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;

				−1463
C1	11156	11156	11156	9693
		−7022		
C2	16715	9693	9693	9693
		+2668	−2051	
C3	9076	11744	9693	9693
		+1838	+402	+225
C4	6945	8783	9185	9410
		+2516	+1076	+1238
C5	4532	7048	8124	9362
C6	9732	9732	9732	9732
			+573	
Non-T	—	—	573	573
Totals	58156	58156	58156	58156

Quota is 9693.

Table 2.3: Test 92/P24C, 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

		+256	+827	+1047	+243
C1	4126	4382	5209	6256	6499
		+191	+167	−5053	
C2	4695	4886	5053	—	—
		+1019	+208	+1120	−693
C3	6081	7100	7308	8428	7735
			−1340		
C4	9075	9075	7735	7735	7735
		+172	+138	+820	+170
C5	5320	5492	5630	6450	6620
		−1638			
C6	9373	7735	7735	7735	7735
				+2066	+280
Non-T	—	—	—	2066	2346
Totals	38670	38670	38670	38670	38670

Quota is 7735.

Table 2.5: Test 92/P22A, Irish 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 trans-

ferred 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.

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].

		+1264	+269	+1075	+140
C1	5407	6671	6940	8015	8155
		−3158			
C2	12008	8850	8850	8850	8850
		+162	+35	−6501	
C3	6304	6466	6501	—	—
		+178	+40	+2558	−216
C4	6290	6468	6508	9066	8850
		+159	+33	+613	+76
C5	7312	7471	7504	8117	8193
			−639		
C6	9489	9489	8850	8850	8850
		+1395	+262	+934	
C7	6288	7683	7945	8879	8879
				+1321	
Non-T	—	—	—	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

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.