other candidates".

I am not seeking to cast any blame. If that usage of the word is widely employed, he is fully entitled to follow it, but a clash of definitions may cause misunderstanding if we do not take great care.

2 AV

Brams and Fishburn [3] give an example of the use of a system called Approval Voting, and they use AV as an abbreviation for it. In this country AV has been used for many years to mean the system called Alternative Vote.

Approval Voting is a system in which a voter uses X-voting for as many candidates as desired, even when there is only one seat to fill. The winner is the one who gets the most Xs. Alternative Vote is what STV reduces to in the single-seat case, voting by preference number, with eliminative counting.

It is not my purpose in this note to examine the relative merits, or lack of merits, of these two systems, but only to warn that they are very different, and that the name AV is, unfortunately, being used for both of them. Again, this may cause misunderstanding if we do not take great care.

3 References

[1] Schulze M. A new monotonic and clone-independent single-winner election method. *Voting matters*, 17, 9-19. 2003.

[2] Woodall D.R. Properties of preferential election rules. *Voting matters*, 3, 8-15. 1994.

[3] Brams S.J. and Fishburn P.C. A nail-biting election. *Social Choice and Welfare*, 18, 409-414. 2001.

A special thanks to David Hill for checking this issue.

Free riding

Markus Schulze
markus.schulze@alumni.tu-berlin.de

1 Introduction

The fact that more and more communities that use proportional representation by the single transferable vote (STV) change from manual count to computer count gives us today the possibility to check hypotheses that have been made in the past about possible voting behaviours. In this paper, I use the ballot data of the 1999 and the 2001 City Council elections and School Committee elections in Cambridge, Massachusetts, to estimate the number of voters who use a voting behaviour that has been predicted e.g. by Woodall [1] and Tideman [2].

2 Woodall Free Riding

Woodall free riding is a useful strategy only for those STV methods where votes of eliminated candidates cannot be transferred to already elected candidates and therefore jump directly to the next highest ranked *hopeful* (i.e. neither yet elected nor yet eliminated) candidate. A *Woodall free rider* is a voter who gives his first preference to a candidate who is believed by this voter to be eliminated early in the count even with this voter's first preference. With this strategy this voter assures that he does not waste his vote for a candidate who is elected already during the transfer of the initial surpluses.

Woodall writes [1]:

> "The biggest anomaly is caused by the decision, always made, not to transfer votes to candidates who have already reached the quota of votes necessary for election. This means that the way in which a given voter's vote will be assigned may depend on the order in which candidates are declared elected or eliminated during the counting, and it can lead to the following form of tactical voting by those who understand the system. If it is possible to identify a candidate W who is sure to be eliminated early (say, the Cambridge University Raving Loony Party candidate), then a voter can increase the effect of his genuine second choice by putting W first. For example, if two voters both want A as first choice and B as second, and A happens to be declared elected on the first count, then the voter who lists his choices as 'A B...' will have (say) one third of his vote transferred to B, whereas the one who lists his choices as 'W A B...' will have all of his vote transferred to B, since A will already have been declared elected by the time W is eliminated. Since one aim of an electoral system should be to discourage tactical voting, this seems to me to be a serious drawback."

However, Woodall free riding can be prevented by restarting the STV count with the remaining candidates whenever a candidate has been eliminated. Actually, the Meek method [3] and the Warren method [4] do this. Therefore, Woodall [1] and Tideman [2] suggest that one of these methods should be used.

A good test for Woodall free riding is an STV election with *write-in options* (i.e. with the possibility for the voters to vote for any person by writing this person's name on the ballot). The City Council and the School Committee of Cambridge, Massachusetts, are elected by an STV method that is vulnerable to Woodall free riding and that has write-in options. In the elections to the 9 seats of the City Council, the voter can vote for up to 9 write-ins. In the elections to the 6 seats of the School Committee, the voter can vote for up to 6 write-ins. Here the optimal Woodall free riding strategy is to give one's first preference to a completely unimportant write-in.

	CC 1999	SC 1999	CC 2001	SC 2001
1	18,613	17,796	17,125	16,488
2	28	26	30	51
3	9	5	12	32
4	0	4	0	2
5	19	17	18	17

Table 2.1: Potential write-in Woodall free riders in the 1999 and the 2001 elections to the City Council and the School Committee of Cambridge, Massachusetts

In table 2, row "1" contains the numbers of voters in the 1999 City Council elections (column "CC 1999"), in the 1999 School Committee elections (column "SC 1999"), in the 2001 City Council elections (column "CC 2001"), and in the 2001 School Committee elections (column "SC 2001") in Cambridge, Massachusetts. Row "2" contains the numbers of voters who cast a first preference for a write-in. Row "3" contains the numbers of voters who have to be subtracted from row "2" because they cast preferences only for write-ins and who are therefore obviously not Woodall free riders. Furthermore, those voters who do not cast at least a valid second and a valid third preference have to be subtracted (row "4") because these voters cannot be Woodall free riders. Therefore, row "5" contains the numbers of voters who could be write-in Woodall free riders.

In all four elections, the number of voters who could be write-in Woodall free riders is about 0.1%. When we investigate these voters in greater detail we observe: Of the 19 potential write-in Woodall free riders in the 1999 City Council elections, only 2 cast a second preference for Galluccio. Of the 17 potential write-in Woodall free riders in the 1999 School Committee elections, only 2 cast a second preference for Turkel. Of the 18 potential write-in Woodall free riders in the 2001 City Council elections, only 5 cast a second preference for Galluccio, 2 for Davis, and one for Murphy. Of the 17 potential write-in Woodall free riders in the 2001 School Committee elections, only 4 cast a second preference for Turkel, one for Fantini, and none for Grassi. Therefore, also these voters seem to be not Woodall free riders because otherwise super-proportionally many of these voters would have cast a second preference for a candidate who reached the quota before candidates had to be eliminated. See table 2.2.

Suppose V is the number of voters. Suppose $V_1(A)$ is the number of voters who cast a valid first preference for candidate A. Suppose $V_2(A)$ is the number of voters who cast a valid first preference for candidate A and at least also a valid second preference. Suppose V(A,B) is the number of voters who cast a valid first preference for candidate A, a valid second preference for candidate B, and at least also a valid third preference.

Woodall free riding is a useful strategy only when one has at least a sincere first and a sincere second preference. A given voter can be a Woodall free rider only when he casts at least a valid first, a valid second, and a valid third preference. When a given voter whose sincere first preference is candidate B uses Woodall free riding then $V_2(B)$ decreases and for some other candidate A, who is eliminated early in the count, V(A,B) increases. Therefore, another good test for Woodall free riding is to calculate V(A,B) for each pair of candidates. If (1) $V(A,B)/V_1(A)$ is large compared to $V_2(B)/V$ and (2) $V(A,B)/V_1(A)$ decreases with increasing $V_1(A)$ for those pairs of candidates where candidate A is eliminated early in the count and candidate B is elected before candidates have to be eliminated then this is evidence that voters use Woodall free riding.

Table 2.2 contains $V_2(B)/V$ for each candidate B who is elected before candidates have to be eliminated. Tables 2.3 to 2.6 contain V(A,B) for each pair of candidates A and B where candidate B is elected before candidates have to be eliminated. Column "$V_1(A)$" contains the numbers of voters who cast a valid first preference for the candidate in column "candidate A". The column "Galluccio" (resp. "Turkel", resp. "Davis", etc.) contains the numbers of voters of column "$V_1(A)$" who cast a valid second preference for Galluccio (resp. Turkel, resp. Davis, etc.) and cast at least also a valid third preference.

In tables 2.3 to 2.6, $V(A,B)/V_1(A)$ rather increases than decreases with increasing $V_1(A)$. Also the prediction that $V(A,B)/V_1(A)$ is large compared to $V_2(B)/V$ is not fulfilled. This is surprising because in so far as Woodall free riding certainly is a useful strategy one would expect that at least some voters use this strategy. A possible explanation why voters do not use Woodall free riding is that they fear that when too many voters give their first preference to candidate A because they believe that he is eliminated early in the count then it could happen that candidate A gets so many votes that he is elected [2, 5, 6]. But this can only explain why $V(A,B)/V_1(A)$ does not decrease so fast with increasing $V_1(A)$; this cannot explain why $V(A,B)/V_1(A)$ increases with increasing $V_1(A)$. A possible explanation why $V(A,B)/V_1(A)$ increases with increasing $V_1(A)$ is that voters are confronted with two problems:

1. It is a useful strategy not to waste one's vote by voting for a candidate B who is elected even without one's vote. However, when too many voters use Woodall free riding and cast a first preference for candidate A because they believe that he is eliminated early in the count even with one's vote then it could happen that candidate A gets so many votes that he is elected.
2. It is a useful strategy not to vote for a candidate A who is believed to be eliminated with a great probability even with one's vote, because otherwise there is the danger that there are not acceptable candidates anymore to whom this voter could transfer his vote when candidate A is eliminated.

Because of problem 2 only those voters who cannot identify themselves with any of the stronger candidates vote for candidates who are believed to be eliminated with a great probability; therefore, $V(A,B)/V_1(A)$ is low for low $V_1(A)$ for those candidates B who are elected before candidates have to be eliminated; therefore, $V(A,B)/V_1(A)$ rather increases than decreases with increasing $V_1(A)$.

3 Hylland Free Riding

Problem 1 can be circumvented by using Hylland free riding instead of Woodall free riding. Hylland writes [7]:

> "Both for groups and for individual voters it could be advantageous not to vote for a candidate who is considered certain of winning election, even if that candidate is one's first choice. Suppose that my true first and second choices are A and B, I am sure A will get many more first preferences than needed for election, but I find B's chances uncertain. If I list A as the first preference on my ballot, its weight is reduced before it reaches B. If I omit A, B gets a vote with full weight."

In short, a Hylland free rider is a voter who omits in his individual ranking completely all those candidates who are certain to be elected. Of course, when too many voters use Hylland free riding then it can happen that the candidate with the cast first preference is elected while the candidate with the sincere first preference is eliminated. However, when a voter uses Hylland free riding then the candidate with the cast first preference is one of this voter's favorite candidates while when this voter uses Woodall free riding then the candidate with the cast first preference is a candidate who this voter does not want to be elected.

Problem 2 can be circumvented by voting only for those candidates who are believed to be in the race until the final count. In so far as a candidate will be in the final count when he has more than V/(S+2) first preferences, where V is the number of voters and S is the number of seats, it is a useful strategy to cast one's first preference only for one of those candidates who are believed to get between V/(S+2) and V/(S+1) first preferences.

This voting behaviour could best be observed in Canada because here the city councils were elected for a one year term and in a single city-wide district so that the voters had very precise information about the support of the different candidates. A consequence of this voting behaviour was that usually almost all first preferences were concentrated on S+1 almost equally strong candidates [8, 9, 10]. Johnston [9] writes that one of the main criticisms of STV was that it was "one of the most common features of PR in Canadian municipal elections" that "the final count closely mirrored the results of the first count". And Pilon [10] writes that the main problem of STV in Canada was that it "did not seem to make much difference in the results. After days of counting, eliminating candidates, and transferring fractions of support from one aspirant to another, there was little difference between the first choice results and the final tally."

4 Summary

Free riding is a very serious problem of STV. The two free riding strategies that have been predicted in the literature are Woodall free riding [1, 2] and Hylland free riding [7]. It is not possible to extract the number of Hylland free riders simply from the ballot data. But with additional assumptions it is possible to extract the number of Woodall free riders.

I used the ballot data of the 1999 and the 2001 City Council elections and School Committee elections in Cambridge, Massachusetts, to estimate the number of voters who use Woodall free riding. I could not find any evidence at all that voters use this strategy. Possible explanations why voters do not use this strategy are:

1. When too many voters cast a first preference for candidate A, not because he is their sincere first preference but because they believe that he will be eliminated early in the count, it could happen that this candidate gets so many votes that he is elected [2, 5, 6].
2. It is not useful to vote for a candidate A who is eliminated with a great probability, because it could happen that there are not acceptable candidates anymore to whom this voter could transfer his vote when candidate A is eliminated.
3. When a voter considers his second favorite candidate to be only slightly worse than his favorite candidate then Hylland free riding [7] is less dangerous than Woodall free riding in so far as a backfire is less severe under Hylland free riding than under Woodall free riding.
4. The political organizations have not yet found a simple way to use Woodall free riding on a larger scale to increase their numbers of seats. Therefore, the voters are usually not pointed to this strategic problem.

5 References

[1] Douglas R. Woodall, *Computer counting in STV elections*, REPRESENTATION, vol. 23, pp. 4-6, 1983 {reprinted in VOTING MATTERS, issue 1, 1994}

[2] T. Nicolaus Tideman, Daniel Richardson, *Better voting methods through technology: The refinement-manageability trade-off in the single transferable vote*, PUBLIC CHOICE, vol. 103, pp. 13-34, 2000

[3] Brian L. Meek, *Une nouvelle approache du scrutin transférable*, MATHÉMATIQUES ET SCIENCES HUMAINES, vol. 25, pp. 13-23, 1969 {reprinted in VOTING MATTERS, issue 1, 1994}

[4] C. H. E. Warren, *Counting in STV Elections*, VOTING MATTERS, issue 1, 1994

[5] R. J. C. Fennell, *Is a feedback method of calculating the quota really necessary?*, VOTING MATTERS, issue 2, 1994

[6] I. D. Hill, *Are better STV rules worthwhile? – A reply to R.J.C. Fennell*, VOTING MATTERS, issue 3, 1994

[7] Aanund Hylland, *Proportional Representation without Party Lists*, pp. 126-153, RATIONALITY AND INSTITUTIONS, eds. Raino Malnes and Arild Underdal, Scandinavian University Press, Oslo, 1992

[8] Joseph P. Harris, *Practical Workings of Proportional Representation in the United States and Canada*, NATIONAL MUNICIPAL REVIEW, vol. 19, pp. 337-383, 1930

[9] J. Paul Johnston, Miriam Koene, *Learning History's Lessons Anew: The Use of STV in Canadian Municipal Elections*, pp. 205-247, ELECTIONS IN AUSTRALIA, IRELAND, AND MALTA UNDER THE SINGLE TRANSFERABLE VOTE, eds. Shaun Bowler and Bernard Grofman, University of Michigan Press, Ann Arbor, 2000

[10] Dennis M. Pilon, *The Drive for Proportional Representation in British Columbia, 1917-23*, Simon Fraser University, Vancouver, 1994

Election	Candidate B	V	V_1(B)	V_1(B)/V	V_2(B)	V_2(B)/V
1999 City Council	Anthony D. Galluccio	18,613	2,705	14.5%	2,515	13.5%
1999 School Committee	Alice L. Turkel	17,796	2,617	14.7%	2,360	13.3%
2001 City Council	Henrietta Davis	17,125	1,713	10.0%	1,645	9.6%
2001 City Council	Brian Murphy	17,125	1,716	10.0%	1,627	9.5%
2001 City Council	Anthony D. Galluccio	17,125	3,230	18.9%	2,947	17.2%
2001 School Committee	Joseph G. Grassi	16,488	2,135	12.9%	1,728	10.5%
2001 School Committee	Alfred B. Fantini	16,488	2,854	17.3%	2,353	14.3%
2001 School Committee	Alice L. Turkel	16,488	2,862	17.4%	2,484	15.1%

Table 2.2: V_2 (B)/V for each candidate B who is elected before candidates have to be eliminated

Candidate A	V_1(A)	Anthony D. Galluccio
Charles O. Christenson	28	2 (7.1%)
Daejanna P. Wormwood-Malone	28	0 (0.0%)
William C. Jones	31	2 (6.5%)
Alan Kingfish Nidle	40	0 (0.0%)
Vincent Lawrence Dixon	44	3 (6.8%)
Jeffrey Jay Chase	102	10 (9.8%)
Dorothy M. Giacobbe	109	22 (20.2%)
James M. Williamson	128	2 (1.6%)
Robert Winters	301	27 (9.0%)
Helder Peixoto	308	46 (14.9%)
David Hoicka	325	7 (2.2%)
Erik C. Snowberg	425	12 (2.8%)
David Trumbull	533	129 (24.2%)
Bob Goodwin	805	296 (36.8%)
David P. Maher	1,030	309 (30.0%)
Katherine Triantafillou	1,167	42 (3.6%)
Michael A. Sullivan	1,321	278 (21.0%)
Kenneth E. Reeves	1,420	149 (10.5%)
Henrietta Davis	1,458	70 (4.8%)
Jim Braude	1,480	50 (3.4%)
Timothy J. Toomey, Jr.	1,497	233 (15.6%)
Marjorie C. Decker	1,642	43 (2.6%)
Kathleen Leahy Born	1,658	100 (6.0%)

Table 2.3: Potential Woodall free riders in the 1999 City Council elections in Cambridge, Massachusetts

Candidate A	V_1(A)	Alice L. Turkel
Shawn M. Burke	212	6 (2.8%)
Jamisean F. Patterson	278	9 (3.2%)
Alvin E. Thompson	373	35 (9.4%)
Melody L. Brazo	471	82 (17.4%)
Donald Harding	698	24 (3.4%)
Elizabeth Tad Kenney	738	134 (18.2%)
Michael Harshbarger	1,550	109 (7.0%)
Nancy Walser	1,894	520 (27.5%)
Susana M. Segat	1,985	480 (24.2%)
Joseph G. Grassi	2,269	97 (4.3%)
Alfred B. Fantini	2,277	55 (2.4%)
Denise Simmons	2,408	506 (21.0%)

Table 2.4: Potential Woodall free riders in the 1999 School Committee elections in Cambridge, Massachusetts

Candidate A	V_1(A)	Henrietta Davis	Brian Murphy	Anthony D. Galluccio	Sum (Gallucio, Murphy, Davis)
James M. Williamson	58	2 (3.4%)	2 (3.4%)	3 (5.2%)	7 (12.1%)
James E. Condit, III	63	6 (9.5%)	0 (0.0%)	5 (7.9%)	11 (17.5%)
Helder Peixoto	69	5 (7.2%)	3 (4.3%)	7 (10.1%)	15 (21.7%)
Vincent Lawrence Dixon	92	2 (2.2%)	3 (3.3%)	7 (7.6%)	12 (13.0%)
Robert L. Hall	153	3 (2.0%)	13 (8.5%)	18 (11.8%)	34 (22.2%)
Jacob Horowitz	155	14 (9.0%)	12 (7.7%)	6 (3.9%)	32 (20.6%)
Steven E. Jens	278	8 (2.9%)	5 (1.8%)	35 (12.6%)	48 (17.3%)
Steve Iskovitz	345	29 (8.4%)	30 (8.7%)	9 (2.6%)	68 (19.7%)
Ethridge A. King	378	43 (11.4%)	46 (12.2%)	25 (6.6%)	114 (30.2%)
David P. Maher	1,017	32 (3.1%)	41 (4.0%)	304 (29.9%)	377 (37.1%)
John Pitkin	1,091	222 (20.3%)	202 (18.5%)	48 (4.4%)	472 (43.3%)
Kenneth E. Reeves	1,141	72 (6.3%)	34 (3.0%)	125 (11.0%)	231 (20.2%)
Michael A. Sullivan	1,315	45 (3.4%)	28 (2.1%)	316 (24.0%)	389 (29.6%)
Denise Simmons	1,339	186 (13.9%)	137 (10.2%)	74 (5.5%)	397 (29.6%)
Timothy J. Toomey, Jr.	1,402	44 (3.1%)	11 (0.8%)	272 (19.4%)	327 (23.3%)
Marjorie C. Decker	1,540	298 (19.4%)	215 (14.0%)	163 (10.6%)	676 (43.9%)
Henrietta Davis	1,713	—	254 (14.8%)	114 (6.7%)	
Brian Murphy	1,716	343 (20.0%)	—	105 (6.1%)	
Anthony D. Galluccio	3,230	137 (4.2%)	90 (2.8%)	—	

Table 2.5: Potential Woodall free riders in the 2001 City Council elections in Cambridge, Massachusetts

Candidate A	V_1(A)	Joseph G. Grassi	Alfred B. Fantini	Alice L. Turkel	Sum (Turkel, Fantini, Grassi)
Vincent J. Delaney	240	23 (9.6%)	29 (12.1%)	5 (2.1%)	57 (23.8%)
Fred Baker	324	28 (8.6%)	62 (19.1%)	9 (2.8%)	99 (30.6%)
Marla L. Erlien	1,193	21 (1.8%)	25 (2.1%)	272 (22.8%)	318 (26.7%)
Susana M. Segat	1,590	61 (3.8%)	107 (6.7%)	619 (38.9%)	787 (49.5%)
Nancy Walser	1,677	42 (2.5%)	68 (4.1%)	596 (35.5%)	706 (42.1%)
Richard Harding, Jr.	1,689	172 (10.2%)	156 (9.2%)	176 (10.4%)	504 (29.8%)
Alan C. Price	1,873	41 (2.2%)	71 (3.8%)	319 (17.0%)	431 (23.0%)
Joseph G. Grassi	2,135	—	698 (32.7%)	94 (4.4%)	
Alfred B. Fantini	2,854	942 (33.0%)	—	158 (5.5%)	
Alice L. Turkel	2,862	97 (3.4%)	133 (4.6%)	—	

Table 2.6: Potential Woodall free riders in the 2001 School Committee elections in Cambridge, Massachusetts

An odd feature in a real STV election

I. D. Hill
No email available.

Although artificial data can be extremely useful in clearly demonstrating difficulties in election rules, there is also much to be said in favour of looking at real data, particularly where anything odd appears to have happened.

A few years ago, there were 23 candidates in an election for 15 seats, and there were 539 votes. The candidates' names have here been coded as A, B, C, etc.

One voter gave preferences, in order, as: M D L R I J C T B E H A O U F etc. Using Newland and Britton (second edition) rules [1], the last candidate elected was F and the runner-up was V. Amazingly, if that one voter had put V instead of F as 15th preference, V would have been elected and F runner-up. In other words, the election result depended upon that one voter's 15th preference.

There is, of course, nothing wrong with a 15th preference being taken into account. If all previous 14 preferences have been excluded it is right that the 15th preference comes through with a value of 1.0 as if it had been a 1st preference. In this case, though, it came through with a value of 1.0 even though 10 of the earlier preferences were elected. Of those 10, 8 had been elected before that vote reached them and, in accordance with the rules, were "leap-frogged". The other 2, J and T are more remarkable; in each case the vote in question was among those that triggered their election and, being part of the last parcel received, was due to be transferred with a transfer value. For both of them, however, there were enough non-transferable votes in the parcel that the transfer value came out as 1.0.

When the final transfer was made, V had 30.31 votes, and F had 30.51, so the additional 1.0 was enough to sway the result. The vote had not had to make any contribution to electing the 10 elected candidates named earlier by the voter.

If Meek rules [2] had been used, that 15th preference would still have been reached, but F would have been ahead of V by almost 4 votes and the value attached to the particular vote, because it would have had to contribute a fair share to electing the earlier 10 candidates, would have been only 0.000000905 and would thus have made no difference.

It is pleasing that, as it happened, the correct result was reached by the actual count, but it could so easily have been the wrong one.

It has sometimes been suggested that messing about with such small fractions of votes, which make no difference to the result, is not worth while. There are two answers to that suggestion. The first is that, if the logic of the Meek method is accepted, then either we can follow that logic through, even if it does result in such "messing about with small fractions", which is easy, or we can put in special rules to stop it doing so, which is much more difficult. We should need to consider not only what special rules to adopt in such cases, but also how to determine when to use them. Obviously it makes sense to do the easy, and correct, thing.

The second answer is that there are cases where such a very small difference can change the answer, so it would be wrong to ignore a 15th preference. If the contest between V and F had reached an exact tie from all the other relevant votes, then the result should, of course, have been settled by what that 15th preference was.

1 References

[1] R A Newland and F S Britton. *How to conduct an election by the Single Transferable Vote.* 2nd edition. Electoral Reform Society. 1976.

[2] I. D. Hill, B. A. Wichmann and D. R. Woodall. Algorithm 123 — Single Transferable Vote by Meek's method. *Computer Journal* Vol 30, pp271-281. 1987.

Full disclosure of data

I. D. Hill
No email available.

The objection to full disclosure made by Otten [1] is valid, but seems to me to be of only minor importance. Considering the huge advantages of disclosure, in giving transparency and allowing anyone who wishes to check the result of the counting, it would be a great pity if Otten's point were allowed to prevail over it.

Disclosure does not in itself give complete transparency of the electoral process, because it takes as given the list of votes and their preferences, but in dealing fully with the second part of the process, the counting of the votes, it is nevertheless of great merit.

Otten's "preferred solution" — to suppress later preferences until there are at least three votes of every published pattern — would undoubtedly be better than not publishing the data at all, but it is a very poor thing compared with full disclosure and would, in many instances, lead to the suppression of the very information that would be of importance.

Taking as an example the election reported on in the preceding paper (Hill [2]), the original votes, which had 531 different preference patterns from the 539 votes, would have been reduced to only 96 different patterns, and these would not have shown the vital information that led to the allocation of the final seat. Indeed the 16 votes that put candidate M first would have been shown as just 13 M . . . and 3 M R . . . The voter whose 15th preference was vital would not have had even a second preference shown.

In an election where political parties were important, it would seem likely that the loss of information would be less severe. Even in the given case, the fact that there were 7 votes starting Q P O S E F H A D J M C B R, and another 3 also starting Q P O S E F, still comes through, indicating obvious collusion between voters (which is not illegal, or even immoral, if that is what they wish to do).

Implementing the Otten procedure is not straightforward, as it is not sufficiently defined. For example, there were 2 votes starting W U A I D, 1 starting W U A I O, 1 starting W U A E. Should these be shown as 4 of W U A . . . , or as 3 W U A I . . . leaving the other 1 to go in with W . . . ? It is not self-evident.

There are many things in life that could be so much simpler if only we could trust everybody, and did not need to bother about fraudsters, but we always need to consider whether a particular fraud is likely, and whether procedures to stop it are doing more harm than good. My personal view is that Otten's suggestion would be doing so.

1 References

[1] Otten J. Fuller disclosure than intended. *Voting matters*, Issue 17, p8. 2003.

[2] Hill I.D. An odd feature in a real STV election. *Voting matters*, Issue 18, p9. 2004.

A note on the use of preferences

B A Wichmann
brian.wichmann@bcs.org.uk

1 Introduction

With STV, the voter is encouraged to specify as many preferences as may be needed to reflect his/her wishes. The number of preferences actually used within the count is quite a different matter which is the main subject of this note.

For the three Irish constituencies for which a trial was undertaken in 2002 of electronic voting, we have full disclosure of the preferences specified by the voters. This provides an opportunity to analyse the use of preferences in a large public election in some depth.

Joe Otten has stated reservations about the full disclosure of preferential voting data on the grounds that it could allow bribery to take place even though the voting is secret [1]. The issue has also been raised by the Irish Commission on Electronic Voting [3].

Here, we consider how the voter's preferences are used and propose alternative solutions to the problem of disclosure.

2 The use of the voter's preferences

It is clear that any preference listed after a continuing candidate cannot be used at that stage of the count. To inspect such a preference would contravene one of the principles of STV. A particular example of this is that those voters who gave their first preference for a candidate who is still a continuing candidate at the end of the count, will not have anything other than their first preference used.

As an example of how preferences are used, consider the 2002 Dáil election for the Meath constituency for which we have full election data. There were 14 candidates for 5 seats (the candidate names have been abbreviated to give only the gender and position in the tables). The election stages were as follows:

Stage 1	Elect M4
Stage 2	Exclude F3 and M11
Stage 3	Exclude M9
Stage 4	Exclude M8
Stage 5	Exclude M10
Stage 6	Exclude M14
Stage 7	Exclude M6
Stage 8	Exclude M7, Elect M2
Stage 9	Elect M1, M5 and F13

Hence the continuing candidate is M12.

Now consider an actual voter whose preferences were as follows:

M9 M8 M7 M10 M12 M11 M14 F3 F13 M1 M4 M2 M6 M5

Consulting the actions of the stages above, it is clear that the preferences for M10 and all those after M12 were never used. In other words, the voter could just as well have voted: M9, M8, M7, M12. The other preferences were *invisible*.

To understand the use of the preferences in more detail, we look at the result sheet in Table 5.1. At the second stage, the surplus of M4 is transferred. To do this, all of the 11,534 votes for M4 are inspected and the number whose second preference is given is found, together with the proportion for each of the remaining 13 candidates. Since 853 votes must be transferred to reduce M4 to the quota, an integer is computed for each candidate giving the correct proportion and total. As an example of a transfer, only one vote is transferred to M11 and that vote is selected at random from those giving M11 as the second preference. This implies that 10,681 votes are inspected for their subsequent preference and a further 853 votes are used in the subsequent stages.

Hence we have two uses of preferences with the Irish rules: those used directly to attempt to elect a candidate and those used indirectly to determine which papers to select at random to transfer. For the Meath election, the number of preferences used directly are those for the first preference (the total vote of 64,081) plus the number of those in the table with a + sign but ignoring those in the non-transferable row. The indirect use, which only arises from a transfer of surplus is therefore only from M4, i.e, the 10,681 mentioned above.

In contrast to this, the Meek method uses all the visible preferences. Our sample ballot paper above had four visible preferences M9, M8, M7 and finally M12. In fact, the Irish rules would use all these preferences.

We can now compute the use of the preferences for the three Irish constituencies, expressed as an average per vote:

Constituency	Irish–direct	Indirect	Meek	All
Meath	1.19	0.17	1.98	4.65
Dublin North	1.33	0.01	2.12	4.98
Dublin West	1.26	0.25	2.11	4.43
Average of 3	1.26	0.14	2.07	4.68

Hence, as a percentage of all the preferences given, the direct use with the Irish rules is 27%, indirect usage is 3%, while Meek uses 44%.

3 Full disclosure?

We can now see that relatively few preferences are actually used in a count. If the voter specifies a large number of preferences, then it is unusual for them all to be used. For an example of a large number of preferences which were used, see [2].

We now have a means of providing an approximation to full disclosure which would nevertheless allow the voter to check the actual count: remove some (or all) of the invisible preferences. For long preference lists, like the one shown above, it would usually be the case that many preferences would be invisible. Hence this strategy of providing full disclosure only of the visible preferences would effectively prohibit the potential problem identified by Joe Otten.

Note that the identification of the invisible preferences depends upon the order of the exclusions and elections which in turn depends upon the particular counting rules being used. Hence, if data were provided with only the visible preferences, then running that data using a different counting rule would not necessarily give the same result as using the actual data.

4 Conclusions

Since many preferences are not used in a count, it is possible to disclose all the used preferences and remove all or part of the unused preferences to avoid any potential breach of confidentiality. The referee made two additional points: it is possible to *add* invisible preferences as well as removing them; and that *any* change to the data implies that a check is not an exact check.

5 References

[1] J Otten. Fuller Disclosure than Intended. *Voting matters*. Issue 17. p 8. 2003.

[2] I D Hill. What would a different method have done? *Voting matters*. Issue 16. p 5. 2003.

[3] Interim Report of the Commission on Electronic Voting on the Secrecy, Accuracy and Testing of the chosen Electronic Voting System. `http://www.cev.ie/htm/report/download_report.htm`

		Surplus M4	Exclude F3+M11	Exclude M9	Exclude M8	Exclude M10	Exclude M14	Exclude M6	Exclude M7
M1	8493	+258 8751	+36 8787	+46 8833	+46 8879	+108 8987	+123 9110	+467 9577	+299 9876
M2	7617	+76 7693	+32 7725	+155 7880	+241 8121	+333 8454	+694 9148	+1733 10881	10881
F3	263	+2 265	−265 —	—	—	—	—	—	—
M4	11534	−853 10681	10681	10681	10681	10681	10681	10681	10681
M5	5958	+61 6019	+52 6071	+68 6139	+126 6265	+374 6639	+737 7376	+1349 8725	+1429 10154
M6	3877	+15 3892	+11 3903	+34 3937	+41 3978	+74 4052	+221 4273	−4273 —	—
M7	3722	+29 3751	+56 3807	+113 3920	+185 4105	+359 4464	+675 5139	+119 5258	−5258 —
M8	1373	+7 1380	+23 1403	+163 1566	−1566 —	—	—	—	—
M9	1199	+3 1202	+42 1244	−1244 —	—	—	—	—	—
M10	2337	+16 2353	+53 2406	+224 2630	+200 2830	−2830 —	—	—	—
M11	180	+1 181	−181 —	—	—	—	—	—	—
M12	6042	+51 6093	+51 6144	+123 6267	+118 6385	+325 6710	+412 7122	+226 7348	+732 8080
F13	8759	+313 9072	+32 9104	+180 9284	+361 9645	+362 10007	+254 10261	+113 10374	+1261 11635
M14	2727	+21 2748	+21 2769	+75 2844	+120 2964	+631 3595	−3595 —	—	—
Non-T	—	—	+37 37	+63 100	+128 228	+264 492	+479 971	+266 1237	+1537 2774
Totals	64081	64081	64081	64081	64081	64081	64081	64081	64081

Table 5.1: Meath, 2002: Quota: 10681. Those elected have their names in italics.

Tie-Breaking with the Single Transferable Vote

Jeffrey C. O'Neill
jco8@cornell.edu

1 Introduction

In tallying the single-transferable vote (STV), ties can occur for several different reasons. With the ERS97 rules [1] for implementing STV, ties can occur when choosing a surplus to transfer (5.2.3), when choosing a candidate to eliminate (5.2.5), and when choosing winners (5.6.2). To illustrate, Table 6.1 shows an example tally with the ERS97 rules. At stage 4, we need to eliminate the candidate with the fewest number of votes, but both C and D are tied for last place.

When ties occur, they need to be broken. One could simply break the tie by lot. However, since there is other information available in an STV count, one can use this information to break the tie. The following are four possible tie-breaking rules.

1. Forwards Tie-Breaking: Choose the candidate who has the most [least] votes at the first stage or at the earliest point in the count where they had unequal votes.

2. Backwards Tie-Breaking: Choose the candidate who has the most [least] votes at the previous stage or at the latest point in the count where they had unequal votes.

3. Borda Tie-Breaking: Choose the candidate with the highest [lowest] Borda score. See [2].

4. Coombs Tie-Breaking: Choose the candidate with the fewest [most] last place votes.

It is possible that after applying one of these tie-breaking rules that the candidates would still be tied. Because of this, it is useful to distinguish between "weak ties" and "strong ties." A weak tie occurs when candidates have the same number of votes at a given stage. A strong tie occurs when candidates are still tied after applying a tie-breaking rule, such as one of the four listed above. A strong tie would be broken by lot.[1]

The ERS97 rules use forwards tie-breaking. The purpose of this paper is twofold. First, to show that backwards tie-breaking is a better solution and to suggest that the ERS97 rules be changed to use backwards tie-breaking instead. Second, to show that substage totals should not be used when breaking ties.

2 Backwards or Forwards Tie-Breaking

In breaking a tie, the ERS97 rules state that one must choose "the candidate who had the greatest vote [or fewest votes] at the first stage or at the earliest point in the count, after the transfer of a batch of papers, where they had unequal votes." This is forwards tie-breaking and is used when choosing a surplus to transfer (5.2.3), when choosing a candidate to eliminate (5.2.5), and when choosing winners (5.6.2).

The difference between backwards and forwards tie-breaking will be illustrated with the example in Table 6.1. In this example, we have to eliminate one candidate at stage 4 and there is a weak tie between candidates C and D. Thus, tie-breaking needs to be used to determine which candidate is to be eliminated. Under ERS97 rules, we break the tie by using forwards tie-breaking. To do this we first look to the counts at stage 1. We see that D has one more vote than C at stage 1. Thus, candidate C is eliminated.[2]

Another alternative is to use backwards tie-breaking. To do this, we look at the previous stage to break ties, and if necessary to preceding stages. Looking at the

[1] Of course one could use another tie-breaking rule if the first tie-breaking rule results in a tie, but this will not be considered here. Borda and Coombs tie-breaking are just presented as available alternatives and will not be discussed further.

[2] If C and D had been tied at stage 1, then we would have looked to subsequent stages. If C and D had been tied at all stages, then we would have had a strong tie which would have been broken by lot.

preceding stage, we see that C is ahead of D at stage 3. Thus, D would be eliminated.

One problem with forwards tie-breaking is that it looks at the stages in an order that is not sequential. In order to determine the candidate to be eliminated at stage 4, we would look at the stages in the following order: 4 1 2 3. Intuitively, this is undesirable. It makes more sense to look at the stages in sequential order. Since one must look first to the current stage, there is only one sequential ordering: 4 3 2 1. This is what backwards tie-breaking would do.

A more important problem, is that forwards tie-breaking does not use the most relevant information to break the tie. The most relevant information to break a tie is the previous stage and not all the way back to the very first stage. By immediately looking to the first stage to break the tie, the ERS97 rules allow the tie-breaking to be influenced by candidates eliminated very early in the process and also by surpluses yet to be transferred. Instead, if we look to the previous stage to break a tie, candidates eliminated early on in the process will have no influence in breaking the tie. In addition, it allows for surpluses to be transferred which gives a more accurate picture of candidate strength.

In Table 6.1, candidate C has more support than candidate D at stage 3. At this point, the surplus of A has already been transferred and candidate F has already been eliminated. Thus, stage 3 is a better measure than is stage 1 as to which candidate should be eliminated at stage 4.

Other implementations of the single transferable vote use backwards tie-breaking instead of forwards tie-breaking: Cambridge, MA STV [3], rules advocated by the Center for Voting and Democracy [4], and rules advocated by the Proportional Representation Society of Australia [5].

3 Elimination of Winning Candidates

An incidental problem related to using forwards tie-breaking is that the ERS97 rules can sometimes eliminate a winning candidate. Consider an example where 31 voters elect one candidate with the following ballots:

4	voters vote	ABC
5	voters vote	BC
5	voters vote	CB
2	voters vote	DABC
4	voters vote	EABC
11	voters vote	F

Table 6.2 shows the results of the tally with ERS97 rules.

At stage 3 of the count, we need to eliminate one or more candidates and candidates B and C are tied with the fewest votes. According to rule 5.2.5(b), both B and C are to be eliminated. However, if instead the tie between B and C was broken by lot, then the other candidate would go on to win the election! In this scenario, suppose candidate C was eliminated by lot at stage three. Then B would be tied with A at stage 4, each with 10 votes. Forwards tie-breaking would be used to break the tie. Candidate A has the fewest votes at stage 1 and would then be eliminated. B would then receive all of A's votes and beat F 20 to 11 in the final stage.

Thus, the ERS97 rules are over-aggressive in eliminating candidates. This is a clear flaw in the ERS97 rules. This flaw arises from the interaction of rule 5.2.5(b) and forwards tie-breaking. This flaw could be fixed in two ways: (1) by changing rule 5.2.5(b), or (2) by using backwards tie-breaking instead of forwards tie-breaking. Since there are already other good reasons for using backwards tie-breaking, the obvious choice is (2).

If backwards tie-breaking were used instead, then both candidates B and C could properly be eliminated at stage 3. If just C were eliminated and B received all of C's votes, then there would again be a tie at stage 4. However, with backwards tie-breaking, B would necessarily have fewer votes than A at the previous stage and would immediately be eliminated.

Backwards tie-breaking would fix this flaw generally, and not just in this specific example. This flaw occurs under specific conditions:[3] (1) a candidate needs to be eliminated and two candidates are tied for last place, (2) the sum of the votes of these two candidates is equal to the candidate with the next fewest number of votes, and (3) after eliminating one of these candidates there would be a subsequent tie with this third candidate. Under these conditions rule 5.2.5(b) requires that the two candidates in last place be eliminated simultaneously. As described above, with forwards tie-breaking a winning candidate could be improperly eliminated. However, with backwards tie-breaking, both of these last-place candidates cannot win and can thus be properly eliminated. The two last-place candidates are guaranteed to lose the second tie because they necessarily

[3] These conditions could be generalized to the case where more than two candidates are tied for last place.

have fewer votes at the previous stage (but they do not necessarily have fewer votes at the first stage).

4 Use of Substages to Break Ties

The word "substage" is not used anywhere in the ERS97 rules, but this terminology is used by people familiar with the rules. Substages can occur when transferring votes from eliminated candidates. Table 6.3 shows an example using ballots from the test T143 where 60 voters are electing two candidates. At stage 3, candidate F is being eliminated. Candidate F has ballots with transfer value 1.00 and ballots with transfer value 0.25 (from the surplus of A). These ballots will be transferred in two substages constituting two different batches. The first substage transfers ballots with value 1.00 and the second transfers ballots with value 0.25.

In stage 4 of this example, we need to eliminate a candidate and candidates C and D are tied for last place. Hence, we need to use forwards tie-breaking. With ERS97 rules, substages must be considered when doing forwards tie-breaking. Candidates C and D are also tied at stage 1 and stage 2, but candidate D is ahead of candidate C at the substage between stages 2 and 3. Thus, candidate C is eliminated.

The problem is that substages are not a good metric for breaking ties. In the example in Table 6.3, either candidate C or D must be eliminated at stage 4. Candidates C and D are tied at stages 4, 1, and 2. Candidate C is ahead at stage 3, but candidate C is eliminated anyway! The reason that C is eliminated is that D has more votes at an intermediary point where only some of candidate F's votes have been transferred. This intermediate point is well-defined but completely arbitrary in terms of fairness. There is no reason to make some of F's votes more important than others. Whether one candidate is ahead of another at this intermediary point is not relevant to which candidate should be eliminated. What is relevant, is what the counts are at each stage of the count, that is after a candidate has been completely eliminated.

5 Conclusions

The ERS97 rules should be changed so that backwards tie-breaking is used instead of forwards tie-breaking. In addition, substage totals should not be considered when breaking ties.

6 Acknowledgments

The author would like to thank David Gamble, Brian Wichmann, and an anonymous reviewer for their valuable suggestions regarding this paper.

7 References

[1] R A Newland and F S Britton. *How to conduct an election by the Single Transferable Vote*. ERS 3rd Edition. 1997. See http://www.electoral-reform.org.uk/votingsystems/stvrules.htm.

[2] Earl Kitchener, Tie-Breaking in STV. *Voting matters*. Issue 11, April 2000.

[3] Massachusetts General Laws, Chapter 54A, Section 9(k).

[4] Choice Voting. The Center for Voting and Democracy. See http://www.fairvote.org/library/statutes/choice_voting.htm

[5] Rules of the Proportional Representation Society of Australia for conducting elections by the quota-preferential method. 2001. See http://www.prsa.org.au/rule1977.htm

		Surplus of A	Eliminate F	Eliminate E	Eliminate C
Stage	1	2	3	4	5
A	23	20.00	20.00	20.00	20.00
B	13	13.00	13.00	15.00	15.00
C	6	6.50	10.00	12.00	2.00
D	7	7.50	9.50	12.00	18.00
E	7	7.50	7.50	-	-
F	4	5.50	-	-	-
Non-Transferable	0	0.00	0.00	1.00	5.00

Table 6.1: Example tally with ERS97 rules where 60 voters are electing two candidates.

		Eliminate D	Eliminate E	Eliminate B & C
Stage	1	2	3	4
A	4	6.00	10.00	10.00
B	5	5.00	5.00	-
C	5	5.00	5.00	-
D	2	-	-	-
E	4	4.00	-	-
F	11	11.00	11.00	11.00
Non-Transferable	0	0.00	0.00	10.00

Table 6.2: Example where the ERS97 rules eliminate a winning candidate.
Thirty-one voters are electing one candidate. Candidate F is the winner.

		Surplus of A	Eliminate F		Eliminate E	Eliminate C
Stage	1	2	substage	3	4	5
A	23	20.00	20.00	20.00	20.00	20.00
B	13	13.00	13.00	13.00	15.00	15.00
C	7	7.50	8.50	10.00	12.00	2.00
D	7	7.50	9.50	9.50	12.00	18.00
E	6	6.50	6.50	6.50	-	-
F	4	5.50	1.50	-	-	-
Non-Transferable	0	0.00	1.00	1.00	1.00	5.00

Table 6.3: ERS97 rules with substage tie-breaking.
Sixty voters are electing two candidates.

ISSN 1745-6231

Voting matters

for the technical issues of STV

published by

The McDougall Trust

Issue 19 November 2004

About the McDougall Trust (reg. charity no. 212151)

The McDougall Trust is a charitable trust formed in 1948. The charity's purposes as stated in its governing scheme of 1959 are to advance knowledge of and encourage the study of and research into:

- political or economic science and functions of government and the services provided to the community by public and voluntary organisations; and
- methods of election of and the selection and government of representative organisations whether national, civic, commercial, industrial or social.

The Trust's work includes the maintenance and development of the Lakeman Library for Electoral Studies, a unique research resource, the production and publication of Representation: The Journal of Representative Democracy, and, of course, this publication **Voting matters**, that examines the technical issues of the single transferable vote and related electoral systems.
For further information on the Trust, please contact:

The Secretary,
McDougall Trust,
6 Chancel Street,
London SE1 0UX, UK.
Telephone: +44 (0)20 7620 1080
Facsimile: +44 (0)20 7928 1528
Email: admin@mcdougall.org.uk
Web: www.mcdougall.org.uk

For further information on this publication, please contact B A Wichmann, the Editor at the above address or by email at: `Brian.Wichmann@bcs.org.uk`

Editorial

Report by Steve Todd

On 9 October this year, New Zealand held a number of STV elections using the Meek counting rules. Several problems arose which delayed the final declaration of the results. It appears that the main problem concerned reconciling the number of voting papers that were scanned into the database with the number that were subsequently sent to the STV calculator.

The realisation that discrepancies were occurring led the local councils and district health boards (DHBs) affected, to call in the Auditor-General's office to audit the entire process. While the computer error was discovered and fixed within a few days, the auditing process meant that it took four weeks to complete all the vote-counting. In contrast, the program which actually performed the count, i.e. the STV calculator, appeared to operate without mishap.

A lesser, but equally frustrating, problem was that the ICR technology used to process the ballot papers was unable to read (with a high level of confidence) a considerably higher percentage of the scanned documents than was expected. This led to much more human intervention than was expected, with a consequent increase in the time taken to process the votes.

The Justice and Electoral select committee of New Zealand's parliament intend to conduct an inquiry into what went wrong. A focus of the inquiry will likely be on why the two Auckland-based companies contracted to process the STV votes in the northern part of the country, did so seemingly without a hitch, and in a timely manner, while the Christchurch and Wellington companies contracted to conduct the remaining STV elections (in respect of 7 of 10 councils and 18 of 21 DHBs) did not.

There has not yet been a full explanation of the problems encountered, but there is a suggestion that the computer systems used by the Christchurch and Wellington companies may not have been completely compatible.

There were also widespread claims of voter confusion (said to have been caused by having FPTP and STV elections on the same A3-size voting documents), leading to many Informal (Invalid) votes (errors) and blank votes (non-participation) being cast, that the select committee will no doubt inquire into.

Informal votes in council areas using STV appear to have been no more than usual — 1.08% in Wellington and 1.49% in Dunedin, for example. However, in the remaining 64 council areas, that used FPTP, the Informal rate in respect of their DHB elections was up as high as 10 to 12%.

A likely explanation for this will be poor voting-document design. There was no bold distinction between FPTP and "tick-voting" for the mayoral and council ward elections, and STV and voting by numbering the candidates in the DHB elections. In fact, apparently due to printing restrictions, the DHB elections were set out under the name of the city or district councils they were associated with! This means that some voters (who did not read the voting instructions carefully) carried on tick-voting into the DHB election — more than one tick for the candidates and the vote was informal.

On the brighter side, the actual ballot data is likely to be made available in respect of most, perhaps all, STV elections and hence it will be possible to 'check' the counts by re-running them.

Voting matters

There are 3 papers in this issue:

- B. A. Wichmann: Tie Breaking in STV. This paper considers a method of handling ties when a computer is used by considering all possible outcomes. It is an unfortunate fact that breaking a tie by a random choice gives an impression that the outcome might be random when this is rarely the case.

- J. Green-Armytage: Cardinal-weighted pairwise comparison. This paper considers the election of a single candidate by adding information to a Condorcet-style count on the strength of the preferences for candidates.

- B. A. Wichmann: A Working Paper on Full Disclosure. This paper attempts to put together major concerns about this issue which have been raised in previous issues of *Voting matters*. The paper was written before the New Zealand election data became available and hence does not mention this.

Tie Breaking in STV

B. A. Wichmann
Brian.Wichmann@bcs.org.uk

1 Introduction

Given any specific counting rule, it is necessary to introduce some words to cover the situation in which a tie occurs. However, such ties are only a practical concern for small elections. For instance, it has been reported that a tie has never occurred with the rules used in the Irish Republic.

Probably the most common form of a tie is when the two smallest first preference votes are the same. Unless both candidates can be excluded, a choice must be made, although in very many cases, the candidates elected will be the same.

This note proposes that when a computer is used to undertake a count, all the possible choices should be examined and that the result is produced by computing the probability of election of the candidates.

2 Ties in practice

It is clear that the propensity to produce a tie will depend largely on the number of votes. However, some estimate can be obtained from a collection of election data that has recently been revised [1]. The data base consists of over 700 ‘elections’, but for this paper we exclude artificial test cases. The figures obtained from the other cases, which are like real elections, with three counting rules ([4, 2, 7]) are as in the table on page 4.

Hence, although with the Church of England rules, only 59 out of 299 involved a tie-break, the average number of tie-breaks in those 59 was actually 9.9. The average number of votes in those 59 cases was 102, while the average for the remaining 240 cases was 12,900. It is important to note that Meek only has ties on an exclusion of a candidate, while the hand-counting rules also have ties on the choice of the candidate whose surplus is to be transferred.

For reasons not relevant to this note, the number of cases run with each rule is different. (Larger cases have only been run with Meek.) It is clear that a small number of votes increases the risk of a tie. Also, given that a tie occurs, the Meek algorithm has only half the risk of a subsequent tie arising, almost certainly due to the higher precision of the calculation.

3 The special case of ties with the Meek algorithm

Brian Meek’s original proposal rests upon the solution of certain algebraic equations. The algorithm given in [7] provides an iterative solution of those equations. The mathematical nature of the equations implies that there is substantial freedom in handling exclusions, since, once a candidate is excluded, it is as if the candidate had never entered the contest. Hence it is not necessary for two implementations of Meek to handle exclusions in the same way — the same candidates will be elected. (In contrast, the hand counting rules need to be specific on exclusions since it affects the result; ERS97 insists on as many as allowable, while CofE insists on only one at a time.)

As an example, David Hill’s implementation of Meek in comparison with my own has revealed differences. We both exclude together all those candidates having no first preferences. David Hill also excludes the next-lowest candidate also (assuming it is safe to do so), while I do not. I will exclude more than one candidate at a time when it is safe to do so, while David Hill sticks to one at a time. Hence both our implementations report a random choice has been made when it is certainly possible to avoid this. Such reporting is undesirable since it might give the impression that those elected have been chosen at random, when this is not the case. Both of us have introduced a tie-breaking rule similar to that in many hand-counting rules based upon the votes in previous stages (but in opposition to that advocated in [5]).

Two other aspects are relevant to the Meek algorithm. The cases reported in [6] indicate that an implementation can report a tie even though in mathematical terms, one candidate is ahead (but by too small an amount to be computed). This situation is not thought to arise in practice. Perhaps somewhat more disturbing is that an algebraic tie can be computed differently, giving one candidate ahead of another. Two implementations of Meek with such a case can even break the actual tie by rounding in different directions. However, since there is a real tie, breaking it by the rounding in the implementation, is not so bad.

4 Results of the proposed method

The only practical method to implement this proposal is to modify software that already implements an existing counting rule. Since I have my own implementation of Meek, I have modified this to analyse all choices when a tie occurs.

The modification works by executing the algorithm once for every possible choice when the rules require a 'random' choice. For my version of Meek, I have provided an option to remove the first-difference rule so that when this rule would otherwise be invoked, a random choice is made[1].

As an example, consider a real (simple) election, R033, having four candidates (A1...A4) for one seat. At the first stage, A2 and A3 have the smallest number of votes: if A2 is excluded, then A1 is elected; if A3 is excluded, then there is a tie between A2 and A4 for the next exclusion. These two alternatives also result in A1 being elected. So the final result is:

Probability from 5 choices from 3 passes.

Candidate	*Excluded?*	*Probability*
A1	no	1
A2	yes	0
A3	yes	0
A4	yes	0

We now know that the election of A1 is not dependent upon the random choices made. The computation involved three election runs. The middle column indicates that the candidates A2, A3 and A4 were all selected in one of the runs for random exclusion.

A more complex example is given by R009, electing 2 from 14 candidates with 43 votes. Here, the final table reads:

Probability from 1364 choices from 264 passes.

Candidate	*Excluded?*	*Probability*
A1	no	1/ 4
A2	yes	0
A3	yes	0
A4	yes	0
A5	no	0
A6	yes	0
A7	yes	0
A8	yes	0
A9	no	1
A10	yes	0
A11	no	3/ 4
A12	yes	0
A13	yes	0
A14	no	0

Here we see that only the candidates A1, A5, A9, A11 and A14 were never subject to random exclusion. Nevertheless, A5 and A14 were never elected.

However, the above result was using the variant of Meek without the first-difference rule. If the first-difference rule had been applied, then A1 would not have been elected in any circumstances. Note that in this case, a large number of passes had to be made due to many of the stages resulting in a tie. Hence this technique is only really possible due to the speed of modern computers.

Given the above election, then there are two possible uses of the outcome: firstly to elect the most probably candidates (A9 and A11), or secondly, to randomly select between A1 and A11 according to the specified possibilities. Since in this paper we are attempting to reduce the random element, we choose the first option.

From the database, 55 cases were selected which correspond reasonably closely to real elections. The results are in the table on page 5. The entry 'Random' gives the number of random choices made with the New Zealand version of Meek which has the first-difference rule. The last three entries are from running the new program. The 'Probs.' column includes the probabilities of election of those candidates who are involved in ties and have nonzero probability of election.

[1] The first-difference rule is a method of breaking a tie by examining the votes in all previous stages, starting at the first stage and selecting the one which has the fewest votes at the first stage at which there is a difference. Of course, if the earlier stages give no difference, then a random method must be used to break the tie.

The three examples with approximate results from the new program took too long to run to completion. Here, the tabulated results are based upon the first few thousand cases executed. The majority ran very quickly and only those with 10,000 or more passes took longer than a minute or two. The case R038 was exceptional in having probabilities of 29/168, 11/35, 29/60, 431/840, 431/1680, 437/1680 and 1 (and none were repeated).

If one was only concerned with the Meek algorithm, then the program could probably be made substantially faster since the ties only arise with an exclusion and Meek is indifferent to the order of the exclusions in the sense that excluding A then B is the same as excluding B then A; this situation will typically be the case when A and B tie on the fewest number of votes. The approach here is a general one that could be applied to any counting rule. It also seemed easier to program the general method presented here.

From the 49 cases which were run to completion, all but 7 reported than the random choice had no effect upon the result.

Election R102 is typical of the situation in which a large number of random choices are made. In fact, 28 exclusions are made before an election. This implies that for all these initial stages, the votes are integers. Given the small size of the election, ties are very common. Unfortunately, this implies that the number of choices is too large to compute them all. However, experimenting with removing those candidates who are excluded early, gives the result shown in the last column.

Followers of the Eurovision Song Contest might like to know that although the official scoring system gave a tie in 1991 between Sweden and France, with Sweden being judged the winner on the basis of having more second (preference) votes, this system gives Sweden a probability of election of 71/288 and hence France the clear winner with a probability of 217/288. According to this system, the UK would have won in 1992 with a probability of 5/6, while the official result declared Ireland as the winner which had a probability of only 1/12.

5 Conclusions

It seems that the provision of this program raises more problems than it solves. If one is prepared to ignore the 14% of cases which question the validity of the random choice, then one can continue the current practice with a clear conscience. On the other hand, when a random choice was made in a real election, it would surely be welcome to show that the result was not in question. However, using this program for that purpose might not give a clear answer when only a fraction of all the possibilities could be executed in a reasonable time (as with the three cases in the table). Of course, in those cases, numerous random choices could be tried, but the object here is to avoid such arbitrariness.

When a candidate has been subject to a random exclusion in an election, he/she could naturally feel aggrieved. One solution to that would be to undertake a re-count without randomly excluding that candidate. If this were undertaken by computer, the number of recounts would be less than the number of candidates and hence very much less than all possibilities which are considered above.

Currently, almost all STV counting rules introduce some rules, like the first-difference rule ([2, 4]) or Borda scores [3], to reduce the need for a random choice to be made. An alternative would be to simplify the counting rules by omitting these provisions, but to use a program like the one presented here to produce a result which is very likely to have no random element.

6 References

[1] B. A. Wichmann. A Guide to an STV data base. July 2004. (Available from the author.)

[2] R. A. Newland and F. S. Britton. *How to conduct an election by the Single Transferable Vote*. ERS 3rd Edition. 1997. See `http://www.electoral-reform.org.uk/votingsystems/stvrules.htm`

[3] Earl Kitchener, Tie-Breaking in STV. *Voting matters*. Issue 11, April 2000.

[4] GS1327: General Synod, Single Transferable Vote regulations 1990 and 1998. (Obtainable from Church House Bookshop, Great Smith Street, London SW1P 3BN.)

[5] J. C. O'Neill: Tie-Breaking with the Single Transferable Vote. *Voting matters* Issue 18. pp 14-17. 2004.

[6] B. A. Wichmann. The computational accuracy using the Meek algorithm. *Voting matters* Issue 12. pp6-7. November 2000.

[7] I. D. Hill, B. A. Wichmann and D. R. Woodall. Algorithm 123 — Single Transferable Vote by Meek's method. *Computer Journal*. Vol 30, pp277-281, 1987.

Rule	*Ties*	*Ties per case*	*Average votes with ties*	*Average votes without ties*
CofE	59 from 299	9.9	102	12900
ERS97	55 from 154	7.1	81	2438
Meek	62 from 587	3.3	12692	44180

Table 1.1: Ties with different election rules

ID	Votes	Candidates	Seats	Random	Choices	Passes	Probs.
M002	131	20	5	1	2	2	1
M112	692	25	6	1	2	2	1
R009	43	14	2	4	1364	264	1/4, 3/4, 1
R012	79	17	2	4	256	48	1
R015	83	19	3	6	32640	3840	1
R017	76	20	2	5	64776	7200	1
R018	104	26	2	11	—	$\approx 6 \times 10^{6}$	1?
R019	73	17	2	5	3876	672	1
R020	77	21	2	5	42184	4572	5/24, 19/24, 1
R027	44	11	2	4	114	30	1
R028	91	29	2	8	—	$\approx 5 \times 10^{6}$	1?
R033	115	4	1	1	5	3	1
R038	9	18	3	3	387	115	*see text*
R040	176	17	5	1	2	2	1
R097	45	17	1	6	283742	31190	1
R100	1031	31	10	1	2	2	1
R102	247	49	10	15	—	$\approx 34 \times 10^{6}$	1/12, 1/4, 1/6, 2*5/6, 2*11/12, 6*1?
S002	16	16	1	1	8	4	2 of 1/2
S003	16	16	1	1	7	5	1
S004	20	20	1	2	12	6	1
S005	18	18	1	1	3	3	1
S006	20	20	1	3	60	18	1
S007	19	19	1	2	46	14	1
S008	19	19	1	3	106	31	1
S009	20	20	1	2	20	10	1
S010	22	22	1	3	448	106	1
S011	21	21	1	4	465	97	1
S012	22	22	1	1	2	2	1
S013	22	22	1	3	3888	624	1
S014	22	22	1	1	176	44	71/288, 217/288
S015	23	23	1	3	646	126	2 of 1/12, 5/6
S016	25	25	4	1	2	2	1
S022	25	25	1	4	1592	329	1
S023	23	23	1	3	288	60	1
S024	17	16	1	2	58	16	1
S025	18	18	1	4	480	96	2 of 1/2
S026	18	18	1	5	39703	6297	1
S027	13	19	1	2	30	12	1
S028	17	18	1	3	229	68	1
S029	18	18	1	2	16	7	1
S030	20	20	1	4	1368	288	1
S031	19	19	1	1	2	2	1
S032	16	19	1	2	16	7	1
S033	22	23	1	2	1132	206	1
S034	25	25	1	4	5774	1072	1
S035	25	25	1	6	70560	10080	1
S036	23	23	1	5	14400	2304	1
S037	24	24	1	2	16	7	1
S038	23	23	1	2	28	10	1
S039	26	26	2	5	17760	2880	1
S047	36	24	1	6	12144	1800	1
S048	24	24	1	6	161280	20160	1

Table 1.2: All results from exhaustive tie-breaking

Cardinal-weighted pairwise comparison

J. Green-Armytage
jarmyta@antioch-college.edu

James Green-Armytage graduated from Antioch College in April of 2004. He is considering a graduate program for an economics PhD.

1 Introduction

This paper introduces a new voting method named **cardinal-weighted pairwise comparison**, or **cardinal pairwise** for short. It is based on Condorcet's method of pairwise comparison, but in addition to asking voters to rank the candidates in order of preference, this method also asks them to rate the candidates, for example on a scale from 0 to 100. The ordinal ranking information is still used to decide the winner and loser of each pairwise comparison, but the cardinal rating information is used to decide the relative strength of the pairwise victories/defeats, which determines how majority rule cycles are resolved if they occur.

Sections 2 through 4 are primarily concerned with definition, and sections 5 through 7 are primarily concerned with analysis and justification. In sections 2, 3 and 4, I define some key terms, define the cardinal pairwise method, and give an example computation. In section 5, I argue that pairwise methods in general are superior to other voting methods when the goal is majority rule. In sections 6 and 7, I discuss the advantages of cardinal pairwise over other pairwise methods, which are as follows: First, it takes into account the relative priority of each pairwise preference to each voter. Second, it may greatly reduce the vulnerability to strategic manipulation that is troublesome for pairwise methods.

2 Preliminary definitions

Pairwise comparison, pairwise defeat, pairwise tie: A pairwise comparison uses ranked ballots to simulate head-to-head contests between different candidates. Given two candidates A and B, there is a pairwise defeat of B by A if and only if A is ranked above B on more ballots than B is ranked above A. If the number of A>B ballots is equal to the number of B>A ballots, then there is a pairwise tie between A and B.

> and = symbols: I use these in two slightly different ways. For example, "A>B" can mean that an individual voter or a specific set of voters ranks A above B, and it can also mean that A has a pairwise victory over B. "A=B" can signify an equal ranking of A and B, or a pairwise tie between A and B. The meaning will be made clear by the context.

Condorcet winner, Condorcet-efficiency, Condorcet criterion: A Condorcet winner, also called a 'dominant candidate,' is a candidate that wins all of its pairwise comparisons. If a voting method always elects a Condorcet winner when one exists, the method is Condorcet-efficient, and passes the Condorcet criterion.

Strong Condorcet winner: A Condorcet winner whose pairwise victories are each supported by more than one half of the ballots.

Majority rule cycle: A circular series of pairwise defeats (e.g. A beats B, B beats C, C beats A) that leaves no single candidate unbeaten.

Condorcet completion method: A voting method that chooses the Condorcet winner when one exists, and is also decisive when there is no Condorcet winner. The following four methods (minimax, ranked pairs, river, and beatpath) are Condorcet completion methods.

Minimax method: The winner is a candidate whose strongest pairwise loss (if any) is the least strong compared to other candidates' strongest losses. Equivalent to a method that drops the weakest pairwise defeat until one candidate is undefeated.

Ranked pairs method: Defeats are considered in descending order of strength. They are locked in place unless they make a cycle with already-locked defeats, in which case they are skipped. The winner will be a candidate who is undefeated after all the defeats have been considered. See Tideman [11].

River method: Similar to ranked pairs, except that it does not lock more than one defeat against the same candidate; once the first has been locked, any others are skipped. See Heitzig [3].

Beatpath method: A beatpath is a series of pairwise defeats that form a path from one candidate to another. For example, if A beats B, and B beats C, then there is a beatpath from A to C. The strength of a beatpath is defined as the strength of its weakest component defeat. If the strongest beatpath from X to Y is stronger than the strongest beatpath from Y to X, then X has a beatpath win over Y. The winner of the beatpath method will be a candidate such that no other candidate has a beatpath win against it. See Schulze [8].

Ordinal pairwise: A shorthand term that I will use to refer to versions of the minimax, ranked pairs, river, and beatpath methods that only use ordinal rankings, and measure defeat strength in terms of a sheer number of votes, whether the number of votes in agreement with a defeat, or the margin between the number of votes in agreement and the number of votes in disagreement.

Minimal dominant set: The smallest set of candidates such that every candidate within the set has a pairwise victory over every candidate outside the set. See Schwartz [10]. The ranked pairs, river, and beatpath methods always choose from the minimal dominant set, whereas the minimax method does not.

Resolvability: A voting method is resolvable if the probability that a random solution will be needed to produce a winner approaches zero as the number of voters approaches infinity.

Mutual majority criterion: If there is a single majority of the voters who rank every candidate in a set S_1 over every candidate outside S_1, then the winner should always be a member of S_1.

3 Definition of the cardinal-weighted pairwise comparison method

3.1 Ballot

1. Voters rank the candidates. Equal rankings are allowed.
2. Voters rate the candidates, e.g. on a scale from 0 to 100. Equal ratings are allowed. If you give one candidate a higher rating than another, then you must also give the higher-rated candidate a higher ranking.

3.2 Tally

1. Determine the *direction* of the pairwise defeats by using the *rankings* for a standard pairwise comparison tally.
2. Determine the *strength* of the pairwise defeats by finding the weighted magnitude as follows. Suppose that candidate A pairwise beats candidate B, and we want to know the strength of the defeat. For each voter who ranks A over B, and *only* for voters who rank A over B, subtract their rating of B from their rating of A, to get the rating differential. The sum of these individual winning rating differentials is the total weighted magnitude of the defeat. (Note that voters who rank B over A do not contribute to the weighted magnitude of the A>B defeat; hence it is never negative.)
3. Now that the direction of the pairwise defeats have been determined (in step 1) and the strength of the defeats have been determined (in step 2), you can choose from a variety of Condorcet completion methods to determine the winner. I recommend the ranked pairs, beatpath, and river methods.

3.3 Optional, additional provisions

These additional provisions are not an essential part of the cardinal-weighted pairwise method, but they may prove helpful.

1. **Maximizing in scale provision**: [1] Once a minimal dominant set has been established by the pairwise tally in step 2, it may be a good idea to max-

imize the voters' rating differentials in scale between the candidates in the set. That is, to change the ratings on each ballot so that the highest-rated minimal dominant set candidate is at 100, the lowest-rated minimal dominant set candidate is at 0, and the rating differentials between the minimal dominant set candidates retain their original ratios. (For example, 50,20,10 would become 100,25,0.) The benefit of this provision is that voters will have equal ballot weight with regard to the resolution of the majority rule cycle in particular. Therefore, voters will not have an incentive against investing priority in preference gaps that are relatively unlikely to fall within the minimal dominant set.

2. **Blank rating option**: This allows voters to give one or more candidates a blank rating, such that if I give some candidate a blank rating, my ballot will still affect the direction of pairwise defeats concerning that candidate, but it will not add to the weighted magnitude of such defeats.

   Another possible way to deal with candidates that voters leave unrated is to determine their ratings using a default formula. For example, a candidate ranked in first place could be given a default rating of 100, a candidate ranked in last place could be given a default rating of 0, and remaining default ratings could be spaced evenly within the constraints imposed by surrounding ratings.

4 An example computation

The notation in the first line below is used to indicate that 26 voters rank the candidates in the order Right > Left$_B$ > Left$_A$, and assign the three candidates ratings of 100, 10, and 0, respectively.

4.1 Example

26: Right > Left$_B$ > Left$_A$ (100,10,0)
22: Right > Left$_A$ > Left$_B$ (100,10,0)
26: Left$_B$ > Left$_A$ > Right (100,90,0)
1: Left$_B$ > Right > Left$_A$ (100,50,0)
21: Left$_A$ > Left$_B$ > Right (100,90,0)
4: Left$_A$ > Right > Left$_B$ (100,50,0)
Direction of defeats (using ranking information):
Right > Left$_B$: 52-48
Left$_A$ > Right: 51-49
Left$_B$ > Left$_A$: 53-47

Weighted magnitude of defeats (using rating information): Right > Left$_B$:
$(26\times(100-10))+(22\times(100-0))+(4\times(50-0)) = 4740$

Left$_B$ > Left$_A$:
$(26\times(10-0))+(26\times(100-90))+(1\times(100-0)) = 620$

Left$_A$>$Right$:
$(26\times(90-0))+(21\times(100-0))+(4\times(100-50)) = 4640$

Completion by cardinal-weighted pairwise with **ranked pairs** or **river**: Consider the defeats in the order of descending weighted magnitude.
4740: Right > Left$_B$ keep
4640: Left$_A$ > Right keep
620: Left$_B$ > Left$_A$ skip (would cause a cycle, Right>Left$_B$>Left$_A$>Right)
Kept defeats produce ordering Left$_A$>Right>Left$_B$; *Left$_A$ wins*.

Completion by cardinal-weighted pairwise with **beatpath**: The strength of a beatpath is defined by the defeat along that path with the lowest weighted magnitude.
beatpath Right → Left$_B$: 4740
beatpath Left$_B$ → Right: 620
beatpath Left$_A$ → Right: 4640
beatpath Right → Left$_A$: 620
beatpath Left$_A$ → Left$_B$: 4640
beatpath Left$_B$ → Left$_A$: 620
Complete ordering is Left$_A$>Right>Left$_B$; *Left$_A$ wins*.

5 Why majoritarian election methods should be Condorcet-efficient

The Condorcet criterion (along with the minimal dominant set, which is a generalization of the same principle) seems to be the most authentic definition of majority rule that is available to us. If there is one candidate who is preferred by some majority over every other candidate individually, it seems inappropriate to call anyone else a majority winner. For example, if candidate A is a Condorcet winner, and a non-Condorcet-efficient method elects candidate B, a majority will prefer A to B. If there was an election just between these two candidates, A should be expected to win that election.

Condorcet efficiency has important practical benefits. First, Condorcet-efficient methods tend toward the political center, which should promote compromise rather than polarization. Second, when a strong Condorcet winner exists with respect to voters' sincere preferences, and another method chooses someone else, the result is unstable in that a majority could have achieved a mutually preferable result if some of them had voted differently.

Condorcet-efficient methods minimize the incentive for the **compromising strategy**, which is insincerely ranking an option higher in order to decrease the probability that a less-preferred option will win. For example, if my sincere preferences are R>S>T, a compromising strategy would be to vote S>R>T or R=S>T, raising S's ranking in order to decrease T's chances of winning. (The drawback is that this often decreases R's chances of winning as well.) All resolvable voting methods that satisfy the mutual majority criterion have a compromising incentive when there is a majority rule cycle. But unlike other methods, such as single-winner STV, voters in Condorcet-efficient methods never have an incentive to use the compromising strategy when there is a Condorcet winner [9]. This is an important property because, in the absence of a majority rule cycle, it allows me to vote my R>S preference without worrying that it will undermine my S>T preference. This is a more complete way of curtailing the "lesser of two evils" problem, that is, decreasing the extent to which voters have to worry about earlier choices drawing support away from later choices. Thus, Condorcet-efficient methods allow more candidates to participate on an equal basis, which should lead in turn to substantially higher levels of responsiveness and accountability.

6 Preference priority and defeat strength

Most Condorcet-efficient methods that have been proposed so far limit voter input to ordinal rankings. Hence, voters can express preferences between candidates, but they cannot express the relative priority of their preferences. If I worship my first three choices, but detest my fourth and fifth choices, I cannot express this on my ballot, and it is not taken into account when the winner is decided.

Ordinal pairwise methods measure defeat strength in terms of a sheer number of ballots. The cardinal pairwise method extends the sensitivity of the process by factoring in a measure of how much priority the voters assign to each ranking. The goal is that the weakest defeat in a majority rule cycle should be the one that has the lowest overall combination of these two factors: 1) the number of voters in agreement with the defeat; 2) the relative priority of the defeat to those voters who agree with it.

It seems almost axiomatic that, when faced with a majority rule cycle, one should drop the defeat(s) in the cycle that are of least importance to the voters. The remaining question is how to define the priority of each defeat to each voter, and how to aggregate these individual priorities. The answer that cardinal pairwise gives to this question is relatively simple. For those who agree with a defeat, we look at the rating differential they express between the two candidates being compared. Then we take the sum of these winning rating differentials to find the overall strength of the defeat.

The idea is that the voters will rate the candidates such that the rating differential between each pair of candidates will reflect the relative priority of their preference between those candidates. The fact that each voter is constrained to the same range of ratings (e.g. 0 to 100) assures that everyone has essentially the same voting "power." The point here is not to do interpersonal comparison of utilities, but rather to allow voters to prioritize their own preferences relative to one another, using a fluid and simple high-resolution scale.

When learning the cardinal pairwise method, one may wonder why it only looks at the rating differentials of those who agree with a particular defeat, rather than subtracting the losing rating differentials from the winning rating differentials. To begin with, I will say that I am more interested in dropping the defeats that are of least importance to the voters overall, rather than the defeats that are the closest in terms of the strength of preference on either side. That is, if there is one pairwise comparison that voters on both sides consider to be a very high priority, I think that it is especially important not to reverse this defeat. Such high-priority defeats should be regarded as crucial within the election, and the cardinal aspect of the method should be used to defend them rather than to undermine them.

In this way, looking at only the winning rating differentials greatly improves the *stability* of the cardinal pairwise method. Because the defeats that voters place the highest priority on are the most difficult to reverse, the cardinal pairwise method is unusually resistant to strategic manipulation. This point will be explored in greater detail in the next section.

7 Strategic manipulation

Although Condorcet-efficient methods minimize the incentive for use of the compromising strategy, they are vulnerable to the **burying strategy**. This strategy entails insincerely ranking an option lower in order to increase the probability that a more-preferred option will win. For example, if my sincere preferences are R>S>T, a burying strategy would be to vote R>T>S or R>S=T, lowering S's ranking in order to increase R's chances of winning. (The drawback is that this often increases T's chances of winning as well.)

Imagine that with respect to voters' sincere preferences in a three-candidate election, A pairwise beats B and C, while B pairwise beats C. A is a sincere Condorcet winner, but it is often possible for supporters of candidate B to gain an advantage by burying A under C, that is, by voting B>C>A instead of B>A>C. This can create an insincere C>A defeat, which can cause a majority rule cycle such that the A>B defeat is the weakest of the three, so that B wins. In this way, it is often possible to overrule a genuine defeat with a fake defeat.

The burying strategy may have the potential to cause substantial trouble in elections that use a Condorcet-efficient method. Some have cited this as a reason not to adopt Condorcet-efficient methods. (Monroe [5]; Richie and Bouricus [6]) Unfortunately, Condorcet-efficient methods cannot be completely invulnerable to the burying strategy, which follows from the fact that Condorcet-efficiency is incompatible with the later-no-help criterion [12]. However, cardinal pairwise may be able to make this vulnerability much less severe.

There are many reasons to think that cardinal pairwise will be more resistant to strategy than most other Condorcet-efficient methods. First, it should tend to prevent the most flagrant strategic incursions. Second, it should tend to balance strategic incentive against strategic ability, so that those who are most interested in changing the result via strategic incursion tend to be those who are least able to do so. Third, it should minimize strategic barriers against the entry of new candidates. Fourth, it should create the possibility of more-stable counterstrategies than those that are available in ordinal pairwise.

7.1 Flagrant strategic incursions

I define a flagrant strategic incursion as one that causes a very high-priority defeat to be overruled by a false defeat. Take example 7.1 below. Sincere votes:
46: A>B>C (100,10,0)
44: B>A>C (100,10,0)
5: C>A>B (100,50,0)
5: C>B>A (100,50,0)

A is a Condorcet winner. Clearly, the primary contest is between A and B, as C is the last choice of 90% of the voters. However, using ordinal pairwise, the B>A>C voters can change the winner to B by voting B>C>A. This is a very flagrant incursion.

In cardinal pairwise, however, this particular type of flagrant incursion does not work. The weighted magnitude of the C>A defeat is 4490, and no defeat with a magnitude greater than $3333^{1}/_{3}$ can be dropped as a result of a three candidate cycle (assuming 100 voters and a 0-100 rating scale).

With larger cycles (four candidates and above, e.g. A>B>C>D>A), the $3333^{1}/_{3}$ limit does not apply, but overruling a high-magnitude defeat is still very difficult. Let's say that there is a candidate B, who is pairwise-beaten by a candidate A. In order for B to win, there must be a chain of defeats from B to A (e.g. B>C>D>A), such that every defeat along that chain has a weighted magnitude that is at least equal to the A>B defeat. The minority who prefer B to A will have a limited amount of weight to distribute along the B>C>D>A chain. A given point of weight can count towards two defeats in this four-candidate chain (e.g. the one-point gap in the vote B>D>C>A (1,1,0,0) counts towards the B>C and D>A defeats), but it cannot count towards more than two.

Cardinal pairwise, unlike ordinal pairwise, does not allow a voter to apply the maximum weight to all of their pairwise preferences. This scarcity of weight produces excellent anti-strategic effects, by placing a limit on the extent to which a strategizing group of voters can build up the weight of multiple pairwise defeats at the same time in order to manipulate the overall result.

In general, flagrant incursions are much less likely to be successful in cardinal pairwise than in ordinal pairwise, because the difficulty of overruling an A>B defeat increases as more voters assign a higher priority to the A>B defeat. I hope that my definition of a flagrant incursion can be seen to have value, and that it can be agreed upon that relatively high-priority defeats should be harder to overrule. Consider that when a defeat of A over B is given a very high priority, we can generally expect B to be very *different* from A (in the eyes of the voters), relative to differences with the other candidates in the election. In order to quantify this difference, we can look at both the average A>B rating differential and the average B>A rating differential for individual voters.

I think it is crucial that we make it as difficult as possible for strategic voters to alter an election result in such a way that the actual winner is considered by the voters to be extremely different from all of the members of the sincere minimal dominant set. Consider how seriously it would undermine the legitimacy of the voting system, if it was found that partisan supporters had pulled off a successful burying strategy which won the election for a candidate who was the ideological polar opposite of the sincere Condorcet winner. Ordinal pairwise unfortunately cannot offer much protection against this disturbing possibility, but cardinal pairwise can.

7.2 Strategic incentive and strategic ability

There are impossibility theorems that show that strategic manipulation cannot be completely avoided in any reasonable election method (Gibbard [2]; Satterthwaite [7]; Hylland [4]), but I'm not aware of a theorem that says that we can't find a method that distributes strategic ability in roughly inverse proportion to strategic incentive.

Let's assume that the intensity of difference that a voter perceives between two candidates tends to be largely independent of their ranking of those candidates, and that the average rating differentials on either side of a defeat will tend to be strongly correlated with one another.

Let's say that there is a candidate A who pairwise beats candidate B. If the incentive for the B>A voters to help B by burying A is particularly strong–that is, if they assign a very high priority to their B>A ranking–then we can expect the A>B voters to assign a high priority to their A>B ranking as well, which will make the A>B defeat very hard to overrule. So, a group of voters' ability to achieve a successful burying strategy generally tends to be smaller in cases where that group has a larger incentive to engage in that strategy.

Conversely, if A and B are considered to be more similar candidates, such that there are low average rating differentials on both sides of the defeat, then it may be more feasible for the B>A voters to help B by burying A, but they would have less to gain by doing so, and more to lose should the strategy backfire.

7.3 Minimizing strategic barriers to candidate entry

In example 4.1 above, Left$_B$ and Left$_A$ can be considered to be relatively similar candidates, in that there is a low average rating differential placed on the comparison between them, going in both directions. If only Left$_A$ and Right were candidates, Left$_A$ would probably win, since he has a pairwise win over Right. In cardinal pairwise, the entry of Left$_B$ does not change this result. However, the winner changes to Right in ordinal pairwise, which defines Right's 49-51 pairwise loss as the weakest in the cycle. In general, it is much harder in cardinal pairwise for the entry of a new, non-winning candidate to do harm to a similar candidate. The reason for this is that if the new candidate beats the similar candidate, but does not win, this defeat will be relatively weak, and hence likely to be overruled in the event of a cycle.

In ordinal pairwise, a voter who would otherwise support a potentially-entering candidate might have some anxiety that this candidate could hurt a similar candidate whom that voter also supports. Because the potentially-entering candidate's support base may feel ambivalent about his presence in the race, entry of the candidate may not occur. Thus, the method retains a certain strategic barrier to entry of new candidates. Cardinal pairwise minimizes this barrier to entry, in that the entry of a new candidate is extremely unlikely to affect the result in opposition to the will of his would-be supporters.

7.4 Stable counterstrategies

If several voters try to coordinate a strategic incursion, and other voters learn about this and consider it to be undesirable, they may attempt to coordinate a counterstrategy, in order to make the initial strategy unsuccessful. One hopes that counterstrategy will rarely or never be needed, but it is nevertheless to the credit of cardinal pairwise that it provides for somewhat more-stable counterstrategies than ordinal pairwise. Actually, this may be important in preventing strategic incursion from achieving a critical mass in the first place.

Example 7.2: Some votes are strategically altered
28: A>B>C (100,60,0)
23: C>A>B (100,40,0)
17: B>A>C (100,60,0)
22: C>B>A (100,40,0)

10: B>C>A (100,100,0) these 10 votes are strategically altered from a sincere ordering of B>A>C

Pairwise comparisons, followed by weighted magnitudes:
A > B: 51-49   C > A: 55-45   B > C: 55-45
A > B: 2040   C > A: 4580   B > C: 3380
Candidate A was a sincere Condorcet winner, but B wins instead using both ordinal and cardinal pairwise, as a result of the B>A>C voters' burying strategy.

There are two basic counterstrategy replies to the burying strategy: the compromising counterstrategy, and the deterrent/burying counterstrategy.

In ordinal pairwise, the **compromising counterstrategy** would entail the C>A>B voters weakening or reversing the defeat against A by voting C=A>B. In cardinal pairwise, a similar effect could be gained by voting C>A>B (100,100,0). Both counterstrategies can return the victory to candidate A. The cardinal pairwise counterstrategy is more stable than the ordinal pairwise counterstrategy, in that it does not risk a change in the winner of the A-C pairwise comparison. This makes it a less perilous choice for the C>A>B voters.

The **deterrent/burying counterstrategy** would entail the A>B>C voters weakening or reversing B's defeat of C, such that the B>A>C voters' burying of A could only backfire by electing C. In ordinal pairwise, this would require some A>B>C voters to equalize or reverse their B>C preference, thus voting A>B=C or A>C>B. In cardinal pairwise, it is possible for the A>B>C voters to get a similar deterrent effect by voting A>B>C (100,0,0).

With the deterrent/burying counterstrategy in general, the counterstrategizers are unlikely to know for sure whether the original strategizers will carry out their incursion or not, until the votes have already been cast. Therefore it is important to have an effective counterstrategy that they can use without severely destabilizing the result, in case the original strategy is not carried out and the counterstrategy punishment is undeserved. In this respect, the cardinal pairwise version of the counterstrategy is preferable, in that it does not alter the direction of any pairwise defeats, and therefore will not interfere with the identification of a Condorcet winner.

Of course, the existence of more-stable counterstrategies in cardinal pairwise does not mean that strategy will never be a problem. However, it suggests to me that the threat of a strategic incursion, should it arise, is less likely to spiral out of control.

8 Conclusion

I believe that voting methods aiming for majority rule should be Condorcet-efficient, and that Condorcet-efficient methods should be improved in two ways. One, they should take the relative priority of voters' pairwise preferences into account; two, they should be more resistant to the burying strategy. I find it serendipitous that the same principle can achieve both benefits simultaneously.

I find both of these potential improvements quite significant, but perhaps the strategic issue is the more pressing of the two, as I suspect that the burying strategy could prove to be a serious problem for Condorcet-efficient methods in contentious elections. It is important to have a method that, in addition to recognizing a Condorcet winner when one is clearly expressed, works to protect sincere Condorcet winners from being obscured by strategic incursion. I believe that cardinal-weighted pairwise accomplishes this to an unusual degree.

So, I do not intend cardinal-weighted pairwise as a frivolous academic exercise or a mathematical curiosity. I intend it as a realistic proposal, and one that I sincerely prefer over other existing proposals. I recognize that it adds an extra layer of complexity, but I feel that the benefits of more-meaningful cyclic resolution and reduced strategic vulnerability far outweigh the cost.

9 Acknowledgments

I thank Nicolaus Tideman, Jobst Heitzig, Chris Benham, and Markus Schulze for their helpful comments on earlier versions of this paper. I am also grateful to many participants of the election methods discussion list for their insight and support.

10 References

[1] Benham, C. (2004) Personal correspondence. June 8.

[2] Gibbard, A. (1973) Manipulation of voting schemes: A General Result. *Econometrica* 41: 587-601.

[3] Heitzig, J. (2004) River method – updated summary. Discussion list entry. October 6. `http://lists.electorama.com/pipermail/election-methods-electorama.com/2004-October/013971.html`

[4] Hylland, A. (1980) Strategy proofness of voting procedures with lotteries as outcomes and infinite sets of strategies. University of Oslo.

[5] Monroe, B. (unpublished) Preferences, votes, and representation: rethinking social choice theory and democratic institutions. Unpublished manuscript. Excerpted at `http://accuratedemocracy.com/archive/condorcet/Monroe/004004MonroeBurt.pdf`

[6] Richie, R. and Bouricus, T. (2004) Draft of CVD analysis about IRV vs. Condorcet voting. `http://groups.yahoo.com/group/instantrunoff/message/1548`

[7] Satterthwaite, M. (1975) Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory* 10: 187-217.

[8] Schulze, M. (2003) A new monotonic and clone-independent single-winner election method. *Voting matters* 17: page 9-19

[9] Schulze, M. (2004) Compromising strategy in Condorcet. Two discussion list entries. September 19. `http://lists.electorama.com/pipermail/election-methods-electorama.com/2004-September/013908.html` `http://lists.electorama.com/pipermail/election-methods-electorama.com/2004-September/013906.html`

[10] Schwartz, T. (1986) The Logic of Collective Choice. New York: Columbia University Press. Page 141.

[11] Tideman, T.N. (1987) Independence of clones as a criterion for voting rules. *Social Choice and Welfare* 4: 185-206.

[12] Woodall, D. (1997) Monotonicity of single-seat preferential election rules. *Discrete Applied Mathematics* 77: 81-98.

A Working Paper on Full Disclosure

B. A. Wichmann
Brian.Wichmann@bcs.org.uk

1 Introduction

This document considers the following problem: given an election in which preferential voting is used and the count is conducted by computer, what information should be disclosed? Running an election consists of several stages, but here we are concerned with the counting process only. This process must not only be trustworthy, but needs to be seen as such by the electorate.

With the manual count, the full result is typically declared by a *result sheet* which contains the great majority of the information gathered during the counting process[1]. If a witnessed count is undertaken, which is, of course, the case with public elections, then all the critical information that would be available to the witnesses appears in the result sheet. The same degree of transparency is needed when a computer count is undertaken.

In the USA, under their Freedom of Information Act, full information of the ballot preferences is available for public elections. Of course, although this information is available, the identity of those who voted in a specific way is not available — ballot secrecy is maintained.

In the case of the experimental use of computers in the Irish Dáil elections in 2002, full information was available for the three constituencies polled by voting machines. It appears that the Republic has a similar Freedom of Information Act to the USA.

There are at least three different types of election in which the full disclosure questions arise: public elections; private elections performed by an independent party; and lastly, private elections performed internally. All three types of election occur with the Single Transferable Vote (and computer counting).

[1] Practices vary in this area. Working calculations should be published but may not be. For some elections, the ballot boxes are opened individually allowing a careful witness some information about the relative strengths of the candidate vote.

2 Data Protection Legislation

Public elections are typically covered by national laws, but private elections would also need to adhere to appropriate national laws. For the EU, this is largely the national laws which enact the European Directive on Data Protection. This gives data subjects the right to information held about them, and for those holding information the need to register and control access to the information.

There are two cases to consider here: those relating to the candidates in an election and those relating to a voter. Assuming that the voter is not specifically identified, then, in effect, no information is held and therefore nothing needs to be disclosed.

For the candidate, it is clear that information is held and therefore the candidate has a right to be told the information held. For a preferential voting system, it has been my opinion (based upon the 1984 Act, which was straightforward to follow), that the candidate should be informed as to how many preferences were recorded against him/her at all the various levels. Of course, the number of first preferences would be available from the result sheet, but the other preferences may not be. Hence, with a computer count, there seems little doubt that more information should be available to candidates than is provided in the result sheet.

The situation is rather more confused when one considers disclosure of more than the above. It is clear that ballot secrecy is paramount and therefore disclosure may be limited by that need. The limitation is surely minimal since ballot secrecy has not been called into question in the USA, where full disclosure takes place.

We consider secrecy in the next section and hence for the moment, we note current practice.

For the 2002 Irish Dáil elections, full disclosure took place. Some reservations have been expressed about this in a recent Irish report [6]. Also, in the context of public elections, Otten has pointed out a means of making bribery effective by the use of an unlikely sequence of preferences [1]. It seems that this problem has not been raised in the USA.

In the case of an independent balloting organisation undertaking a count, it is not immediately clear who 'owns' the ballot data. If it is the balloting organisation, then disclosure rests with them, otherwise it rests with their client.

Currently, Electoral Reform Services maintain that full disclosure is not possible even when the client requests it. I cannot understand this position and I am not alone in this.

3 Secrecy

Less that 150 years ago it was argued by some that secret voting was not desirable, but nowadays everyone seems to accept that secrecy is paramount. Given that, then the question arises as to whether this imposes some restrictions in applying the principle of full disclosure.

Secrecy has an important limitation. If the entire electoral process is clothed in secrecy, then the validity of the result will be open to question. Hence public elections are open to substantial external scrutiny. In our context, we are concerned with elections in which the count is undertaken by computer. It is far from clear how the process of validating a *count* should be undertaken under such circumstances. Again, we are assuming that the other parts of the electoral system perform the intended function in a manner acceptable to the electorate. The integrity of the count was part of the concern in the report on the Irish system [6].

One means to overcome part of this problem is full disclosure. Then anybody can use the data to repeat the count in order to confirm the result. (Counting software is needed, but that is readily available for almost all counting rules.) This is a *stronger* validation method than the traditional method of a witnessed manual count. When an STV manual count has been checked afterwards by using a computer, some errors are almost always found — sometimes even affecting the result!

Is ballot secrecy compatible with full disclosure? There are two possible problems: firstly, elections with a small number of votes, and secondly, the problem of a long preference list which can act as a signature for the voter.

3.1 An example — census data

It seems to me that there is a good analogy between the problem here and that in handling census data. Complete disclosure occurs after 100 years. People can also request their own data. However, substantial statistical information is made available without restriction — a clear need for Government planning. The apparent conflict is overcome by grouping information into sufficiently large numbers so that an individual return cannot be identified.

It is my understanding that the protocol that the Office of National Statistics uses was agreed with the Royal Statistical Society.

It is my contention that a similar protocol needs to be agreed for preferential election data.

4 Technical measures to ensure secrecy

It seems that there is no concern about the information available from a result sheet. I have been informed of an example in which the result sheet could be regarded as problematic. This was for the 1999 North Tipperary local election in which a candidate got no first preference votes. One could envisage a situation in which such a candidate was then hostile to his/her friends, family, employees, etc.

The preferences themselves can be revealed. Let us say one is voting in an election in which your preferences are A, B, C and finally D. It is not possible to exclude the possibility that the existence of such a voting pattern will be evident from the result sheet. For an actual example in which a long preference list was evident, see [2], which was evident due to full disclosure.

In practice, the percentage of preferences actually used in an election is quite small, so it is usual for long preference lists to consist mainly of unused preferences (see [4]). It is therefore possible to provide a form of disclosure in which some of the preferences on the ballot papers are omitted or changed, but still provide data which confirms the result of the count. In other words, there is plenty of room to provide a form of disclosure which allows for count validation but nevertheless ensures ballot secrecy.

The statistical analogy to the census data problem would perhaps be to disclose a fraction of the ballot papers. This is not a good method, since the data would then not provide a means of validating the count. I have written a program myself to make a number of changes to ballot data so that both the election and the candidates could not be identified. Unfortunately, such changes make it impossible to perform some reasonable forms of analysis, like determining if there is an alphabetic bias in the voting data.

It is certainly true that if ballot data is provided only for some forms of statistical research that a sampling method could be effective. Such a form of disclosure would be of use, but only to a very limited audience.

I am unclear how small any election should be before full disclosure could not reasonably be undertaken. If full disclosure is not provided, then the issue of count validation remains.

Finally, it should be noted that once any public form of disclosure takes place, the use to which it is put is uncontrolled. Here, we are not concerned with making information available under some form of non-disclosure agreement that might restrict its use for research purposes.

5 Conclusions

From the above, I make the following conclusions:

1. In the interests of openness and the validation of computer counting, full disclosure should be the default.
2. Legal advice should be obtained on any caveats to full disclosure as a result of the Data Protection Directive.
3. Technical measures should be agreed on how full disclosure should be implemented, given the paramount importance of ballot secrecy.
4. Purists may well object to anything other than making the ballot data available without change, but disclosure which is sufficient for count validation is surely required.

6 Postscript

Drafts of this paper have been sent to several people who I know are interested in this subject. I have tried to reflect the views of those who commented on the drafts, but this has not always been possible. Those who provided comments include: James Gilmour, Steve Todd, Joe Otten, Colin Rosenstiel, Anthony Tuffin, Jeffery O'Neill and David Hill.

David Hill was strongly of the view that no change should be made to the ballot preferences. I would prefer that, but think that it is better to have effective disclosure (in which small changes are made), rather than no disclosure which is the position with the majority of STV computer counts at the moment.

7 References

[1] Otten J. Fuller disclosure than intended. *Voting matters*, Issue 17, p8. 2003.

[2] Hill I. D. An odd feature in a real STV election. *Voting matters*, Issue 18, p9. 2004.

[3] Hill I. D. Full disclosure of data. *Voting matters*, Issue 18, p10. 2004.

[4] Wichmann, B. A. A note on the use of preferences. *Voting matters*, Issue 18, p9. 2004.

[5] Hill I. D. What would a different method have done? *Voting matters*. Issue 16. p5. 2003.

[6] Interim Report of the Commission on Electronic Voting on the Secrecy, Accuracy and Testing of the chosen Electronic Voting System. http://www.cev.ie/htm/report/download_report.htm

ISSN 1745-6231

Voting matters

for the technical issues of STV

published by

The McDougall Trust

Issue 20 June 2005

About the McDougall Trust (reg. charity no. 212151)

The McDougall Trust is a charitable trust formed in 1948. The charity's purposes as stated in its governing scheme of 1959 are to advance knowledge of and encourage the study of and research into:

- political or economic science and functions of government and the services provided to the community by public and voluntary organisations; and
- methods of election of and the selection and government of representative organisations whether national, civic, commercial, industrial or social.

The Trust's work includes the maintenance and development of the Lakeman Library for Electoral Studies, a unique research resource, the production and publication of Representation: The Journal of Representative Democracy, and, of course, this publication **Voting matters**, that examines the technical issues of the single transferable vote and related electoral systems.
For further information on the Trust, please contact:

The Secretary,
McDougall Trust,
6 Chancel Street,
London SE1 0UX, UK.
Telephone: +44 (0)20 7620 1080
Facsimile: +44 (0)20 7928 1528
Email: admin@mcdougall.org.uk
Web: www.mcdougall.org.uk

For further information on this publication, please contact B A Wichmann, the Editor at the above address or by email at: `Brian.Wichmann@bcs.org.uk`

Editorial

There are 4 papers in this issue:

- I. D. Hill and C. H. E. Warren: Meek *versus* Warren.

  This article compares two computer-based STV counting algorithms. Although the Meek version seems to be the only version which is widely used, readers of *Voting matters* should surely appreciate the differences and draw their own conclusions.

- I. D. Hill and Simon Gazeley: Sequential STV — a further modification.

  This paper considers a variant of STV in which later preferences are used to exclude candidates. The modification described here has proved necessary due to two issues which are described in the paper.

- Earl Kitchener: A new way to break STV ties in a special case.

  This short paper considers one special case in which the proposal is surely non-controversial. This is followed by summary and moderated debate on breaking ties produced by the editor with assistance from those listed.

- P Kestelman. Apportionment and Proportionality: A Measured View.

  The author's abstract reads: Apportionment (allocating seats to multi-member constituencies equitably) can illuminate proportionality (allocating seats to parties fairly) and its quantification. Sainte-Laguë (Webster) is the fairest method of apportionment — and electoral principle. Several disproportionality measures have been proposed: among which the Loosemore-Hanby Index straightforwardly measures Party total over-representation. UK general elections (First-Past-the-Post) have clearly proved non-PR; and even nominally PR elections of British MEPs and Regional Assemblies have yielded only semi-PR ('broad PR'). Allowing for vote transferability, multimember STV in Ireland has mediated full PR (despite low District Magnitude); while Alternative Voting in Australia has arguably proved semi-PR.

The New Zealand STV elections

A Parliamentary investigation (Justice and Electoral Committee) is under way into the delays in producing the results. It has not yet reported.

Steve Todd reported in the last issue that the ballot data should be available. In fact, the electoral officers were divided on the provision of this data so that complete data is only available for 15 of the 79 elections. (There were 81 STV elections, but two were not contested.) A table giving the availability of the data is available on `http://stv.sourceforge.net/`.

The British Columbia Referendum for STV

The Referendum produced a majority for STV, but not the 60% to ensure that the necessary legislation will be passed. It is unclear at this stage what will happen.

Meek *versus* Warren

I. D. Hill and C. H. E. Warren
No email available.

1 Where we agree (I.D. Hill and C.H.E. Warren)

We admire traditional STV methods (Newland and Britton rules [1] and other similar methods) as being a good approximation to what STV is trying to achieve, while being easy enough to do by hand within a reasonable length of time, but in this electronic age, we ought to do better than that. Of course we accept that the ability to count by hand is an advantage; but does such an advantage justify the consequence that, quite often, the set of candidates who best meet the voters' wishes are not elected? We think not. But if we seek to campaign for something better, we need to agree on the better thing that we should support.

We agree that fairness is of prime concern in a voting system, but it is a tricky concept — one only has to listen to politicians all claiming that taxation, for example, must be fair ("and must be seen to be fair" as if that addition helped), while totally disagreeing with each other about what is fair and what is not.

The Meek method [2] and the Warren method [3] are very similar to each other but, in deciding how much of each vote is retained by an elected candidate and how much is passed on to the next choice, the Meek method uses multiplicative 'keep values' but the Warren method uses additive 'portions apportioned'. We here denote the Meek keep value and the Warren portion apportioned for candidate C as c_m and c_w respectively. These quantities have a value between 0 and 1, and they are calculated so that, if a candidate has a surplus, their use reduces the vote for that candidate to just the quota. The calculation of these quantities so that they meet this requirement is a mathematical problem, usually requiring a computer. All that we need to know in this paper is that they can be calculated.

With the Meek method c_m is defined as the proportion of the vote that is passed to candidate C which candidate C retains, so that $(1 - c_m)$ is the proportion of that vote that is passed on. In the case of a ballot that reads ABC...

the portion of vote which A retains is a_m
the portion of vote which A passes on to B is $(1 - a_m)$
the portion of vote which B retains is $(1 - a_m)b_m$
the portion of vote which B passes on to C is $(1 - a_m)(1 - b_m)$
the portion of vote which C retains is $(1 - a_m)(1 - b_m)c_m$
the portion of vote which C passes on is $(1 - a_m)(1 - b_m)(1 - c_m)$
and so on.

From the above statements we see why the Meek keep values are called multiplicative.

With the Warren method c_w is defined as the portion of a vote that is apportioned to candidate C if such apportionment is possible. In the case of a ballot that reads ABC...

the portion of vote which is apportioned to A is a_w
if $a_w + b_w > 1$, the portion of vote which is apportioned to B is $(1 - a_w)$ and nothing is apportioned to C and beyond
if $a_w + b_w \leq 1$, the portion of vote which is apportioned to B is b_w
if $a_w + b_w \leq 1$ and $a_w + b_w + c_w > 1$, the portion of vote which is apportioned to C is $(1 - a_w - b_w)$ and nothing is apportioned beyond
if $a_w + b_w + c_w \leq 1$, the portion of vote which is apportioned to C is c_w
and so on.

From the above statements we see why the Warren portions apportioned are called additive.

Although a Meek keep value c_m may, in some circumstances, turn out to have the same value as a Warren portion apportioned c_w, in general their numerical values are different.

The methods are equally easy to program for a computer and, for real voting patterns as distinct from test cases, they nearly always produce the same answers, not in numerical terms but in terms of which candidates are elected and which are not. In those circumstances, we agree that it does not matter too much which is used, so it is preferable to support the one that is better in principle — but which one is that?

We recognise that impossibility theorems, such as Woodall's theorem [4], show that to seek an absolute ideal is a 'wild-goose chase'. It follows that it will always be possible to produce particular examples that tell against any given method. Unlike proving a proposition in pure mathematics, where one counter-example is enough to demonstrate that we have failed, here we always need to look at examples in a comparative sense, not an absolute sense, deciding which faults to allow for the sake of avoiding others.

2 Why I prefer the Meek method (I.D. Hill)

To my mind the essence of STV is this — if we have a quota of 7, and 12 identical votes putting A as first preference and B as second (with no others for A) then 7 votes must be held for A as a quota while the other 5 are passed to B and, from that point on, behave exactly as if they had originally been 5 votes for B as first preference. The fact that those voters had A as first preference, and A has been elected, has been fully allowed for in holding 7 votes back and the other 5 votes are now simply B votes.

In practice, we never get such identical votes, so the only fair way of doing things is, instead of holding 7 complete votes back and passing on 5 complete votes, to hold back $\frac{7}{12}$ of each vote and pass on $\frac{5}{12}$ of each vote, but the principle, that the 12 votes each of value $\frac{5}{12}$ should together have the same power as 5 complete votes, remains the same. This principle is fulfilled by the Meek method, but not by the Warren method. Because perfection is impossible, it could be that some advantage could be shown by the Warren method that would outweigh this disadvantage, but I am not aware that any advantage has been claimed for it that is strong enough to do so.

If, at the next stage, we have 5 votes with B as first preference, plus our 12 votes each now of value $\frac{5}{12}$, we have 10 votes altogether pointing at B. Only 7 are needed for a quota so $\frac{7}{10}$ needs to be retained allowing $\frac{3}{10}$ to be passed on, so the 5 votes are passed on with a value of $\frac{3}{10}$, giving them a total power of $1\frac{1}{2}$ votes. If the 12 votes are passed on with a value of $\frac{5}{12}$ times $\frac{3}{10}$, that gives them a total power of $1\frac{1}{2}$ votes too, showing that 12 each of value $\frac{5}{12}$ are being treated just like 5. To get that effect necessarily requires a multiplicative rule, not an additive rule.

To look at it from a slightly different angle, the rule should be that the proportions of the total vote for a candidate that come from different sources, and are used in deciding that the candidate can now be elected, should be maintained in the amounts of vote retained and transferred. Thus, in the same example, the votes from the AB voters and from the B voters that are used to decide to elect B are in proportion 1 to 1, whether the Meek or the Warren method is used. With Meek, the votes retained from the two groups are $3\frac{1}{2}$ and $3\frac{1}{2}$, also 1 to 1, and those transferred are $1\frac{1}{2}$ and $1\frac{1}{2}$, also 1 to 1. With Warren, the votes retained are $4\frac{16}{17}$ and $2\frac{1}{17}$, or 2.4 to 1, and those transferred are $\frac{1}{17}$ and $2\frac{16}{17}$, or 1 to 50, devoid of all the proportionality that I believe they should have.

The Meek method is able to promise voters that once their first n choices have all had their fates settled, either as excluded or as elected with a surplus, a fair share of their vote will be passed to their $(n + 1)$th choice, unless no more transfers are possible because all seats are now filled. How much is a fair share may, perhaps, be arguable (though I do not personally see it as such) but it cannot possibly be zero, which the Warren method often makes it.

Thus the basis of STV in Meek mode is that everything has to be done in proportion to the relevant numbers at the time. This means that if we have 1 ballot paper of value 1 pointing at XY, and n ballot papers each of value $\frac{1}{n}$ pointing at XZ, and X's papers are to be redistributed, then what happens to Y and to Z from those papers should be identical.

Suppose 8 candidates for 7 seats, counted by Newland and Britton rules. If there are 40 votes reading 5 ABCG, 5 ABCH, 5 ABDG, 5 ABDH, 5 ABEG, 5 ABEH, 5 ABFG, 5 ABFH, it is evident from the symmetry that ABCDEF must be elected but the final seat is a tie between G and H. If, however, there is a 41st vote reading BH, that ought to settle it in favour of H, but those rules declare it still to be a tie between G and H to be settled at random. Either Meek or Warren counting would have awarded the seat to H.

However, suppose the 41st vote, instead of being just BH reads BCDEFH. Again Newland and Britton rules fail to discover that the symmetry has been broken, and incorrectly call it a GH tie. But now so do Warren rules. With Meek rules, only 0.012 of the vote gets through as far as H, but that is enough to tilt the balance to get the right result.

In the past, when Hugh Warren and I have argued about this, each of us has, from time to time, put forward an example with an 'obviously right' answer which the other one's preferred method failed to find. However, with those examples, the other one of us never accepted that the answer in question was 'obviously right'. It was therefore necessary to produce something where the answer could not be denied. I claim to have done this with the example: 4 candidates for 3 seats, and just 3 votes: 1 ABC, 1 BC, 1 BD. Without even knowing anything about STV, it must be clear that ABC is a better answer than ABD. Meek does elect ABC, but Warren says that C and D tie for the third seat and a random choice must be made between them. Unless something equally convincing can be found that points the other way, that seems to me to be conclusive.

So far as I am aware, the only actual advantage claimed for Warren over Meek is that it is supposed to give consistency when some voters change the order of two candidates both of whom are elected anyway. This seems to me to be only a very slight advantage, and Warren rules do not always succeed even in that. With 5 candidates for 4 seats and votes 9 ABCD, 8 BD, 8 CE, 7 D, 7 E, either Meek or Warren elect ABCD. But if the ABCD votes had been ACBD instead, either Meek or Warren would elect ABCE.

The difference arises from the fact that one quota of votes is necessarily ineffective and changing the order of some preferences can change which votes those are and thus, in marginal cases, affect the result. I suggest that in practice any such inconsistency would never be noticed and is of very minor importance compared with making the count so that everything is kept in proportion to the numbers concerned.

I am less convinced than I was even that such behaviour can be called an anomaly. If two candidates are both elected anyway, it would seem at first sight that, if some voters change the order of those two, it ought not to affect who else gets elected, but is that really a good rule? In this example, there is some connection between B and D, and between C and E. We do not know what the connection is, but it is clearly there since every voter putting B first puts D second, while every voter putting C first puts E second. The second choice of the A supporters is then saying what they think about the feature that gives the connection. In such circumstances, it does not seem unreasonable that if the A voters prefer B to C that helps D, but if they prefer C to B that helps E, particularly when the first preferences for D and E are tied.

Overall, while accepting that the Warren method works quite well, it does not seem to me to have any real advantage over the Meek method, and its failure to meet what I regard as basic requirements can sometimes lead to a result that I would think unfortunate. Given how wrong it seems, I am surprised that it works as well as it does.

3 Why I prefer the Warren method (C.H.E. Warren)

I prefer the Warren method because I consider it to be based on a better principle.

The main principle behind the Warren method (given as the second principle in [3]) can be stated as: if a voter votes for candidates A, B, C in that order, and if candidates A and B each have a surplus of votes above the quota, then, on principle, no portion of the vote for ABC shall be credited to candidate C unless the voter has contributed, as far as he is able, the same portion of his vote to the election of candidate B as other voters who have contributed to the election of candidate B.

The main principle behind the Meek method (given as principle 2 in [2]) can be stated as: if a voter votes for candidates A, B, C in that order, and if candidates A and B each have a surplus of votes above the quota, then, on principle, a portion of the vote for ABC shall be credited to candidate C.

These different principles lead to the different rules as set out in paragraphs 3 to 8 of section 1.

I think that whether one prefers the Meek method to the Warren method, or vice versa, should be based on principle, and I prefer the principle upon which the Warren method is based. As stated in paragraph 8 of section 1, because of the impossibility theorems, it will always be possible to produce particular examples that tell against any given method. So I prefer to rest my case on the matter of principle, rather than on seeking examples of where the Warren method gives a 'better' result than the Meek method. Nevertheless, an example will be given, not with the object of showing that one method gives a better result than the other, but of showing how the two methods can give different results.

Consider the following election for 3 seats by 39996 voters, for which the quota is 9999.

10000	vote	ABC
100	vote	AE
10000	vote	BD
9998	vote	C
9898	vote	D

The numbers have been chosen so that, unlike the situation in real elections, the count can be done manually.

Under the Meek method the count can be portrayed as follows:

Voter	Number of such voters	Portion of vote contributed by each voter to each candidate A	B	C	D	E
Keep value		0.99	0.99	1	1	1
ABC	10000	0.99	0.0099	0.0001	0	0
AE	100	0.99	0	0	0	0.01
BD	10000	0	0.99	0	0.01	0
C	9998	0	0	1	0	0
D	9898	0	0	0	1	0
Total vote for each candidate		9999	9999	9999	9998	1

Under the Warren method the count can be portrayed as follows:

Voter	Number of such voters	Portion of vote contributed by each voter to each candidate A	B	C	D	E
Portion apportioned		0.99	0.9899	1	1	1
ABC	10000	0.99	0.01	0	0	0
AE	100	0.99	0	0	0	0.01
BD	10000	0	0.9899	0	0.0101	0
C	9998	0	0	1	0	0
D	9898	0	0	0	1	0
Total vote for each candidate		9999	9999	9998	9999	1

We see from these tables that the Meek method elects candidates A, B, C, whereas the Warren method elects candidates A, B, D.

We observe that the Meek and Warren methods are in agreement as to the portion of vote that each of the ABC voters and the AE voters contribute to candidate A, which is in keeping with the Warren principle that all contributors to the election of a candidate should contribute the same portion of their vote.

We observe that the Meek and Warren methods differ in the portion of vote that each of the ABC voters, and each of the BD voters, contribute to candidate B. Both methods ask the BD voters to contribute closely 99% of their vote to candidate B, and ask the ABC voters to contribute only closely 1% to candidate B. The Warren method accepts this difference, because, although it would have preferred that all groups of voters contributed the same portion, it recognises that the ABC voters did use up all that was left of their vote after contributing to candidate A, and could not contribute more.

The Meek method is desirous that, if a voter votes for a candidate who is elected with a surplus, then that voter should not be asked to contribute so much of his vote to that candidate that he has nothing to pass on. Accordingly, although each ABC voter is contributing only closely 1% of his vote to the election of candidate B, compared with the 99% that each BD voter is contributing, Meek's principle requires that the ABC voters shall contribute slightly less than 1% of their vote to the election of candidate B in order that a portion, which amounts to about one ten-thousandth of a vote, shall be passed to candidate C.

This shows what the difference between the Meek and Warren methods amounts to. In my opinion the difference raises the question as to whether the ABC voters, who have contributed only closely 1% of their vote to the election of candidate B, whereas the BD voters have contributed closely 99% towards the same end, merit the right, in these circumstances, to pass on a portion of their vote to candidate C, as Meek's principle requires, at the expense of expecting the BD voters to bear even more of the burden of electing candidate B. If one thinks that the right should be afforded, then one should prefer the Meek method. But if one thinks that it would not be fair to afford this right, then one should prefer the Warren method.

4 References

[1] R.A. Newland and F.S. Britton. *How to conduct an election by the Single Transferable Vote*. 2nd edition. Electoral Reform Society. 1976.

[2] B.L. Meek. A new approach to the Single Transferable Vote. *Voting matters*, issue 1, 1–11. 1994.

[3] C.H.E. Warren. Counting in STV elections. *Voting matters*, issue 1, 12–13. 1994.

[4] D.R. Woodall. An impossibility theorem for electoral systems. *Discrete Mathematics*, 66, 209–211. 1987.

Sequential STV — a further modification

I.D. Hill and Simon Gazeley
simon.gazeley@btinternet.com

1 Introduction

We had hoped that our earlier paper [1] would be the final version of the Sequential STV system, but we have found two examples since then that seem to call for further amendment.

The aim is to find a system that will be noticeably like ordinary STV but: (1) will correct unfairness, if any, to candidates excluded by the reject-the-lowest rule; (2) will automatically reduce to Condorcet's method rather than Alternative Vote when there is only a single seat.

It seeks to find a set of n candidates that observes Droop Proportionality [3], which we regard as an essential feature of any worthwhile voting system, and is preferred by the largest majority of voters to any other possible set of n. Tideman's CPO-STV [2] has similar objectives. The successful set will usually be such that any set of n+1 candidates, consisting of those n and 1 more, will result in the election of those n when an STV election is performed and in this case we refer to the successful set as a Condorcet winning set.

In a small election, or when n=1, it would be relatively easy and quick to do a complete analysis, as CPO-STV does. The challenge is to find a way that will work in a reasonable time in large elections, where such a complete analysis would be impracticable. We recognise that the meanings of 'a reasonable time' and 'impracticable' are open to dispute, and that what is practicable will change as computers continue to get faster. As Tideman and Richardson say "We are not yet at a point where computation cost can be ignored completely".

In cases where it is practicable to do a complete analysis of all sets of n+1, n+2, etc., it might be possible to find a solution that, in some sense, is preferable to that produced by this system that (after an initial stage) looks only at sets of n+1 and only at some of those. We think, however, that it would be hard to claim a severe injustice to any non-elected candidate after this system had been used, and it does keep things within manageable limits. It would be interesting to compare the performance of Sequential STV and CPO-STV, but this has not been done yet.

Of the two worrying examples, one showed that the system, as previously given, could fail to preserve Droop Proportionality, while the other showed that we were a little over-optimistic in claiming that, if the special procedure to deal with a Condorcet paradox had to be invoked, "most of the original candidates will be either excluded or certainties, [so] there is no need to fear an astronomical number of tests needing to be made". This second example was highly artificial and the optimism was probably justified for any real voting pattern that is at all likely to occur, but even artificial patterns ought not to cause trouble.

To cure the first of these troubles it is necessary, when the special procedure is used, to let it exclude just one candidate before restarting the main method, instead of continuing to use the special procedure. To cure the second, the special procedure has been much simplified, to calculate a value for each continuing candidate based upon Borda scores, and to exclude the one with the lowest score. We emphasise that in real elections, as distinct from specially devised test cases, Condorcet loops rarely occur and so the special procedure is rarely called into use.

Borda scores on their own, as an electoral method, we regard as a very poor option. Those elected are far too dependent upon whether or not other (non-winning) candidates are standing, and the method is much too open to tactical voting; but as a method of helping to sort out a Condorcet paradox, they can be useful. Where a paradox arises, we know that there cannot be a good result because, whoever is elected, it is possible to point to some other option that a majority of the voters preferred; so the best that can be done is to try for a not-too-bad result and, for this limited purpose, Borda scores can serve.

2 Revised version of Sequential STV

All STV counts mentioned are made by Meek's method. It would be possible to use a similar system with some other version of STV but, since many counts are to be made using the same data, to try it other than by computer would make little sense. If a computer is required in any case, Meek's method is to be preferred.

An initial STV count is made of all candidates for n seats, but instead of dividing into those elected and those not elected, it classifies those who would have been elected as probables, and puts the others into a queue, in the reverse order of their exclusion in that STV count, except that the runner-up is moved to last place as it is already known that an initial challenge by that candidate will not succeed. Having found the probables and the order of the queue, further rounds each consist of n+1 candidates, the n probables plus the head of the queue as challenger, for the n seats. Should a tie occur during these rounds, between a probable and a challenger, it is resolved by maintaining the current situation; that is to say, the challenger has not succeeded.

If the challenger is not successful, the probables are unchanged for the next round and the challenger moves to the end of the queue, but a successful challenger at once becomes a probable, while the beaten candidate loses probable status and is put to the end of the queue. The queue therefore changes its order as time goes on but its order always depends upon the votes.

This continues until either we get a complete run through the queue without any challenger succeeding, in which case we have a solution of the type that we are seeking, or we fall into a Condorcet-style loop.

A loop may have been found if a set that has been seen before recurs as the probables. If the queue is in the same order as before then a loop is certain and action is taken at once. If, however, a set recurs but the queue is in a different order, a second chance is given and the counting continues but, if the same set recurs yet again, a loop is assumed and action taken.

In either event the action is the same, to exclude all candidates who have never been a probable since the last restart (which means the start where no actual restart has occurred) and then to restart from the beginning except that the existing probables and queue are retained instead of making a new initial STV count.

If there is no candidate who can be so excluded, then a special procedure is used, in which each continuing candidate, other than any who has always been a probable since the last restart, is classified as 'at-risk'. Taking each continuing candidate, a Borda score is calculated, as the sum over all votes of the number of continuing candidates to whom the candidate in question is preferred, taking all unmentioned continuing candidates as equal in last place. A continuing candidate who is not mentioned in a particular vote is given, for that vote, the average score that would have been attained by all those unmentioned. In practice it can help to give 2 points instead of 1 for each candidate beaten, because all scores, including any averages required, are then whole numbers.

The at-risk candidate with the lowest score (or a random choice from those with equal lowest score) is then excluded and the main method restarted from the beginning, except that the existing probables and queue order are retained instead of making the initial STV count. If the newly excluded candidate was one of the queue, he or she is merely removed from the queue, but if the candidate was a probable, the candidate at the head of the queue is reclassified as a probable and removed from the queue. Then a restart is made from the beginning except that the existing probables and queue are retained instead of making a new initial STV count.

3 Proof of Droop Proportionality compliance

The 'Droop proportionality criterion' says that if, for some whole numbers k and m (where k is greater than 0 and m is greater than or equal to k), more than k Droop quotas of voters put the same m candidates (not necessarily in the same order) as their top m preferences, then at least k of those m candidates will be elected.

We know that a normal STV count is Droop Proportionality compliant so, in Sequential STV, for k and m defined as above, at least k of the m will be probables at the first count. If on a later count a challenger takes over as a probable then, because that was also the result of an STV count, there will still be at least k of the m among the probables, even if the replaced candidate was one of the m. This ensures compliance if no paradox is found.

If a paradox is found, at least k of the m will have been probables at some time since the last restart, so excluding all who have not been probables must leave at least k. If the special procedure, using Borda scores, is required, then if only k exist, k will have always been probables since the last restart, and so are not at risk of exclusion, but if there are more than k, the exclusion of just one of them must leave at least k. This ensures compliance where a paradox is found.

4 Examples

Example 1

This is the example that showed the old version of Sequential STV to fail on Droop Proportionality. With 9 candidates for 3 seats, votes are

10	ABCDEFGH	10	BCDAFGHI
10	CDABGHIE	11	DABCHIEF
19	EFGHIDAB	19	FGHIEBCD
1	GHIEFCDA		

41 votes (more than 2 quotas) have put ABCD, in some order, as their first choices so, to satisfy Droop Proportionality, at least 2 of them must be elected. The old version elected DEF but the new version elects ADE.

Example 2

This is the example that showed the old version of Sequential STV not always to finish within a reasonable time. With 40 candidates for 9 seats, votes are

69	ABCDE	94	BCAED	98	CBAED
14	DEBAC	60	ECBDA	64	FGJHI
43	GIFJH	42	HJIGF	97	IHGJF
33	JIHGF	32	KLMNO	44	LMNOK
56	MNOKL	76	NOKLM	90	OKLMN
18	PQRST	91	QRSTP	69	RSTPQ
21	STPQR	76	TPQRS	36	UVWXY
78	VWXYU	99	WXYUV	29	XYUVW
4	YUVWX	64	abcde	35	bcdea
69	cdeab	98	deabc	16	eabcd
40	fghij	44	ghijf	79	hijfg
42	ijfgh	68	jfghi	13	kmnop
64	mnopk	83	nopkm	30	opkmn
33	ponmk				

This new version of Sequential STV terminates after 835 STV counts, whereas the old version would, we estimate, have required over 177,000 counts. We emphasise again that the voting pattern is highly artificial — in a real election, with 40 candidates for 9 seats, more than 60 counts would be very unusual.

Example 3: "Woodall's torpedo"

With 6 candidates for 2 seats, votes are

11	AC	9	ADEF	10	BC
9	BDEF	10	CA	10	CB
10	EFDA	11	FDEB		

Sequential STV elects CD even though AB form the unique Condorcet winning set. Examining why this happens, it is found that A and B are always elected by STV from any set of 3 in which they are both present, but neither A nor B is ever elected if one of them is there but not the other. Meanwhile C is always elected if present in a set of 3 except for the one set ABC. D, E and F form a Condorcet loop. CD, CE or CF would be a second Condorcet winning set if the other two of D, E and F were withdrawn.

Such a strange voting pattern is unlikely to arise in practice. It shows that Sequential STV cannot be guaranteed to find a Condorcet winning set even where one exists but it does not shake our belief that Sequential STV is a good system; it would be hard to deny that C is a worthier winner than either A or B in this example.

5 Acknowledgements

We thank Douglas Woodall for devising example 3, and the referee for useful comments on earlier versions of this paper.

6 References

[1] I.D. Hill and Simon Gazeley, Sequential STV — a new version. *Voting matters*, 15, 13–16. 2002.

[2] T. Nicolaus Tideman and Daniel Richardson, Better voting methods through technology: the refinement-manageability trade-off in the single transferable vote. *Public Choice*, 103, 13–34. 2000.

[3] D.R. Woodall, Properties of preferential election rules. *Voting matters*, 3, 8–15. 1994.

A new way to break STV ties in a special case

Earl Kitchener
ekk@talk21.com

1 Proposal

The simplest example of a particular type of tie has three votes, AB, BA, CA, for one place. The quota is 1.5, and so, under the normal rules, one candidate is selected at random for exclusion, giving the chance of election as 2/3 for A and 1/3 for B. If it is B, A will be justifiably aggrieved, and opponents of STV will argue that a random choice has given a perverse result.

A general rule to cover cases of this type would be to say that when all continuing candidates are tied (whether for exclusion or for election), they are all to be excluded, but only for the current preferences, all later preferences being unaltered. If voting is seen as a process of cutting off the top preference of each vote as soon as the fate, election or exclusion, of the candidate concerned has been decided, and reducing the value of the vote in the case of election, then this proposal introduces a new type of exclusion in which the top candidate is cut off in the normal way, but the candidate is not removed from any other votes.

The above votes, but with two places to be filled, give an example of a tie for election. Under the normal rules, whichever candidate is elected first, each of the other two has an equal chance of second place. So each of the three candidates has 2/3 of a chance of being elected. Under the proposals, A wins with 2 votes, B is elected with 1, and C gets none.

A possible objection is that the proposal violates the rule that later preferences must never be looked at until the fate of earlier ones has been decided, and there is a danger that it might discourage sincere voting, but this seems unlikely, and is out-weighed by its advantages if voting is sincere.

If Borda's method of counting votes is used for tie-breaking, this proposal would not be necessary; but it has the advantage of being less of a departure from the present system.

This tie is very unlikely except in small elections, but it might well occur if partners are voting for a senior partner. If the proposal is considered too sweeping, it could be restricted to the case where the voters are the same as the candidates, and they each vote first for themselves. This would still give most of the benefits.

A powerful test of any proposed change to vote counting is, "Would it, compared with other rules, make any voters or candidates justifiably aggrieved, or lead to insincere voting?" This proposal gains on the first test, and only loses slightly on the second. Allowing parties to put up more candidates than they can hope to get in, and discouraging tactical voting, are also important, but not likely to be affected by changes in tie-breaking rules.

2 Editorial notes on tie breaking

The question of ties with STV has arisen several times in *Voting matters*. The previous material can be summarised as follows:

- Earl Kitchener in Issue 11 of *Voting matters* advocates the use of Borda scores [1].
- David Hill in Issue 12 argues against the use of Borda scores [2].
- Jeff O'Neill in Issue 18 notes that many rules use a first-difference rule, but he advocates a last-difference rule [3].
- Wichmann considers the use of computers in Issue 19. Here, the suggestion is that no specific rule is needed and that the computer can try all options and the result taken can be the most likely one [4].
- Earl Kitchener has returned to the subject with an alternative proposal to Borda scores in a special case which appears above.

2.1 Existing rules

The ERS rules [6] and the Church of England rules use the first-difference method in an attempt to break a tie.

The Meek algorithm [7] uses a deterministic algorithm based upon a random number generator to break a tie. No manual intervention is used. The New Zealand variant uses a similar method.

When the Church of England rules are applied using a computer, then the software must break the ties without manual intervention in a manner which is not defined (by the rules).

For Ireland, the manual rules are being computerized and have been used for three trial constituencies in 2002. Here, tie-breaking invokes a manual procedure, ie, the computer software does not break the tie.

A curiosity is that in the Irish rules if when allocating surplus remainders there is a tie of the fractional part, the surplus vote is given to the candidate with the largest total number of papers from that surplus; if that is also tied then first difference is used.

It seems that a Condorcet comparison has been used to resolve a strong tie between A and B (i.e. tie can't be broken by first/last difference) in very small manual counts i.e. examine the papers to see how many times A is ahead of B compared to vice versa.

2.2 Discussion

This section was produced as a result of an email debate; those contributing included: James Gilmour, David Hill, Michael Hodge, Joe Otten, Joe Wadsworth and Douglas Woodall.

A number of issues arise from tie-breaking:

Are tie-breaking rules needed? Surely better to have a rule than toss a coin?

If a rule like first-difference, fails to break the tie, then drawing lots or some computer equivalent is needed unless we allow later preferences to be looked at. But the disadvantages remain formidable as we are then unable to promise that later choices cannot upset earlier ones. These extra tie-breaking rules complicate the counting process, since ties can arise in more than one way. It seems that just drawing lots would be adequate.

If we are saying that for:

1	AB
1	BA
1	CA

fairness demands A is elected, the same would apply to

1000	AB
1000	BA
1000	CA

So what about

999	AB
1000	BA
1000	CA

Or even

1000	AB
1001	BA
1000	CA

It seems that if the logic of looking at later preferences is sound and compelling, then they should be considered in these later examples. They are all almost identical with almost the same support for A, yet B wins with probability 1/3, 1 and 1/2 respectively. If the 1/3 should be 0, on the grounds of later preferences, perhaps the 1 or 1/2 should be reduced too?

There seems nothing in the logic of the argument that limits it to ties. Why not judge all exclusions on the basis of 'probability of election' in some sense given an analysis of all later preferences, limited only by a 'probably-later-no-harm' principle defined statistically?

This would be a rival to STV, to be considered on its merits, without muddying the waters by introducing features of it to STV for extremely marginal benefits. The claim being made here is that we want the Condorcet winner (or a similar result in the multi-seat case) rather than the AV winner. The argument is quite separate from tie breaking as such, and Condorcet-type rules need paradox breakers as well as tie breakers. If anything of the sort is to be considered, then Sequential STV [8] could be the starting point.

If rules are used, what criteria are appropriate?

There is significant opposition to using later preferences in breaking a tie, see [2], for instance. One can argue against this on the grounds that it is hard to observe the difference between *any* tie-breaking logic and a random choice.

There was significant support for using the last-difference rule as opposed to the first-difference rule. One correspondent wrote of the latter, "It would be a bit like requiring the Speaker, in the event of a tied vote in the House, to cast his vote not in favour of the status quo, but in favour of the outcome that more closely resembled the very earliest legislation ever passed on that question." But it can also be argued that any such rule is arbitrary and, if it is not necessary to change, it is necessary not to change.

The first-difference rule can have the effect of giving preference to first-preference votes as opposed to transfers — this seems against the spirit of STV. With a computer, one can experiment with different procedures for breaking a tie. A reasonable criterion would be the method that most reliably resulted in the election of the candidates with the highest probabilities of being elected from breaking the ties in all the possible ways. The special case that Kitchener uses would always give the optimal result, but it is unclear how often that special case arises.

The use of Borda scores is not liked by the supporters for STV, but it is unclear if similar perverse results could be obtained if Borda scores were introduced only to break ties.

The issue of voter satisfaction has been raised. It certainly seems unsatisfactory that all the existing rules will report a random choice for elections in which the choice does not change the candidates elected. This is quite common with candidates with very low numbers of first-preferences. However, the following could be proposed to measure voter satisfaction in a tie-breaking rule:

- the method which maximizes the voters contributing to those elected;

  Maximising voters seems to accord to the *inclusive* view of STV which allows voters to be added to those supporting an already elected candidate as occurs with the Meek rules.

  The conventional approach of the manual rules is *exclusive* in which voters are not added to the list of those supporting an already elected candidate.

- the method which minimizes the non-transferable votes.

  The conventional practice with the manual rules is to minimise the non-transferable votes by considering transferable votes first when transferring surplus. In contrast, the Meek rules do not do this. However there are those who would claim that any proposal artificially to reduce non-transferables is immoral, in that it distorts what the voters have asked for.

3 References

[1] Earl Kitchener. Tie-Breaking in STV. *Voting matters*. Issue 11, pp 5–6. (see `http://www.mcdougall.org.uk/VM/ISSUE11/P3.HTM`)

[2] I. D. Hill. Tie-breaking in STV. *Voting matters*. Issue 12, pp 5–6. (see `http://www.mcdougall.org.uk/VM/ISSUE12/P2.HTM`)

[3] J. C. O'Neill. Tie-Breaking with the Single Transferable Vote. *Voting matters*. Issue 18, pp 14–17. (see `http://www.mcdougall.org.uk/VM/ISSUE18/I18P6.PDF`)

[4] B. A. Wichmann: Tie Breaking in STV. *Voting matters*. Issue 19, pp 1–5. (see `http://www.mcdougall.org.uk/VM/ISSUE19/I19P1.PDF`)

[5] GS1327: General Synod, Single Transferable Vote regulations 1990 and 1998. (Obtainable from Church House Bookshop, Great Smith Street, London SW1P 3BN.)

[6] R. A. Newland and F. S. Britton. *How to conduct an election by the Single Transferable Vote*. ERS 3rd Edition. 1997. (see `http://www.electoral-reform.org.uk/votingsystems/stvrules.htm`)

[7] I. D. Hill, B. A. Wichmann and D. R. Woodall. Algorithm 123 — Single Transferable Vote by Meek's method. *Computer Journal*. Vol 30, pp277–281, 1987.

[8] I.D. Hill and Simon Gazeley. Sequential STV — a further modification. *Voting matters*, Issue 20, pp6–8.

Apportionment and Proportionality: A Measured View

P Kestelman
No email available.

1 Introduction

Collins (2003) **English Dictionary** defined 'Proportional Representation' (PR) as: "representation of parties in an elective body in proportion to the votes they win". Few elections translate every Party Vote-fraction into the same Seat-fraction, thereby mediating exact PR; and raising the question of when to describe an election as full PR, semi-PR ('broad PR') or non-PR.

According to Gallagher, Marsh and Mitchell [11], "Ireland uses the system of proportional representation by means of the single transferable vote (PR-STV) at parliamentary, local, and European Parliament elections (the president, too is elected by the single transferable vote)". Presidential single-member STV is **Alternative Voting** (AV), which also elects the Australian House of Representatives.

Is AV therefore a PR electoral system? The Independent Commission on the Voting System [13] — the Jenkins Report — maintained that AV alone "is capable of substantially adding to ['First-Past-the-Post' (FPP)] disproportionality". The more recent Independent Commission on PR [12] affirmed that "AV can produce a hugely disproportionate result".

How should we compare the Party disproportionality of different electoral systems? Which is the fairest — most proportional — electoral system? In other words, how should disproportionality — departure from exact PR — be quantified?

2 Apportionment

First consider the analogous question of the fairest method of apportionment. Collins (2003) **English Dictionary** defined 'apportionment' as: "*U.S. government.* the proportional distribution of the seats in a legislative body, esp. the House of Representatives, on the basis of population".

The USA has long wrestled with the problem of the most representative apportionment; trying various methods (Balinski and Young [1]). **Table 4.1** gives the apportionment of 105 Seats among 15 States in the first (1791) House of Representatives, applying the main five Divisor methods. For the five most and least populous States, proportionality is measured as the ratio between their aggregate Seat-fractions and Population-fractions ($S\%/P\%$).

Adams, Dean and Hill yield the same apportionment: slightly *under*-representing the five most populous States ($S\%/P\% = 0.99$); while *over*-representing the five least populous States ($S\%/P\% = 1.09$). These methods produce a **Relative Bias** of + 10 percent (Bottom/Top third $S\%/P\% = 1.09/0.99 = 1.10$).

On the other hand, Jefferson *over*-represents the top five States ($S\%/P\% = 1.02$); and *under*-represents the least populous States ($S\%/P\% = 0.89$): a Relative Bias of – 13 percent (Bottom/Top third $S\%/P\% = 0.89/1.02 = 0.87 = 1 - 0.13$). With the lowest Relative Bias (– 2 percent), Webster yields the fairest 1791 Apportionment.

Requiring at least one Seat per State usually over-represents the least populous States. Eliminating that constraint — so quantifying method-specific bias more precisely — **Table 4.1** (bottom panel) gives the Mean Bias for all 22 USA apportionments (1791–2000). The Webster (Sainte-Laguë) Method proved the least biased overall (averaging 0.1 percent); whereas Adams (Smallest Divisor) and Jefferson (d'Hondt) were the most biased (over 20 percent).

3 Apportioning England

Nearer home, **Table 4.2** apportions 71 MEPs between the nine English Regions, applying the five Divisor methods to their 1999 Electorates. Adams and Dean coincided but, despite identical Bottom/Top third Relative Bias, differed slightly from Hill and Webster. Which apportionment is fairer?

The *European Parliament (Representation) Act 2003* prescribes that: "the ratio of electors to MEPs is as nearly as possible the same in each electoral region". In testing fairness, the Electoral Commission [7] accepted a measure that "involves calculating for each region the difference between the number of electors per MEP for that region and the overall number of electors per MEP, and adding up all these differences (having ignored minus signs). The smaller this total is, the more equitable the outcome".

A little mathematical notation helps here. The overall number of Electors per MEP, $E/S = \sum E_R/S_R$, where $\sum$(Sigma) denotes 'Sum' (over all Regions); E_R is the number of electors in a Region; and S_R is the corresponding number of seats. Each Regional deviation is the absolute difference (that is, ignoring negative signs) between its E_R/S_R and E/S ; and

$$\begin{aligned}\text{Total Deviation } &= \sum \mid E/S - E_R/S_R \mid \\ &= E/S \sum \mid 1 - (E_R/E)/(S_R/S) \mid \\ &= E/S \sum \mid 1 - E_R\%/S_R\% \mid ,\end{aligned}$$

where E_R % and S_R % are the Regional Elector- and Seat- fractions *(percent)*, respectively.

For any given apportionment, total Electors and Seats — and thus E/S — are fixed: hence Regional MEP apportionment is required to minimise $\sum \mid 1 - E_R\%/S_R\% \mid$. The UK statutory criterion implies the Dean Method (Balinski and Young [1]).

Nonetheless, for the June 2004 European Elections, the Electoral Commission [7] recommended the Webster (Sainte-Laguë) Method, making the ratio of *MEPs to electors* as nearly as possible the same in each Region (beyond the statutory minimum of three MEPs). Based on December 2002 Regional electorates, Dean and Webster apportionments coincided.

We may therefore define a **Dean Index** $= \sum \mid 1 - E_R\%/S_R\% \mid$; and a **Webster Index** $= \sum \mid 1 - S_R\%/E_R\% \mid$. **Table 4.2** (bottom panel) confirms that the Dean Method minimises the Dean Index; and the Webster Method minimises the Webster Index.

4 Paradox and Proportionality

Overall measures of malapportionment (like the Dean and Webster indices defined above) are better than partial measures (like Bottom/Top third Relative Bias). The Webster Method minimises total *relative* differences between Regional Elector-fractions and Seat-fractions:

$$\begin{aligned}\text{Webster Index } &= \sum \mid 1 - S_R\%/E_R\% \mid \\ &= \sum \mid E_R\% - S_R\% \mid /E_R\%.\end{aligned}$$

Total *absolute* differences between Regional Elector-fractions and Seat-fractions are minimised by the **Hamilton** Method (Largest Remainders: LR–Hare).

This Quota Method allocates to each Region the integer part of its proportional entitlement (number of Hare Quotas: one Hare Quota = National Electors/National Seats). Any residual seats are then allocated to the regions with the largest fractional parts (remainders) of Hare Quotas.

We may therefore define a **Hamilton Index** $= \sum \mid E_R\% - S_R\% \mid$; minimised by the Hamilton Method. Applied to all 22 USA apportionments (without seat minima), Hamilton averages a (Bottom/Top third) Relative Bias of – 0.3 percent: differing insignificantly from Webster (– 0.1 percent).

Unlike Webster, the Hamilton Method of apportionment is vulnerable to paradox: notably the Alabama Paradox. The 1880 USA Census disclosed that, if total House size were *increased* from 299 to 300 seats, then the Hamilton apportionment to Alabama would have *decreased* from eight to seven seats (Balinski and Young [1])!

That consideration excludes Hamilton as a method of apportionment; though not necessarily for evaluating malapportionment. So how best to quantify malapportionment — or disproportionality?

5 Party Disproportionality

Gallagher [10] concluded that each PR method "minimizes disproportionality according to the way it defines disproportionality". However, Lijphart [14] argued that LR-Hare (Hamilton) and Sainte-Laguë (Webster) mediate "inherently greater proportionality" than d'Hondt (Jefferson); thereby justifying proportionality measures "biased in favour of LR-Hare".

LR-Hare minimises the **Loosemore-Hanby Index** (Loosemore and Hanby, [15]):

LHI *(percent)*$= \frac{1}{2} \sum \mid V_P\% - S_P\% \mid$,
where $V_P\%, S_P\% =$ Party Vote–, Seat–fractions *(percent)*.

Compare the Hamilton Index $= \sum \mid E_R\% - S_R\% \mid$, as defined above. Halving the sum ensures that LHI ranges 0–100 percent.

LHI is the 'DV score' mentioned by the Independent Commission on the Voting System [13]; and as defined by the Independent Commission on PR [12]. LHI complements the **Rose Proportionality Index** (Mackie and Rose, [16]) *percent*:

$= 100 - \frac{1}{2} \sum \mid V_P\% - S_P\% \mid = 100 -$LHI *(percent)*.

Table 4.3 illustrates the calculation of LHI and RPI for the 2004 European Parliamentary Election in Britain. Over-represented and under-represented Party Total Deviations are necessarily equal and opposite (± 14.7 percent in **Table 4.3**); and Party total over-representation is simply the Loosemore-Hanby Index (LHI = 14.7 percent).

6 Debate

As a measure of Party disproportionality, the Loosemore-Hanby Index (LHI) has been criticised on three main grounds: for violating Dalton's Transfer Principle (Taagepera and Shugart [22]); for being vulnerable to paradox (Gallagher [10]); and for exaggerating the disproportionality of PR systems involving many parties (Lijphart [14]).

Dalton's Transfer Principle states that transferring wealth from a richer to a poorer person decreases inequality, decreasing any inequality index (Taagepera and Shugart [22]). However, transferring seats between over-represented parties (or between under-represented parties) leaves LHI unchanged.

Thus in the 2004 European Election in Britain (**Table 4.3**), imagine the Conservatives (from 27 to 25 seats) losing two seats to Labour (from 19 to 21 seats). Then both Party deviations would converge ($S_P\% - V_P\% =$ from + 9.3 to + 6.6 percent, and from + 2.7 to + 5.4 percent, respectively); decreasing GhI (from 8.3 to 7.7 percent), leaving LHI unchanged (14.7 percent). However, Party total over-representation remains unchanged: so why should overall disproportionality change?

Minimised by LR-Hare (Hamilton), LHI is susceptible to the paradoxes of that Quota method (Gallagher [10]). Because Sainte-Laguë (Webster) is the least biased Divisor method — and immune to paradox — Gallagher [10] recommended a **Sainte-Laguë Index** "as the standard measure of disproportionality":

SLI *(percent)* $= \sum (V_P\% - S_P\%)^2 / V_P\%$.

However, in a single-member constituency, if the winner receives under half of all votes, then SLI exceeds 100 percent (unlike LHI, which measures unrepresented — wasted — votes).

Nowadays, Gallagher [10] is mainly cited for his 'Least Squares Index':

GhI *(percent)*$= \sqrt{\frac{1}{2} \sum (V_P\% - S_P\%)^2}$.

Also minimised by LR–Hare, GhI is subject to the same paradoxes as LHI. Gallagher [10] saw GhI as "a happy medium" between LHI and the **Rae Index** (Rae [18]):

RaI *(percent)* $= \sum \mid V_P\% - S_P\% \mid / N$,
where N = Number of parties ($V_P\% > 0.5$ percent).

Thus RaI measures *average* deviation per Party; whereas LHI measures (half) *Total* Deviation. Yet why hybridise such conceptually distinct measures in one measure (GhI)?

Taagepera and Grofman [21] have attributed the recent shift, from LHI towards GhI, "to sensitivity to party system concentration"; based on the intuition of Lijphart [14] that a few large deviations ($V_P\% - S_P\%$) should be evaluated as more disproportional overall than many small deviations with the same Total Deviation (and hence LHI). It remains unclear why larger Party deviations should be potentiated; and smaller ones attenuated.

For example, in the 2004 European Election in Britain, exact GhI was 8.3 percent. However, aggregating unrepresented parties ($S_P\% = 0.0$ percent: **Table 4.3**) increases GhI to 10.7 percent; leaving LHI unchanged (14.7 percent). In the process, Party total under-representation has not changed: so why should Total Disproportionality change? Likewise, in single-member constituencies, GhI depends on the division of votes among losing candidates.

Monroe [17] proposed an inequity index rather similar to GhI:

$$\mathbf{MrI}(percent) = \sqrt{\frac{\sum (V_P\% - S_P\%)^2}{1 + \sum (V_P\%/100)^2}}$$

LR-Hare also minimises MrI; which falls below 100 percent for extreme disproportionality involving more than two parties (like GhI, but unlike LHI).

Taagepera and Shugart [22] mentioned an electoral analogue of the widespread Gini Inequality Index, with several examples; but without defining any **Gini Disproportionality Index** (GnI). It turns out that **GnI** *(percent)*:

$$= \sum \sum \mid (V_P\% \times S_Q\%) - (S_P\% \times V_Q\%) \mid /200$$

Thus GnI sums the absolute differences between the $S_P\%/V_P\%$ ratios of every pair of parties, weighted by the product of their vote-fractions $(V_P\%/100)$. This complex GnI satisfies Dalton's Transfer Principle; and aggregating unrepresented parties ($S_P\%$ = 0.0 percent) leaves GnI unchanged (like LHI and SLI).

Taagepera and Grofman [21] evaluated 19 disproportionality indices against 12 criteria, sustaining five measures: LHI; GhI; SLI ('chi-square'); MrI; and GnI. Nonetheless, they overlooked both a Farina Index (FrI) and a Borooah Index (BrI).

Woodall [24] cited JEG Farina for a vector-based measure of Party Total Disproportionality: the angle between two multidimensional vectors, whose coordinates are Party vote and seat numbers. Its fraction of a right angle defines a **Farina Index**, FrI (*percent*) =

$$\arccos\left[\frac{\sum(S_P\% \times V_P\%)}{\sqrt{\sum S_P\%^2 \times \sum V_P\%^2}}\right] \times 100/90^\circ$$

ranging 0–100 percent (instead of 0–90 degrees).

Borooah [2] proposed an electoral analogue of the Atkinson Inequality Index, depending on "society's aversion to inequality" (like Gini, originally measuring income inequality). Establishing national 'Societal Aversion to Disproportionality' seems arbitrary; while a moderate value (SAD = 2) defines a **Borooah Index**,

BrI *(percent)* $= 100 - 1/[\sum(S_P\%/100)^2/V_P\%]$, ranging 0–100 percent.

7 Correlations

For 82 general elections in 23 countries (1979–89), Gallagher [10] reported high correlations between LHI, GhI and SLI. Graphing high correlations between LHI, GhI, SLI and FrI, Wichmann [23] noted that central placement reinforced LHI.

Table 4.4 gives the correlations between all seven indices in the last 44 UK general elections (1832–2005). Most notably, LHI proved very highly correlated with GnI; GhI with MrI; and SLI with BrI ($R > 0.99$). Indeed, LHI and GnI were highly correlated ($R > 0.95$) with all other measures of Party Total Disproportionality.

8 Proportionality Criteria

The Independent Commission on the Voting System [13] observed that "full proportionality ... is generally considered to be achieved as fully as is normally practicable if [LHI%] falls in the range of 4 to 8". More generously, we might allow LHI under 10 percent to characterise **full PR**. LHI ranging 10–15 percent could then encompass **semi-PR** ('broad PR'); with LHI over 15 percent constituting **non-PR**.

In **UK** general elections (FPP) since World War I, LHIs have ranged from 27 percent (1918); to only four percent (1951) — ironically, when the Conservatives won fewer votes, but more seats, than Labour (Rallings and Thrasher [19]). In the last nine general elections (1974–2005: **Table 4.5**), LHIs have ranged 15–24 percent, averaging 20 percent: clearly *non-PR*.

What of the nominally PR elections, introduced in Britain since 1997? In the 1999 and 2004 European Parliamentary elections, Regional d'Hondt yielded LHIs of 14.1–14.7 percent (between Party List votes and MEPs) nationwide: barely *semi-PR*. Likewise applied regionally, either Sainte-Laguë (LHI = 6.1–8.4 percent), or LR–Hare (LHI = 6.1–5.4 percent), would have mediated *full PR*. So the method used here can make a considerable difference.

In the 1999 and 2003 **Scottish Parliament** and **National Assembly for Wales** elections, between Party List votes and Total (FPP Constituency + Additional Regional) seats, LHIs ranged 11–14 percent. The 2000 and 2004 **London Assembly** elections (also FPP-plus, but with a five percent Party Vote Threshold) yielded similar Party List LHIs of 14–15 percent. Thus all three British Regional Assemblies remain *semi-PR* at best (Independent Commission on PR [12]).

In contrast, both 1998 and 2003 **Northern Ireland Assembly** elections (multi-member STV) mediated First Preference LHIs of only 6.0–6.4 percent: *full PR*. **Table 4.6** ranks UK national and regional election LHIs over the past decade (1995–2005).

9 Vote Transferability and District Magnitude

Transferable voting complicates evaluating the disproportionality of both AV and multi-member STV. First Count LHI is *not* the sole criterion; though Final Count LHI over-estimates Party proportionality (Gallagher [9]). For comparing transferable voting with other electoral systems, averaging First and Final Count LHIs appears reasonable.

Under Alternative Voting (**AV**), in the last nine general elections in **Australia** (1983–2004), First Count LHI ranged 12–20 percent, averaging 16 percent (**Table 4.5**): practically *non-PR*. Final Count LHI ranged 5–13 percent, averaging eight percent (PR); while mean First + Final Count LHI averaged 12 percent: *semi-PR* overall (compare **Table 4.6**).

So much for empirical claims that AV "is capable of substantially adding to [FPP] disproportionality" (Independent Commission on the Voting System, [13]). FPP votes — involving tactical considerations — should not only be compared with AV First Preferences.

Taagepera and Shugart [22] called AV 'semi-PR'; and attributed any 'semi-PR effect' in multi-member STV elections to low **District Magnitude** (M = Number of Seats per Constituency). As Gallagher [9] noted: "the smaller the constituency [M], the greater the potential for disproportionality"; and reported decreasing LHI with increasing STV District Magnitude in 16 Irish general elections (1927–1973).

Table 4.7 gives national aggregate LHI, by District Magnitude and Count, in the last 13 Irish general elections (1961–2002). Between such low District Magnitudes (M = 3–5), disproportionality might be expected to fall steeply: so the relative insignificance of all LHI differences is remarkable.

Overall, First Count LHIs ranged 3–13 percent (averaging seven percent); Final Count LHIs ranged 1–7 percent (averaging three percent); and mean First + Final Count LHI averaged only five percent (and 6–7 percent for M = 3–5). Virtually regardless of District Magnitude, multi-member STV mediates *full PR*.

10 Conclusions

Sainte-Laguë (Webster) is the most equitable method of apportionment — and the most proportional electoral principle. The d'Hondt (Jefferson) Method over-represents the most populous regions (and the most popular parties).

Not much has changed since Gallagher [10] lamented "surprisingly little discussion of what exactly we mean by proportionality and how we should measure it". Certainly, Party disproportionality indices have proliferated; among which the Loosemore-Hanby Index (LHI) — straightforwardly measuring Party total over-representation — remains the most serviceable. Moreover, such absolute disproportionality is what matters politically [14, 21].

Continuing debate on the 'best' measure of disproportionality may distract attention from the main task of evaluating the relative disproportionality of different electoral systems. Taagepera and Grofman [21] marginally preferred the Gallagher Index (GhI); allowing that its advantages over LHI were debatable.

LHI fails Dalton's Transfer Principle; yet transferring seats between over- (or under-) represented parties should arguably not change a measure of Total Disproportionality. LHI, GhI and MrI alike remain vulnerable to the paradoxes of the Largest Remainders (LR-Hare/Hamilton) Method.

The Sainte-Laguë Index (SLI) is unsuitable for measuring Party Total Disproportionality. Fortunately highly correlated with LHI, the Gini Disproportionality Index (GnI) is rather complicated to explain and calculate (virtually necessitating computerisation). Interestingly, Riedwyl and Steiner [20] traced the LHI concept back to Gini (1914–15).

Settling for the most elementary LHI clearly demonstrates that, in recent UK general elections, FPP has proved *non-PR*. Even nominally PR elections in Britain have barely mediated *semi-PR*. Yet in both Northern Ireland Assembly elections, multi-member Single Transferable Voting has yielded *full PR* of Party First Preferences.

Allowing for vote transferability, STV has also mediated full PR in recent Irish general elections; hardly affected by District Magnitude (between three and five seats per constituency). Likewise in Australia, Alternative Voting has arguably proved semi-PR; and certainly no more disproportional than First-Past-the-Post.

11 References

[1] Balinski, M. and Young, H. P. (2001): *Fair Representation*. Meeting the Ideal of One Man, One Vote. Brookings Institution Press, Washington DC.

[2] Borooah, V. K. (2002): 'The Proportionality of Electoral Systems: Electoral Welfare and Electoral Inequality': *Economics and Politics* **14**, 83–98.

[3] Chief Electoral Officer for Northern Ireland (1998): *The New Northern Ireland Assembly Election 25 June 1998: Election of 6 Members to*

each of the 18 (NI) Westminster Constituencies. Belfast.

[4] Dáil Éireann (1962–2003): *Election Results and Transfer of Votes in General Election*: October 1961 – May 2002. Stationery Office, Dublin.

[5] Electoral Commission (2001): *Election 2001: The Official Results*. London.

[6] Electoral Commission (2003a): *Combination of Gibraltar with a European electoral region in England and Wales*. Consultation paper, June 2003.

[7] Electoral Commission (2003b): *Distribution between electoral regions of UK MEPs*. Recommendation, October 2003.

[8] Electoral Office for Northern Ireland (2004): *Northern Ireland Assembly Elections 26 November 2003 Results*.

[9] Gallagher, M. (1986): 'The Political Consequences of the Electoral System in the Republic of Ireland': *Electoral Studies 5*, 253–75.

[10] Gallagher, M. (1991): 'Proportionality, Disproportionality and Electoral Systems': *Electoral Studies 10*, 33–51.

[11] Gallagher, M, Marsh, M. and Mitchell, P. *eds* (2003): *How Ireland Voted 2002*. Palgrave Macmillan, Basingstoke.

[12] Independent Commission on PR (2004): *Changed Voting Changed Politics*. Lessons of Britain's Experience of PR since 1997. Constitution Unit, School of Public Policy, UCL, London.

[13] Independent Commission on the Voting System (1998): *The Report of the Independent Commission on the Voting System*. Cm 4090–I. The Stationery Office, London.

[14] Lijphart, A. (1994): *Electoral Systems and Party Systems*. Oxford University Press.

[15] Loosemore, J. and Hanby, V. (1971): 'The Theoretical Limits of Maximum Distortion: Some Analytic Expressions for Electoral Systems': *British Journal of Political Science 1*, 467–77.

[16] Mackie, T. T. and Rose, R. (1991): *The International Almanac of Electoral History*. Macmillan, Basingstoke.

[17] Monroe, B. L. (1994): 'Disproportionality and Malapportionment: Measuring Electoral Inequity': *Electoral Studies* **13**, 132–49.

[18] Rae, D. (1967): *The Political Consequences of Electoral Laws*. Yale University Press, New Haven.

[19] Rallings, C. and Thrasher, M. *eds* (2000): *British Electoral Facts 1832–1999*. Ashgate, Aldershot.

[20] Riedwyl, H. and Steiner, J. (1995): 'What is Proportionality Anyhow?': *Comparative Politics* **27**, 357–369.

[21] Taagepera, R and Grofman, B (2003): 'Mapping the Indices of Seats-Votes Disproportionality and Inter-Election Volatility': *Party Politics* **9**, 659–677.

[22] Taagepera, R. and Shugart, M. S. (1989): *Seats and Votes*. The Effects and Determinants of Electoral Systems. Yale University Press, New Haven.

[23] Wichmann, B. A. (2003): 'Proportionality Revisited': *Voting matters* Issue 16, 6–8.

[24] Woodall, D. R. (1986): 'How Proportional is Proportional Representation? *The Mathematical Intelligencer* **8** : 4, 36–46.

Table 4.1: State Population, Seat Apportionment and Relative Bias

(**Bottom/Top third** most populous States), by Divisor Method: House of Representatives, **USA:** 1791 Apportionment; and 1791–2000 **Mean Bias** (22 Apportionments, without seat minima).

State of Union	Population (P)	**Divisor Method: Number of Seats (**S**)** Adams	Dean	Hill	Webster	Jefferson
Total (USA)	3,615,920	105	105	105	105	105
Virginia	630,560	18	18	18	18	19
Massachusetts	475,327	14	14	14	14	14
Pennsylvania	432,879	12	12	12	13	13
North Carolina	353,523	10	10	10	10	10
New York	331,589	10	10	10	10	10
Maryland	278,514	8	8	8	8	8
Connecticut	236,841	7	7	7	7	7
South Carolina	206,236	6	6	6	6	6
New Jersey	179,570	5	5	5	5	5
New Hampshire	141,822	4	4	4	4	4
Vermont	85,533	3	3	3	2	2
Georgia	70,835	2	2	2	2	2
Kentucky	68,705	2	2	2	2	2
Rhode Island	68,446	2	2	2	2	2
Delaware	55,540	2	2	2	2	1
Top third (5)	2,223,878	64	64	64	65	66
Bottom third (5)	349,059	11	11	11	10	9
Seat/ Population fraction (S%/P%)	Top third	0.99	0.99	0.99	1.01	1.02
	Bottom third	1.09	1.09	1.09	0.99	0.89
1791 **Relative Bias**, *percent* *		+10	+10	+10	–2	–13
1791–2000 **Mean Bias**, *percent* *		+20.3	+7.0	+ 5.0	–0.1	–20.7

* **Relative Bias:** Percentage deviation from unity of ratio between Seat/Population (or $S\%/P\%$) ratios of **Bottom/Top third** most populous States.

Data Source: Balinski and Young [1].

Table 4.2: Regional Electors, Seat Apportionment and Relative Bias

(**Bottom/Top third** most populous Regions) and **Malapportionment Index**, by Divisor Method: MEPs, **England**, 1999.

Region	Electors (E)	**Divisor Method: Number of Seats (S)** Adams	Dean	Hill	Webster	Jefferson
Total (England)	37,079,720	71	71	71	71	71
South East	6,023,991	11	11	12	12	12
North West	5,240,321	10	10	10	10	10
London	4,972,495	10	10	9	9	10
Eastern	4,067,524	8	8	8	8	8
West Midlands	4,034,992	8	8	8	8	8
Yorkshire & Humber	3,795,388	7	7	7	7	7
South West	3,775,332	7	7	7	7	7
East Midlands	3,199,711	6	6	6	6	6
North East	1,969,966	4	4	4	4	3
Top third (3)	16,236,807	31	31	31	31	32
Bottom third (3)	8,945,009	17	17	17	17	16
Seat-/Electorate-fraction ($S\%/E\%$)	Top third	0.997	0.997	0.997	0.997	1.029
	Bottom third	0.993	0.993	0.993	0.993	0.934
Relative Bias, *percent* *		–0.46		–0.46		–9.24
Malapportionment Index *(percent)* †	Dean	30.96		30.98		50.01
	Webster	31.22		31.07		45.05

* **Relative Bias:** Percentage deviation from unity of ratio between Seat/Electorate (or $S\%/E\%$) ratios of Regions with **Bottom/Top third** most electors.

† **Malapportionment Index:**
Dean Index *(percent)* $= \sum \mid 1 - E_R\%/S_R\% \mid \times 100$; and
Webster Index *(percent)* $= \sum \mid 1 - S_R\%/E_R\% \mid \times 100$:

Data Source: Electoral Commission [6].

Table 4.3: Analysis of Party Votes and Seats

Number, Fraction and **Loosemore-Hanby Index**: European Election (d'Hondt Regional Closed Party Lists): **Britain**, June 2004.

Party	Number		Fraction, *percent*		Seat–Vote Fraction
	Votes (V_P)	Seats (S_P)	Votes ($V_P\%$)	Seats ($S_P\%$)	**Deviation**, *percent* ($S_P\% - V_P\%$) *
Total (Britain)	16,448,605	75	100.0	100.0	0.0
Conservative	4,397,090	27	26.7	36.0	+9.3
Labour	3,718,683	19	22.6	25.3	+2.7
UK Independence	2,650,768	12	16.1	16.0	–0.1
Liberal Democrat	2,452,327	12	14.9	16.0	+1.1
Green	1,028,283	2	6.3	2.7	–3.6
Scottish National	231,505	2	1.4	2.7	+1.3
Plaid Cymru	159,888	1	1.0	1.3	+0.4
Others (unrepresented)	1,810,061	0	11.0	0.0	–11.0
Over-represented *	10,959,493	61	66.6	81.3	+14.7†
Under-represented	5,489,112	14	33.4	18.7	–14.7

* Over-represented Party $S_P\% > V_P\%$ (under-represented $S_P\% < V_P\%$).

† **Loosemore-Hanby Index** (LHI) = Party total over-representation $= \frac{1}{2}\sum \mid V_P\% - S_P\% \mid =$ **14.7 percent**.

Rose Proportionality Index (RPI) = Complement of Party total over-representation = 100.0 – 14.7 = **85.3 percent**.

Data Source: Guardian, 16 June 2004.

Table 4.4: Correlations between Seven Party Total Disproportionality Indices

UK (FPP: 44 general elections), 1832–2005.
Values as percentages.

Index	LHI	GhI	SLI	MrI	GnI	FrI	BrI
LHI	–	96.4	91.0	97.7	98.1	96.5	91.4
GhI		–	84.8	99.8	94.0	96.4	86.1
SLI			–	86.4	92.3	84.7	99.5
MrI				–	95.4	97.2	87.5
GnI					–	94.7	93.0
FrI						–	85.6
Mean Index	11.5	9.2	11.4	11.2	13.4	11.8	9.6

Data sources: Electoral Commission [5]; Rallings and Thrasher [19] and Guardian, 7 May 2005.

Table 4.5: Loosemore-Hanby Index

Last Nine General Elections in **UK** (FPP), 1974–2005;
and **Australia** (AV), 1983–2004.

UK : Election	**FPP**: **LHI**, *percent*	**Australia:** Election	**AV** Count: **LHI**, *percent* First	Final *
Feb 1974	19.9	1983	15.2	11.2
Oct 1974	19.0	1984	11.8	7.9
1979	15.3	1987	13.6	9.8
1983	24.2	1990	17.1	5.0
1987	20.9	1993	14.1	7.4
1992	18.0	1996	18.8	12.6
1997	21.2	1998	20.5	6.4
2001	22.1	2001	18.2	4.9
2005	20.7	2004	15.8	6.6
1974–2005 Mean	20.1	1983–2004 Mean	16.1	8.0
		(**First + Final**)	(**12.0**)	

* AV Final Count: Two-Candidate Preferred (excluding few non-transferable votes: in Australia, valid voting necessitates rank-ordering all AV preferences).

Data Sources: Rallings and Thrasher [19]; Electoral Commission [5]; and Australian Electoral Commission (personal communications, 1988–2005).

Table 4.6: Loosemore-Hanby Index

By Assembly, Electoral System and Election (Year): **UK**, 1995–2005.

Assembly	**Electoral System**	Year	**LHI**, *percent*
House of Commons (UK MPs)	**FPP**	2001	22.1
	(First-Past-the-Post)	2005	20.7
European Parliament (British MEPs)	**CPL** (Closed Party List:	1999	14.1
	Regional d'Hondt)	2004	14.7
London Assembly	**FPP** + 44% **CPL**	2000	14.8
	(Party List $V_P\% > 5\%$)	2004	13.6
National Assembly for Wales	**FPP** + 33% **CPL**	1999	11.2
	(Regional d'Hondt)	2003	14.1
Scottish Parliament	**FPP** + 43% **CPL**	1999	10.5
	(Regional d'Hondt)	2003	12.5
Northern Ireland Assembly	**STV** (Six Seats	1998	6.0 to 3.8*
	per Constituency)	2003	6.4 to 5.4*

* First to Final count (excluding non-transferable votes).

Data Sources: Chief Electoral Officer for Northern Ireland [3]; Electoral Commission [5]; Electoral Office for Northern Ireland [8]; Rallings and Thrasher [19]; **Guardian**, 6 May 2000, 3 May 2003 and 7 May 2005; **Times**, 12 June 2004.

Table 4.7: National Aggregate Loosemore-Hanby Index

By STV District Magnitude, Count and Election:
Irish Republic, 1961–2002.

Election	**District Magnitude** (Seats per STV Constituency): **LHI**, *percent* (First to Final Count*)			
Year (Month)	Total	3	4	5
1961	8.4 to 3.4	9.4 to 4.5	10.7 to 7.1	9.7 to 4.7
1965	3.2 to 2.3	3.2 to 2.0	6.0 to 5.8	4.2 to 2.1
1969	7.1 to 4.5	7.3 to 4.6	7.5 to 4.5	4.3 to 2.0
1973	4.3 to 1.2	4.5 to 2.4	4.6 to 2.6	7.3 to 8.9
1977	7.4 to 4.1	7.3 to 6.0	9.7 to 4.1	8.5 to 1.1
1981	5.8 to 2.4	4.6 to 2.3	10.2 to 2.6	5.3 to 4.0
1982 (Feb)	3.4 to 1.9	2.6 to 2.0	4.4 to 2.8	4.2 to 1.1
1982 (Nov)	4.2 to 1.9	2.6 to 3.8	7.2 to 3.0	4.7 to 3.4
1987	9.9 to 1.3	10.5 to 7.3	10.9 to 2.8	10.1 to 2.2
1989	7.1 to 2.4	6.0 to 3.9	8.9 to 2.5	7.8 to 2.6
1992	8.2 to 3.7	9.8 to 3.6	10.5 to 5.6	8.5 to 3.9
1997	12.9 to 5.1	14.9 to 6.9	16.2 to 6.7	13.2 to 5.7
2002	12.6 to 6.6	15.8 to 10.4	14.2 to 6.3	11.4 to 5.6
1961–2002 Mean	7.3 to 3.1	7.6 to 4.6	9.3 to 4.3	7.6 to 3.6
(**First + Final**)	(**5.2**)	(**6.1**)	(**6.8**)	(**5.6**)

* Final Count: Excluding non-transferable votes.

Data source: Dáil Éireann (1962–2003).

ISSN 1745-6231

Voting matters

for the technical issues of STV

published by

The McDougall Trust

Issue 21 March 2006

About the McDougall Trust (reg. charity no. 212151)

The McDougall Trust is a charitable trust formed in 1948. The charity's purposes as stated in its governing scheme of 1959 are to advance knowledge of and encourage the study of and research into:

- political or economic science and functions of government and the services provided to the community by public and voluntary organisations; and
- methods of election of and the selection and government of representative organisations whether national, civic, commercial, industrial or social.

The Trust's work includes the maintenance and development of the Lakeman Library for Electoral Studies, a unique research resource, the production and publication of Representation: The Journal of Representative Democracy, and, of course, this publication **Voting matters**, that examines the technical issues of the single transferable vote and related electoral systems.
For further information on the Trust, please contact:

The Secretary,
McDougall Trust,
6 Chancel Street,
London SE1 0UX, UK.
Telephone: +44 (0)20 7620 1080
Facsimile: +44 (0)20 7928 1528
Email: admin@mcdougall.org.uk
Web: www.mcdougall.org.uk

For further information on this publication, please contact B A Wichmann, the Editor at the above address or by email at: `Brian.Wichmann@bcs.org.uk`

Editorial

The delay in producing this issue is due to the lack of material. An issue is produced when about 20 pages of articles are available.

There are 3 papers in this issue:

- Jeff O'Neill: *Fast Algorithms for Counting Ranked Ballots*.

  Many years ago, the speed of undertaking a computer count was an issue. Computers are now fast enough for this not to be a serious concern. This paper shows that comparatively modest changes in the way a program operates can make significant changes to the speed of counting.

- Brian Wichmann: *Changing the Irish STV Rules*.

  The Republic of Ireland has used STV since its independence, but used a counting rule in which the order of the ballot papers could potentially change the result, albeit rather infrequently. This paper considers a change to the Meek rules which is assessed by means of computer simulation.

- Franz Ombler: *Booklet position effects, and two new statistics to gauge voter understanding of the need to rank candidates in preferential elections*.

  The use of STV in New Zealand is a very welcome development. The New Zealand elections randomised the order in which candidates were listed in ballot papers for some elections, but not in an accompanying booklet given to all voters. This paper demonstrates effects of the booklet and proposes measures of voter understanding of the importance of ranking their chosen candidates.

We have an innovation with this issue which is actually some additional material under the heading **Internet Resources** on the McDougall web site. The additional material is in the form of links to papers or references that are being used in *Voting matters* contributions. Hypertext links are typically too long to handle easily by means of printing, and therefore present a problem in producing *Voting matters*. There is also an additional hazard with such links as they can be removed or their position changed. The web site should be able to record changes and record material that has been lost.

Lastly, a report on electronic voting produced by the Irish Commission should be available shortly on their web site at: `http://www.cev.ie/`.

TV voting

There is an increasing use of popular voting associated with TV programmes, which, unfortunately, does not include preferential voting. With a programme like BBC's *Big Read*, one wonders what the result would have been. For instance, if one could (somehow) arrange preferential voting in which is the voters had read the books in their list, how would *War and Peace* have compared with *Harry Potter*?

Fast Algorithms for Counting Ranked Ballots

Jeffrey C. O'Neill
jco8@cornell.edu

1 Introduction

This paper shows how some vote-counting methods can be implemented significantly faster by organizing ranked-ballot data into a tree rather than a list. I will begin by explaining how the tree data structure works and then apply it to Meek's method and Condorcet voting.

2 Tree-Packed Ballots

The most basic way of storing ballots is in a list. For example, suppose Alice, Bob, and Cindy are candidates and we have ten voters. The votes could be stored in a list, where each line corresponds to a ballot, and within each line, the candidates are listed in order of preference. I call this raw or unpacked ballot data, and an example is shown in Figure 1.1.

In this example, as is inevitable in any real election with ranked ballots, some voters will cast the exact same ballot. Instead, one could store only one copy of duplicate ballots along with the number of times the ballot occurred. I call this list-packed ballots. Figure 1.2 shows the same ballots from Figure 1.1 packed into a list.

Many vote-counting methods can use list-packed ballots instead of raw ballots and save computations. For example IRV, ERS97 STV, and Meek's method can all use list-packed ballots but Cambridge and Irish STV cannot. The reason Cambridge and Irish STV cannot is that the outcome is dependent on the order of the ballots, and order information is lost with list-packed ballots.

The ballots, however, can be packed even more densely into a tree, what I call tree-packed ballots. Figure 1.3 shows the same ballots packed into a tree. The root of the tree lists the total number of ballots, which is ten. From the root, branches go downward corresponding to the first-ranked candidates. The subsequent nodes list the number of times that candidate was ranked first on a ballot. Note that these three numbers add up to ten. The second level corresponds to the second-ranked candidates listed after the corresponding first-ranked candidates. Note that no candidate is ever ranked second after Bob. Further, note that four ballots have Alice first, but only two ballots list a candidate second after Alice. This is because two of the four voters who listed Alice first did not rank a candidate second.

For the three data structures, the size of the data struc-

Alice, Cindy
Cindy
Cindy, Alice
Bob
Bob
Alice
Cindy, Alice
Alice
Alice, Bob, Cindy
Bob

Figure 1.1: Raw ballots.

3	Bob
2	Cindy, Alice
2	Alice
1	Cindy
1	Alice, Cindy
1	Alice, Bob, Cindy

Figure 1.2: List-packed ballots.

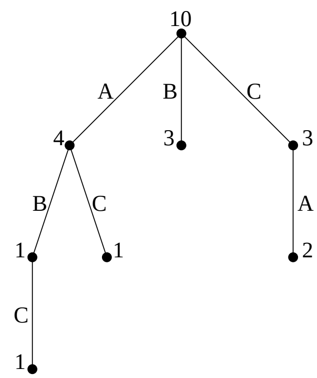


Figure 1.3: Tree-packed ballots.

Data Structure	Time
Load and No Packing	17.7
Load and List Pack	26.7
Load and Tree Pack	31.1

Table 1.1: Comparison of loading and packing times (in seconds).

Method	List	Tree
SNTV	0.6	
IRV	1.2	
ERS97 STV	5.5	
BC STV	4.7	
Meek STV	32.8	5.9 (18%)
Warren STV	30.8	3.0 (10%)
Condorcet	13.3	7.7 (59%)

Table 1.2: Timing of vote-counting methods with list-packed and tree-packed ballots (in seconds). The percentages in parenthesis indicate the computation time of the tree-packed implementation relative to the list-packed implementation.

ture corresponds to the number of entries, which is the number of times that candidate names are stored. For example, the size of the data structure in Figure 1.1 is 15, the size of the data structure in Figure 1.2 is 10, and the size of the data structure in Figure 1.3 is 7 (the root node isn't counted). Table 1.3 shows the sizes of the three data structures for the ballots from eight elections. B is the number of ballots, C is the number of candidates, and S is the number of seats to be filled.

List-packed ballots are 65% of the size of raw ballots. Tree-packed ballots are 45% of the size of list-packed ballots and 29% of the size of raw ballots. I expect the computation time of a particular implementation to be roughly proportional to the size of the data structure used. Thus, I expect the computation time with tree-packed ballots to be about 45% of the computation time with list-packed ballots. The more complicated data structures will also add some overhead that will increase the computation time to some extent.

Before presenting the details of implementing vote-counting methods with the different data structures, I will present the timing results with the different data structures. The timing results should only be considered in a rough sense since the efficiency of the particular implementations may vary. All timing results are cumulative for the above eight elections and are in seconds. First, the times in seconds for loading, loading and list packing, and loading and tree-packing are shown in Table 1.1.

Next I compare the computation times for a number of vote-counting methods using list-packed and tree-packed ballots. Because the relationship between raw and list-packed ballots is obvious, those times are not compared in this paper. [1] Further, only the slower methods are implemented with tree-packed ballots because these are the only ones that are in need of improvement. The methods are single non-transferable vote (SNTV), instant runoff voting (IRV), Electoral Reform Society STV (ERS97 STV), STV rules proposed for British Columbia in 2005 (BC STV), Meek STV, Warren STV, and Condorcet. [2] The computation times are shown in seconds in Table 1.2. The percentages in parentheses indicate the computation time of the tree-packed implementation relative to the list-packed implementation.

While we expected the computation times with tree-packed ballots to be 45% of the times for list-packed ballots, they are much faster for Meek and Warren STV. Why this is so will be explained below.

[1] Implementing a particular method with raw or list-packed ballots uses nearly the same code. The code iterates over the raw ballots or iterates over the list-packed ballots. The computation time is simply proportional to the number of loop iterations. In contrast, with tree-packed ballots, the code needs to be rewritten from scratch as is discussed below.

[2] The timing for Condorcet is only for computing the pairwise comparison matrix. Computing the Condorcet winner from the pairwise comparison matrix is generally much faster than computing the pairwise computation matrix.

3 Meek STV with Tree-Packed Ballots

I will now give the details of how to implement Meek STV using tree-packed ballots. The process is very similar for Warren STV. A full description of Meek STV is beyond the scope of this paper [1, 2, 3]. Instead, I will present the details most relevant to the fast implementation.

In each stage of counting votes with Meek STV, all the votes must be counted from scratch. This is distinct from other STV methods where some votes are simply transferred from one candidate to another and a full recount is not necessary at each round. With Meek STV, each candidate is assigned a fraction, f[c], where c denotes the candidate. At the beginning of the count, all the fractions are 1.0, and the fractions remain 1.0 as long as a candidate is under the quota. When a candidate has more than a quota, the fraction essentially discounts the value of that candidate's votes to bring the candidate back down to a quota. With a discount less than 1.0, the subsequently ranked candidates on a ballot will receive a portion of the vote.

In each round of a Meek STV count, the fractions f[c] will be updated and the ballots recounted. The following is a segment of Python pseudo-code for counting ballots for one round of a Meek count. Note that it uses list-packed ballots. The ith packed ballot is b.packed[i] and the corresponding weight of that packed ballot is b.weight[i].

```
# Iterate over all of the ballots.
for i in range(nBallots):
  # Each ballot is worth one vote.
  remainder = 1.0
  # Iterate over the candidates on this ballot.
  for c in b.packed[i]:
    # If the candidate is already eliminated
    # then skip to the next candidate on the
    # ballot.
    if c in losers:
      continue
    # This candidate gets a portion
    # of this ballot.  For the first non-losing
    # candidate on the ballot, the remainder will
    # be 1.0. If the candidate is under quota,
    # then f[c] is also 1.0 and this candidate
    # gets all of the ballot.  Otherwise the
    # candidate gets less than the full value,
    # and will share the ballot with
    # subsequently ranked candidates.
    count[c] += remainder * f[c] * b.weight[i]
    # Calculate how much of this ballot remains,
    # if any, to be counted for subsequently
    # ranked candidates.
    remainder *= 1 - f[c]
    # Stop if this ballot is used up.
    if remainder == 0:
      break
```

This code can be rewritten to use tree-packed ballots. The computations are exactly the same as before, they are just done in a different order so that similar computations can be done together. Consider the ten ballots presented above. Alice is ranked first on four ballots. With list-packed ballots, it would take three loop iterations to count these three ballots, but with tree-packed ballots all the first-place votes for Alice are counted at the same time, thus saving computations.

The code is more complicated, because it involves a depth-first traversal of the tree. The following shows how the nodes of the tree are accessed and also the order of a depth-first traversal.

```
tree[n] = 10
tree[Alice][n] = 4
tree[Alice][Bob][n] = 1
tree[Alice][Bob][Cindy][n] = 1
tree[Alice][Cindy][n] = 1
tree[Bob][n] = 3
tree[Cindy][n] = 3
tree[Cindy][Alice][n] = 2
```

A convenient way to implement the depth-first traversal is to use a recursive subroutine. Note that the subroutine calls itself by passing one branch of the tree, which is just a smaller tree, and possibly a diminished value for the remainder.

```
def updateCountMeek(tree, remainder):
  # Iterate over the next possible candidates.
  for c in tree.nextCands():
    # Copy the remainder for each iteration.
    rrr = remainder
    # Skip over losing candidates.
    if c not in losers:
      # Count the votes as before but weight with
      # the tree-packed data instead of the
      # list-packed data.
      count[c] += rrr * f[c] * tree[c][n]
      # Calculate how much of this ballot remains,
      # if any, to be counted for subsequently
      # ranked candidates.
      rrr *= 1 - f[c]
    # If there are any candidates ranked after
    # the current one and this ballot is not used
    # up, then recursively repeat this procedure.
    if tree[c].nextCands() != [] and rrr > 0:
      updateCountMeek(tree[c], rrr)
```

The initial call to the subroutine uses the base of the tree, and as before, the initial value of the remainder is 1.0

```
updateCountMeek(self.b.tree, 1.0)
```

Now that I have explained the fast algorithm, I can explain why it works much faster than expected. The unexpected speed increase arises from the fact that in any STV election, it is overwhelmingly the top choices on the ballots that are counted. In the first round of a Meek election, only the first-ranked candidates are counted. Consider the ballots for the Dublin North 2002 election. With list-packed ballots, one needs to count the 138,647 weighted ballots, but with tree-packed ballots, one needs to count only the twelve nodes of the tree corresponding to the first rankings of the twelve candidates. As the rounds progress, more and more nodes in the tree will be needed for the count, but generally this will be far less than the total number of nodes in the tree and even further less than the number of list-packed ballots.

Readers who understand the differences between Meek STV and Warren STV will immediately realize why Warren STV is much faster than Meek STV with the tree-packed ballots: Warren STV is less likely than Meek STV to use lower-ranked choices on a ballot.

4 Condorcet with Tree-Packed Ballots

Tree-packed ballots can also be used to compute the pairwise comparison matrix in a Condorcet election. The pairwise comparison matrix, pMat[c][d], counts the number of times that candidate c is ranked higher than candidate d on the ballots. Computing the pairwise comparison matrix is straightforward with list-packed ballots:

```
# Iterate over all the ballots.
for i in range(nBallots):
  # Copy the list of candidates.
  remainingC = candidates[:]
  # Iterate over the candidates the ballot.
  for c in b.packed[i]:
    # Get list of lower-ranked candidates.
    remainingC.remove(c)
    # Iterate over all lower-ranked candidates.
    for d in remainingC:
      # c is ranked higher than d.
      pMat[c][d] += b.weight[i]
```

This code can also be rewritten to use tree-packed ballots. As before it involves the depth-first traversal of the tree.

```
def ComputePMat(tree, remainingC):
  # remainingC is a list of candidates not higher in
  # the ballot than the current candidate.  Initially
  # it is a list of all the candidates.
  # Iterate over the next possible candidates.
  for c in tree.nextCands():
    # Copy the list of remaining candidates.
    rc = remainingC
    # Remove candidate from remaining list.
    rc.remove(c)
    for d in rc:
      # Current candidate is ranked higher than
      # candidates in remaining list.
      pMat[c][d] += tree[c][n]
    # Continue if more candidates.
    if tree[c].nextCands() != []:
      ComputePMat(tree[c], rc)

# First call is with entire tree and list of all
# candidates.
ComputePMat(tree, allCands)
```

Computing the pairwise comparison matrix is faster with tree-packed ballots, but the improvement is not nearly as great as for Meek STV. The reason for this is that computing the pairwise comparison matrix requires traversing the entire tree, thus the computation times are roughly proportional to the relative sizes of the data structures. The overhead involved with using tree-packed ballots makes the implementation with tree-packed ballots a little slower than expected.

5 Conclusions

Using tree-packed ballots instead of other data structures can greatly increase the speed of some vote-counting methods. Such speed improvements need to be weighed against the time needed to create the tree-packed ballots and the cost of maintaining more complex code. Meek and Warren STV are approximately five and ten times faster, respectively, with tree-packed ballots than with list-packed ballots. Such enormous speed improvements clearly outweigh the costs. In contrast, with Condorcet voting, the time saved is about equal to the time required for tree-packing the ballots so any benefits are minimal. Other methods, such as ERS97 STV and BC STV, are so fast with list-packed ballots that tree-packed ballots are clearly not beneficial.

My implementation of all of the vote counting methods mentioned in this paper (and others) is available for download at `http://stv.sourceforge.net.`

6 References

[1] Nicolaus Tideman, *The Single Transferable Vote*, Journal of Economic Perspectives, Vol. 9, No. 1, pp. 27-38 (1995).

[2] B. L. Meek, *A New Approach to the Single Transferable Vote*, Voting matters, Issue 1, p. 1 (Mar. 1994).

[3] I. D. Hill, B. A. Wichmann, and D. R. Woodall, *Algorithm 123 – Single Transferable Vote by Meek's Method*, Computer Journal, Vol. 30, No. 3, pp. 277-281 (1987).

Election	B/C/S	Raw	List	Tree
Dublin North 2002	43,942/12/4	218,933	138,647	57,568
Dublin West 2002	29,988/9/3	132,726	69,860	23,730
Meath 2002	64,081/14/5	298,106	174,737	74,105
Cambridge 1999 City Council	18,777/29/9	106,585	90,816	47,813
Cambridge 2001 City Council	17,126/28/9	95,440	79,385	40,566
Cambridge 2001 School Committee	16,489/16/6	66,254	33,860	12,907
Cambridge 2003 City Council	20,080/29/9	115,232	98,055	54,182
Cambridge 2003 School Committee	18,698/14/6	66,389	29,637	9,764
Total		1,099,665	714,997	320,635

B/C/S = Ballots/Candidates/Seats

Table 1.3: Sizes of the three data structures for the eight elections. The size of a data structure is the number of entries. See the text for more details.

Changing the Irish STV Rules

Brian Wichmann
brian.wichmann@bcs.org.uk

1 Introduction

For elections to the Dáil, the Irish Government has been using a form of STV which has remained essentially unchanged since the state was formed, in spite of small adjustments [1]. The counting rules have a significant flaw: they use a method of transferring surpluses that makes a random choice of the votes to be transferred [2]. Specifically, the rules require that the papers are placed in a random order. When a transfer is undertaken, all the relevant papers are examined in order to determine how many of them should be transferred to each continuing candidate, but the actual papers chosen for transfer depend on the random order. This method can affect the result if transferred papers are transferred again later in the count.

With the advent of computer-based counting (which is likely to be introduced shortly), the dependence upon the (random) order of the papers will become apparent. In the case of the three constituencies for which computer-based counting was used in 2002, the full ballot data was placed on the Internet (with the papers ordered as for the official count). In those three cases, the results were not order dependent, but order-dependence is bound to arise at some stage in the future. If a candidate could have been elected but was not, it is clear that a legal challenge to the result would be possible (especially if, considering all possible random orders of the papers, the aggrieved candidate was more likely to be elected than one of the candidates who actually was!).

This paper presents a study of the likely effect of changing the STV Rules for the Dáil to use the Meek method [3]. As with all modern counting rules, the Meek method has no order-dependence.

2 A method for simulating Irish voting patterns

For three Dáil elections held in 2002 we have the complete ballot data as noted above. This implies that many forms of analysis can be undertaken, for instance, the use of preferences as below:

Constituency	*Average used (Meek)*	*Average used (Irish)*	*Average given*	*Seats/ Candidates*
Dublin North	2.12	1.34	4.98	4/12
Dublin West	2.11	1.49	4.43	3/9
Meath	1.98	1.43	4.65	5/14

Here we use the data in another way. A previous paper [4] describes a way of generating simulated ballot data from a conventional STV result sheet using a simple statistical technique [5]. We wish to tailor this method to Irish voting patterns, which we can do by making the simulated ballot data more closely resemble the actual ballot data in the three Dáil elections for which the latter are known. To that end, the following changes have been made to the method described in [4]:

1. a proportion of the papers with only one or two preferences are ignored, since otherwise there would be too many such papers;
2. an appropriate proportion is added of strict party votes — all the preferences being for one party;
3. additional votes are added in which the final preferences are in ballot paper (or reverse) order because such are observed in the actual data. This is done by taking some of the generated papers which listed between a half and three quarters of the available candidates and inserting the remaining candidates;
4. for those candidates having a very small number of first preference votes, there is an adjustment to

ensure that the number of second preferences for them is also low.

The best possible outcome would be if the generated papers looked as if they came from the same population as the actual papers for the three constituencies. If fact, the results were as follows:

First preference test. This compares the distribution of first preferences for the actual and generated papers. The program construction should ensure that this test passes.

First two preferences test. Each pair of candidates is considered and also each candidate singly where no second preference is expressed. For the pairs the order of the two candidates is disregarded, counts for AB and BA being put together. The distributions formed from the actual and generated papers are then compared. It is not very surprising that this test fails because much of the necessary information about the relationships between candidates is missing in result sheets, and hence the generator's random selection will not produce a good fit. For Dublin North, for instance, the Labour and Green Party candidates appear to have a common following giving a high count to papers containing these as the first two choices. The result sheet for this election shows the high transfers at count 7 from the (elected) Green candidate to Labour, but does not show the reverse. In general, so many of the second preferences are unknown that the test cannot be expected to perform well.

Length test. This test considers the distribution of the number of preferences specified. Those that specify every candidate, and those that specify every candidate except one, are merged as their meanings are regarded as identical. This test is not passed, but does not fail so badly as to indicate a need to modify the program.

Rank test. This considers the ranking of the candidates against the ballot paper order. It passes with one of the three constituencies, and does not appear to warrant further program modification.

It is clear that the three available constituencies have different statistical properties, not all of which can be related to the differing numbers of seats (3, 4 and 5). Hence, the generator cannot be expected to obtain a good match for all of them. It is thought that any further change to the generator would be unlikely to make much improvement.

3 Generating data to match two Dáil elections

For each of the constituencies for the 1992 and 1997 Dáil elections, the result sheet is used, together with the generator described in the previous section, to produce three (related) sets, making 246 in total. The total number of candidates to be elected was 993. This ballot data could then be processed using the Irish rules and Meek. The observed differences were in 17 constituencies, 16 giving a difference of one candidate and one a difference of two. Hence the differences were in 1.8% of the candidates elected. (The difference in candidates was 18/993, while that in constituencies was 17/246, but the former is taken since that is the number which influences the Dáil.)

In all of the 17 constituencies, on completing the count with both rules, there was only one continuing candidate. In 13 of these, the set of those elected plus the continuing candidate was the same — the difference between the two rules was in the choice of the last candidate to elect.

We now need to consider ways of determining what should be the 'correct' result for these 17 cases. Two general methods are considered:

Order-dependence. We need to consider whether the Irish count was influenced in the final outcome by the order of the ballot papers. The papers were initially in random order and hence would not be expected to favour a specific candidate.

In theory, it should be possible to compute the probability of each possible outcome from the ballot papers. However, this seems rather difficult and hence the approach taken is to determine the two candidates whose position is different with the two rules. A program is then used to re-order the papers to favour the Meek outcome. Then the Irish rules are applied to the re-ordered papers to see if a different result is obtained. If a different result is produced, then it is clear that the papers *are* order-dependent, even if the probabilities of the different outcomes are not known. However, if the same result is produced, it is not possible to be sure that there is no order-dependence in the result, unless transferred surplus votes are not subsequently transferred again.

If the papers are order-dependent, then the Irish result is certainly questionable. In all such cases,

Test	*Seats*	*Withdrawn test*		*Order*
		Cands.	*Result*	*Depend.*
92/P19A	4	5	Meek	Yes
92/P22A	4	6	Irish	No?
92/P22B	4	6	Irish	No?
92/P23A	4	5	Meek	Yes
92/P24B	5	6	Meek	Yes
92/P24C	5	6	Irish	Yes
92/P26C	4	6	Meek	Yes
92/P27C	5	6	Meek	Yes
92/P35A	5	6	Meek	No?
92/P35B	5	6	Meek	Yes
92/P35C	5	7	Irish	Yes
92/P43A	4	5	Meek	No?
92/P43B	4	5	Meek	No?
97/P18C	3	4	Meek	No
97/P35B	3	4	Meek	No
97/P46B	4	5	Meek	No
97/P46C	4	5	Meek	No

Table 2.1: The differences analysed

reordering the papers can produce the Meek result.

Withdraw no-hopers. All the candidates who were neither elected nor a continuing candidate with either rule can be considered as having no hope of election. Under such circumstances, with STV, it is reasonable to assume that withdrawing these no-hopers from the count would not change the result. With the Meek rules, we know that this test *will* produce the same result, but the Irish result is uncertain. In the 17 cases under consideration, when running the Irish rules (with the papers in the same order), the result is either as with the original election, or else changes to the Meek result, as indicated in the Table 2.1.

In Table 2.1, the 6 cases in which the *withdrawn* test gives the Meek result and where there is also order-dependence, we regard as showing that the Meek result is superior. This leaves another 11 cases to consider in more depth.

The last four results in Table 2.1 are *not* order-dependent because the votes transferred after a surplus are not subsequently transferred. It is instructive to consider the first one of these further. The first stages of both Meek and the Irish rules are to exclude the five no-hopers. Hence, after these exclusions, the votes for the remaining five candidates are the same for both rules. (The stages are out of step as the Irish rules exclude two in one stage, while Meek rules do not.) The *withdrawn* test shows that if the Irish rules were applied starting from this point, then the Meek result would have been produced. However, the two actual outcomes can be summarised in Table 2.2.

Candidate	*Meek, Stage 6* *Irish, Stage 5*	*Meek, Stage 7* *Irish, Stage 6*	*Result*
C1	7241	7621	Elected
	7241	7317	
C3	7875	7614	Elected
	7875	7939	Elected
C5	7411	7592	
	7411	7472	Elected
C8	8316	7614	Elected
	8316	8111	Elected

Table 2.2: Test 97/P18C Analysis

(Meek results rounded to integers.)

With the Irish rules, since the quota is calculated once at the start, C8 is elected with 639 (8111-7472) more votes than C5. The reduced quota with Meek means that many more of those people who voted first for C8 had a fraction of their vote transferred to their next preference. Moreover the 205 votes that were transferred from C8 all came from the excluded candidate C6. With Meek, all the votes for C8 are considered and an appropriate fraction retained while the rest of the votes are passed to the next preference. In our opinion, Meek can be seen to be fairer, although it requires more work to examine each vote at each stage.

All the other three cases for 1997 are similar.

We now consider the case 92/P24C in which the *withdrawn* test still produces the Irish result but we know that reordering the papers can produce the Meek result. Also, the *withdrawn* test is very simple in that only one candidate needs to be excluded. We give the result sheet for each rule in Tables 2.3 and 2.4. The elected candidates are in italics and underlined.

Comparing these two result sheets reveals the key differences as follows:

1. at the second stage, the Irish rules transfer the surplus of C2, while Meek transfers the surpluses of C1, C2 and C6. With the Irish rules, the surplus of C6 is never transferred;

C1	11156	11156	11156	−1463 9693
C2	16715	−7022 9693	9693	9693
C3	9076	+2668 11744	−2051 9693	9693
C4	6945	+1838 8783	+402 9185	+225 9410
C5	4532	+2516 7048	+1076 8124	+1238 9362
C6	9732	9732	9732	9732
Non-T	—	—	+573 573	573
Totals	58156	58156	58156	58156

Quota is 9693.

Table 2.3: Test 92/P24C, Irish rules

C1	4126	+256 4382	+827 5209	+1047 6256	+243 6499
C2	4695	+191 4886	+167 5053	−5053 —	—
C3	6081	+1019 7100	+208 7308	+1120 8428	−693 7735
C4	9075	9075	−1340 7735	7735	7735
C5	5320	+172 5492	+138 5630	+820 6450	+170 6620
C6	9373	−1638 7735	7735	7735	7735
Non-T	—	—	—	+2066 2066	+280 2346
Totals	38670	38670	38670	38670	38670

Quota is 7735.

Table 2.5: Test 92/P22A, Irish rules

C1	11156	10692	9017
C2	16715	9732	9005
C3	9076	10832	9020
C4	6945	7983	8906
C5	4532	6142	9002
C6	9732	11121	9011
Non-T	—	1654	4195
Totals	58156	58156	58156
Quota	9693	9417	8993

Table 2.4: Test 92/P24C, Meek rules

C1	4126	5084	5821	6997
C2	4695	5008	—	—
C3	6081	7129	7985	7070
C4	9075	7649	8178	7040
C5	5320	5587	6291	6790
C6	9373	7650	8207	7059
Non-T	—	563	2188	3714
Totals	38670	38670	38670	38670
Quota	7734	7621	7296	6991

Table 2.6: Test 92/P22A, Meek rules

2. the quota reduction of 700 votes with Meek is much larger than the difference of only 48 votes between the last two candidates (C4 and C5) under the Irish rules;

3. the number of non-transferable votes is very much larger with Meek. The reason for this is that all votes are treated the same way, while the Irish rules only transfer votes which have subsequent preferences specified (given that there are sufficient votes to do this). Some people might see this as a weakness of the Meek method, but for an opposing view, that it is a good feature of the method, see [6]— this point is considered further later.

With the possible exception of the issue of handling of non-transferable papers, the Meek result cannot be criticized, while the obvious imperfections in the Irish rules gives cause to doubt the result.

We now consider case 92/P22A (92/P22B is essentially the same). Again, for simplicity, we consider the *withdrawn* test rather than the full election. The two result sheets are presented in Tables 2.5 and 2.6.

It would be reasonable to ask why a further simplification could not be made by removing candidate C2, excluded by both rules. C2 is there as the continuing candidate with the Irish rules for the full election. Hence the candidate cannot be regarded as a no-hoper.

One can analyse the Irish results for evidence of order-dependence. The 191 and then 167 votes transferred to C2 are then transferred again and thus depend upon the choice of votes made. This total of 358 is greater than the 121 vote-difference between the last two candidates (C1 and C5). Hence the question mark remains: it might be possible to obtain the Meek result by a suitable re-ordering.

The number of non-transferable votes is high in both cases. Meek can compensate for this by reducing the quota, while with the Irish rules, an excessive number of papers remain with the three leading candidates. This excess amounts to about 2,000 votes, while the key difference is that C1 leads C5 by 207 votes with Meek, but by C5 leads C1 by 121 votes with the Irish rules.

Hence the primary source of the difference is the high number of non-transferable votes arising when C2 is excluded. The Meek logic is clearly superior in this case.

The three cases 92/P35A, 92/P43A and 92/P43B are all similar in having a weak order-dependence which cannot change the result by re-ordering the papers. However, in all these cases, the *withdrawn* test gives the Meek result. It is regrettable when the presence of a no-hope candidate changes an election result.

The last case, 92/P35C, is the most extreme since the closeness of the voting and the difference in the rules gives a difference of two seats. This is also exhibited by the election with the no-hopers removed, which is shown in Tables 2.7 and 2.8.

The order-dependence in this case arises from the 162 and 35 votes transferred to C3 which are subsequently transferred again and hence are subject to random sampling. However, an attempt to obtain a different result by changing the order failed (with the no-hopers removed), in spite of the original election being order-dependent (see Table 2.1).

The striking difference is that the Irish rules exclude C3 whom Meek rules eventually elect. However, the choice between C3 and C4 is close with both rules — 7 votes in favour of C3 for the Irish rules against 1 in favour of C4 with Meek. The quota reduction undertaken by Meek is enough to make the change, although this is again a consequence of the short lists logic.

C1	5407	+1264 6671	+269 6940	+1075 8015	+140 8155
C2	12008	−3158 8850	8850	8850	8850
C3	6304	+162 6466	+35 6501	−6501 —	—
C4	6290	+178 6468	+40 6508	+2558 9066	−216 8850
C5	7312	+159 7471	+33 7504	+613 8117	+76 8193
C6	9489	9489	−639 8850	8850	8850
C7	6288	+1395 7683	+262 7945	+934 8879	8879
Non-T	—	—	—	+1321 1321	1321
Totals	53098	53098	53098	53098	53098

Quota is 8850.

Table 2.7: Test 92/P35C, Irish rules

C1	5407	6846	7595	8041	8532
C2	12008	8796	9227	8756	8560
C3	6304	6497	8950	8678	8543
C4	6290	6496	—	—	—
C5	7312	7495	8131	8223	8324
C6	9489	8796	9307	8793	8569
C7	6288	7850	8458	8907	8577
Non-T	—	322	1430	1700	1993
Totals	53098	53098	53098	53098	53098
Quota	8850	8796	8611	8566	8517

Table 2.8: Test 92/P35C, Meek rules

4 Conclusions

It is possible to generate ballot data based upon Irish result sheets which is sufficiently similar to actual data to give a basis for comparing two counting rules. The analysis of the Irish rules shows that order-dependence is a significant problem, confirming the result in [2].

The Meek counting rule overcomes the order-dependence, as do all the modern counting rules (such as the Gregory rules used in Northern Ireland).

The analysis here shows that the property of Meek that the exclusion of no-hope candidates is the same as if those candidates had never entered the election is also important. Surely the intervention of such candidates should not influence the result? Other commonly used counting rules do not have this property.

The analysis also reveals that Meek usually has a much higher number of non-transferable papers than the Irish rules. It is the author's view that Meek is correct in this regard since every vote is handled in an identical fashion, while in the Irish rules (as with most of the hand-counting rules), the logic is dependent upon the other votes. This can easily have the effect of totally ignoring the wishes of those votes which gave few preferences in the sense that no transfer to non-transferables is undertaken. Whatever the reader might conclude on this point, this is a smaller effect than those arising from order-dependence and the influence of no-hope candidates noted above.

Although the difference in those elected is quite small (1.8% of the candidates elected), such a difference could be critical in the Dáil. The two major parties are frequently very nearly tied, so that the proportion of seats to them is critical in the formation of a Government. An actual counting error of 1.8% would be correctly regarded as quite unacceptable.

It might be maintained that the 'complexity' of using the Meek algorithm is not justified in view of the small differences observed in this analysis. However, in Ireland, when computers are being used, the complexity is not what it seems. An implementation of the Irish rules in Java amounts to around 2,000 lines of code [7], while the author's implementation of Meek in Ada is less than half that. There are a lot of exceptional cases in the Irish rules but virtually none in the Meek rules.

5 Acknowledgements

The paper is based upon a joint work with David Hill [8].

A significant fraction of this work would not have been possible without the ability to run a program of Joe Otten that implements the Irish rules [1].

6 References

[1] Electoral Act 1992 as amended by the Electoral (Amendment) Act 2001. Republic of Ireland.

[2] M. Gallagher and A. R. Unwin. Electoral Distortion under STV Random Sampling Procedures. *B.J.Pol.S.* Vol 16, pp243-268. 1986.

[3] I. D. Hill, B. A. Wichmann and D. R. Woodall. Algorithm 123 — Single Transferable Vote by Meek's method. *Computer Journal*. Vol 30, pp277-281, 1987.

[4] B. A. Wichmann. Producing plausible party election data, *Voting matters* Issue 5 pp. 6-10, January 1996.

[5] B. A. Wichmann. A simple model of voter behaviour, *Voting matters*. Issue 4. pp3-5. August 1995.

[6] I D Hill. Are non-transferables bad? *Voting matters*, Issue 8. p4. May 1997.

[7] Secrecy, Accuracy and Testing of the Chosen Electronic Voting System. Commission on Electronic Voting (Ireland). December 2004. (Appendix 2E gives details of the 'CD' implementation — the author has been told this is 2,000 lines.) See: http://www.cev.ie/

[8] David Hill and Brian Wichmann. STV in the Republic of Ireland. December 2003.

Booklet position effects, and two new statistics to gauge voter understanding of the need to rank candidates in preferential elections

Franz Ombler
franz@franzo.com

1 Introduction

In 2004 the Single Transferable Vote (STV) method replaced plurality for the election of members of New Zealand's District Health Boards (DHBs) [1]. While being unable to assess ballot position effects due to unrecorded random ordering of candidates' names on each ballot paper this article demonstrates effects that may be explained by the order of candidates' names in an accompanying booklet of the candidates' profiles. Such effects undermine the intended benefits from randomly ordering candidates' names on ballot papers, but prove useful in questioning voter understanding of the need to rank candidates. Two new statistics are proposed to better gauge voter understanding of a preferential voting method: the percentage of plurality style informal ballots and a rank indifferent percentage.

2 The elections

Two elections are considered: the Canterbury DHB election and the Otago DHB election. In both cases seven candidates were to be elected. Ballot papers were sent to voters by post. The ballots for the DHBs were printed with candidates' names randomly ordered such that each ballot paper might be unique. An accompanying booklet with candidates' profiles listed the candidates alphabetically [2]. It seems likely that few candidates for the elections were previously known to voters and the election would seem relatively non-partisan. Voters were allowed to rank order any number of candidates and a ballot was deemed informal if there was no 'unique first preference' indicated on the ballot [3].

2.1 Canterbury

The Canterbury DHB election was run alongside other territorial elections including those for the Christchurch City Council mayor, ward councillors and Canterbury Regional Council. These other elections continued to use plurality, so the voter had to contend with two methods in their ballot papers. There were 29 candidates. Of 117,852 non-blank ballots, 8,986 (7.6%) were deemed 'informal' and removed from the count. Of these, 7,579 (84.0% of informal votes, or 6.4% of total votes) marked all of the candidates for whom they voted as a first preference (either with a tick, or by writing '1'), presumably unaware of the need to rank candidates and thus voting as if it were a plurality election.

2.2 Otago

The Otago DHB election was run alongside territorial elections like those for Canterbury, but all elections were conducted using STV. There were 26 candidates. Of 65,389 non-blank ballots, 3,016 (4.6%) were deemed 'informal' and removed from the count. Of these, 1,315 (43.6% of informal votes, or 2.0% of total votes) marked candidates as if it were a plurality election.

As can be seen from the second-last row of Table 3.1, Canterbury DHB voters were over three times more likely to waste their vote by treating the election as a plurality election (6.4% versus 2.0%). This is probably because the Otago DHB election voters were more familiar with STV due to its use for all the elections on the Otago ballot papers. To better gauge voter understanding of preferential elections the percentage of plurality style informal ballots could be reported alongside the more usually reported total number of informal ballots.

	Canterbury	Otago
Number of seats	7	7
Candidates	29	26
Non-blank ballots	117,852	65,389
Formal ballots	108,866	62,373
Informal ballots	8,986 (7.6%)	3,016 (4.6%)
Informal ballots with multiple first preferences only (plurality-style)	7,579 (6.4%)	1,315 (2.0%)
Rank indifferent (see below)	5.1%	2.9%

Table 3.1: The Canterbury and Otago DHB elections

3 Ballot position effects

The voter burden of ordering the candidates is higher when the candidates are unfamiliar to voters, when there are so many candidates (29 for Canterbury, 26 for Otago), and where the district magnitude is high (seven) [4]. Furthermore, due to the lack of familiarity with candidates, position effects are probably greater [4], and these effects have greater consequences when voters are required to rank order candidates [5]. These effects may also be expected to be amplified by voters' lack of experience in rank ordering candidates, especially when they have to contend with multiple methods on their ballot papers as in the Canterbury election.

Candidates' names were randomly ordered during ballot paper printing, presumably to prevent ballot position effects, that is, where the positions of the candidates' names on the ballot affect voters' selection or ranking of the candidates. Randomising candidate name order should certainly have reduced the effect of 'donkey votes': ballots in which the voter ranked all the candidates in the order in which they appeared on the ballot. However, the number of donkey votes cannot be assessed due to the absence of information as to the order in which the candidates were listed on each ballot sheet. For the same reason, other ballot position effects cannot be assessed either.

4 'Booklet position effects'

Due to voters' lack of familiarity with the candidates many voters would have relied heavily on the booklet of candidates' profiles to draft their selections and rankings. The booklet listed the candidates alphabetically. We might call ensuing effects 'booklet position effects', which will dilute the intended benefits from randomly ordering the candidates' names on ballot papers; indeed it is interesting to consider (although not demonstrated here) whether booklet position effects may be greater than ballot position effects in elections in which voters are less familiar with the candidates. Certainly, the cost-effectiveness of randomising ballot paper candidate name order is questionable if the order of candidates' profiles in an accompanying booklet is not also randomised.

Assigning the candidates numbers according to their positions in the booklet (alphabetically) helps compare the rankings of candidates on each ballot with the order in which they appear in the booklet. The real ballot '2 10 14 17 19 24 26', where this voter has ranked candidate number 2 first (that is, they wrote the number one beside the candidate who appeared second in the booklet), candidate number 10 second and so on, may be described as perfectly ordered as it lists the candidates in the same order in which they appeared in the booklet. Similarly, '9 6 14 19 21 24 27' seems near perfectly ordered.

Spearman's rank correlation coefficient (r_s) may be used to assess the correlation of two rankings. We can apply this to each ballot, finding the r_s of the rankings of candidates in the ballot and the same ballot with candidates re-ordered alphabetically. For example, the r_s of the ballot '2 10 14 17 19 24 26' with its ordered self (the same ballot) is exactly 1.0, showing a perfect positive correlation; while r_s for '9 6 14 19 21 24 27' and its ordered self ('6 9 14 19 21 24 27') is 0.96.

The average r_s of each formal ballot's ranking of candidates with its ordered self is only 0.06 for Canterbury and 0.03 for Otago, showing such weak positive correlations that one might be tempted to infer an absence of booklet position effects. This is likely to draw criticism that it proves nothing due to 'failure to randomly assign groups of voters to different name orders' [4]. Indeed it would be consistent with this bare analysis to claim that position effects were present to a large degree and that if the booklets had been printed randomly that we would have seen a lower average r_s. This might be true to some extent but we are unable to assess it properly due to the absence of information about the order of names on each ballot; however, even without this information, booklet position effects can be demonstrated.

If we assess the frequency of the various values of r_s for the ballots, we find inordinately high numbers of perfectly ordered and near perfectly ordered ballots. Figure 3.1 (the data for which is presented in Table 3.2) shows such an analysis of the 51,730 ballots that listed exactly seven candidates in the Canterbury election. In light grey is the exact distribution of r_s for N=7, as would be approximated by randomly ordering these same ballots. Clearly there is a heavy tail on the right for the real ballots. Focussing on the rightmost bar, these 1,286 ballots (2.49%) are listed perfectly in order, but the expected number of ballots to be found in order for these 51,730 voters is only ten (0.02%) if preferences are randomly distributed.

Analyses of ballots listing other numbers of candidates (but more than 1) also find a notably higher than expected number of perfectly ordered ballots, 2,962 more than expected in total (see Table 3.3).

The Otago DHB election shows a similar but less prominent pattern (Figure 3.2). Given the similarity of the elections in other respects, this difference might be best explained by the use of STV in all of the elections on the Otago ballot papers and therefore greater voter awareness and understanding of the method.

Booklet position effects are apparent, but there are other potential explanations. It is conceivable that some voters are strongly biased towards candidates whose names start with letters nearer the beginning of the alphabet and admittedly booklet position effects cannot be distinguished from alphabetic effects in this election [4]. It is also possible that a group of candidates may actually be preferred in alphabetical order, perhaps by a small group of voters, perhaps following how-to-vote cards with candidates ordered alphabetically. However, as discussed above, the Canterbury voters would have been less aware of STV, they were more than three times more likely than Otago voters to vote as if the election were being run as a plurality election, and the charts show a greater percentage of perfectly ordered ballots for the Canterbury election. I contend that the charts' heavy tails primarily demonstrate ignorance of, or indifference towards, the ranking of candidates.

5 A measure of voter indifference to ranking

Where booklet or ballot position order can be assessed it may be worthwhile reporting a 'rank indifferent' statistic alongside the percentage of informal votes usually reported in elections. However, it isn't easy to say how many voters are rank indifferent.

Considering the Canterbury DHB election, it certainly seems reasonable to assert that most of the 1,286 voters who listed seven candidates in perfect order were rank indifferent: all but the ten expected, perhaps (refer Table 3.3). It would also seem true of the remaining 151 who listed more than seven candidates in perfect order, as the probability of this occurring is so low. It is less compelling to argue that 38 of the 2,526 voters who listed only two candidates in perfect order should also count, as the probability of this occurring by chance is so much greater. The appropriateness of this measure would then depend on some aspects of the election: if the number of candidates is low or if there are few candidates with popular support, sincere preferences are far more likely to happen to accord with ballot or booklet position and this may result in an inordinate number of perfectly ordered or near perfectly ordered ballots.

One way to avoid this problem is to count the higher than expected number of ordered ballots only when the probability of this occurring is extremely low, below 1% perhaps, which would only assess ballots listing five or more candidates. The Canterbury DHB election would then have a statistic of 2%. However, this seems conservative given the significantly more than expected number of near perfectly ordered ballots shown in the second-to-rightmost bar in Figure 3.1. Therefore one might also consider those ballots with an r_s, such that, say, less than 1% of ballots are to be expected to be found with this r_s or higher. The appropriate choice of r_s will then depend on the number of candidates in the ballot.

Taking this approach encapsulates the above in which we ignored ballots with less than five candidates, as with fewer than five candidates, there are fewer possible values of r_s and the probability of finding ordered ballots is greater than 1%. For example, where a ballot ranks only two candidates, there are only two possible arrangements resulting in an r_s (with its ordered self) of either 1 or -1, and with a probability of 50% either way. With three candidates there are only four possible values of $r_s : -1, -0.5, 0.5$ and 1, and the expected number of ballots having an r_s of 1 is one in six (16.7%) [6]. For four candidates, the expected number of ballots with an r_s of 1 is 4%. It is not until we reach five candidates that the expected number of ballots with an r_s of 1 drops below 1%. For six candidates, the expected number of ballots with an $r_s \geq 0.94$ (an r_s of either 0.94 or 1) is less than 1%, so we now count near perfectly ordered ballots as well as perfectly ordered ballots.

The appropriate values to use for r_s are thus the critical values to be found tabulated in textbooks. The expected number of ballots can be calculated from the probability of an r_s greater than or equal to the critical value: this might be assumed to be 1%, but it varies due to the discrete nature of r_s. Thus we also need to look up the probability of this value of r_s and calculate the number of ballots that may be expected to have this r_s if the ballots were randomly ordered. Critical values for the number of candidates in the ballot from 5 through 50 and the probabilities of finding these values are listed in Table 3.4.

Thus one can step through each ballot that ranks five or more candidates, correlating the ballot with its ordered self, and counting those that are 'highly ordered', that is, those with an r_s greater than the critical value for its number of candidates. One can then subtract the expected number of highly ordered ballots, which can be simply calculated by counting the number of ballots with each number of candidates and multiplying this by the probabilities listed in Table 3.4. Dividing this difference by the total number of formal ballots provides an accessible statistic. This statistic may be interpreted as the percentage of voters that were almost certainly rank indifferent. For the Canterbury DHB election this is 3.8% and for the Otago DHB election it is 1.9%.

However, the probability of a voter being rank indifferent can be expected to be unrelated to the length of the ballot even though we cannot identify rank indifference in shorter ballots with confidence. This seems reasonable when one considers that there is no reason to believe that voters who ranked fewer candidates might have had any greater understanding of STV than those who listed five or more candidates. Therefore, we should really divide the difference by the number of formal ballots that listed five or more candidates. For the Canterbury DHB election the rank indifferent statistic is then 5.1% and for the Otago DHB election it is 2.9% (see Table 3.5 for working).

6 Conclusions

Booklet position effects should be considered when assessing the cost-effectiveness of randomising the order of candidates' names on the ballot paper, especially if voters are unfamiliar with the candidates or if the need to rank candidates might be poorly understood.

Two new statistics may be reported to better gauge voter understanding of preferential voting: first, the percentage of plurality-style informal ballots, that is, ballots in which the voter marked all of the candidates (for whom they voted) with a tick or a '1'; and second, for elections where voters might be expected to rank order five or more candidates, the percentage of voters that were almost certainly rank indifferent. However, in interpreting the rank indifferent percentage one should be wary of other potential causes of perfectly ordered or near perfectly ordered ballots such as how-to-vote cards.

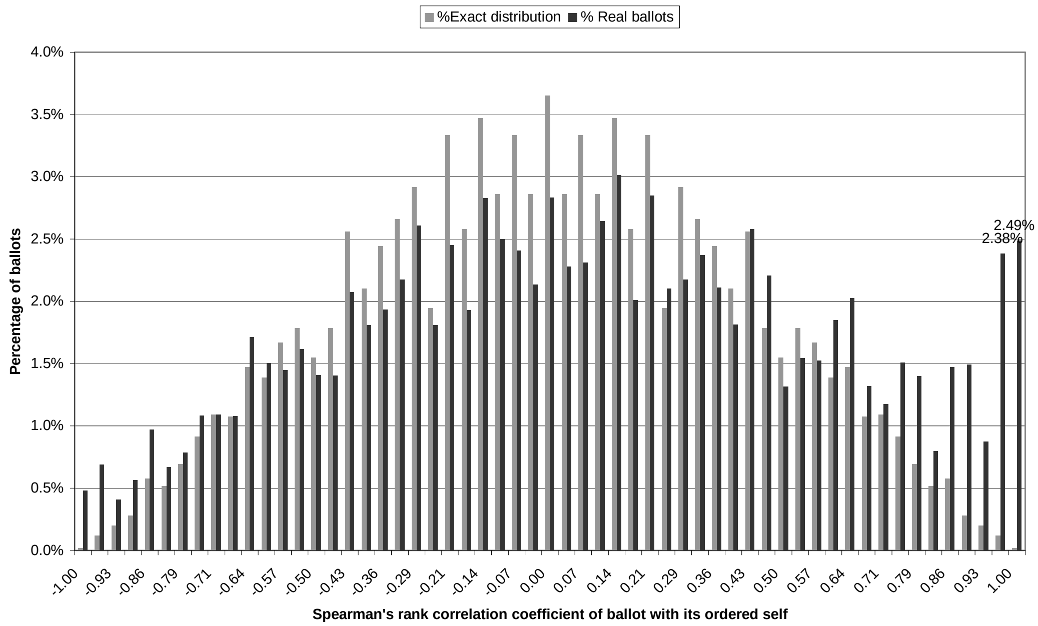


Figure 3.1: Canterbury DHB: frequency of ballots for Spearman rank-order correlation coefficients of voters' ballots with their ballots ordered alphabetically, for ballots listing seven candidates.

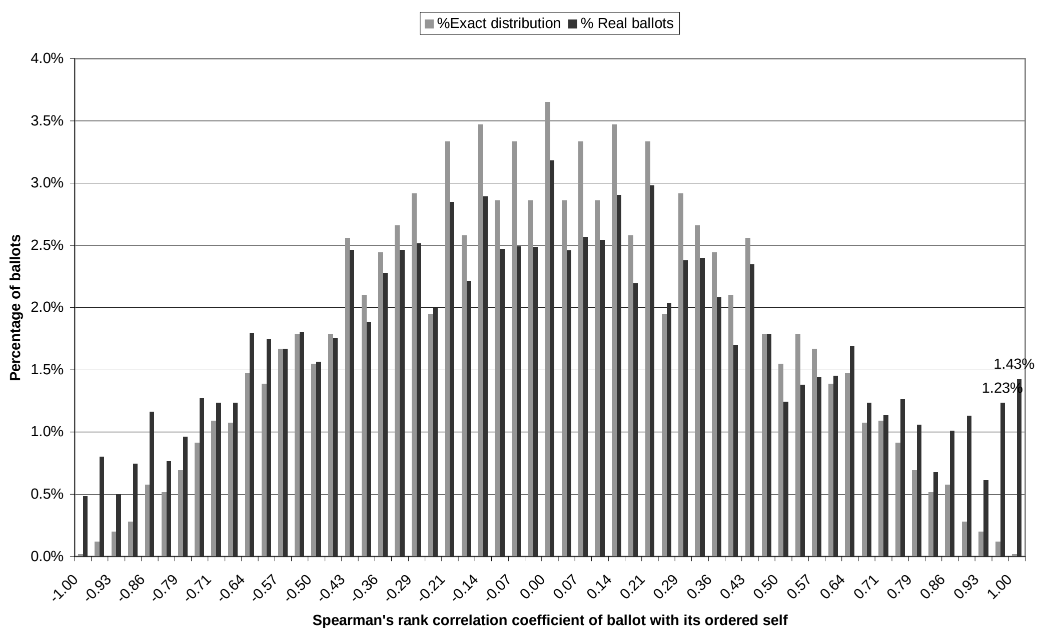


Figure 3.2: Otago DHB: frequency of ballots for Spearman rank-order correlation coefficients of voters' ballots with their ballots ordered alphabetically, for ballots listing seven candidates.

Canterbury DHB data				
Spearman's rank correlation coefficient	**%Exact distribution**	**Real ballots**	**% Real ballots**	**Exact distribution**
-1.00	0.02%	249	0.48%	10
-0.96	0.12%	357	0.69%	62
-0.93	0.20%	211	0.41%	103
-0.89	0.28%	292	0.56%	144
-0.86	0.58%	501	0.97%	298
-0.82	0.52%	346	0.67%	267
-0.79	0.69%	406	0.78%	359
-0.75	0.91%	559	1.08%	472
-0.71	1.09%	564	1.09%	565
-0.68	1.07%	557	1.08%	554
-0.64	1.47%	885	1.71%	760
-0.61	1.39%	778	1.50%	718
-0.57	1.67%	749	1.45%	862
-0.54	1.79%	836	1.62%	924
-0.50	1.55%	729	1.41%	801
-0.46	1.79%	726	1.40%	924
-0.43	2.56%	1072	2.07%	1324
-0.39	2.10%	935	1.81%	1088
-0.36	2.44%	1000	1.93%	1262
-0.32	2.66%	1123	2.17%	1375
-0.29	2.92%	1348	2.61%	1509
-0.25	1.94%	936	1.81%	1006
-0.21	3.33%	1268	2.45%	1724
-0.18	2.58%	999	1.93%	1334
-0.14	3.47%	1463	2.83%	1796
-0.11	2.86%	1293	2.50%	1478
-0.07	3.33%	1245	2.41%	1724
-0.04	2.86%	1103	2.13%	1478
0.00	3.65%	1465	2.83%	1889
0.04	2.86%	1179	2.28%	1478
0.07	3.33%	1196	2.31%	1724
0.11	2.86%	1367	2.64%	1478
0.14	3.47%	1559	3.01%	1796
0.18	2.58%	1041	2.01%	1334
0.21	3.33%	1473	2.85%	1724
0.25	1.94%	1086	2.10%	1006
0.29	2.92%	1124	2.17%	1509
0.32	2.66%	1227	2.37%	1375
0.36	2.44%	1090	2.11%	1262
0.39	2.10%	939	1.82%	1088
0.43	2.56%	1333	2.58%	1324
0.46	1.79%	1141	2.21%	924
0.50	1.55%	681	1.32%	801
0.54	1.79%	799	1.54%	924
0.57	1.67%	788	1.52%	862
0.61	1.39%	957	1.85%	718
0.64	1.47%	1048	2.03%	760
0.68	1.07%	682	1.32%	554
0.71	1.09%	606	1.17%	565
0.75	0.91%	779	1.51%	472
0.79	0.69%	724	1.40%	359
0.82	0.52%	412	0.80%	267
0.86	0.58%	761	1.47%	298
0.89	0.28%	771	1.49%	144
0.93	0.20%	453	0.88%	103
0.96	0.12%	1233	2.38%	62
1.00	0.02%	1286	2.49%	10
	100.00%	51730	100.00%	51730

Otago DHB Data				
Spearman's rank correlation coefficient	**%Exact distribution**	**Real ballots**	**% Real ballots**	**Exact distribution**
-1.00	0.02%	123	0.48%	5
-0.96	0.12%	204	0.80%	30
-0.93	0.20%	126	0.50%	50
-0.89	0.28%	189	0.74%	71
-0.86	0.58%	295	1.16%	146
-0.82	0.52%	194	0.76%	131
-0.79	0.69%	244	0.96%	176
-0.75	0.91%	322	1.27%	232
-0.71	1.09%	314	1.24%	277
-0.68	1.07%	313	1.23%	272
-0.64	1.47%	455	1.79%	373
-0.61	1.39%	443	1.74%	353
-0.57	1.67%	423	1.67%	423
-0.54	1.79%	457	1.80%	453
-0.50	1.55%	397	1.56%	393
-0.46	1.79%	445	1.75%	453
-0.43	2.56%	625	2.46%	650
-0.39	2.10%	479	1.89%	534
-0.36	2.44%	578	2.28%	620
-0.32	2.66%	625	2.46%	675
-0.29	2.92%	638	2.51%	741
-0.25	1.94%	507	2.00%	494
-0.21	3.33%	723	2.85%	846
-0.18	2.58%	562	2.21%	655
-0.14	3.47%	734	2.89%	882
-0.11	2.86%	627	2.47%	725
-0.07	3.33%	632	2.49%	846
-0.04	2.86%	631	2.49%	725
0.00	3.65%	808	3.18%	927
0.04	2.86%	624	2.46%	725
0.07	3.33%	652	2.57%	846
0.11	2.86%	645	2.54%	725
0.14	3.47%	737	2.90%	882
0.18	2.58%	557	2.19%	655
0.21	3.33%	757	2.98%	846
0.25	1.94%	518	2.04%	494
0.29	2.92%	604	2.38%	741
0.32	2.66%	609	2.40%	675
0.36	2.44%	528	2.08%	620
0.39	2.10%	431	1.70%	534
0.43	2.56%	595	2.34%	650
0.46	1.79%	453	1.78%	453
0.50	1.55%	315	1.24%	393
0.54	1.79%	350	1.38%	453
0.57	1.67%	365	1.44%	423
0.61	1.39%	369	1.45%	353
0.64	1.47%	429	1.69%	373
0.68	1.07%	313	1.23%	272
0.71	1.09%	288	1.13%	277
0.75	0.91%	321	1.26%	232
0.79	0.69%	269	1.06%	176
0.82	0.52%	172	0.68%	131
0.86	0.58%	257	1.01%	146
0.89	0.28%	287	1.13%	71
0.93	0.20%	156	0.61%	50
0.96	0.12%	313	1.23%	30
1.00	0.02%	362	1.43%	5
	100.00%	25389	100.00%	25389

Table 3.2: Data for Figures 3.1 and 3.2: the numbers of ballots for each possible value of r_s and the exact distribution (as would be approximated by randomly ordered ballots) for ballots ranking seven candidates [7]

Candidates in ballot (*n*)	Ballots (*b*)	Perfectly ordered (*p*)	Probability of being in order (1/*n*!)	Expected ballots in order (*b*/*n*!)	% found in order (*p*/*b*)	Number of times more than expected (*p*/(*b*/*n*!))
1	5691	5691	1.000000	5691	100.00%	1.00
2	4977	2526	0.500000	2489	50.75%	1.02
3	8483	1766	0.166667	1414	20.82%	1.25
4	8030	817	0.041667	335	10.17%	2.44
5	8639	514	0.008333	72	5.95%	7.14
6	5857	229	0.001389	8	3.91%	28.15
7	51730	1286	0.000198	10	2.49%	125.29
8	3331	55	0.000025	8.3E-02	1.65%	665.75
9	2224	39	2.8E-06	6.1E-03	1.75%	6363.45
10	2721	27	2.8E-07	7.5E-04	0.99%	36007.94
11	1107	12	2.5E-08	2.8E-05	1.08%	4.33E+05
12	1170	3	2.1E-09	2.4E-06	0.26%	1.23E+06
13	503	6	1.6E-10	8.1E-08	1.19%	7.43E+07
14	507	4	1.1E-11	5.8E-09	0.79%	6.88E+08
15	361	1	7.6E-13	2.8E-10	0.28%	3.62E+09
16	294	0	4.8E-14	1.4E-11	0.00%	0.00
17	166	2	2.8E-15	4.7E-13	1.20%	4.29E+12
18	131	0	1.6E-16	2.0E-14	0.00%	0.00
19	91	0	8.2E-18	7.5E-16	0.00%	0.00
20	112	0	4.1E-19	4.6E-17	0.00%	0.00
21	68	0	2.0E-20	1.3E-18	0.00%	0.00
22	50	0	8.9E-22	4.4E-20	0.00%	0.00
23	37	0	3.9E-23	1.4E-21	0.00%	0.00
24	47	0	1.6E-24	7.6E-23	0.00%	0.00
25	33	0	6.4E-26	2.1E-24	0.00%	0.00
26	49	0	2.5E-27	1.2E-25	0.00%	0.00
27	45	0	9.2E-29	4.1E-27	0.00%	0.00
28	47	0	3.3E-30	1.5E-28	0.00%	0.00
29	2365	2	1.1E-31	2.7E-28	0.08%	7.48E+27

Table 3.3: Perfectly ordered ballots in the Canterbury DHB election

Number of candidates selected on ballot	Minimum r_s	Probability of finding such a ballot	Number of candidates selected on ballot	Minimum r_s	Probability of finding such a ballot
5	1.000	0.00833	28	0.440	0.01
6	0.943	0.00833	29	0.433	0.01
7	0.893	0.00615	30	0.425	0.01
8	0.833	0.00769	31	0.418	0.01
9	0.783	0.00861	32	0.412	0.01
10	0.745	0.00870	33	0.405	0.01
11	0.709	0.00910	34	0.399	0.01
12	0.678	0.00926	35	0.394	0.01
13	0.648	0.00971	36	0.388	0.01
14	0.626	0.00953	37	0.383	0.01
15	0.604	0.00973	38	0.378	0.01
16	0.582	0.00999	39	0.373	0.01
17	0.566	0.00983	40	0.368	0.01
18	0.550	0.00986	41	0.364	0.01
19	0.535	0.01	42	0.359	0.01
20	0.520	0.01	43	0.355	0.01
21	0.508	0.01	44	0.351	0.01
22	0.496	0.01	45	0.347	0.01
23	0.486	0.01	46	0.343	0.01
24	0.476	0.01	47	0.340	0.01
25	0.466	0.01	48	0.336	0.01
26	0.457	0.01	49	0.333	0.01
27	0.448	0.01	50	0.329	0.01

Table 3.4: Critical values and probabilities for r_s

[6, 7]

Canterbury DHB				
Candidates in ballot (*n*)	Ballots	Expected highly ordered	Found highly ordered	Difference
1	5691			
2	4977			
3	8483			
4	8030			
5	8639	72.0	514	442.0
6	5857	48.8	229	180.2
7	51730	318.1	2972	2653.9
8	3331	25.6	237	211.4
9	2224	19.1	169	149.9
10	2721	23.7	222	198.3
11	1107	10.1	81	70.9
12	1170	10.8	78	67.2
13	503	4.9	45	40.1
14	507	4.8	33	28.2
15	361	3.5	30	26.5
16	294	2.9	27	24.1
17	166	1.6	12	10.4
18	131	1.3	8	6.7
19	91	0.9	6	5.1
20	112	1.1	4	2.9
21	68	0.7	3	2.3
22	50	0.5	5	4.5
23	37	0.4	2	1.6
24	47	0.5	5	4.5
25	33	0.3	4	3.7
26	49	0.5	3	2.5
27	45	0.5	3	2.6
28	47	0.5	0	-0.5
29	2365	23.7	76	52.4
		576.8	4768	4191.2
Total	108866		Rank indifferent *n* >= 1	3.8%
Total *n* >= 5	81685		Rank indifferent *n* >= 5	**5.1%**

Otago DHB				
Candidates in ballot (*n*)	Ballots	Expected highly ordered	Found highly ordered	Difference
1	4323			
2	4196			
3	6047			
4	5515			
5	4573	38.1	157	118.9
6	4100	34.2	86	51.8
7	25389	156.1	831	674.9
8	1905	14.6	115	100.4
9	1112	9.6	62	52.4
10	1470	12.8	90	77.2
11	502	4.6	25	20.4
12	577	5.3	26	20.7
13	225	2.2	9	6.8
14	282	2.7	10	7.3
15	148	1.4	7	5.6
16	117	1.2	8	6.8
17	54	0.5	1	0.5
18	56	0.6	3	2.4
19	34	0.3	1	0.7
20	56	0.6	2	1.4
21	23	0.2	1	0.8
22	26	0.3	0	-0.3
23	11	0.1	1	0.9
24	21	0.2	2	1.8
25	81	0.8	1	0.2
26	1530	15.3	79	63.7
		301.7	1517	1215.3
Total	62373		Rank indifferent *n* >= 1	1.9%
Total *n* >= 5	42292		Rank indifferent *n* >= 5	**2.9%**

Table 3.5: Manual calculation of rank indifferent statistic

Further information and computer programs to automate the production of these statistics are available from the author on request.

7 References

[1] Department of Internal Affairs, 'STV Information', 2004, URL on web site.

[2] Personal communication with Christchurch City Council, 13 April 2005.

[3] New Zealand Government, 'Part 4 - Conduct of Elections and Polls using Single Transferable Voting Electoral System', *Local Electoral Amendment Regulations 2002.*

[4] Jon A. Krosnik, Joanne M. Miller, and Michael P. Tichy, 'An unrecognized need for ballot reform', in Ann N. Crigler, Marion R. Just, and Edward J. McCaffery (eds.), *Rethinking the Vote : The Politics and Prospects of American Election Reform* (New York: Oxford University Press, 2004), pp. 52, 53, 63.

[5] Susan Banducci, Michael Thrasher, Colin Rallings and Jeffrey A. Karp, 'Candidate appearance cues in low-information elections', 2003, Paper presented at the Annual Conference of the American Political Science Association (Philadelphia), URL on web site. p. 14.

[6] Sidney Siegel and N. John Castellan, Jr., *Nonparametric statistics for the behavioral sciences,* second edition (New York: McGraw-Hill, 1988), pp. 242, 360-361.

[7] Mark van de Wiel, Computer program `spearman.c` for calculating the 'Exact distribution of Spearman's rank statistic', available from URL on web site.

ISSN 1745-6231

Voting matters

for the technical issues of STV

published by

The McDougall Trust

Issue 22 July 2006

About the McDougall Trust (reg. charity no. 212151)

The McDougall Trust is a charitable trust formed in 1948. The charity's purposes as stated in its governing scheme of 1959 are to advance knowledge of and encourage the study of and research into:

- political or economic science and functions of government and the services provided to the community by public and voluntary organisations; and
- methods of election of and the selection and government of representative organisations whether national, civic, commercial, industrial or social.

The Trust's work includes the maintenance and development of the Lakeman Library for Electoral Studies, a unique research resource, the production and publication of Representation: The Journal of Representative Democracy, and, of course, this publication **Voting matters**, that examines the technical issues of the single transferable vote and related electoral systems.
For further information on the Trust, please contact:

The Secretary,
McDougall Trust,
6 Chancel Street,
London SE1 0UX, UK.
Telephone: +44 (0)20 7620 1080
Facsimile: +44 (0)20 7928 1528
Email: admin@mcdougall.org.uk
Web: www.mcdougall.org.uk

For further information on this publication, please contact B A Wichmann, the Editor at the above address or by email at: `Brian.Wichmann@bcs.org.uk`

Editorial

There are five papers in this issue, the first three being:

- Jonathan Lundell: *Random tie-breaking in STV*.

  Although *Voting matters* has had several papers about tie-breaking, one can see that there is still more to be said on the matter.

- David Hill: *Implementing STV by Meek's method*.

  David Hill has provided an implementation of Meek's method for many years. This implementation has been taken as the 'definition' of the method for the New Zealand elections. In this paper, the details of this implementation are described and contrasted with that of the original *Computer Journal* article.

- Robert Newland: *Computerisation of STV counts*.

  Although Robert Newland died in August 1990, readers may well be surprised at the relevance of this paper for today. Up to his death, he was the leading technical expert on STV within ERS. This paper was located by David Hill and since it was never published, printing it here seemed appropriate. It is hoped that readers will respond to the suggestions made.

The final two papers have a common theme: the form of STV proposed by British Columbia and now being considered for the Scottish local elections to be held next year.

- Jeff O'Neill: *Comments on the STV Rules Proposed by British Columbia*.

  This paper presents the details of an implementation of the British Columbia rules which has been available on the Internet for some time. It is a very simple version of STV in computer terms. Several issues arose from this work which are detailed in the paper.

- James Gilmour: *Developing STV Rules for manual counting to give effect to the Weighted Inclusive Gregory Method of transferring surpluses, with candidates' votes recorded as integer values.*

  The paper is a complete contrast to the previous one. Like the previous paper, the aim is to transfer surpluses by considering all papers, not just the last batch that gave rise to the surplus. The contrast is in its presentation as a manual counting process and the provision of the conventional result sheet. One novelty is (at least within the UK) that the calculations are undertaken with high precision, but the results are presented as integers.

  James Gilmour has produced a proposal and sent it to the Scottish Executive. This proposal, slightly modified, is now on the McDougall web site. Hence the article provides the rationale and background to the proposal.

It is hoped that the contrast between the two methods above will clarify the choices to be made for the Scottish elections. The final choice will be awaited with interest.

Two other items may be of interest to readers. Firstly, the final report on electronic voting in Ireland is due out shortly and will be found at: `http://www.cev.ie/`. Secondly, it has come to my attention that the British Computer Society elect their council by STV but do not provide a result sheet to their electorate — only a list of those elected. Since the transfer of votes will not be visible, this seems to me to be STV in name only. Do readers have other examples of this?

Random tie-breaking in STV

Jonathan Lundell
jlundell@pobox.com

1 Introduction

The resolution of ties in STV elections is not a settled question. On the contrary, it remains a topic of lively discussion, with several papers published on the subject in these pages; see Earl Kitchener's note, "A new way to break STV ties in a special case" [1] for a summary.

Ties can arise in any STV election during exclusion. With some methods ties can arise at other stages as well; Jeffrey O'Neill [2] lists the cases.

O'Neill also lists four tie-breaking methods. Two methods use the first or last difference in prior rounds to break a tie, and two methods use later preferences—Borda scores or most (fewest) last-place preferences. Brian Wichmann [3] proposes to examine all possible outcomes.

None of these tie-breaking methods is guaranteed to break a tie, since they can themselves result in a tie, or in the case of [3] become so computationally expensive as to be impractical. These cases (strong ties) are typically broken randomly. Some election methods, eg, the Algorithm 123 version of Meek's method [4], rely exclusively on random tie-breaking.

Objections to random tie-breaking fall into two categories. One is a worry that voters and candidates will object to election decisions being made by chance instead of by voter preference. Thus Wichmann [3]: "When a candidate has been subject to a random exclusion in an election, he/she could naturally feel aggrieved." Other objections adduce examples in which it appears intuitively preferable to break a tie based on some measure of voter preference.

All STV election methods rely on random tie-breaking (or at least tie-breaking based on some consideration other than voter preference) to break strong ties. (Ties in first-past-the-post elections are often broken randomly as well, by coin toss, drawing straws, or drawing a high card.)

2 Prior-round tie-breaking

The rationale for forwards tie-breaking (using O'Neill's terminology) appears to be that it gives greatest weight to first preferences. O'Neill [2] argues for backwards tie-breaking:

> A more important problem, is that forwards tie-breaking does not use the most relevant information to break the tie. The most relevant information to break a tie is the previous stage and not all the way back to the very first stage. By immediately looking to the first stage to break the tie, the ERS97 rules allow the tie-breaking to be influenced by candidates eliminated very early in the process and also by surpluses yet to be transferred. Instead, if we look to the previous stage to break a tie, candidates eliminated early on in the process will have no influence in breaking the tie. In addition, it allows for surpluses to be transferred which gives a more accurate picture of candidate strength.

Carrying O'Neill's argument to its logical conclusion, however, the "most relevant information" is not in any prior round, but rather in the current round—and the current round declares a tie.

Prior-round tie-breaking encourages insincere voting. Consider this election fragment, with two candidates to exclude:

5	A
4	B
1	CB

Excluding C, we have:

5 A
5 B

and must now break the tie. Prior-round tie-breaking requires that we exclude B, since A led B 5-4 in the previous round. So voter CB, believing that the first choice (C) is likely to be excluded, is encouraged to insincerely vote B (or BC) so as not to jeopardize B's chances in the event of an A–B tie.

Prior-round tie-breaking is especially troublesome in the context of Meek rules, since it violates Meek's Principle 1: If a candidate is eliminated, all ballots are treated as if that candidate had never stood. But if C had never stood, A and B would have been tied.

3 Later-preference tie-breaking

Kitchener [5] points out a problem case for random tie-breaking:

> An extreme case can arise where there is one seat and the electors are the same as the candidates; for example, if a partnership is electing a senior partner. Each candidate may put himself first, and all, except candidate A, put A second. Under most present rules, one candidate then has to be excluded at random, and it may be A. There is no way of getting over this unreasonable result without looking at later preferences. . . .

The smallest such election:

1 A
1 B A
1 C A

Prior-round tie-breaking methods are of no help in the first round, and a random choice excludes A, the consensus choice, one third of the time. Kitchener proposes to use Borda scores to break the tie; we must still randomly break a strong B-C tie, but A survives and is elected.

This case is related to a problem with STV in general, pointed out by Meek [6]. "A related point, and probably the strongest decision-theoretic argument against STV, is the fact that a candidate may be everyone's second choice but not be elected."
. . . and also related to the general problem of premature exclusion.

Kitchener concedes that there is a problem with Borda tie-breaking, as there is with any tie-breaking method that relies on later preferences.

> It is a fundamental principle of STV that later preferences should not affect the fate of earlier ones; this encourages sincere voting, but means that some arbitrary or random choice must be made to break ties, which can give unreasonable results.

Responding to the Borda tie-breaking suggestion, David Hill [7] objects: "What matters is that tactical considerations have been allowed in, where STV (in its AV version in this case) is supposed to be free of them."

This point is crucial. In any election system, the rules, including the method of breaking ties, must of course be specified in advance. When we look at the partnership election example above, we interpret the ballots as the sincere expression of the voters, and so read the ballots as favoring A. But as both Hill and Kitchener observe, once later-preference tie-breaking is introduced, we must expect insincere voting. In the face of later-preference tie-breaking, B and C, to maximize their chances of winning (after all, each is their own first choice) must resort to bullet voting (American English—one might say characteristically AmE—for plumping). The ballots would then read,

1 A
1 B
1 C

. . . and we're forced to resort to a random choice. This seems a shame, since it does appear from the presumably sincere ballots in the initial profile that both B and C prefer A to the other. The partners might be well advised to adopt a special rule forbidding each to vote for herself. In that case, we would have:

1 abstain
2 A

. . . and A wins outright.

4 Random tie-breaking

An advantage claimed by Meek [6] for STV is that "There is no incentive for a voter to vote in any way other than according to his actual preference." One of Meek's motivations for proposing a new STV method is to come closer to that ideal. Likewise Warren [8], "It is one of the precepts of preferential voting systems that a later preference should neither help nor harm an earlier preference."

Any election method relies for its properties on the implicit assumption that voters will vote sincerely, that is, that their ballots will reflect, within the limitations of the specific method, their true preferences. Without sincere votes, any election method fails to reflect the will of the electorate, on the principle of garbage in, garbage out. It is perverse to use tie-breaking methods that reintroduce incentives for voters to vote insincerely. Hill and Gazeley [9]:

> In considering this, we need to take into account, among other things, that the true aim of an election should not be solely to match seats as well as possible to votes, but to match seats to the voters' wishes. Since we do not know the wishes we must use the votes as a substitute, but that makes it essential that the votes should match the wishes as far as possible. That, in turn, makes it desirable that the voters should not be tempted to vote tactically.

5 Voter psychology

One might counter that, except in small elections, the chances of a tie are sufficiently small that a voter ought to ignore the possibility of a tie altogether and vote sincerely. This argument is problematic on two fronts. First, our methods should work with small elections as well as large ones (and the line between small and large elections is not well defined). Second, especially in a high-stakes election, the voter's estimation of the risk associated with voting sincerely is likely to be wrong.

Computer security authority Bruce Schneier, interviewed in *CSO Magazine* [10], comments:

> Why are people so lousy at estimating, evaluating and accepting risk?... Evaluating risk is one of the most basic functions of a brain and something hard-wired into every species possessing one. Our own notions of risk are based on experience, but also on emotion and intuition. The problem is that the risk analysis ability that has served our species so well over the millennia is being overtaxed by modern society. Modern science and technology create things that cannot be explained to the average person; hence, the average person cannot evaluate the risks associated with them. Modern mass communication perturbs the natural experiential process, magnifying spectacular but rare risks and minimizing common but uninteresting risks. This kind of thing isn't new—government agencies like the [US] FDA were established precisely because the average person cannot intelligently evaluate the risks of food additives and drugs—but it does have profound effects on people's security decisions. They make bad ones.

For our purposes, read *tactical voting decisions* for *security decisions*. Rational insincere voting is bad enough; insincere voting based on faulty information or poor tactics is even worse.

6 A note on weighting votes in later-preference tie-breaking

Consider this election profile (BC rules, two to be elected, quota 10):

12	AB
7	BC
9	C
2	D

A is elected, and D is excluded, leaving B and C tied with nine votes each in the third round. If we break the tie with Borda scores:

A	36	(elected)
B	24+21 =	45
C	14+27 =	41
D	6	(excluded)

C is excluded, and B is elected as the last candidate standing for the second seat.

Notice in particular that while B receives only the two transferable votes from the AB voters (a quota of 10 being retained by A, who is elected), B gets full credit for all 12 AB votes in the Borda tiebreaker.

I suggest that the AB voters, having elected A, must carry only the transferable weight of their votes in calculating the tie-breaking Borda score. Otherwise these voters *double dip*, not only electing A, but also participating disproportionately and decisively in the tie-breaking elimination of C and subsequent election of B.

If we calculate the Borda scores using the weight of transferable votes (that is, votes currently allocated to hopeful candidates), we have:

A		(elected)
B	4+21 =	23
C	14+27 =	41
D		(excluded)

Calculated with the vote weights that give rise to the tie itself, the Borda score now breaks the tie to eliminate B, and C is elected.

The same argument applies to any method that breaks ties with later preferences. Votes committed to already-elected candidates should not be counted again in breaking subsequent ties.

7 A better later-preference tie-breaking method

The chief problem with STV tie-breaking with Borda scores is that it violates the principle of later-no-harm, and it does so in an especially egregious way. Suppose that six candidates are in the running, that I have voted ABC, and that B and C are tied for elimination. The Borda scores for B and C pick up four and three points, respectively, from my ballot. If the three points that my ballot contributes to C's Borda score is the margin for C's victory over B in the Borda tiebreaker, then my later mention of C has led directly to the defeat of B, even though I prefer B to C.

Consider an alternative later-preference tiebreaker. For the sake of simplicity, I will describe it for two-way ties, and then extend it to *n*-way ties. To break a tie, compare the ballots that prefer B to C to the number of ballots that prefer C to B, weighted as described in the note above. Exclude the less-preferred candidate. Break strong ties randomly.

This method, like all later-preference methods, violates later-no-harm, but it preserves a property that I will call *later-no-direct-harm*. My ranking of ABC will not harm B's chances in a BC tie. In the case of a BC tie, my ballot will either have no effect (the margin of B over C or vice versa without my ballot is sufficient that my ballot makes no difference), or it will cause the BC tie to be broken in favor of B, my preferred candidate in the tie (B and C are strongly tied without my ballot), or my ballot will convert a one-vote C advantage (without my ballot) to a strong tie (with my ballot), giving B an even chance in a random tiebreak.

That is, my ABC ballot either has no effect on breaking a BC tie, or it benefits B.

By *later-no-direct-harm*, I mean that the fact that I have ranked the later preferences BC will not harm my favorite in the potential tie between B and C. Later-no-harm is not avoided; my ABC preference could break a tie in favor of B, and B could subsequently defeat my first preference, A, whereas A might have prevailed had C won the BC tiebreaker. Any harm to A, however, will come indirectly, in a later round—and it would be rude for me to complain that the BC tie was broken on the basis of my preference for B over C.

Generalizing to breaking an *n*-way tie for exclusion:

1. Find the first mention of any member of the tied set of candidates on each ballot, and calculate the total such mentions for each of the candidates, using the transferable weight of each ballot. Ignore ballots that do not mention at least one tied candidate.

2. If all *n* candidates are still tied, exclude one tied candidate at random; *finis*.

3. Otherwise, remove from consideration for exclusion the candidate (or a random choice from the tied set of candidates) with the highest score from step 1.

4. If only one candidate remains, exclude that candidate; *finis*.

5. Otherwise, *n* is now the remaining number of tied candidates (that is, less the reprieved candidates from step 3); continue at step 1.

If the tie is for a winner rather than an exclusion, then remove from consideration the candidate with the lowest rather than the highest score. This is simply single-winner STV (AV or IRV) with weighted ballots, and suggests an alternative to the proposed algorithm for breaking a tie for exclusion: break an n-way tie for exclusion by counting an STV election (again with weighted ballots) with n candidates and $n-1$ winners; exclude the single loser.

It's worth noting that a similar procedure based on lowest preferences (along the lines of Coombs tie-breaking) does not satisfy the principle of later-no-direct-harm. For example, if candidates X, Y and Z are tied for exclusion and I have ranked those candidates XYZ, it's possible that my preference for Y over Z is decisive in favor of Y, and that Y but not Z beats X in a head-to-head tiebreaker; thus my preference for Y over Z decides the tiebreak in favor of Y over X, contrary to my preferences.

Likewise, Condorcet ranking is equivalent to the proposed method for two-way ties, but violates later-no-direct-harm in the general *n*-way-tie case.

The proposed tie-breaking method—let's call it *weighted first preference*—differs from prior-round tie-breaking methods in that it considers the preferences of all voters (suitably weighted), and not only voters who have ranked the tied candidates first (after elections and exclusions) in a prior round.

Hill and Gazeley [9] observe, in the context of Sequential STV:

> With this new version, should it be recommended for practical use? That depends upon whether the user is willing to abandon the principle that it should be impossible for a voter to upset earlier preferences by using later preferences. Many people regard that principle as very important, but reducing the frequency of premature exclusions is important too. We know that it is impossible to devise a perfect scheme, and it is all a question of which faults are the most important to avoid.
>
> In considering this, we need to take into account, among other things, that the true aim of an election should not be solely to match seats as well as possible to votes, but to match seats to the voters' wishes. Since we do not know the wishes we must use the votes as a substitute, but that makes it essential that the votes should match the wishes as far as possible. That, in turn, makes it desirable that the voters should not be tempted to vote tactically.
>
> They would not be so tempted if they felt confident that later preferences were as likely to help earlier ones as to harm them, and if they could not predict the effect one way or the other. At present, we see no reason to doubt that these requirements are met.

The proposed method for breaking ties satisfies the same criteria: later preferences are as likely to help earlier ones as to harm them, and voters cannot predict the effect one way or the other. This is not the case for other preference-based tie-breaking methods discussed in these pages.

Whether this slight opening of the door to a violation of later-no-harm is justified by the benefit of breaking ties non-randomly (in most cases) is, in David Hill's words [7], a matter of judgment.

8 Summary

Arguments for various nonrandom tie-breaking implicitly assume sincere voters. But the introduction of those very methods undermines that crucial precondition, and without sincere voters the arguments fail.

When O'Neill argues [2] that "forwards tie-breaking does not use the most relevant information to break the tie," and that later rounds reflect better information, the logical conclusion of his argument is that the most relevant information is not in a prior round at all, but rather in the current round that gives rise to the tie. That information is, simply, that the candidates have equal support, by the means we've chosen to measure that support.

Meek [6] drives this point further home with his Principle 1: "If a candidate is eliminated, all ballots are treated *as if that candidate had never stood*." Prior-round tie-breaking typically, though not exclusively, depends on preferences for candidates who have been excluded in the tie-breaking round. To consider those preferences violates Meek's Principle 1.

Later-preference tie-breaking (eg. Borda or Coombs) encourages insincere voting by violating the later-no-harm principle.

The encouragement of insincere voting is too high a price to pay for partially excluding chance from STV election methods. We should prefer random tie-breaking in all cases.

If preferences must be considered in breaking ties, then ties should be broken on the basis of overall earliest preferences, using transferable ballot weights.

9 References

[1] Earl Kitchener. A new way to break STV ties in a special case. *Voting matters*. Issue 20, pp 9–11.

[2] J C O'Neill. Tie-Breaking with the Single Transferable Vote. *Voting matters*. Issue 18, pp 14–17.

[3] Brian Wichmann. Tie Breaking in STV. *Voting matters*. Issue 19, pp 1–5.

[4] I. D. Hill, B. A. Wichmann and D. R. Woodall. Algorithm 123—Single Transferable Vote by Meek's method. *Computer Journal*. Vol 30, pp 277–281, 1987.

[5] Earl Kitchener. Tie-breaking in STV. *Voting matters*. Issue 11, pp 5–6.

[6] B L Meek. A New Approach to the Single Transferable Vote. *Voting matters*. Issue 1, pp 1–7.

[7] I D Hill. Tie-breaking in STV. *Voting matters*. Issue 12, pp 5–6.

[8] C H E Warren. An example showing that Condorcet infringes a precept of preferential voting systems. *Voting matters*. Issue 5, pp 5–6.

[9] I D Hill and Simon Gazeley. Sequential STV - a new version. *Voting matters*. Issue 15, pp 13–16.

[10] The Evolution of a Cryptographer, *CSO Magazine,* September 2003. Available from the McDougall web site.

Implementing STV by Meek's method

I.D. Hill
d.hill928@btinternet.com

1 Introduction

At the time of the original implementation of STV by Meek's method [1] we were feeling our way. Later thought has shown that, in some respects, the details can be improved while keeping the overall plan. Thus my own later implementation, as part of a suite of programs to deal with the whole election process rather than just the vote counting, and to include other versions of STV as well as the Meek version, made some changes from that original implementation. The aim of this paper is to describe those changes and the reasons for them.

My program is written in the Pascal computer language. While designed to be used under the MS-DOS operating system, it can also be easily accessed from Windows XP.

In [1] Woodall gave mathematical proof that the Meek formulation has a unique solution for any given voting pattern, and that the method necessarily converges upon that solution. Strictly speaking that proof assumes infinite mathematical precision. In this paper I refer to that proof even though my implementation has only finite precision. Provided that the degree of precision is adequate, the approximation to Woodall's proof will be close enough for practical purposes.

2 Terminology

In [1] we used the term 'weight' for the fraction, of each vote or part of a vote received, that a candidate retains. This has now become known as the candidate's 'keep value', to be in accordance with the traditional term 'transfer value'.

We also used 'excess' for the amount of vote remaining after all candidates mentioned in the voter's preferences have received their shares. The more traditional, but longer, term 'non-transferable' is now used for this.

3 Arithmetic

In [1] the numbers of votes and the keep values were declared as 'real' variables in the computer sense. These would be represented in the computer in floating-point form, which is necessarily only approximate and there is no guarantee that exactly the same approximations will be used on different computer systems. Given the robustness of the Meek method, it is highly improbable that a different candidate would ever be elected because of this, except perhaps in the case of a tie, but it is thought wise to avoid even the possibility.

It is therefore better to make sure that the numbers are so represented that, although still approximate because only a finite number of decimal places is used, the results are necessarily identical on all computers. To achieve this, floating-point methods are avoided altogether, each 'real' number being represented by a pair of integers, integer arithmetic on computers being exact.

Assuming 32-bit integers to be available, the maximum allowable integer is 2147483647 so to allow 9 decimal places for the fractional part is safe and convenient. Thus a number such as 123.456, for example, is represented as a pair of integers with 123 as the value of its integral part and 456000000 as its fractional part. Adding or subtracting such numbers is simple enough, the integral parts are added or subtracted, and the fractional parts are added or subtracted. If the resulting fractional part exceeds 999999999, then 1000000000 is subtracted from it and 1 is added to the integral part. Similarly, if the resulting fractional part is negative, then 1000000000 is added to it and 1 is subtracted from the integral part. There is no need to worry about the whole number, rather than just its fractional part, ever being negative; that never happens within the Meek method.

Multiplication and division are not so simple, and special routines are necessary to enable them to be performed with no risk of overflow.

In principle, a fixed number of significant figures might be preferable to a fixed number of decimal places, but all that really matters is that the precision should be great enough as to ensure that the use of more precision would be virtually certain not to change the outcome. The fixed 9 decimal places undoubtedly satisfies this and is convenient.

4 Quota definition

Meek's formulation [2] used the integral part of $1+T/(s+1)$, where T is the total number of active votes and s is the number of seats to be filled. He obviously intended that the initial 1 of this formula should be replaced by 1 in the last decimal place used, when not working solely in integers. An alternative approach is that of the second edition of Newland and Britton [3] in ignoring the initial 1 altogether if the calculation comes out exactly, while adding extra rules to ensure that no more than s candidates can be elected even in exceptional cases. In [1] we adopted the Newland and Britton approach (with the necessary extra rules) because the number of decimal places that would be used by a floating-point implementation was unknown.

When working solely in integers, or to only 2 decimal places as in Newland and Britton rules, there are advantages in their formulation, but those advantages are minimal where greater precision is used. For my implementation, therefore, I have included the addition of 0.000000001 to the quota, so that no extra rules are needed, while it is very hard to believe that such a tiny increment will ever cause any disadvantage.

5 Output

In [1], mainly because we were still feeling our way at that time, more output was given than now seems sensible, producing two tables at each stage of the iteration, one to say, in effect, "Where are we now?", the other to say "What are we going to do about it?" There is really no need for any output for those iterations that do not elect or exclude any candidate, so immediate output has been cut down to just showing the names of candidates elected or excluded as those events occur, with storage in computer files of enough information to allow various forms of table to be easily produced when wanted.

There is also provision for an animated form of output, showing coloured lines on the screen performing the transfers of votes. This is deliberately slowed down to make it easy to watch.

6 Ties

In the event of a tie, where a candidate must be excluded and two or more are exactly equal in last place, [1] gave only a pseudo-random choice as the solution. In my implementation, I was persuaded by ERS Technical Committee to include the traditional 'ahead at first difference' criterion as a first tie-breaker, with a pseudo-random choice only if that did not solve it.

Strictly speaking this is contrary to Meek's stated principles on which his method is based, and was somewhat against my will, but it is unreasonable to expect to win every argument, and it does no real harm, particularly as ties hardly ever occur in real elections.

The pseudo-random method used is similar except that [1] calculated random numbers only if and when required. I have found it more convenient to assign such numbers to the candidates in the first instance and thus to have them already available if wanted. However I change the assigned numbers at each stage so that, if A is randomly preferred to B on the odd stages, then B is preferred to A on the even stages.

7 Election

In [1] candidates were not deemed elected until the end of an iteration. The keep values having converged, it was then considered whether any additional candidate had achieved the quota. Further thought has shown that it is absolutely safe to elect as soon as a candidate reaches the quota during the iterations and at once to start adjusting that candidate's keep value, along with those of any others already elected. This follows from Woodall's proof, given as part of [1], that if there is a feasible vector, then there is a unique solution vector — see that proof for the definitions of those terms.

8 Convergence

Both in [1] and my present implementation, the overall plan consists of iterations within iterations, the outer iterations being the operations up to and including the exclusion of a candidate, the inner iterations being the successive adjustments of keep values.

In [1] the inner iterations were taken as having converged when each elected candidate's votes were individually close enough to the current quota. This has been simplified to saying that the sum of the current surpluses of all the elected candidates must be no greater than 0.0001. It is almost certain in any case that, if such a small sum of all surpluses is ever reached, the lowest candidates are tied and further iterations would not separate them. Because of the short-cut exclusion rule mentioned below, however, it hardly ever happens that iterations need to proceed so far.

9 Short-cut exclusion rule

During the iterations, if it is found that the lowest candidate's current votes plus the total surplus of the elected candidates is less than the current votes of the next lowest candidate, it is certain that, if the iterations were continued all the way to convergence, that lowest candidate would necessarily still be the lowest and would have to be excluded. It is therefore safe to exclude the candidate at once. The next iterations will then start from a different point than would otherwise have been the case, but it follows from Woodall's proof that the next solution vector will still be the same, so the eventual result must be unchanged.

To see that, in these circumstances, the lowest candidate cannot catch up, it should be noted that the total number of votes remains unchanged and the effect of reducing the keep values of elected candidates is to pass their surplus votes to other candidates or, possibly, to non-transferable. If all the surpluses are passed to the lowest candidate, that candidate would necessarily, given the conditions, remain the lowest. If some are passed to other candidates that is even worse for the lowest, even if some of those candidates become elected.

The only point that needs more thought is to consider what happens if some surplus becomes non-transferable, resulting in a reduction of the quota. If n votes become non-transferable, the extra surplus created thereby is $mn/(s+1)$ where m is the number of elected candidates so far, and s is the number of seats. We know that m is less than s, because otherwise all seats are filled and the whole election is over. Therefore $mn/(s+1)$ is less than n, which shows that the amount that could have gone to the lowest candidate has been reduced.

Similar arguments show that, if two or more lowest candidates have a total number of votes that, together with the current surplus, is less than the votes of the candidate next above, it is safe to exclude them all at once, provided that enough would remain to fill all seats. I have not implemented this (except in the special case where several lowest candidates have zero votes) believing it to be simpler to explain what is going on if only one at a time is excluded.

With traditional style STV it is important that rules are firmly laid down as to whether or not multiple exclusions are to be made, because it can change the result. Thus, for example, Newland and Britton rules [3] insist that multiple exclusions must be made when possible, whereas Church of England rules [4] insist on only one at a time. With Meek rules, however, it is optional, as the result is necessarily the same either way. The fact that I exclude only one at a time is not intended to suggest that there is anything wrong, within a Meek system, with multiple exclusions if others wish to use them.

10 Equality of preference

Meek [2] suggested allowing voters to express equality of preference where desired. In [1] this option was not included. My program does include the option but there are some difficulties involved, as explained in detail in [5]. I continue to hold the conclusion expressed there that "the complications may be too many to be worth it ... [but] the facility is strongly valued by a significant number of electors".

11 Constraints

Not proposed by Meek, the program also allows constraints, whereby a maximum number, or a minimum number, may be laid down for certain categories among those elected. I dislike such constraints in principle [6], but they are necessary in certain circumstances in the Church of England [4] and, if the Church ever wished to update its procedures to use Meek-style STV, it would be necessary to demonstrate that it could cope with this additional complication.

At present the main thing for which constraints may be wanted is the filling of casual vacancies, where this is done by recounting the original votes with the late occupier of the vacant seat withdrawn. The constraint that is then necessary is to disallow exclusion of any existing seat holder.

12 STV in New Zealand

Those working on the introduction of STV for certain elections in New Zealand, having decided that the Meek rules were what they wanted, had my implementation available to them, and most of its details given above, such as the 9-decimal place working, and the figure of 0.0001 for the total surplus to indicate convergence, have been incorporated into their Act of Parliament [7].

There is, of course, no objection to these details having been used, but I hope that it will not become 'folklore' that they must be used and that Meek has not been properly implemented otherwise.

13 Acknowledgements

I thank both the editor and the referee for suggestions that led to substantial improvements in this paper.

14 References

[1] I.D. Hill, B.A. Wichmann and D.R. Woodall. Algorithm 123 – Single Transferable Vote by Meek's method. *Computer Journal*, 30, 277-281, 1987.

[2] B.L. Meek. A new approach to the Single Transferable Vote. *Voting matters*, Issue 1, 1-11, 1994.

[3] R.A. Newland and F.S. Britton. How to conduct an election by the Single Transferable Vote, 2nd edition. Electoral Reform Society, 1976.

[4] GS1327: General Synod. Single Transferable Vote regulations 1998.

[5] I.D. Hill. Difficulties with equality of preference. *Voting matters*, issue 13, 8-9, 2001.

[6] I.D. Hill. STV with constraints. *Voting matters*, issue 9, 2-4, 1998.

[7] New Zealand Government. Local Electoral Regulations 2001. Schedule 1A.

Editorial Postscript

I (Brian Wichmann) also have an implementation of the Meek algorithm, written in Ada 95. I have not made this generally available for several reasons: firstly, it lacks any system for preparing the data and even any adequate diagnostics on incorrect data (and hence is just a counting program); secondly the program has a number of extensions written to aid some investigations (typically reported in *Voting matters*); thirdly the program does not perform the arithmetic exactly correctly. There are a number of small differences between my Ada 95 version and the version in this paper; ties are broken differently and I will exclude several candidates together having the same number of votes provided it is safe to do so.

In 2000, I did perform a check of the Meek implementation described in this paper against the original version published in 1987 [1]. The report of this validation can now be found on the McDougall website. One interesting finding was that a test (M135) was actually a tie between two candidates for exclusion. However, both programs performed slightly different calculations in approximating the solution in such a way that neither reported a tie and the differences in the rounding resulted in a different exclusion. This is not considered a fault as, where there is really a tie, either result is acceptable.

Computerisation of STV counts

Robert Newland
(deceased)

This note, located by David Hill, appears not to have been published. It is dated February 1983. It is unclear why it was not published. Since it raises many interesting issues, it is reproduced here. Readers may wish to comment on the proposals. We hope to include their comments in a subsequent issue of *Voting matters* — Editor.

(1) It has often been suggested that STV counts should be computerised to save time/money. I have always regarded that view as unrealistic. Much of the time of any election count is taken up with preliminaries, such as envelope-slitting in postal ballots, unfolding voting papers, checking their authenticity, and, in public elections, reconciliation of numbers of papers issued.

With computerised counts, input would be time-consuming, whether by operators working in pairs to ensure accuracy, or whether by special equipment reading special voting papers presented in succession. Voting machines capable of accepting preferences seem an unlikely investment for infrequent public elections.

The time required for manual STV counts can be exaggerated, while any saving in time/money in computerised counts is doubtful or marginal. Unless there are other positive advantages to be gained from the computerisation of STV counts, it seems wrong to deprive candidates and others of the opportunity of witnessing manual counts.

(2) As Stephen Freeland said in his recent paper, COUNTING STV BY COMPUTER, "the existing 1976 procedures for counting STV elections represent a balance between technical refinement and speed of counting". Indeed, the 1976 procedures included improvements over earlier procedures both in technical refinement *and* in speed of counting. The current (1976) procedures are probably the best that can be achieved in manual counts.

Although little can be said in favour of computerisation of STV counts if the objective is merely the supposed saving of time/money, nevertheless, if computerisation is intended, the opportunity can be taken of incorporating improved counting procedures into STV which are not practicable in manual counts.

One minor improvement is obvious. It would be absurd to write a computer program restricting the calculation of quota, $V/(N+1)$, and of transfer values, to two decimal places. Using more decimal places would, on occasion, lead to a different, better, result. Since the results of manual and computer counts would then no longer be comparable, it would be sensible to make other improvements to achieve even better, different, results.

(3) In my COMPARATIVE ELECTORAL SYSTEMS where I was concerned primarily with the comparison of systems employing manual counts, I indicated briefly in section 7.8(c), Further Refinements, two areas of improvement not practicable in manual counts, viz., (i) the recommencement of counts from the beginning after exclusions, and (ii) the transfer of voting papers to next preferences even though already elected.

Stephen Freeland discusses the first of these in his paper. Following exclusion, often some voting papers are non-transferable. In consequence, towards the end of the count, candidates are elected without the quota: votes are of unequal effect.

The remedy is to re-commence the count ab initio after each exclusion. (A)

Non-transferable papers showing preferences only for excluded candidates would be discarded, and a new, lower, quota would be calculated. Eventually all candidates would be elected on attaining the same (lowest) quota: votes would be of equal effect.

Non-transferable papers showing preferences for already elected candidates would now be used to help elect those candidates: there would be fewer non-transferable papers.

Moreover, a well-known tactical voting ploy would be pre-empted. Suppose that in an election with quota 9, candidate A has 10 voting papers: 9 AB, 1 AC. The count proceeds thus:

A	10	−1	9
B	-	+0.9	0.9
C	-	+0.1	0.1

Under current rules, the elector who voted AC can maintain his support for A, but increase his support ten-fold for C by voting ZAC, where Z is not the elector's genuine first choice, but is believed to have little or no support. The count proceeds:

A	9		9
B	-		-
C	-	+1	1
Z	1	−1	-

There is an inherent danger that many such tactical voters might elect Z unintentionally.

Such tactical voting is pre-empted if the count is re-commenced after the exclusion of Z:

A	9	
B	-	
C	-	
Z	1	excluded.

New start:

A	10	−1	9
B	-	+0.9	0.9
C	-	+0.1	0.1

(4) In manual counts, it is standard practice, in transferring a consequential surplus, only to examine, and where appropriate transfer, those papers, all of one value, last received, which gave rise to the surplus. It is sometimes suggested that *all* the papers of an elected candidate should be examined and where appropriate transferred, since they all contributed to the existence of the surplus. This is an apparently attractive argument, but such a procedure, by itself, is unsound.

Suppose that in an election with quota 8, candidate A has 10 papers marked ABCD, B has 8 papers, and C has 7 papers. The count proceeds:

A	10	−2	8	
B	8		8	
C	7	+2	9	−1

It would clearly be unsound to examine and transfer any of the original 7 papers for C while the larger number of 8 papers for B have no further effect on the count. The 8 papers for B remain unexamined because B had already attained the quota, and the surplus of A was transferred, passing over B, direct to C.

The remedy is to transfer voting papers to next preferences even if already elected, thereby enabling all voting papers of an elected candidate to be examined when a consequential surplus is transferred. (B)

Electors would then be more equally represented.

Suppose in an election with quota 10, preferences for candidates A, B, C are shown on 30 voting papers: 20 AB, 10 BC. The count proceeds under existing rules thus:

A	20	−10	10
B	10		10
C	-		-
NT	-	+10	10

But if the surplus of A is transferred to the next preference B, the count proceeds:

A	20	−10	10		10
B	10	+10	20	−10	10
C	-		-	+10	10

The 30 electors with three quotas of votes have now elected three representatives.

The practical difficulty with this desirable procedure is that if part of the surplus of a candidate A is transferred to a candidate B, who is already elected, or may thereby be elected, part of B's surplus may be transferred to A, and then part of A's surplus to B, and so on indefinitely.

Brian Meek examined the problem in some detail in EQUALITY OF TREATMENT OF VOTERS AND A FEEDBACK MECHANISM FOR VOTE COUNTING, papers published in 1969 and 1970 in *Mathematiques et Sciences Humaines* (English language versions available).

Douglas Woodall also discusses the problem in COMPUTER COUNTING IN STV ELECTIONS in the current issue (Winter 1982-83 issue) of *Representation*.

To illustrate the effect of transferring votes between elected candidates, suppose that in an election with quota 12, candidate A has 18 papers, and candidate B has 10 papers. The papers for candidate A are marked: in case (i) 18 ABC (ii) 15 ABC, 3 A

(iii) 6 ABC, 12 A In each case the 10 papers for B are marked BAD.

Under existing rules, except for non-transferable differences, the result in each case is the same. The consequential surplus of B is transferred entirely to C, and D receives nothing:

A	18	−6	12		12
B	10	+6	16	−4	12
C	-		-	+4	4
D	-		-		-

If voting papers are transferred between A and B however, D receives votes in each case; fewest votes in case (i) when most papers show a (third) preference for C; most votes in case (iii) when fewest papers show a preference for C. In case (iii) the transfers soon terminate, but in the other two cases there is a theoretically unending alternation of transfers as the votes credited to A and B gradually converge to the quota. In practice, the calculations are terminated when a desired degree of accuracy is attained.

Details are appended. In case (iii) the transfers are worked out fully. In cases (i) and (ii) only the early alternations are shown [1].

It may be noted that I have followed principles which differ in some respects from both Meek and Woodall.

(5) If STV counts are to be computerised, it would be foolish not to include remedy (A), since to recommence the count after each exclusion requires only a little more computer time. If satisfactory computer programs can be devised, it would also be appropriate to include remedy (B), incorporating the procedures as illustrated.

A manual STV count is already immensely superior to any other method of election, votes being of nearly equal effect. Remedies (A) and (B) are designed to treat voting papers equally, and to ensure that votes are of exactly equal effect.

(6) This paper makes no suggestion to change the apparently obvious criterion of successively excluding candidates with fewest votes. I know of no better criterion.

The procedures described above will ensure that at most a quota of voters is not represented. Different criteria for exclusion would merely result in the non-representation of a different quota of voters.

[1] These details have been omitted here because Newland changed his mind later. When the members of ERS Technical Committee were arguing between three alternative ways of doing the job: Newland, Meek and Warren, he had another look at it and switched to supporting the Meek method as better than what he had proposed in this paper, so it is fairer to him to ignore his proposed method.

Comments on the STV Rules Proposed by British Columbia

Jeffrey C. O'Neill
jco8@cornell.edu

1 Introduction

In May 2005, the Canadian province of British Columbia conducted a referendum to decide whether to adopt the single transferable vote (STV) to elect the members of its legislative assembly. Although 57% of the electorate voted in favor of adopting STV, the measure was not adopted as a super majority of 60% was required for adoption. A Citizens' Assembly drafted a proposed set of STV rules, which will henceforth be called BC-STV. These rules are set forth in pages 17-20 of a Technical Report drafted by the Citizens' Assembly [1] and are also included as an appendix to this article.

The purpose of this article is to clarify the details of the BC-STV implementation and provide some insight into the rationale underlying the rules. Much of the information presented in this article has been gleaned from email conversations with James Gilmour, Jonathan Lundell, Brian Wichmann, and Joe Wadsworth. I have implemented the BC-STV rules in the software package called OpenSTV.[6]

2 Unitary and Inclusive Philosophies

The primary difference between different STV rules is in how surplus votes are transfered. The different methods for transferring surplus votes can be grouped into two different categories, what I call the unitary and inclusive philosophies of transferring surplus votes.

Before describing these two categories, a distinction must be made between an initial surplus of votes and a secondary surplus of votes. An initial surplus arises when a candidate has more than a quota of first choices, i.e., a surplus after the first stage of counting. A secondary surplus occurs when a candidate does not have an initial surplus but later goes over the quota after receiving votes from other elected or excluded candidates.

Consider an election where the quota is 100. Suppose candidate A has 140 votes after the first stage and thus an initial surplus of 40 votes. Suppose candidate B has 90 votes after the first stage and 110 votes after the second stage, after receiving 20 votes of A's surplus. At the second stage, candidate B has a secondary surplus of 10 votes.

Under the unitary philosophy of surplus transfers, only whole votes are transferred. With candidate A, 40 of her votes transfered at full value, while the other votes remain with A at full value. Similarly with candidate B, 10 votes are transferred at full value. A common practice is to take these 10 votes from the 20 that B received during the second stage.

Under the inclusive philosophy of surplus transfers, a portion of each of a candidate's votes is transferred. With candidate A, each of her votes will be transferred to their second choices at a transfer value of 40/140. The total value of the votes transferred is 40. The transfer is inclusive because each of A's votes takes part. With candidate B, the idea is the same, except that one could (and should) account for the fact that some of the votes that B received in the second stage could already have a value of less than one.[1]

Some STV rules can be clearly classified as exemplifying one of these two philosophies, while others employ a hybrid of these two philosophies. I will now consider several STV rules in addition to BC-STV: Cambridge STV (Massachusetts, USA), Dail STV (Ireland), Northern Ireland STV, Malta STV, Tasmania STV (Australia), Australian Capital Territory or ACT STV, and Meek STV (New Zealand).

Cambridge and Dail STV are examples of the unitary philosophy. With Cambridge STV, the votes selected for transfer are chosen at random. With Dail

[1] Under a method used in Australia, all votes are treated the same even if some of them were received at less than full value. In contrast, BC-STV appropriately weights the votes received at less than full value [4].

STV, the votes selected for transfer are chosen in a manner that proportionally represents the following choices on the ballots but does not seek to proportionally represent later choices on the ballots. Both of these methods are ballot order dependent – the outcome is not guaranteed to be the same if the votes are recounted with the ballots in a different order – a fact that some people find highly undesirable. David Robinson has proposed an interesting unitary STV rule that is ballot order independent (or nearly so).[5]

Northern Ireland, Malta, Tasmania, and ACT STV employ a hybrid of the two philosophies and each is an example of the long-established Gregory method of STV counting. The idea underlying these methods appears to be to exemplify the unitary philosophy to the extent possible but to also ensure that the rules are ballot order independent. With these rules, the method of surplus transfer is different for an initial surplus and a secondary surplus. An initial surplus is transfered according to the inclusive philosophy. While not impossible, it is difficult to transfer an initial surplus in a unitary fashion that is also ballot order independent. The method for transferring secondary surpluses is still hybrid, but much closer to being unitary. For secondary surpluses, only the last batch of received votes is considered. This last batch could arrive from a previous transfer of surplus votes or from the exclusion of a candidate. For example, consider candidate B from above. The last batch of votes has a total value of 20 and the surplus is 10. Each of the votes in this last batch is transfered to the next candidate on the ballot with a transfer value of 10/20.[2] The transfer is thus inclusive among the last batch but much more unitary than a completely inclusive transfer.

BC-STV and Meek STV are examples of the inclusive philosophy. For both initial and secondary surpluses, a portion of each vote is transfered to its next choice. The primary difference between BC-STV and Meek STV is the following: with BC-STV votes are transfered only to unexcluded candidates with less than a quota while with Meek STV votes are transfered to all unexcluded candidates. Meek STV is clearly a better method than BC-STV, but Meek STV requires a computer program to count the votes while BC-STV can be counted by hand.

3 Provenance of the BC-STV Rules

Over the years, rules similar to the BC-STV rules have been considered in numerous places. The Proportional Representation Society of Australia urged Australia to replace an existing STV method with a method similar to BC-STV[4]; Douglas Amy's book includes a method similar to BC-STV[2]; and the model statute on the website of the Center for Voting and Democracy (a United States organization) is similar to BC-STV. Rules similar to BC-STV rules have likely been independently derived numerous times, and I present two possible derivations.

Among people familiar with the different STV rules, Meek STV is generally regarded as the "best" set of rules for STV. The greatest difficulty with Meek STV is that it cannot be counted by hand. The most obvious simplification to Meek STV to make it hand countable is to not allow vote transfers to elected candidates. With this modification, Meek STV becomes very similar to BC-STV.

The Gregory method is another well-known method for counting STV elections, which has been used for more than a century. As described above, for secondary surpluses with the Gregory method only the last received batch of votes is considered. Some may regard this as unfair since the last batch of votes may be quite different from previous batches of votes.[4] Intuitively, it seems desirable to change the transfer of secondary surpluses so that all of the candidate's votes are considered and not just the last batch. With this modification, the Gregory method becomes very similar to BC-STV.

Farrell and McAllister used the term "weighted inclusive Gregory method" to refer to rules like the BC-STV rules, and the drafters of the BC-STV rules also used this terminology.[3] While this terminology is perhaps descriptively correct, I find it misleading in that it overstates the relationship between the BC-STV and Gregory methods. Using only the last batch of votes in transferring secondary surpluses is a distinctive feature of the Gregory method. Without last-batch transfers, the similarity with the Gregory method is mostly lost. The BC-STV rules could also be described as "hand-countable Meek" or "Meek without transfers to elected candidates." A more accurate description of the BC-STV rules is simply "inclusive STV."

[2] For the sake of simplicity, I am assuming that each of the votes has a valid next choice.

[3] Farrell and McAllister appear to have coined this terminology.[4]

4 Corrections to the BC-STV Rules

Several people have pointed out ambiguities and errors in the BC-STV rules. I believe that they are all straightforward to address, and I will briefly do so.

First, the BC-STV rules necessarily entail computations with fractions. The rules do not say if these computations are to be performed exactly or through precisely-specified rounding techniques. While this is an important detail, it is one that can easily be resolved. In my implementation of the BC-STV rules, I round to eight decimal places to approximate an exact solution [6].

Second, there is one clear error in the rules, but this error has a simple and obvious fix. In the appendix, the underlined text has been added to fix this error.

Third, in two places, the rules need to be generalized. First, in part 8 of "Counting procedure rules," the rules acknowledge that it is possible for one candidate to be elected with less than a quota of votes. In reality, it is possible that multiple candidates could be elected with less than a quota of votes. One possible correction would be to delete the second sentence in part 8 and replace it with the following: "When the total number of elected and remaining candidates is equal to the number of members to be elected, then all the remaining candidates are elected even if they have less than a quota of votes." Second, part 3 of "Provisions for tied votes" explains how a tie between two candidates is to be broken, and this needs to be generalized to break a tie among three or more candidates.

Fourth, the BC-STV rules do not precisely specify how to transfer surplus votes. Suppose that two candidates have a surplus on the first count, that after transferring the largest first-count surplus a third candidate is elected, that after transferring the second first-count surplus a fourth candidate is elected, and that the fourth winner has a larger surplus than the third. The rules do not indicate which of the two remaining surpluses is to be transferred first. One could choose the largest surplus (that of the fourth winner) or the earliest surplus (that of the the third winner). In accordance with common practice, I chose to always transfer the largest surplus.

5 Advantages and Disadvantages of the BC-STV Rules

I see four advantages of the BC-STV rules: (1) the rules are very simple, (2) votes can be counted by hand, (3) the rules employ the inclusive philosophy, and (4) the rules avoid the unfairness of transferring only the last batch for secondary surpluses. Only the fourth advantage requires more explanation. Consider candidate B, described above. He received 90 first place votes and later received 20 votes that had been transfered as part of candidate A's surplus. It is quite possible that the latter 20 papers represent quite different views than the first 90 papers, yet only the latter 20 papers have further effect. This hardly seems fair to the 90 voters who ranked B first. Farrell and McAllister cite such a dispute arising from an Australian election where the Gregory method was used.[4]

I see one main disadvantage of BC-STV rules. The outcome of the count is not continuous in the sense that changing only one vote can dramatically affect the outcome. For example, consider the following two sets of ballots for electing three candidates:

Set 1	Set 2
4501 ABC	4500 ABC
2499 BD	2500 BD
1200 C	1200 C
1800 D	1800 D

The quota is 2500, and the two sets of ballots differ by just one vote. I now count these ballots using BC-STV rules.

With Set 1, candidate A is elected and has a surplus of 2001 votes. Since candidate B is second on all of these ballots and candidate B has less than a quota, candidate B receives all of these 2001 votes. Now B has a total of 4500 votes and a surplus of 2000 votes. For these 4500 votes, 2001 rank C next (the ballots transfered from A) and 2499 rank D next. Thus,

$$\frac{2000}{4500} \times 2001 = 889.3$$

ballots of the surplus go to candidate C, and

$$\frac{2000}{4500} \times 2499 = 1110.7$$

ballots of the surplus go to candidate D. Candidate D is elected with 2910.7 votes and candidate C loses with 2089.3 votes.

Now consider Set 2. Candidate A is elected and has a surplus of 2000 votes. Since candidate B is also elected, A's surplus of 2000 votes goes directly to candidate C. Thus, candidate C wins with 3200 votes and candidate D loses with 1800 votes. Although there is only one different ballot in these two sets, the outcome differs by more than 1000 votes.

In comparison, with all of the other STV counting methods mentioned in this paper, there is no such discontinuity with these two sets of ballots. For example, let us count the two sets of ballots with the Gregory method. With Set 1, A's surplus of 2001 votes goes to candidate B. B now has a surplus of 2000 votes. Only votes from the last batch are further transfered, so 2000 votes are now transfered to candidate C who wins with 3200 votes. With Set 2, A's surplus of 2000 votes goes directly to candidate C who again wins with 3200 votes. Here, the change in one ballot produced a similarly small change in the outcome.

6 Conclusions

In considering the relative merits of BC-STV and Gregory methods, there is no clear winner. With the Gregory method, one can argue that it is unfair to use only the last batch of received votes in transferring secondary surpluses. With BC-STV, the outcome is not necessarily continuous with small changes in the ballots. The clear solution to this conundrum is to use Meek STV, assuming that computer counts are possible, which does not suffer from either of these disadvantages.

7 References

[1] Making Every Vote Count. The Case for Electoral Reform in British Columbia. Technical Report. December 2004. Available via McDougall web site.

[2] Douglas J. Amy, *Behind the Ballot Box* (2000).

[3] Center for Voting and Democracy. A Sample Statute for Choice Voting. Available via McDougall web site.

[4] David M. Farrell and Ian McAllister, *The 1983 Change in Surplus Vote Transfer Procedures for the Australian Senate and its Consequences for the Single Transferable Vote*, Australian Journal of Political Science 38:479-491 (2003).

[5] Dave Robinson. Courageous STV. Description available via McDougall web site.

[6] OpenSTV – Software for computing the Single Transferable Vote. Available at http://stv.sourceforge.net.

Appendix: The Recommended BC-STV Electoral System

[Author's note: James Gilmour pointed out a small but important error in the counting rules. This has been fixed with the addition of the underlined text. I have also corrected the incorrect numbering in the section "Provisions for tied votes."]

This section describes the recommended BC-STV system. It provides guidelines to be used in drafting a new election act and in making changes to the current Electoral Boundaries Commission Act.

In addition to choosing an electoral system that incorporates its basic values, the Citizens' Assembly on Electoral Reform wanted a system that is open to public scrutiny and whose results can be reviewed and validated. Consequently, BC-STV is designed to use paper ballots which are available for recount, if required.

General

1. BC-STV is a system of proportional representation by the single transferable vote (STV) method.

2. The members of the Legislative Assembly of British Columbia will be elected from multi-member electoral districts.

3. The number of members in each district will vary from two (2) to seven (7). Given that achieving proportional electoral outcomes is a primary reason for recommending BC-STV, using larger rather than smaller numbers of members per district should always be preferred when drawing district boundaries. While some very sparsely populated areas may require districts with as few as two members, the principle of proportionality dictates that, in the most densely populated urban areas, districts should be created at the upper end of the range.

4. The "Droop quota" will be the formula for calculating the number of votes required by a candidate for election in a district. The quota formula is:

$$\left(\frac{\begin{array}{c}\text{total number of valid}\\ \text{ballots cast in the district}\end{array}}{1 + \begin{array}{c}\text{number of members}\\ \text{to be elected}\end{array}} \right) + 1$$

   Fractions are ignored.

5. The method of distributing surplus votes from those candidates with more than the minimum number of votes needed to be elected will be the "Weighted Inclusive Gregory method" (see below, as well as Appendix: Glossary [Author's note: the Glossary is not included.]).

The ballot paper

1. The ballot paper will display the names of all the candidates contesting seats for a district. The names will be grouped according to party affiliation.

2. Candidates who do not indicate a party affiliation, and candidates who do not indicate that they are running as an independent, will be grouped together.

3. Parties with only one candidate, and each candidate running as an independent, will each have their own group.

4. Groupings with more than one candidate in a district will have the rank order of the candidates' names rotated at random so that each candidate has an equal chance of being placed in every position within the grouping.

5. The rank order of groupings appearing on the ballot will be rotated at random so that each grouping has an equal chance of being placed in every position on the ballot paper.

6. The ballot paper will not provide the option of voting for all the candidates of one group by marking a party box (this is the so called "above the line" option used in some Australian elections).

Valid ballots

1. Voters will indicate their preference for the candidates listed on the ballot paper by putting the numbers 1, 2, 3, 4, etc. next to candidates' names.

2. A ballot paper must include a first preference for the ballot to be counted as a valid ballot. The number of subsequent preferences marked on the ballot is at the discretion of the voter.

3. In the case of a ballot paper with gaps or repetitions in the sequence of numbers beyond a first preference, the preferences are valid up to the break in the sequence.

4. If a voter puts a mark next to only one candidate's name, and that mark makes the voter's intention clear, the mark will be accepted as the expression of a single preference for that candidate and the ballot will be counted as a valid ballot.

Counting procedure rules

1. Once the total number of valid ballots is established in each multi-member district, the minimum number of votes required for a candidate to be elected is calculated using the Droop quota formula.

2. All ballots are counted and each ballot is allocated as a vote to the candidate against whose name a first preference (i.e., "1") is shown on the ballot.

3. If a candidate(s) on the first count has a number of first preference votes exactly equal to the minimum number of votes needed to be elected, then that candidate(s) is declared elected and the counted ballot papers indicating that candidate(s) as a first preference are put aside and the other preferences recorded on the ballots are not examined.

4. If a candidate on the first count gains more than the minimum number of votes needed to be elected, the candidate is declared elected, and the number of votes in excess of the number of votes needed to be elected (the surplus) is recorded. All of the elected candidate's ballots are then re-examined and assigned to candidates not yet elected according to the second preferences marked on the ballots of those who gave a first preference vote to the elected candidate. These votes are allocated according to a "transfer value." The formula for the transfer value is:

$$\frac{\text{surplus votes cast for the elected candidate}}{\text{total number of votes received by the elected candidate}}$$

5. If two or more candidates on the first count gain more than the minimum number of votes needed to be elected, all of those candidates are declared elected. The ballots of the candidate with the largest number of first preference votes will be re-examined first and assigned (at the transfer value) to candidates not

yet elected according to the second preferences marked on that candidate's ballots, or the next available preference, if the second preference candidate has already been elected. The ballots of the other elected candidate(s) will then be re-examined and their surpluses distributed in order according to the number of first preference votes each candidate received.

6. If a candidate reaches more than the minimum number of votes needed to be elected as the consequence of a transfer of votes from an elected or excluded candidate, the number of votes in excess of the number of votes needed to be elected (the surplus) will be transferred to other candidates. This transfer will be to the next available preference shown on all of this candidate's ballots. These ballots now include 1) the candidate's first preference ballots, and 2) the parcel(s) of ballots transferred to the candidate from one or more elected or excluded candidates. The transfer value for the candidate's first preference ballots is:

$$\frac{\text{surplus votes cast for the elected candidate}}{\text{total number of votes received by the elected candidate}}$$

The transfer value for each parcel of ballots transferred to the candidate from one or more elected or excluded candidates is:

$$\left(\frac{\text{surplus votes cast for the candidate}}{\text{total number of votes received by the candidate}}\right) \times \left(\text{the transfer value of the parcel of ballots received by the candidate}\right)$$

7. If no candidate has a number of votes equal to or greater than the minimum number of votes needed to be elected, the candidate with the smallest number of votes is excluded. All of that candidate's ballots–both first preference ballots and any parcel or parcels of ballots transferred from other candidates–are transferred to candidates who have not been elected or excluded according to the next available preference shown on the excluded candidate's ballots. The excluded candidate's first preference ballots are transferred to the second (or next available) preferences at full value. Ballots received from previously-elected (or excluded) candidates are transferred at the transfer value at which the ballots were received.

8. Counting continues in the described sequence: the surplus of elected candidates is assigned until no more candidates are elected, then the ballots of excluded candidates are assigned until another candidate is elected. When all but one of the candidates to be elected from the district have been elected, and only two candidates remain in the count, the candidate with the most votes is declared elected, even though the candidate may not have reached the minimum number of votes (the quota) needed to be elected.

9. If, during the transfer of preferences, a ballot paper does not indicate an available preference, the ballot is put aside as "exhausted." This can occur because:

- the voter only indicated one, or a small number of preferences;
- all the preferred candidates have already been elected or excluded; or
- there are gaps or repetitions on the ballot in the sequence of numbering preferences.

Provisions for tied votes

1. Where two or more candidates have the same number of first preference votes at the end of the first count, and this number is more than the minimum number of votes necessary for election, then the candidate whose surplus is distributed first will be decided by lot.

2. Where no candidate has a number of first preference votes equal to or greater than the number of votes necessary for election at the end of the first count, and two or more candidates have the same number of first preference votes, this number being the smallest number of first preference votes gained by any candidate, then the candidate who is excluded first will be decided by lot.

3. If, at any stage of the count other than during the first count, two candidates have the same number of votes, the candidate who is declared elected first, or who is not excluded will be:

a) the candidate with the larger number of votes in the previous or immediately next preceding count where there is a difference in the votes between the two candidates; or

b) the candidate whose name is drawn by lot, where there is no difference in the number of votes between the candidates at any preceding count.

By-elections

The single transferable vote method (preferential voting) is to be used for by-elections where a candidate is to be elected to fill a single casual vacancy in a district. The BC-STV method is to be used where candidates are to be elected to fill two or more casual vacancies in a district.

Developing STV Rules for manual counting to give effect to the Weighted Inclusive Gregory Method of transferring surpluses, with candidates' votes recorded as integer values

James Gilmour
jgilmour@globalnet.co.uk

The Local Governance (Scotland) Act 2004 [1] makes provision for councillors in Scotland to be elected by the single transferable vote (STV) from wards returning either three or four councillors. The first elections under these new provisions will be held in May 2007. The Act does not specify any STV counting rules, but requires Scottish Ministers to make such rules by order.

1 Proposal to use WIGM

When the Local Governance (Scotland) Bill [2] was introduced in the Scottish Parliament it included most (but not all) of the STV counting rules used for District Council elections in Northern Ireland [3]. Among those included were the provisions for the transfer of surplus votes by the Gregory Method, applied only to the 'last parcel' of ballot papers for a consequential surplus [4]. During the Stage 1 consideration of the Bill by the Local Government and Transport Committee of the Scottish Parliament, several MSPs questioned the use of the Gregory Method and suggested that the 'last parcel' provision treated some voters unfairly (eg see [5] at col 380). The Committee also discussed the possibilities of using electronic counting which was attractive because the elections for the Scottish Parliament (by a regional version of the Additional Member System) would be held on the same day.

In their Report [6] on the Stage 1 consideration of the Bill, the Committee said, in relation to technical issues surrounding the counting of votes:

> "The Committee: Concludes that the method set out in the Bill is the most appropriate one for local government elections in Scotland at this time, given the currently available counting technology;
> Believes that its preferred alternative, the 'weighted inclusive Gregory method', is, theoretically, the most effective counting method as it ensures that the preferences expressed by all voters are counted; but notes manual counts using this system would be unrealistically time consuming; and
> Recommends that the 'weighted inclusive Gregory method' be introduced to replace the system set out in the Bill when electronic counting becomes available."

Several technical amendments to the STV counting rules were discussed during the Stage 2 debate on the Bill, but the Gregory Method and the 'last parcel' provision were retained for the transfer of surpluses. However, at the Stage 3 debate on the Bill, on the floor of the Parliament Chamber, the Scottish Executive Minister tabled amendments that had the effect of removing all the detailed STV counting rules, and these amendments were passed [7,8].

The second Newsletter of the 2007 Elections Steering Group [9] announced: "Scottish Executive Ministers have agreed that work should go forward on the possibility of introducing e-counting for the 2007 local government elections." The invitation to tender for the provision of e-counting facilities was issued in August 2005 [10]. (The award of this contract to DRS Data Services Ltd was announced in February 2006 [11].)

The tender document issued to interested contractors [12] specified that the STV counting rules were to be based on the "Weighted Inclusive Gregory Method" (WIGM) of transferring surpluses. The tender document included a description of STV rules incorporating WIGM, based on the incomplete and defective description given in the Technical Report of the British Columbia Citizens' Assembly on Electoral Reform [13].

2 Definition of WIGM

The term "Weighted Inclusive Gregory Method" appears to have been coined by Farrell and McAllister [14], where they give the following description of the procedure for determining the transfer value for a candidate's surplus votes:

"For those votes that the candidate has received at full value, TV = s/v, where v is the candidate's total vote. For those votes that the candidate has received from another candidate's surplus, TV = $(s/v)\beta$, where β is the TV that was applied in the transfer of the surplus votes to the previous candidate."

(The definitions of "TV" and "*s*" were given earlier in the paper: "TV" = transfer value; "*s*" = candidate's surplus.)

The Weighted Inclusive Gregory Method has not yet been implemented anywhere in the world and so there is no working legislative language available. However, a legislative description of WIGM was included in the Electoral Legislation Amendment Bill 2003 presented to the Legislative Assembly of the Parliament of Western Australia [15]:

> "Unless all the vacancies have been filled, the surplus votes (if any) of any candidate elected under clause 4, or elected subsequently under this clause, shall be transferred to the continuing candidates as follows —
>
> (a) the number of surplus votes of the elected candidate shall be divided by the number of votes received by him and the resulting fraction shall be the surplus fraction;
>
> (b) in relation to any particular ballot papers for surplus votes of the elected candidate, the surplus fraction shall be multiplied by the transfer value at which those ballot papers were transferred to the elected candidate, or by one if they expressed first preference votes for the elected candidate, and the product shall be the continued transfer value of those particular ballot papers;
>
> (c) the total number of ballot papers for surplus votes of the elected candidate that each
>
> (i) express the next available preference for a particular continuing candidate; and
>
> (ii) have a particular continued transfer value,
>
> shall be multiplied by that transfer value, the number so obtained (disregarding any fraction) shall be added to the number of votes of the continuing candidate and all those ballot papers shall be transferred to the continuing candidate,
>
> and if on the completion of the transfer of the surplus votes of the elected candidate to a particular continuing candidate that candidate has received a number of votes equal to or greater than the quota, that candidate shall be elected."

(The Bill received a first and second reading, but was withdrawn in November 2003 for reasons not related to the proposed change to the STV counting rules.)

This legislative description introduces the term "surplus fraction" for Farrell and McAllister's calculated "*s/v*", which is then applied to each parcel of ballot papers with a different current value, Farrell and McAllister's "β", ie the "transfer value" at which those ballot papers were received by the candidate with the current surplus. The Western Australian Bill used the term "continued transfer value" for the value at which the ballot papers would be transferred from the candidate with the current surplus. In UK STV rules we prefer the term "current value" for whatever value a ballot paper may have when a calculation is made and "transfer value" for the value at which the ballot paper will be transferred to the next available preference.

3 Putting WIGM into UK legislation

The terminology of the Western Australia Bill is helpful in that it distinguishes (and names) the two steps in the process of calculating correctly weighted transfer values when a candidate has a surplus and all of that candidate's ballot papers are transferred. This legislative language does not, however, provide 'ballot-paper-by-ballot-paper' handling instructions of the kind usually found in UK rules for the conduct of STV counts (eg [3]). It was with this in mind that I prepared the detailed rules in the document that has been deposited on the McDougall website [16]. That document has been through several drafts and I am grateful to Brian

Wichmann, David Hill, John Curtice and the anonymous Referee of this paper for corrections and helpful comments. It has been made widely available to those who are involved in the preparation of the secondary legislation that will be required for the 2007 elections.

Although the intent was that e-counting would be used for the 2007 elections, and the Local Government and Transport Committee of the Scottish Parliament recommended the use of WIGM only if e-counting were to be introduced, there was nothing to indicate that manual counting by WIGM rules should not be undertaken if this were demanded or necessary. A manual count by WIGM rules would take longer than a manual count by (classical) Gregory Method rules because more ballot papers have to be sorted and counted more times, but it would not be impracticable for a public election as an exceptional requirement. It seemed appropriate, therefore, to devise first the WIGM rules for a manual count. Once these had been determined as coherent and unambiguous, it would be a smaller task to adapt the manual rules for e-counting. As explained in the preamble [16], the rules were written to fit into a more comprehensive legislative document and follow the conventions of UK secondary legislation (eg [3]).

4 Consequential issues

The essential description of WIGM is quite simple, but its adoption raises several issues that affect other aspects of the STV counting rules.

Because surpluses are to be spread across all the ballot papers then held by the candidate from whom the surplus is being transferred, each ballot paper will, in most cases, carry forward a smaller vote value. In the Northern Ireland rules [3], transfer values are calculated to two decimal places and any remainder ignored. The votes transferred to successive preferences are similarly calculated to two decimal places and the totals of votes credited to candidates are shown to two decimal places on the result sheet. If the WIGM calculations were similarly truncated at two decimal places, substantial numbers of ballot papers would quickly have no recordable value. The precision of calculation must, therefore, be greater when WIGM rules are applied. To ensure reproducibility no matter how the count is undertaken, it is necessary also to specify the precision of each step of each calculation. As explained in the preamble to the rules, the precision was set at seven decimal places on pragmatic and practical grounds. (The information about the precision of the transfer value calculations in the STV elections to the Australian Federal Senate taken from the AEC website and given in an earlier paper [17] was incorrect [18]. For those STV elections the precision is not limited at all [19], but this has no consequences because of the 'value averaging' method that is used in those rules to calculate transfer values *de novo* for each surplus.)

As noted in the document deposited on the McDougall website, these rules do not make any provision to overcome the anomaly that arises with WIGM when votes are not transferred to already elected candidates. This will be the subject of a separate paper.

5 Integer vote values

It is a feature of the Australian STV rules that use an 'inclusive' method of transferring surplus votes that only whole numbers of votes are credited to candidates when transfers are made [20]. The Commonwealth Electoral Act 1918 prescribes the flawed "Inclusive Gregory Method" and not the Weighted Inclusive Gregory Method, but the Western Australian WIGM Bill [15] included the same provision (see sub-paragraph (c) in the text quoted above). This approach has much to commend it, as it will simplify the result sheet and so aid public comprehension. (It would probably be of benefit if it were adopted more widely for STV counting rules.) Apart from its presentational advantages, this approach avoids acceptability problems that could arise in WIGM elections from candidates being separated by minute fractions of votes. With integer vote totals, candidates will either be separated by at least one vote or have the same number of votes.

Of course, the fractional parts of the vote totals that are not transferred to the candidates cannot be ignored; they must be accounted for properly. These fractional parts are shown separately on the Australian integer result sheets as 'Lost by fraction'. I prefer the term 'Vote fraction not transferred' because it is more correctly descriptive and does not convey the idea that any votes can be "lost".

This truncation to an integer value is applied only to the total value of all the parcels and sub-parcels being transferred to any one candidate; it is not applied to the values of the individual parcels and sub-parcels before the candidate's transferable total is calculated. There is only one truncation for each candidate to whom votes are transferred in any one stage. That way the 'Vote fraction not transferred' is minimised.

Note that when a multiple exclusion occurs, the 'Vote fraction not transferred' can be negative. This happens when the sum of the values of the ballot papers, **including all the fractional parts**, held by the excluded candidates exceeds the sum of the integer votes credited to the excluded candidates. Thus previously 'non transferred' votes can be brought back into play. This is another reason for preferring a term other than "lost".

6 Non-transferable votes

When an 'inclusive' transfer of a surplus is effected, the transfer values are calculated taking into account **all** the votes then credited to the elected candidate and **all** the ballot papers are transferred. Ballot papers with no 'next available preference' are set aside as 'non transferable' and take with them as 'non-transferable' the proportionate share of the surplus. This approach is wholly consistent with the 'inclusive' concept that is given effect by the requirement to examine and transfer all parcels of ballot papers held by the candidate with the surplus.

7 Deferred surpluses

It could be argued that the 'inclusive approach' that underlies WIGM would require the transfer of **all** surpluses, ie that there should be no provision to defer the transfer of any surplus, no matter how small. However, if there is to be any possibility of manual counting, it would be best to retain the 'deferred surplus' provision so that the handling of large numbers of ballot papers of extremely small values could be avoided except when the votes on those ballot papers would affect what has to happen next.

8 Sub-stages during exclusions

STV counting rules that use the Gregory Method of transferring surpluses usually provide for sub-stages during exclusions, in which the transfer of a parcel of ballot papers of the same value constitutes a sub-stage. The transfer of first preference ballot papers before the transfer of other ballot papers of value 1 vote also constitutes a separate sub-stage in the Northern Ireland rules [3]. If any candidate attains the quota at the end of a sub-stage, that candidate is 'deemed elected' and no further transfers are made to that candidate. This is consistent with the 'exclusive approach' to STV that seeks to keep the voters in discrete, 'exclusive' groups so far as possible. Although it is clearly not directly related to WIGM, the sub-stage approach to handling exclusions seems incompatible with the 'inclusive' approach that underlies WIGM. I have, therefore, made no provision for sub-stages during exclusions.

9 Publication of results

I have taken the opportunity to specify fully what must be published once an STV count has been completed. This rectifies a deficiency in the Northern Ireland rules [3].

10 Casual vacancies

The suggested rules do not include any provisions relating to the filling of casual vacancies because policy decisions on casual vacancies are required before the relevant election rules can be devised. Should it be decided that a by-election must be held when a single vacancy occurs, I would commend the use of the special purpose STV rules published by the Electoral Reform Society [21]. I codified these rules in their present form in 1978, working under the guidance of Frank Britton and Robert Newland.

11 'Inclusive' and 'exclusive' representation

A discussion of the 'inclusive' and 'exclusive' approaches to proportional representation and STV counting rules will be the subject of a separate paper.

12 References

[1] Local Governance (Scotland) Act 2004. Royal Assent 29 July 2004. Available via the McDougall web site.

[2] Local Governance (Scotland) Bill. Introduced 21 November 2003. Available via the McDougall web site.

[3] The Local Elections (Northern Ireland) Order 1985. Statutory Instrument 1985 No. 454.

[4] Nanson, E. J. (1900) The real value of a vote and how to get it at the coming Federal elections. Melbourne, J. T. Picken. quoted by: McLean, I. (2002) Australian electoral reform and two concepts of representation. Available via the McDougall web site.

[5] Local Government and Transport Committee Official Report 9 December 2003. Available via the McDougall web site.

[6] Local Government and Transport Committee 2nd Report 2004 (Session 2). Stage 1 Report on the Local Governance (Scotland) Bill Volume 1: Report. Available via the McDougall web site.

[7] Local Governance (Scotland) Bill: Marshalled List of Amendments selected for Stage 3. Available via the McDougall web site.

[8] Scottish Parliament Official Report (Session 2) 23 June 2004, Col. 9392

[9] Second Newsletter of the Local Elections 2007 Steering Group – July 2005. Available via the McDougall web site.

[10] UK-Edinburgh: data capture services 2005/S 171-170043 Contract Notice Scottish Executive Tender for the Provision of Electronic Counting Services for Elections. Available via the McDougall web site.

[11] UK-Edinburgh: data capture services 2006/S 36-038769 Contract Award NoticeProvision of Electronic Counting Services for Elections. Available via the McDougall web site.

[12] Scottish Executive (undated). Invitation to Tender for the Provision of Electronic Counting Facilities for Elections. Contract Reference: QLC 17/1Schedule 2: Statement of Requirements and Appendixes D and E.

[13] Making Every Vote Count : the case for electoral reform in British Columbia : the British Columbia Citizens' Assembly on Electoral Reform technical report. December 2004. Available via the McDougall web site.

[14] Farrell, D. M. and McAllister, I. (2003) The 1983 Change in Surplus Vote Transfer Procedures for the Australian Senate and its Consequences for the Single Transferable Vote. *Australian Journal of Political Science, Vol. 38, 479–491.* Available via the McDougall web site.

[15] Western Australia: Electoral Legislation Amendment Bill 2003. Section 39. Available via the McDougall web site.

[16] Gilmour, J. (revised 1 July 2006) STV Rules for manual counting to give effect to the Weighted Inclusive Gregory Method of transferring surpluses, with candidates' votes recorded as integer values. Available via the McDougall web site.

[17] Gilmour, J. (2003) Calculation of Transfer Values - Proposal for STV-PR Rules for Local Government Elections in Scotland. *Voting matters*, Issue 17, October 2003, 20-24

[18] Pugh, R. (March 2006) (personal communication) Australian Electoral Commission, Research Section.

[19] Australian Electoral Commission (2006). The Senate Voting System: Senate Theory Training. 7 June 2006, pp 40.

[20] Australia : Commonwealth Electoral Act 1918 - Section 273: Scrutiny of votes in Senate elections : paragraph (9)(b). Available via the McDougall web site.

[21] Electoral Reform Society : Regulations for the Election of One Person by means of the Single Transferable Vote. Available via the McDougall web site.

ISSN 1745-6231

Voting matters

for the technical issues of STV

published by

The McDougall Trust

Issue 23 February 2007

Contents

About the McDougall Trust (reg. charity no. 212151)

The McDougall Trust is a charitable trust formed in 1948. The charity's purposes as stated in its governing scheme of 1959 are to advance knowledge of and encourage the study of and research into:

- political or economic science and functions of government and the services provided to the community by public and voluntary organisations; and
- methods of election of and the selection and government of representative organisations whether national, civic, commercial, industrial or social.

The Trust's work includes the maintenance and development of the Lakeman Library for Electoral Studies, a unique research resource, the production and publication of Representation: The Journal of Representative Democracy, and, of course, this publication **Voting matters**, that examines the technical issues of the single transferable vote and related electoral systems.
For further information on the Trust, please contact:

The Secretary,
McDougall Trust,
6 Chancel Street,
London SE1 0UX, UK.
Telephone: +44 (0)20 7620 1080
Facsimile: +44 (0)20 7928 1528
Email: admin@mcdougall.org.uk
Web: www.mcdougall.org.uk

For further information on this publication, please contact B A Wichmann, the Editor at the above address or by email at: `Brian.Wichmann@bcs.org.uk`

Editorial

There are 5 papers in this issue, all of which are comments or reviews of other work:

- David Hill: *Comments on Newland's paper*

  Here, David Hill responds to some specific technical points in Newland's paper.

- *Edited comments on Robert Newland's suggestions*.

  Robert Newland's article, written in 1983 made many suggestions which were thought to be an appropriate topic of a moderated email discussion. A heavily edited version of this discussion appears here. It points to a number of topics which could well be the subject of future papers in *Voting matters*.

- Brian Wichmann: *Review of The Machinery of Democracy*

  The report reviewed here is one undertaken by leading experts in the US to show what is needed to avoid some of the problems that occurred during the Presidential election of 2000.

  Parts of this report are relevant to the use of scanning machines for the Scottish local elections to be held later this year. The US Freedom of Information Act ensures that electoral data is open to public scrutiny, whereas the position in Scotland is uncertain at this point. This implies that the transparency of the Scottish STV elections might be less than those of Northern Ireland for which manual procedures are used.

- Jonathan Lundell: *Review of the Second Report of the Irish Commission on Electronic Voting*

  The Irish Commission has completed its work with its second report. It is unclear at this stage what action the Government will take. This report has some similarities with the previously mentioned US report which makes for some interesting comparisons.

- David Hill: *Review of Collective Decisions and Voting by Nicolaus Tideman*

  The book reviewed here is central to many of the issues covered in *Voting matters*, and hence this review should be of interest to many of our readers.

Scotland

The final stages of the legal process for the local STV elections in Scotland have been agreed. The counting method is based upon the Weighted Inclusive Gregory Method, but is as simple as it could be in computer terms. Hand counting using this logic is possible, but would take longer than current manual counts because of the need to examine all of the elected candidate's papers when a surplus is transferred. It is interesting to contrast this with the Meek method, which is more complex, since the quota is recomputed and transfers are made to elected candidates. In electoral terms, Meek has the advantage that the intervention of a no-hope candidate cannot change the choice of the elected candidates — a failing of all the rules used for current hand-counting STV methods.

The Order approved by the Scottish Parliament at the end of January will require the Returning Officers to publish much fuller details about votes and transfers of votes at each stage of the count than the corresponding legislation for STV elections in Northern Ireland. However, the rules strangely include a requirement to publish the numbers of non-transferable papers at each stage but not the numbers of non-transferable votes. That vote is needed because, with WIGM, the non-transferable papers will have different values when they become non-transferable.

Because the ballot papers will be scanned and counted electronically, there is a new requirement for one copy of the electronic information so obtained to be kept for four years after the count, while the paper records need to be kept for only one year, as usual. However, it is most regrettable that the release of any of the electronic information, even in anonymous form, is specifically prohibited. One ray of hope for a more enlightened approach is that the Scottish Executive has given an undertaking to consult on this. I certainly hope that full preferential data will be made available because that would be in everyone's best interest.

Comments on Newland's paper

I.D. Hill
d.hill928@btinternet.com

1 Introduction

Like all work published posthumously, if there are any faults in this paper [1], the author should not be blamed for them because, had he lived longer, he might well have revised it, or even withdrawn it. The paper is important as showing Newland supporting some of the main features of the Meek method. It is a pity that he did not support all of them, but his disagreement with the Meek method of handling short votes gets no mention here.

It is easy to agree with him that to think of saving time or money, as a result of computer counting, is unrealistic, but he fails to mention other advantages of counting by computer, even if the rules remain those of hand-counting methods. These advantages are that, given a correct computer program: (i) anybody can carry out an STV election without having to understand the rules; (ii) the results are much more likely to be correct, provided that due care is taken in converting the ballot paper information to a computer file. Such evidence as is available suggests that STV hand-counts, even by experienced staff, usually have errors in them.

His saying that "It would be absurd to write a computer program restricting the calculation ... to two decimal places" is therefore not correct. Where existing systems require the two-decimal place restriction, doing it by computer, for the sake of a correct result within those rules, is worth while.

He says that "Using more decimal places would, on occasion, lead to a different, better, result". Although the words "on occasion" need to be noticed, I take his meaning to be that on occasion there will be a difference but, if there is, it will necessarily be a difference for the better. Whether that is so depends upon how "better" is defined. In the hope of avoiding controversy, let us take it to mean, in the context of Newland's paper, "more like the result that would have been obtained by adopting remedies (A) and (B) of the paper". Such work as I have done on it suggests that merely more precision in the calculations does not help to that end.

2 Remedies (A) and (B)

Newland's "Remedy (A)" is to re-commence the count *ab initio* after each exclusion; his "Remedy (B)" is to transfer voting papers to next preferences even if already elected. He says that "If STV counts are to be computerised, it would be foolish not to include remedy (A)". He appears not to have realised that to include (A) without (B) can be troublesome. I take it that he was thinking in terms of the rules of Newland and Britton 2nd edition [2] and adding remedy (A) to those, so I shall do so in the following examples.

2.1 What is wrong with Remedy (A) on its own

Example 1

Suppose 8 candidates for 5 seats, with votes

25	ACDF..
24	BCEF..
7	D..
5	E..
2	F..
6	G
3	HBC

We get a quota of 12 and the count proceeds as:

A	25	−13	12		12		12
B	24		24	−12	12		12
C		+13	13		13		13
D	7		7		7		7
E	5		5	+12	17	−5.00	12
F	2		2		2	+4.80	6.80
G	6		6		6		6
H	3		3		3		3
n/t						+0.20	0.20

For this publication, see www.votingmatters.org.uk

Exclude H and restart:

A	25		25	−13	12		12		12
B	27	−15	12		12		12		12
C		+14.85	14.85		14.85		14.85	−2.85	12
D	7		7	+13	20	−8	12		12
E	5		5		5		5	+2.64	7.64
F	2		2		2	+8	10		10
G	6		6		6		6		6
n/t		+0.15	0.15		0.15		0.15	+0.21	0.36

Exclude G and restart. There are now 6 fewer valid votes, so the quota becomes 11:

A	25		25	−14	11		11
B	27	−16	11		11		11
C		+15.93	15.93		15.93		15.93
D	7		7	+14	21	−10	11
E	5		5		5		5
F	2		2		2	+10	12
n/t		+0.07	0.07		0.07		0.07

Thus E was deemed elected in the first count, and had a surplus transferred, but had to be unelected and take back that surplus for the second count. Finally E fails to get even half a quota and loses. It might be said that there is no need to say that anyone has been elected until the final result is known, but then how can the surplus transfer be explained, for without it F would have been excluded first instead of H?

Example 2

Suppose 8 candidates for 5 seats, with votes

25	ACDF..
24	BCEH..
7	D..
5	E..
2	F..
6	G
3	HBC

These are identical votes to example 1 except that 24 BCEF.. has been changed to 24 BCEH..

Following through the election in a similar way, those elected are found to be ABCDE. Thus E succeeds if those 24 vote BCEH but E fails if those 24 vote BCEF. So their choice of a later preference has upset the fate of their earlier preference. My memory of Robert Newland says that he would have hated that.

3 Conclusions

We must always remember that it is mathematically impossible to find a faultless system, so these faults of remedy (A) on its own are not necessarily conclusive, but they tell strongly against it. What would be safe would be to restart after each exclusion, provided that no candidate had yet been deemed elected.

4 Acknowledgement

I thank the referee for some very helpful comments.

5 References

[1] R.A. Newland. Computerisation of STV counts. *Voting matters*, issue 22, 11-13, 2006.

[2] R.A. Newland and F.S. Britton. How to conduct an election by the Single Transferable Vote, 2nd edition. Electoral Reform Society, 1976.

Edited comments on Robert Newland's suggestions

Editor
Brian.Wichmann@bcs.org.uk

1 Introduction

A moderated email discussion was held based upon the questions raised by Robert Newland [1] about 23 years ago, but only published in 2006. Those participating in the discussion were (in alphabetical order): Bernard Black (BB), James Gilmour (JG), David Hill (IDH), Michael Hodge (MH), Chris Jerdonek (CJ), Henry Kitchener (HK), Jonathan Lundell (JL), Michael Meadowcroft (MM), Joe Otten (JO), Colin Rosenstiel (CR), Markus Schulze (MS), Nicolaus Tideman (NT), and Paul Wilder (PW).

Although the discussion was initially concerned with ten questions, it soon diverged into other, related, topics. It was agreed that the editor should attempt to edit the material rather than relying upon using only the original email text.

2 The questions and discussion

The questions and the discussion that arose from each are enumerated in the following sub-sections. Not surprisingly, some respondents said the questions were wrong and answered a slightly different point.

Questions raised in 1983 are not necessarily appropriate for today. A count in 1983 would probably have needed a main-frame while today any office computer could do a count in a few seconds.

Direct input to a computer (DRE - Direct Recording Electronic voting) would not typically have been envisaged in 1983, nor was the capability to read ballot papers using OCR as well developed — the questions need to be phrased in a manner suitable for today. On the other hand CR had a counting program working on a ZX81 in 1981.

For this publication, see www.votingmatters.org.uk

2.1 Does computerising STV counts save time/money?

BB: This is of no consequence; the right result is all important. IDH: Not to any noticeable extent, unless a recount is necessary to fill a casual vacancy or for some other purpose. Then it is very substantial. (A point repeated by MH.)

JL: Probably. Certainly, if ballots are cast in a computer-readable form (DRE or optical scan, say). Other considerations are probably more significant.

In particular, Newland's comment that, "Voting machines capable of accepting preferences seem an unlikely investment for infrequent public elections," is probably wrong today, at least in the United States, where Federal law mandates machinery that, as a happy side effect, is capable of implementing STV, given the requisite laws, programming and certification.

On the other hand, the widespread practice of voting by mail will continue to require voting machinery in which the primary ballot is paper. In my county (San Mateo, just south of San Francisco), more than half the ballots cast in the June primary election were cast by mail.

NT: This is an empirical question, so its final resolution will presumably be determined by experience. However, if voting is done on a computer screen, as seems increasingly likely, I cannot imagine how it could happen that a computerised count would not save time and money in elections with more than 100 or so voters. Even if voting is not by computer, as long as voters produce scanable ballots, I would expect computer counting to save time and money. If the votes are made public, as I am inclined to think they ought to be, then there will be programs in the public domain to count them, so it will be a good idea to use a computer to count them, to avoid consequential human errors in the counting process. The availability of such programs, along with the votes cast, will make it possible for anyone who wishes to do so to verify that the accepted program elects the candidates that officials say are elected.

JG: As someone else has already suggested, this question should now be answered by reference to the data available from recent computerised counts in large scale elections. Modern high-speed scanning of paper ballots and intelligent OCR have almost certainly changed this out of all recognition since Robert wrote his note in 1983.

CR: I agree that when we introduced full computer counting into Liberal Democrat elections it made little difference in time and effort. However, from long experience it is now clear to me that we made a considerable gain in accuracy because copying ballot paper data are inherently simpler than interpreting preferences when making transfers.

2.2 How important is witnessing a manual count?

BB: The opportunity to view the count should be available to candidates or their agents. IDH: Not very. It can appear much more meaningful than it actually is, because witnesses can rarely see much that is really relevant. Having systems that actually get the right answer is much more important, but convincing the public that it has been properly done is vital.

MH: I regard it as vital that candidates (or their representatives) can witness counts, whether manual or computer.

JL: To digress slightly, California law requires a manual count of 1% of the ballots (county by county) as a check on the automated count. This raises obvious problems for STV in general and computation-intensive STV methods in particular.

I witnessed a manual recount recently (city council, at large plurality election for three seats). I had a lot more confidence in the result as a consequence of seeing the count, even though the margin was very small. That is good, albeit somewhat subjective.

I agree with David Hill that, "Having systems that actually get the right answer is much more important, but convincing the public that it has been properly done is vital." That is to say, a witnessed manual count is but a means to an end.

NT: Fairly important, I would say.

JG: I suspect this does not happen in most private elections. It appears to be important in public elections for two reasons; Firstly, it is the only means by which candidates and their agents can have any assurance that the ballot papers have been counted correctly; Secondly, it is the only means by which candidates and their agents can collect some information about voting patterns that they consider useful for future campaigning.

Auditing

Apart from a witnessed count, another method to gain confidence in the result are auditing procedures. There was a lengthy discussion on this which is summarised below.

JL: Have reformers settled the question of the extent to which STV algorithms should be replicable "by hand"? To me this question has primacy over questions of representation and "inclusiveness" because it is about trusting the validity of the tally itself. Some answers may limit which algorithms can be considered.

If proper procedures are followed, it seems to me that no replicability by hand is needed. In the United States there is a manual tally process for machine counted elections that involves manually checking the ballots in 1% of precincts selected at random. (Whether this is implemented correctly in practice is another matter.) It seems that no replicability by hand is needed if (1) the ballot rankings are publicly and digitally released, and arranged by some grouping (e.g. by precinct), (2) the digital data are manually checked against the physical ballots in some fraction of those groupings (e.g. 1% of them), and (3) the voting algorithm is fully specified to the public. This would be enough for any organization or member of the public to verify the tally.

JO: It seems that no replicability by hand is needed if (1) the ballot rankings are publicly and digitally released, and arranged by some grouping (e.g. by precinct), (2) the digital data are manually checked against the physical ballots in some fraction of those groupings (e.g. 1% of them), and (3) the voting algorithm is fully specified to the public.

I agree that simplicity of the rules is important. Meek rules I find the simplest, other rules tending only to appear simple when details about the order in which things are done and so forth are glossed over. However while their simplicity is an advantage, their impracticality for hand-counting is not.

JG: With regard to transparency, so far as the imminent (2007) elections in Scotland are concerned, you should remember that the conventional STV paper ballots will be scanned and the counting all done within a computer program. So the tally-men and tally-women will not be at all able to tally the papers or the votes. Indeed, the STV (local government) and AMS (Scottish Parliament) ballot papers will possibly be scanned together — the software separates the votes.

If DRS stick with the scanning procedure they demonstrated, and *if* the Scottish Executive allow the publication of one of the very useful reports that program produced, it will give the parties and others a great deal of information about the STV preferences, ballot box by ballot box. The report I have in mind shows the numbers of preferences at each level (1, 2, 3, 4 etc) for each candidate. It does not show the patterns of transfers, but it does provide very valuable information for the candidates and their agents, and it does it painlessly. I have written to the Scottish Executive and to lots of others saying this is *one* part of the open reporting we need to have in the Scottish procedure.

PW: Transparency in procedures and counting methods in all elections is important, but in public elections it is crucial to maintaining confidence in and the legitimacy of those elected.

[There was a discussion about the US style of auditing and its potential application to Scotland. This has not been included.]

2.3 Are the ERS76 rules the best for a manual count?

Respondents were given an opportunity to consider ERS97 in their response.

BB: Neither. All possible improvements were not made in the 97 version. IDH: Given that all manual counts are only approximations, for reasons of practicability, the ERS rules are probably almost as good as can be got, though I am still waiting for a proper description of the reduced quota feature of ERS97.

NT: The rules could probably be improved a little, here and there, but the improvements would not add much value to the existing rules. I would guess that 98% or 99% of what could be achieved by the best manual-count rules could be achieved by the existing rules. So the important thing is to get STV in use, and then consider refinements.

JG: To answer this question you must first define "best".

I would suggest there are six sets of rules that *could* be used for manual counts: Dáil Éireann, Northern Ireland, ERS73 (not quite identical to the NI rules), ERS76, ERS97, and my version of WIGM STV. (I exclude the Australian Federal Senate rules based on the Inclusive Gregory Method because the transfer value averaging procedure in those rules means that they do not comply with "one person, one vote" [3].)

Exclusive versus Inclusive rules

Farrell and McAllister [3] use the term "inclusive" to characterise a variant of STV which uses more votes in a transfer thus ensuring that more voters are involved in the election of subsequent candidates. Hence one could characterise a rule as "exclusive" if it minimises the voters involved.

JG: I think it is important that any and all discussions of computerisation of STV counts and of the counting procedures that computerisation might make practicable, should take fully into account the effects of the various procedures in relation to the "exclusiveness" or "inclusiveness" of representation. This essential context is missing from almost all these questions.

You may define "best" in terms of the "exclusiveness" or "inclusiveness" of the procedures in different sets of STV rules; there is a diversity of views on which is "best" in this respect. You may define "best" in terms of practicality; there is likely to be less diversity of view on that.

If maximum "exclusiveness" is your definition of "best", you will choose the Dáil Éireann rules. If any element of chance is completely unacceptable, you will exclude the Dáil Éireann rules from any further consideration.

If maximum "inclusiveness" is your definition of "best", you will choose my WIGM STV rules [4]. If you want the maximum "exclusiveness" without any element of chance, you will choose the NI rules or ERS73.

If you want to maximise the practicality you would probably choose ERS76 or ERS97.

Interestingly, in revising ERS76 to ERS97 some "exclusive" features were dropped, but this does not appear to have been done with any conscious intent of making the rules more "inclusive".

MM: Maybe some rules have defects, but the crucial difference with the rules for Dáil Éireann elections and for those in Northern Ireland, is that they are already entrenched in law and have been used successfully in many elections.

CR: What about the Cambridge, Mass, rules which could be described as more exclusive (I do not really buy the simple linear scale model of inclusiveness/exclusiveness anyway because there are other, more political factors to weight various counting rules by).

Cambridge has no derived surpluses at all. If a candidate reaches the quota during a transfer they are leapfrogged by further votes in that round. The only surpluses they have are first stage ones. They are randomly selected for transfer or not, see [8].

JG: The Dáil Éireann rules have a principled structure, which come at the "exclusive" end of the spectrum (called "exclusive" only because it is the opposite of the "inclusive" variants). The Cambridge, MA. rules certainly present a simplification compared with the Dáil rules, but I don't think their arbitrary handling of what would otherwise be consequential surpluses in any way enhances the "exclusiveness" of the representation they deliver.

2.4 Given a computer count, should improved counting procedures be used?

BB: Yes. IDH: Yes. It is absurd to be stuck with approximations where they are unnecessary. NT: Yes.

MH: No, due to the desire to allow a manual count using the same rules — the procedure adopted by the Church of England.

JG: As noted above, the wording of this question reveals the questioner's prejudice and it presents no context for the assessment of "improved".

2.5 Given a computer count, should more than two decimal places be used?

BB: Yes. IDH: Yes, but merely that without other changes does not help much. NT: Yes.

JG: Before considering the number of decimal places that should be used for calculations within STV procedures, I would strongly recommend that all STV counting rules for public elections should prescribe that when votes are transferred, candidates should be credited with only integer numbers of votes. That would greatly simplify the presentation of the results and would aid public understanding and acceptance. This, however, is not a matter of "rounding for presentation" - that way lies disaster. As in the Australian Federal Senate rules, the candidates are credited with only the integer part of the total vote to be transferred and appropriate procedures have to be specified to deal with the "vote fractions not transferred". I have not tried to apply this "integer only" approach to Meek STV, but it can be applied to all other versions of STV rules, from Northern Ireland rules to my WIGM rules for manual counting. Dáil Éireann STV is already integer only.

Once the practicality of result sheet presentation has been separated from internal calculation (by adopting integer transfers), determining the number of decimal places to be used in calculations becomes essentially an exercise in numerical analysis. We should certainly use more than 2 decimal places because of the significant vote loss than can occur with such truncation, as explained in my paper [5]. Where the possibility of a manual count has to be retained alongside computerised counting, I have recommended 7 decimal places for practical reasons associated with the use of pocket electronic calculators [4].

2.6 Given a computer count, restart after an exclusion?

BB: Yes. IDH: Yes, provided that other changes are made to make it work properly. Merely to do that without other changes is disastrous, see [9]. NT: Yes.

JG: I presume by this you mean "go back to the beginning and start the count again as though the excluded candidate had never stood". This presumably reduces the total valid vote by the number of votes for the excluded candidate that are not transferable (no next available preference) and so reduces the quota for the "new" count. That could have all sorts of interesting effects.

2.7 Given a computer count, transfer to already elected candidates?

BB: Yes. IDH: Yes.

JL: The benefits of Meek's method are compelling, if we use computers for the count. However, a manual count, or recount, or verification, becomes impossible, and while publication of the ballots would make independent computer counting possible, there are significant ballot secrecy concerns associated with such publication.

Moreover, manual verification requires another step prior to the (computerized) count, namely verifying that the ballots in the ballot file represent the will of the individual voters. In California, that's likely to mean examining a voter-verified paper copy of an electronic ballot, another area for ballot secrecy concerns, and one in which truncation of unused preferences will not help (they are already on the paper).

NT: Yes.

JG: This is an illogical question because the decision whether or not to transfer votes to already elected candidates does not depend on computerisation, but on the STV procedures you are using. It would, of course, be impractical for public elections without the use of a computer, but that is a separate issue.

As Robert Newland showed in this 1983 note [2], it would be wrong to transfer votes to already elected candidates if you are using the Gregory Method of fractional transfers with last parcel only. Robert also showed that, to give coherent results, transfers to already elected candidates are required if you are transferring all ballot papers, as in WIGM and Meek.

Consideration of Meek

The use of the Meek algorithm arose several times within the debate on the main questions, but the issues raised are collected here.

NT: To my mind, the answer to improving and simplifying is the Meek rules. These rules have been around for nearly 40 years now. They eliminate some limitations of the Newland-Britton rules that are very distressing to voting theorists. They have a very straightforward explanation. It would generally take too long to count by these rules by hand, but confirming a count by hand-calculator is reasonably straightforward, if rather time-consuming. The rules have been written into "legislation" by the Royal Statistical Society (and in New Zealand law: Editor).

To make the Meek rules even more acceptable, I would propose that someone write a computer program with even more auditing than the present program. In particular, I would suggest that the program should produce an audit trail that shows the allocation of each vote at each stage of the count.

If you feel that the Meek rules are too complicated, then the rules now in use in Northern Ireland (a slight variation on Newland-Britton) might be considered. Voting theorists will be concerned of the ease with which strategy can be employed against them.

CR: Interestingly, Robert Newland's article, in a few short sentences, shows why Weighted Inclusive Gregory treatment of surpluses is such a nonsense.

This discussion also needs to consider more political aspects of different STV variants. My main objection to Meek (and implicitly to some of Robert's ideas) is that they reduce the effective value of votes of less well-informed voters, those who do not express full preference lists. These voters are likely to be politically skewed, with effects on party representation and on the acceptability of STV to our potential supporters.

JL: During a manual recount in California, witnesses must be permitted. They are generally representatives of the candidates. So, independent of whether a computer is making the primary count, ballots are visible to the (semi-) public during the recount. Is this an issue? Perhaps not; recounts are expensive and rare, and as you say, could be implemented without any one person seeing the entire ballot.

With Meek's method, though, a hand count is not practical. So a "manual recount" must be replaced by some other process, presumably a manual verification of the ballot file, and then making the ballot file available for an independent count, and it is not clear to me that truncation (say) could be part of either step.

I am not particularly concerned about the secrecy problem at this step in the process. Again, just looking at the California process, there are secrecy issues already in a manual recount; a vote-seller could "prove" his ballot by casting a distinctive write-in in an irrelevant race. Worse, our vote-by-mail system, used by a large percentage of the electorate, is wide open to both vote-selling and coercion. That is not a good thing, of course, but introducing STV is not going to make things appreciably worse.

On the other hand, jurisdictions with a stronger commitment to ballot secrecy are likely to have a problem implementing STV, maintaining secrecy, and making counting transparent.

HK: Many voters will only know enough about the candidates to put a few at the top of their list. There may be a "party" in whom they have confidence, and who they would like to use to complete their paper. I have found this with the Friends of the National Trust, and with the ERS Support Group. Adding Party Lists would eliminate, or at least reduce, short votes, which would meet the objection some people have to the way Meek treats short votes.

CR: My political concern, especially about Meek but it could also apply to WIG, is that the votes of people who express short preference lists can be devalued. As it is expecting a lot of voters in mass elections to have enough valid information to make informed preference choices for all candidates this could give some voters an advantage.

2.8 Given a computer count, should all candidates be elected with the same number of votes?

BB: Yes. IDH: Yes, in principle, but it is not necessary in practice to do extra work to reach that, once it known for certain which candidates are elected and which are not.

JL: I like the principle, but I am doubtful that it is practical, if we mean to (say) reduce the quota until all seats are filled at the original quota. If quota q

fills one too few seats (without reducing the quota), and quota $q' < q$ fills all the seats, is there a quota q'' between q and q' that also fills all the seats, but with different winners?

In Green Party (California and US) internal STV elections, we require that a candidate reach the quota to be deemed elected, and leave seats empty if necessary, another way (not always appropriate or practical) to answer this question in the affirmative.

MM: Clearly the search for improvements to the operation of STV is on-going, and the advent of the computer opens up new possibilities, but the nature of STV and the relatively complex (for the average elector) concept of the quota and redistribution according to preferences etc, lends itself to caricature by its opponents.

It is interesting to note that the various arithmetical formulae relating to the distribution of list seats does not attract the same attack.

NT: Yes, provided that there is a restart after exclusions. The quota should be lowered as votes become non-transferable.

JG: It is difficult to imagine why anyone would want to do this. It could be achieved only by a complex iterative procedure with an ever-diminishing quota and a series of transfers among the already known winners until all the winners were credited with an equal number of votes. The purpose of the election is to identify the unique set of winners to fill a stated number of seats. When you reach the stage at which you can do that (according to the rules you are using), there is little point in proceeding further.

If you are using a Droop quota and you have filled all the vacancies and there are some votes (less than one quota) then credited to the runner-up, I can see no useful purpose in transferring those votes, much less any useful purpose in going on to equalise the numbers of votes credited to each of the already elected candidates.

CJ: I can see doing this in cases where a "countback" may be used later on to fill a vacancy. In one version of countback, vacancies are filled with STV using all votes that went to elect the vacating candidate(s) in the last election or countback, together with the exhausted votes. If candidate totals are not first equalized, then some voters will not have a fair say in the countback result. For example, if one candidate has a large surplus at the conclusion of the election and some other candidate vacates, the countback would not be fair to the voters who have votes in that pile with surplus. If the tally had continued and surpluses cleared, a lot of those votes could have wound up in the exhausted pile (affecting the result of the countback).

2.9 Given a computer count, should all papers be considered for transfer of a consequential surplus?

BB: Yes. IDH: Yes, all *relevant* papers.

JL: Yes (Meek)

NT: Yes.

JG: Like several other questions, this question has nothing to do with computer counting but everything to do with the type of STV rules you are implementing. As Robert Newland has shown [2], for rules that are to be internally consistent, you must take only the last parcel for Dáil Éireann, Northern Ireland, ERS73, ERS76 and ERS97 rules. In contrast, for internal consistency in WIGM and Meek, you must transfer all papers. So the real question is, once again, do you want "exclusive" or "inclusive" representation, and by how much?

2.10 Is excluding the lowest candidate the best?

BB: Yes.

IDH: If we stick to the principle that later preferences must not under any circumstances upset earlier ones, it appears to be the only sensible rule available, though it is sometimes unsatisfactory. If we are prepared to abandon that absolute principle then I believe "Sequential STV" to be better, see [10].

JL: Here we presumably mean lowest number of first-place votes. I want to preserve later-no-help/harm, and so am reluctant to consider any but first-place votes, so: yes. I think so.

The attractions of Condorcet methods (for single-seat elections) and Sequential STV (otherwise) are undeniable, but the value of being able to unconditionally assure the voter that subsequent preferences will not harm earlier ones is very valuable, not to be give up lightly.

NT: If exclusions are to be done one by one, I prefer a rule of excluding the candidate who would not be elected if the number to be elected were one less than the total not excluded yet. This rule excludes at each stage the candidate with the least apparent claim to inclusion with the others. This rule is not ideal. Its weakness is apparent in the fact that if just one candidate is to be elected, the rule can exclude a Condorcet winner. But even though the rule is not ideal, it is an improvement on eliminating the candidate with the fewest votes.

If a better exclusion rule is desired, then my recommendation is to not exclude candidates one by one, but rather employ a rule that takes account of the comparisons of all possible outcomes (sets of elected candidates) with one another, see [6].

MS: An alternative STV method is also available [8].

JG: Here again, it depends on what you mean by "best". Some of us like to give electors an absolute guarantee that a later preference can *never* harm an earlier preference. If you regard this as an important principle, to be upheld in all circumstances, you have no option but to exclude the lowest candidate (or pair, or three, etc). Those who come from a social choice background are concerned (or horrified) that a Condorcet winner could be excluded by this procedure and criticise STV for this effect. But if you once open the door to taking later preferences into account to decide the fate of earlier preferences in any circumstances, you will have opened the door to tactical voting in STV. In public elections, with large numbers of anonymous voters, tactical voting is impossible under the present "lowest candidate exclusion" rules and it would have very serious implications to make any change in that.

3 References

[1] R. Newland. Computerisation of STV counts. *Voting matters* Issue 22, pp11-13.

[2] R. Newland. Consequential Surplus: Transfer of the last parcel. *Representation* No. 99, Vol 25, summer 1985, pp 18-19.

[3] D. A. Farrell and I. McAllister. The 1983 Change in Surplus Vote Transfer Procedures for the Australian Senate and its Consequences for the Single Transferable Vote. *Australian Journal of Political Science*, vol 38, No 3, November pp479-491. Available via the McDougall web site.

[4] J. Gilmour. Developing STV Rules for manual counting to give effect to the Weighted Inclusive Gregory Method of transferring surpluses, with candidates' votes recorded as integer values. *Voting matters*. Issue 22. pp21-25.

[5] J. Gilmour. Calculation of Transfer Values — Proposal for STV-PR Rules for Local Government Elections in Scotland. *Voting matters* Issue 17, pp20-24.

[6] N Tideman and D. Richardson. "Better Voting Methods through Technology: The Refinement-Manageability Trade-Off in the Single Transferable Vote", *Public Choice* 103 (2000) pp. 13-34.

[7] ELECTION OF CERTAIN CITY AND TOWN OFFICERS BY PROPORTIONAL REPRESENTATION OR PREFERENTIAL VOTING. Massachusetts General Laws, Chapter 54A. http://www.cambridgema.gov/Election/mgls.html

[8] M Schultze. Free Riding and Vote Management under Proportional Representation by the Single Transferable Vote. http://m-schulze.webhop.net/schulze2.pdf

[9] I. D. Hill. Comments on Newland's paper. *Voting matters*. Issue 23, pp1-2.

[10] I. D. Hill and Simon Gazeley: Sequential STV — a further modification. *Voting matters*. Issue 20, pp6-8.

Review— The Machinery of Democracy, Protecting Elections in an Electronic World

Brian Wichmann
Brian.Wichmann@bcs.org.uk

1 Introduction

The document being considered here [1] is a highly significant report which deserves careful study by those nervous about the security aspect of using computers for elections. The report is from a Task Force with many experts with established reputations in the field. Moreover, many others clearly performed studies for the Task Force, including the National Institute for Standards and Technology (NIST).

Equally important to the work were reviews and comments made by those professionally responsible for elections across the USA — Registrars and Auditors.

There are important limitations to the study, namely that the only voting systems considered were ones available at the time, and that postal voting was not considered. For the UK, this last restriction is important, since a recent legal case has indicated fundamental weaknesses in the UK postal voting system [2].

Lastly, this report is specifically written to address problems in the US system, and hence its application to other jurisdictions is for readers to decide.

2 The context

The US has thousands of electoral jurisdictions — many more than one per state. The number of jurisdictions that make their own decisions about voting procedures and equipment is smaller, but runs into hundreds. Hence the issues to be addressed are large and diverse due to the different technologies used. The report divides the electronic voting systems into three classes:

DRE Direct Recording Electronic. A DRE machine directly records the voter's selections in each contest, using a ballot that appears on a display screen. There are at least 9 types of machine like this.

DRE w/VVPT A DRE with Voter-Verified Paper Trail captures a voter's choice both internally in electronic form, and contemporaneously on paper. There are at least 5 machines of this type.

PCOS Precinct Count Optical Scan. PCOS voting machines allow voters to mark paper ballots, typically with pencils or pens, independent of any machine. Voters then carry their sleeved ballots to a scanner. At the scanner, they unsleeve the ballot and insert into the scanner, which optically records the vote. There are at least 3 systems of this type.

Note that all three types of voting systems need to be configured for a specific election. Undertaking this task implies access to the machine that could lead to security issues.

3 The methodology

Given the scale of the problem in the US, a methodology was needed to provide a framework for the work and ensure that the result could be understood without too much difficulty.

From existing electoral statistics from 10 states, an artificial state called Pennasota, was devised. The 10 states were all marginal making them potential targets for an electronic attack. The main analysis was for the Governor of Pennasota with the following voting pattern:

Candidate	Party	Total Votes	Percentage of Votes
Tom Jefferson	Dem-Rep	1,769,818	51.1
Johnny Adams	Federalists	1,689,650	48.8

In addition to the overall figures above, the split of the votes amongst the precincts and polling stations and voting machines was produced.

For this publication, see www.votingmatters.org.uk

The next stage of the methodology was to produce a list of potential threats — 120 in all. These 120 were then analysed to identify the most important ones. The key to this part of the analysis was noting how many people would be needed to undertake a successful attack. The main conclusion from this was that threats against individual polling stations would be unlikely to be successful due to the number of stations needed to swing the Pennasota vote — 40,000 votes out of over 3 million.

There are two forms of analysis — one a generic one concerned with the nature of PC-based equipment, the other arising from the most important of the 120 identified threats.

Basing voting machines on PC technology has obvious problems due to the known security issues with both Windows and Linux. It seems that all the equipment considered use either of these two operating systems. Personally, I consider this inappropriate for polling station equipment since it would be difficult to ensure adequate security both at the polling stations and during storage and transport between elections.

Of course, validation and checking is undertaken of voting machine software. However, it seems this is limited to the software written for the purpose, rather than the entire system (which could be very large). This seems to imply that using the operating system to subvert the voting machine software is a credible line of attack. This supports my own contention that polling station machines should be like other embedded software systems — such as the systems used to control the engine of modern cars.

Another generic issue to be faced with all the equipment is the need to customise it for a specific election. For this purpose, ballot definition files are used. Hence an issue to be considered is whether changes to such a file could be undertaken with a view to changing the election result. Here the threat seems less credible.

3.1 Threat analysis

By way of illustration, we take the most credible attack on each of the three systems.

For the **DRE** system, this attack is a Trojan Horse inserted into the operating system. To remain undetected, it would probably have to be activated carefully so that testing prior to the election would not reveal the Trojan Horse, nor would the limited validation undertaken immediately prior to the election. To me, this attack seems very credible which is why I believe such machines should have embedded software and not rely upon a conventional operating system.

For the **DRE w/VVPT** system, a Trojan Horse again seems to be the most credible form of attack. The difference here is that there is a much more complex task since a paper trail needs to be produced as well. Since this paper record can be checked by the voter it probably means that success would depend upon the voter making no such check, which is usually the case. This threat seems much less credible than the previous one.

For the **PCOS** systems, a memory card is used to record the votes, and hence an attack on this is credible, as is the Trojan Horse attack yet again.

As another example of this analysis, consider the system to be used in Scotland for this year's local elections. Here, there are a small number of counting centres to which the ballot boxes are transported. Hence the security problem for **PCOS**-style machines at these centres is much easier to manage than having equipment at each polling station. Moreover, the process of transport and handling ballot boxes is well established. Hence, although an attack is not impossible it seems very much less credible than in the US context.

4 Conclusions

A large number of recommendations arise from the study: for instance, that no use should be made of wireless components due to the potential security threat. A feature of the analysis is the nature of counter-measures that would be effective against specific threats. Here, statistical analysis of results could reveal unusual voting patterns which could indicate an attack, or perhaps faults in equipment.

There is substantial evidence in this report that the validation, checking and counter-measures against a security threat were inadequate in practice. It seems unlikely that all of the detailed recommendations in the report could have been acted upon for the elections in November 2006.

For the position in Scotland using scanning equipment, the key issue would be how many informed participants it would take to perform a successful attack.

For those with any direct responsibility for elections involving electronic equipment, the report should be studied carefully — it is impossible to summarise the 147 pages adequately here — in any case, the key issues will depend upon the type of system being used.

(Further reports have been issued by the Brennan Center on Usability, Access and Cost of voting systems — these are not reviewed here.)

5 References

[1] THE MACHINERY OF DEMOCRACY: PROTECTING ELECTIONS IN AN ELECTRONIC WORLD. The Brennan Center Task Force on Voting System Security. www.brennancenter.org, 28th June 2006. (File date 8th August 2006)

[2] Dominic Kennedy and Jill Sherman. Postal voting is an invitation to fraud, says judge. *Times* 23rd March 2005. http://www.timesonline.co.uk/article/0,,2-1537754,00.html

Review— Second Report of the Irish Commission on Electronic Voting

Jonathan Lundell
jlundell@pobox.com

1 Introduction

> . . . the Commission concludes that it can recommend the voting and counting equipment for use at elections in Ireland, subject to further work it has also recommended, but that it is unable to recommend the election management software for such use.

So reads the conclusion of the Irish Commission on Electronic Voting [1].

The government of Ireland chose an electronic voting system for use beginning with the local and European Parliamentary elections of 11 June 2004. Responding to public criticism, the government established the Independent Commission on Electronic Voting and Counting at Elections in March 2004 [2]. In April 2004, the Commission issued an interim report recommending against using the chosen system for the 2004 elections, citing concerns over secrecy, accuracy and testing. The Commission issued its First Report in December 2004, and its Second (and final) Report in July 2006; the Commission was dissolved in September 2006. Except for a limited pilot test in 2002, the system has not been deployed.

In addition to recommending further work on the voting equipment, and replacement of the election management software, the Commission recommended changes to the overall operation of the elections system, including better physical security for the machinery itself, and noted that more testing will be required:

> The testing of the system as a whole carried out to date, as well as the investigation, analysis and independent testing and certification of its individual components, is insufficient to provide a secure basis for the use of the system at elections in Ireland. There is thus a need for comprehensive, independent and rigorous end-to-end testing, verification and certification by a single accredited body of the entire system as proposed for use in Ireland. While the Commission's work has laid the foundations for this process, more work will be required in this area ([1] p8).

The Second Report runs to more than 350 pages, not including much supplementary information available on the Commission's website: public submissions, technical information on the chosen system, and more. An adequate summary of the report is beyond the scope of this review, but the report itself is quite readable; the interested reader would do well to begin with the report's summary and conclusions ([1] Part 7).

This review generally confirms the judgment of the Commission, but, based on additional information, questions the Commission's conclusion that the chosen system can be made acceptable with further work.

2 The chosen system: hardware

The voter sees a series of up to five paper ballots behind transparent plastic. Each paper ballot lists up to 14 candidates, and beside each candidate is a button and a numeric LED display. In an STV election, the voter presses the candidate buttons in order of preference, and the numeric displays reflect the preference order. When all preferences have been entered, the voter presses another button to record the ballot in a removable nonvolatile memory (Ballot Module) installed in the Voting Machine.

A small LCD screen provides feedback and instructions to the voters. A cable connects the Vot-

For this publication, see www.votingmatters.org.uk

ing Machine to a separate control unit, used by the polling station staff to control the Voting Machine and monitor its operation.

After the close of voting, the Ballot Module is physically transferred from the Voting Machine to a Programming and Reading Unit (PRU) connected to a PC that runs software to read the ballot data and transfer it to a CD for consolidation with ballot data from other machines to be counted.

(The PRU is also used before the election to write information to the Ballot Module that the Voting Machine uses to configure itself, including a description of the layout of the paper ballots affixed to the Voting Machine, with the names of the candidates, which are also displayed to the voter on the LCD screen as voting buttons are pressed.)

The CDs containing ballot information are transported to a central facility where they are read, aggregated, and counted ([1] Part 3.2).

3 The chosen system: software

The Voting Machine software, written in ANSI C, runs on the PRU as well as the Voting Machine.

The "Integrated Election Software" (IES) runs on a "hardened" PC running Microsoft Windows 2000. Written in Delphi, Borland's Object Pascal, IES consists of modules for STV counting, election management, and management of the PRU. In addition, IES uses several third-party tools and libraries, including the Microsoft Access database system.

The Voting Machine software comprises some 25,000 lines of code, while IES approaches 100,000 lines, of which some 40,000 lines are devoted to the counting module ([1] Part 3.2).

4 Public comments

The Commission invited submissions from the public, and has published them on its website. Submissions were received from a variety of sources, including private individuals, opposition parties, voting-system advocacy groups, and the Irish Computer Society. Common to most of the submissions is an insistence on a voter-verifiable audit trail (VVAT).

5 Vendor comments

The Second Report includes an extensive response from Nedap NV, the Dutch vendor of the chosen system. Nedap generally takes the position that the chosen system as supplied conforms to their contract, and that it is trustworthy and secure. Nedap argues that a voter-verifiable paper audit trail (VVPAT) is not just unnecessary but actually undesirable, and argues that an open-source voting system (ie, one in which the details of the hardware and software implementations are made public) is undesirable as well.

Nedap cites a paper by Selker and Goler [3] criticizing VVPAT. However, the paper in question actually advocates VVAT but considers VVPAT inferior to alternative approaches to VVAT (Selker advocates a voter-verified audio audit transcript trail (VVAATT) in which the voter verifies an audio transcript of his or her choices; the audio transcript is recorded for use in a possible audit [4]).

Nedap and their Irish branch, Powervote Ireland LTD, assert that the system has already been adequately tested:

> The hardware and software of the VM, PRU and BM were analysed and tested by the accredited German "Physikalisch Technische Bundesanstalt" who is the body that is appointed by German law to analyse and test electronic voting systems before they can be deployed in Germany ([1] p290).

With respect to the Integrated Election Software,

> The Integrated Election Software can be divided into 3 main sections:
>
> 1. Preparation and Administration
> 2. Programming and reading in ballot modules
> 3. The Count
>
> Sections 1 and 2 have been in use in other countries for many years. Millions of votes have been processed and counted without incident or challenge. These 2 crucial sections are therefore very well proven in practice and form part of the Irish version.
>
> Unlike Sections 1 and 2, Section 3 was developed specifically for Ireland. This was subjected to extensive testing by the Department prior to its deployment at the Dáil election and the Nice referendum. IES is a mature and stable design. Adaptations and enhancements are inevitable for each new country. Changes to electoral practices are common and require software which can be

> readily adapted to meet these changing requirements in a very timely way. Each time a change is introduced requires testing to be carried out.
>
> Once testing is completed satisfactorily then that particular build number is not allowed to be changed and is issued for use ([1] p362).

6 "We don't trust voting computers"

Since the Commission's Second Report was issued, the Dutch group "Wij vertrouwen stemcomputers niet" ("We don't trust voting computers") has demonstrated the ability to compromise the Nedap voting equipment used in the Netherlands [5]. In response, the Dutch government has mandated security changes to their voting machines in advance of their November elections [6]. The Dutch voting equipment is essentially similar to Ireland's chosen system, and it's likely that the chosen system has similar vulnerabilities.

7 Comparative assessment against paper voting

The Irish government added to the Commission's tasks a "comparative assessment of the security and accuracy of the current system (ie, the paper-based system) for voting at elections and referenda." ([1] p147). The Commission found that the paper system is "moderately superior overall" to the chosen system as it currently exists, but that if all the concerns of the Commission could be addressed, the chosen system as improved would be superior to the paper system.

Not addressed is the question of whether the potential benefits of the chosen system outweigh its cost of acquisition and ongoing overhead, as well as the less tangible cost of the potential loss of confidence of Ireland's voters in its elections, a consequence suggested by the public comments.

8 VVAT

A voter-verifiable audit trail (VVAT) is intended to provide a means, independent of the integrity of the voting machinery in use, 1) to determine whether the election was accurately recorded and reported and 2) to provide an independent means of recounting the election should the accuracy of the electronic voting machinery be called into question.

A VVAT is typically accomplished by printing a paper record of each voter's ballot in such a way that the voter can verify that the paper record is correct, while not permitting the voter to retain a copy (which would be contrary to the secrecy requirement). The paper record is then used to spot-check the electronic results and, if necessary, to serve as the basis of a recount.

Implementation of an effective VVPAT is nontrivial, requiring among other things that an adequate proportion of voters actually check the paper record in detail, so that discrepancies are detected, and that a statistically adequate sample of paper ballots be counted to have good assurance that the electronic count is correct. Selker [4] advocates a "voter-verifiable audio audit transcript trail" (VVAATT) instead of a paper trail, but this approach has drawbacks of its own, being more difficult to audit.

9 NIST Discussion Draft

In 2002, US federal legislation [7] effectively mandated electronic voting equipment as a means of correcting election-systems deficiencies that came to light in the 2000 US presidential election, as well as of allowing more disabled voters to vote without assistance. The law charged the National Institute of Standards and Technology (NIST) with assisting in the development of technical guidelines for voting systems. In November 2006, NIST issued a draft document concerned with the upcoming 2007 update of the US federal guidelines. The NIST draft is unequivocal in its opinion of electronic voting systems without independent audit trails.

> One conclusion drawn by NIST is that the lack of an independent audit capability in DRE [direct record electronic] voting systems is one of the main reasons behind continued questions about voting system security and diminished public confidence in elections. NIST does not know how to write testable requirements to make DREs secure, and NIST's recommendation . . . is that the DRE in practical terms cannot be made secure [8].
>
> One of the central themes in the debate over voting system approaches such as the DRE is whether the level of certainty in the DRE is still adequate to ensure that the records have been recorded correctly. . . . Trust in an election outcome relies heavily upon trusting the correctness of the DRE's software and upon trusting that the DRE software has not been replaced nor tampered with. But, assuring software correctness and security is very difficult and expensive, and techniques for doing this are still an open

research topic. ...Simply put, the DRE architecture's inability to provide for independent audits of its electronic records makes it a poor choice for an environment in which detecting errors and fraud is important ([8] p7).

Are there ways to improve DREs so that they can be made secure and fully auditable? NIST and the STS do not know how to write testable requirements to satisfy that the software in a DRE is correct. The use of COTS [commercial off-the-shelf] software in DREs causes additional problems; having, for example, a large opaque COTS operating system to evaluate in addition to the voting system software is not feasible ([8] p9).

(In the context of the chosen system, "COTS" includes Microsoft Windows, the Microsoft Access database system, and the Borland Delphi software development environment.)

According to the NIST, 35 of 50 US states use voter-verifiable paper records entirely, and another 10 states use them on a county-by-county basis. Only five states now use DRE with no paper trail statewide.

10 Commentary

My own background is in the design and manufacture of computer systems, and I find the Commission's conclusions on hardware and software quality all too plausible, though the proprietary nature of the chosen system's software makes it impossible for me to independently verify the Commission's conclusions.

The Commission suggests that the defects of the chosen system could be remedied, in part by completely rewriting the IES election management and counting software. It seems likely that the Commission, had its remit included a determination of best practices, would have seriously considered a requirement for a VVAT of some kind.

The Irish government's selection of an electronic voting system of any kind was in retrospect premature. Such systems have received much attention recently, especially in the US, and the technology is in flux. In any case, the Commission's comparison of the chosen system with paper ballots does not make a compelling case for a change to electronic voting.

One of the difficulties in completely auditing the chosen system lies in being able to guarantee that the software running in binary form on each voting machine, as well as the IES systems, corresponds exactly to the software examined in source form by the auditors. It must be possible for a signed and certified copy of the original source code to be compiled independently into a signed and certified binary copy of the code, and in turn to be able to guarantee that the software running on the voting systems is in fact a faithful copy of the certified binary. This is complicated by the fact that the IES is critically dependent on third-party software such as Microsoft Windows and the Microsoft Access database system, as well as the Borland Delphi software development environment, none of which has been independently audited.

While some of these difficulties can be mitigated, and others entirely corrected, it is impractical, if not impossible, to be able to guarantee that any electronic voting system is completely trustworthy and, as important, is seen to be trustworthy. The fact that a company with the resources of Microsoft has not been able to guarantee the security of its own web browser (let alone the entire Windows operating system) despite years of effort and large incentives, suggests that a fully secure and trustworthy electronic voting system may be an unattainable goal, especially given the complexity of the overall system and the incentives for subverting it, making an effective independent VVAT mandatory.

11 Options

The Irish government is left with several options for moving forward.

Adopt the Commission's recommendations. Improve the voting machine and its software, improve procedures during and between elections, and replace the IES with alternative software that can meet the Commission's standards.

Adopt the Commission's recommendations as above, but require the vendor to provide a voter-verifiable audit trail (VVAT), and adopt appropriate procedures for taking advantage of the VVAT.

Abandon the chosen system, begin a process to define new criteria for a voting system, and then identify and acquire such a system.

Abandon the chosen system and continue to use the existing paper-based system, perhaps with procedural improvements, leaving open the option of considering an electronic voting system at some future time.

The Sunday Business Post (Dublin) reports that the government is leaning toward option 1, estimating the cost of complying with the Commission's recommendations to be approximately € 500K, compared with a sunk cost of some € 60M. The € 500K figure is disputed, however, and regardless of the cost of option 1, the cost of option 2 would be substantially higher [9].

My advice? Choose option 4, and establish a new commission that would, with public participation, recommend improvements to the present paper-ballot system, monitor the experience and (dis)satisfaction of other users of electronic voting systems, and develop criteria for the eventual selection of a system for Ireland. The world of electronic voting is evolving rapidly, and Ireland is in a fine position to take advantage of the experience (including the bad experience) of others before taking such an important step.

12 References

[1] *Second Report of the Commission on Electronic Voting on the Secrecy, Accuracy and Testing of the Chosen Electronic Voting System*. Independent Commission on Electronic Voting, July 2006, p7. http://www.cev.ie/htm/report/download_second.htm

[2] Terms of Reference and the several Reports of the Commission are available at http://www.cev.ie.

[3] Ted Selker & Jon Goler, *Security Vulnerabilities and Problems with VVPT*, CalTech/MIT Voting Technology Project White Paper #16, April 2004. Available via the McDougall web site.

[4] Ted Selker & Sharon Cohen, *An Active Approach to Voting Verification*, CalTech/MIT Voting Technology Project White Paper #28, May 2005. Available via the McDougall web site.

[5] Rop Gonggrijp et al, *Nedap/Groenendaal ES3B voting computer: a security analysis*, the “We do not trust voting computers” foundation, October 2006. Available via the McDougall web site.

[6] Aaron Gray-Block, “Fraud concerns over ballot computers”, Expatica, 18 October 2006. Available via the McDougall web site.

[7] *Help America Vote Act of 2002*, United States Public Law 107-252. http://www.fec.gov/hava/law_ext.txt

[8] *Requiring Software Independence in VVSG 2007: STS Recommendations for the TGDC*, William Burr, John Kelsey, Rene Peralta, John Wack, National Institute of Standards and Technology, November 2006, p4. Available via the McDougall web site.

[9] Adam Maguire, “Fine Gael accuses Ahern on e-voting costs”, Sunday Business Post Online, 29 October 2006.

Review— *Collective Decisions and Voting* by Nicolaus Tideman

I.D. Hill
d.hill928@btinternet.com

This is a very worthwhile book containing a wealth of useful information.

I have seen it said that, when making a speech, it should be divided into three parts: (1) tell them what you are going to tell them; (2) tell it to them; (3) tell them what you have told them. This book certainly follows that plan, not only overall but also within each chapter. It is divided into two parts — Collective Decisions, chapters 1 to 6, and Voting, chapters 7 to 16, before a short summing up in chapter 17. I feel that chapter 16 should really be included in part 1, rather than part 2. Chapters 1 to 6 and 16 are really more suitable for review in economics journals rather than in *Voting matters*, and I shall therefore concentrate here on chapters 7 to 15.

The book seems a little unbalanced in the degree of mathematical knowledge expected of the reader, who is expected to cope happily with $\int$, with ! (in its mathematical usage), with *ln*, with *iff*, etc., so it is surprising that $\prod$ and $\sum$, as multiplying and adding operators, apparently need explaining. Certainly anyone who struggles with mathematical notation will have to skip some parts but could still gain a lot from reading the surrounding plain text; it is unfortunate that those struggling to understand the notation will run into some misprints, that will make their understanding harder because they may not recognise them as being misprints but suspect that the fault is theirs.

I also found it unbalanced in having an 80 page chapter discussing various rules for electing to a single seat, yet only a 26 page chapter for the multi-seat case, which surely deserved more than that.

There are detailed discussions and proofs of how voting cycles can arise, of Arrow's theorem, and of the Gibbard-Satterthwaite theorem. It is useful to have these together for reference. Even those who do not wish to go into the detail of the proofs will gain knowledge of the facts that it is impossible to have a voting system without unsatisfactory features, and impossible to have one that is immune to strategic voting. Personally I find it a pity that Woodall's theorem [1] is not also given a place. I have found Woodall rather than Arrow to be the more convincing, both to myself and to explain to others. However part of this preference is because Arrow deals with trying to form an overall ranking of options whereas Woodall is more specifically about dividing candidates into those elected and those not elected. The book does deal with that point, giving a variation of Arrow's theorem to deal with it.

I also regret that there is no mention, to go with Gibbard-Satterthwaite, of the work of Bartholdi and Orlin [2] who show theoretically that STV is remarkably strategy-proof. This is certainly known in practice by those who vote using it for multi-seat elections. Careful study of the votes after the event may sometimes show where strategic voting could have succeeded, but to know what to do, other than vote honestly, at the time of voting, is virtually impossible.

There is discussion of properties used to evaluate the various proposed methods, under the headings of Domain, Consistency, Responsiveness, Stability and Qualitative Attractiveness: 18 different properties altogether. It would help in reading the book if short definitions of these properties were available on a separate card that could be kept handy. Then those who, for example, do not know their Smith consistency from their Schwartz consistency, or who wish to be reminded of exactly what is implied in this context by Homogeneity, would find things easier. I felt this in particular when finding a mention of non-negative responsiveness. Looking in the index it was not there, so where is it to be found? I found positive responsiveness and had to make the obvious guess from that.

For this publication, see www.votingmatters.org.uk

Many of the particular methods discussed for a single seat elect the dominant option (often called the Condorcet winner) if there is one, while the differences between those methods apply only when seeking to sort out whom to elect when, because of cycles, there is no dominant option. It is a pity that the casual reader might not realise that, in real elections, there usually is a dominant option, and much of the detail of what to do when there is not is then irrelevant. I have too often seen Condorcet voting dismissed as a useful method because this fact is not understood.

Among the methods discussed there is no mention of Supplementary Vote, as now used in Britain to elect town mayors. Perhaps it is thought too silly to deserve serious discussion by adults. If so I agree, but it would be worth just a sentence or two to say so.

Another reference that I should have liked to see is to Moulin's devastating work [3], showing that any system that elects the dominant option if there is one cannot also guarantee that turning out to vote at all is going to be helpful. It is unlikely in practice that abstaining could be better, but the fact that it is theoretically possible is worrying.

In evaluating the methods the author uses both technical considerations and, where preferences are used, a practical look at the voting patterns in a collection of real elections, mostly from the ERS, conducted by STV. In particular he uses these to evaluate the frequency of cycles. It is recognised that to take multi-seat elections and use the data as if for a single seat may not always be realistic. He is wrong in saying that in these elections voters are asked to rank all candidates. It is standard doctrine within ERS that voters should have total freedom to rank as many or as few as they wish.

At the end of the long chapter on single-seat methods, there are 5 pages headed "Summary". This is surely the wrong heading; a summary should refer briefly to what the chapter has already said, not introduce new material. Yet here we find the author's recommendations on the comparative value of the methods. These do not seem to me to concentrate enough on what I believe to be the main point to consider — namely whether one wishes to preserve a promise to voters that putting in later preferences cannot upset the chances of their earlier preferences, or whether one is willing to forego that promise so as to avoid the problems caused by successive eliminations. In the first case it is doubtful whether anything better than Alternative Vote is available; in the second case it makes sense to go for electing the dominant option if there is one, while what to do in the event of a cycle for top place, while it must be decided, is really a secondary matter as such cycles are rare.

The evaluations are mainly in objective terms of whether or not a method possesses each particular property, but for the properties contained in the Qualitative Attractiveness category the evaluations are necessarily subjective and it is easy to disagree with some of them. It is always difficult to find names for such features that will not be misunderstood but, for example, under "ease of use" the author appears to be considering only the relative difficulty of marking a cross against one candidate compared with recording a preference ranking against all candidates, and not to take into account the different degree of strategic thinking that may be needed for properly thought-out votes. Surely that is also a considerable part of ease of use.

Turning to multi-seat elections the author is wrong in saying that "European systems of proportional representation of the party-list type all have added features to give voters some voice in the selection of representatives within parties". British voters in European Parliament elections are not given any such voice.

The main discussion in this section is of STV, mostly well done, but I find the eventual preference for Warren counting rather than Meek counting surprising. Taking the example given, carefully devised so that Newland & Britton, Warren and Meek give three different answers, there are 5 candidates (R, S, T, U, V) for 3 seats. Meek elects R, S, T where Warren elects R, S, U. It is clear from this that, in this case, V is just a nuisance candidate and a useful comparison can be made by treating V as withdrawn [4]. If that is done Warren switches to the Meek result. Furthermore using the author's own CPO-STV method, he finds that the Meek result is the dominant outcome. These facts are not in themselves conclusive because they relate to only one example and it may well be possible to find another example that does the opposite. But I suggest that they are enough to call for further thinking from the author. His view seems to be only that "the Warren variation ... accords with my conception of fairness" rather than any detailed technical analysis. Fairness is a difficult concept and my own view of it points strongly in the reverse direction.

In considering the problems caused by eliminations he includes a mention of a suggestion that I made nearly 20 years ago and regards it as "too *ad hoc* to be satisfying". So do I. But he ignores the fact that it was a very tentative suggestion that was subsequently developed to become Sequential STV

[5]. I should love to see his views on that, even if unfavourable, but it gets no mention.

In considering the refinement-comprehensibility trade-off, he appears to think that more refinement always leads to less comprehensibility. When merely tinkering with rules in minor ways, this is usually correct, but when a major rethink occurs, such as the move from methods designed for hand-counting to the Meek method, I do not believe it to be true at all. Meek is not only more refined but also far more comprehensible. Those who promote hand-counting methods, and claim them to be easy to understand, usually pass over the messy details in their descriptions of them. He also claims that the Meek rules are faster, which is not so in my experience, but it is in any case unimportant. Compared with the time, trouble and expense of conducting an election, what are a few extra seconds in calculating the result?

In the end he favours a hybrid system of allowing STV preferences only for a maximum of perhaps 10 or 12 candidates, followed by a party-list for the rest. I think that this is very inferior to STV throughout and, to echo back his own words, is too *ad hoc* to be satisfying.

On the whole the book is well set out and easily readable, but I do dislike the modern custom of putting footnotes at the end of the chapter, where they have to be searched for, rather than in their proper footnote place.

But for all my criticisms, I should like to end by repeating my first sentence and say again that this is a very worthwhile book containing a wealth of useful information.

1 References

[1] D.R. Woodall (1987) An impossibility theorem for electoral systems. *Discrete Mathematics*, 66, 209-211.

[2] J.J. Bartholdi III and J.B. Orlin (1991) Single transferable vote resists strategic voting. *Social Choice and Welfare*, 8, 341-354.

[3] H. Moulin (1988) Condorcet's principle implies the no show paradox. *Journal of Economic Theory*, 45, 53-64.

[4] I.D. Hill (1994) The comparative steadiness test of electoral methods. *Voting matters*, issue 3, 5.

[5] I.D. Hill and S. Gazeley (2005) Sequential STV a further modification. *Voting matters*, issue 20, 6-8.

Collective Decisions and Voting
by Nicolaus Tideman
Ashgate Publishing Ltd., Aldershot, 2006
ISBN-13: 978-0-7546-4717-1
ISBN-10: 0-7546-4717-X