Anomalies of Instant Runoff Voting

BY Eivind Stensholt

DISCUSSION PAPER

NHH

Institutt for foretaksøkonomi
Department of Business and Management Science

FOR 06/2020
ISSN: 1500-4066
June 2020

16.06.20

Eivind Stensholt
Norwegian School of Economics

Anomalies of Instant Runoff Voting

Abstract

Struggles over the single-seat preferential election method IRV (Instant Runoff Voting) go on in public arenas and scientific journals, with focus on two "anomalies".

"Monotonicity failures" **are preference distributions that allow a startling strategic voting called Pushover or its reverse. Analysis shows how a Pushover action works, and why it is hard to predict and exploit an opportunity. While not rare, monotonicity failures should be seen as a minor nuisance.**

"No-Show paradoxes" **are alarms. The IRV tally eliminates a very clear Condorcet winner in a realistic, but somewhat unusual preference structure, too close to Duncan Black's Single- Peak condition: Too many YXZ-ballots let Z win instead of a very clear Condorcet-winner X who is eliminated; this harms IRV's legitimacy. Baldwin's elimination rule when three candidates remain is a suggested remedy.**

Preference distributions with the same IRV-tally are grouped together and analyzed with "pictograms" as a tool. That allows a generalization of Black's Single-Peak condition; real cases are close to "Perfect Pie-sharing", which explains why Condorcet cycles are rare.

AMS subject classification 91B12 JEL classification D72

Introduction: Non-monotonicity and constellations in 3-candidate IRV

IRV (Instant Runoff Voting), a.k.a. Alternative Vote and Ranked-Choice Voting, is one of many single seat preferential election methods. Every ballot ranks the candidates. In each of several rounds, the candidate with the smallest number of top-ranks is eliminated from all ballots; *here a tiebreak rule is tacitly assumed*. A ballot counts for its top-ranked remaining candidate. A candidate who reaches > 1/2 of the top-ranks becomes IRV-winner. First we focus on the round with three candidates, assuming each has < 1/2 of the top ranks.

Notation **Consider an IRV tally with N voters, where 3 candidates, X, Y, and Z remain. Each ballot contains one of six orderings: XYZ, XZY, ZXY, ZYX, YZX, or YXZ. Let |X| voters rank X on top; let |XYZ| of them have ranking XYZ. Thus,**

(0.1) $$N = |X| + |Y| + |Z|, \quad |X| = |XYZ| + |XZY|, \quad \text{etc.}$$

(0.2) ***Deficiencies*** **δ_X, δ_Y, δ_Z tell how many top-ranks X, Y, Z are away from 50%:**

$$|X| + \delta_X = |Y| + \delta_Y = |Z| + \delta_Z = N/2. \text{ Thus,}$$

$$\delta_X + \delta_y + \delta_z = N/2,\ |X| = \delta_Y + \delta_z,\ |Y| = \delta_Z + \delta_X,\ |Z| = \delta_X + \delta_Y.$$

The supporters of X decide the *pairwise comparison* in {Y, Z}:

Equality occurs when $|Y| + |XYZ| = |Z| + |XZY|$, i.e. $|XYZ| = \delta_Y$, $|XZY| = \delta_z$, etc.

When three candidates remain, the IRV tally uses two social preference relations. One relation orders the candidates by number of top-ranks. The other is the *Condorcet relation* (the relation of pairwise comparisons). Both may be any of the six orderings, but the Condorcet relation may also be one of two cycles, for short denoted XYZX and XZYX.

A main theme in this paper is effects on the outcome, wanted or not, by actionists who change the size |XYZ| of a single voter category, i.e. by participating or skipping, or who move their vote from one category to another, e.g. in some strategic voting or its reverse.

(0.3) ***Strategic voting*** **In single-seat preferential elections, three kinds of *strategic* (also called *tactical*) *voting* get most attention. With three candidates, X, Y, and Z, they are as follows:**

1) ***"Compromise"*:** **original ballot ranking XYZ lets Z win; new ranking YXZ lets Y win.**
2) ***"Burying"*:** **original ballot ranking XYZ lets Y win; new ranking XZY lets X win.**
3) ***"Pushover"*:** **original ballot ranking XYZ lets Z win; new ranking YXZ lets X win.**

Together, 1) and 3) form an anti-Z strategy sometimes available to avoid election of Z.

Standard labelling **When a 3-candidate preferential election is a reference throughout a discussion, it is convenient to label the candidates according to how they fare in an IRV tally of the reference election. With two rounds left in IRV, three candidates, A, B, and C, remain:**

(0.4) **C is *eliminated* because |C| < |A| and |C| < |B|;**
B is *runner-up* because |B| + |CBA| < |A| + |CAB|;
A is *IRV-winner*.

With (x, y) = (|ACB), |BCA|), information ignored in the IRV tally, the *vote vector* is:

(0.5) **(|ABC|, |ACB|, |CAB|, |CBA|, |BCA|, |BAC|)**
= (|A|–x, x, |CAB|, |CBA|, y, |B|–y).

Monotonicity failure and Frome 2009

Our main reference is an IRV election in Frome, South Australia, analyzed in section 1. The Electoral Commission published more data, in effect revealing x = 3801 in the vote vector (0.5). An estimate, y = 2748, then gives the vote vector in row 1 of the table in (0.6), visualized in the middle "pictogram" of Figure (1.3). In both, the labelling is as in (0.4). Thus:

After elimination of C, A defeated B.

Two different changes of row 1 are ingredients in the Pushover strategy of (0.3):

(*) 100 new voters join the election of row 1 in CBA. C passes A in top-ranks:
After elimination of A, B defeats C.
(**) 100 voters leave the election of row 1 from BCA. Alone, (**) keeps A as winner:
After elimination of C, A defeats B.

If (**) happens alone, it gives row 2 in the table (0.6); the original tally is repeated with adjusted numbers, and A wins. Constant electorate size is obtained by balancing (*) with (**). Together, (**) and (*) give row 3. With adjusted numbers, the tally goes as if (*) happens alone; B wins.

(0.6)	|ABC|	|ACB|	|CAB|	|CBA|	|BCA|	|BAC|	|A|	|C|	|B|
	1761	3801	4425	1107	2748	5467	5562	5532	8215
(**)	1761	3801	4425	1107	*2648*	5467	5562	5532	*8115*
(**) & (*)	1761	3801	4425	*1207*	*2648*	5467	5562	*5632*	*8115*

To construct Pushover, glue together (**) and (*): *Identify* those who leave BCA in (**) with those who enter CBA in (*). Now, concerning rows 1 and 3, the narrative becomes as follows:

***Down-ranking* of B in 100 ballots *is the only change*, and B wins instead of A;**
in reverse: *Up-ranking* of B in 100 ballots *is the only change*, and A wins instead of B.

A single winner preferential election method is *monotonic* if up-ranking winner W cannot make W a non-winner [down-ranking a non-winner L cannot make L a winner], *"while nothing else is altered in any ballot"*. Thus, IRV is a *non-monotonic method*, and this term is a misnomer. Focus on (**), i.e. less top-ranks to B, is misleading. It has nothing to do with B's victory. C competes with A to challenge B in the final tally round: The explanation is (*), i.e. more top-ranks to C.

The expression "down-ranking of B" deflects readers' attention away from the decisive (*) by hiding it as a chosen but not mentioned concomitant to (**); see Example (0.1). By construction, Pushover or its reverse is *possible* in some preference distributions called *"monotonicity failures"*, receiving an attention it is hard to ignore, e.g. Gierzynski (2009, 2011), Ornstein and Norman (2014), Miller (2017), and Supreme Court treatment: Minnesota (2009), Maine (2017).

If some voters with a BCA-ranking see Pushover as a realistic way to win, they have an incentive to perform it, but how realistic is it? Are there reasons for practical concern? Points in case are

- how frequently opportunities to win a 3-candidate IRV-election by Pushover occur;
- which structural features in a preference distribution that may allow Pushover;
- how easy it is to detect an opportunity before election and perform Pushover;
- how strong the incentives to join a Pushover action really are.

Empirical evidence for Pushover actions in IRV is hard to find. Instead, the tally mechanism itself illuminates possibilities and incentives. The conclusions are, in short version:

• quite frequent; • $N/4 < |A| < N/3$ and $|A|+|ACB| < N/2 - 1$; • difficult and risky; • very weak.

EXAMPLE (0.1) A website for election science claimed about Frome 2009: *"..., the Liberal Party* [B] *lost because some voters ranked him too high".*[1] But obviously, if B wins an IRV tally with no tie-break, each elimination is repeated without tie-break if an extra ballot has B on top.

In row 1 of (0.6), each BCA-ballot has changed exactly one account, viz. |BCA|. Still, the context makes it clear that the writer had in mind the reverse of Pushover. *Theoretically*, row 1 of (0.6) may include 100 (say) BCA voters who had consciously moved from a planned CBA to BCA in their ballots and changed an imagined row 3 to row 1. If so, they might regret that they thereby caused A to replace C in the final tally and even win it. Without evidence, a claim that this *really* happened, and that |CBA| *really* had been significantly larger, is creative accounting.

[1] Out of context, the claim is that a "strong" *No-Show accident* occurred; an accident is strong [weak] if too many XYZ-ballots cause Y or Z to win instead of X [Z to win instead of Y]. By Theorem (0.2), IRV allows only the weak kind. If there really had been a weak No-Show, then Frome 2009 was cyclic; see Figure (3.4). In reality, B-supporters just missed a win by Pushover.

No-Show Arguably, a more unfortunate possibility than Pushover and its reverse is the *No-Show Paradox,* which also may hit row 1 of (0.6): Let 100 new voters enter the election and vote CAB. Then |C| increases to 5632, elimination of A follows and B defeats C. The new voters would have gotten a better winner according to their own ballots, i.e. A instead of B, by not showing up at the poll-site to participate. This *"No-Show accident"* changes only one component of the vote vector: There is no creative accounting. Evidence is in the new vote vector: Reduction of the |CAB|-account by 100 leads back to row 1 of (0.6), and restores A as winner instead of B.

Standard tally In 3-candidate IRV with labelling (0.4), also the tally will be called *standard* if

(0.7) (|A|, |B|, |CAB|, |CBA|) is the information revealed.

However, the unknown (x, y) in (0.5) determines how action by a voter group may cause a change of winner, e.g. by means of Pushover or a No-Show accident.

Constellations

The constellation of two social relations, i.e. ranking by top-ranks and Condorcet's pairwise comparison, is central to Pushover and No-Show. A *constellation diagram* is a tool for visualization of the relations' interplay and for natural reasoning.

(0.8) *Definition* In Figure (0.1), eight *constellations* are shown in *diagrams* (3x3-tables), labelled *i, ii, iii, iv, v, vi, $i_{(cyclic)}$* and *$iii_{(cyclic)}$*. There is one candidate in each column and one in each row. Number 1 (*Plurality winner*), 2, and 3 in top-ranks are in column 1, 2, and 3, respectively. In pairwise comparison, the candidate in row 1 [2] beats the one in row 2 [3]. In the cyclic cases, i.e. *$i_{(cyclic)}$* and *$iii_{(cyclic)}$*, the candidate in row 3 beats the one in row 1.

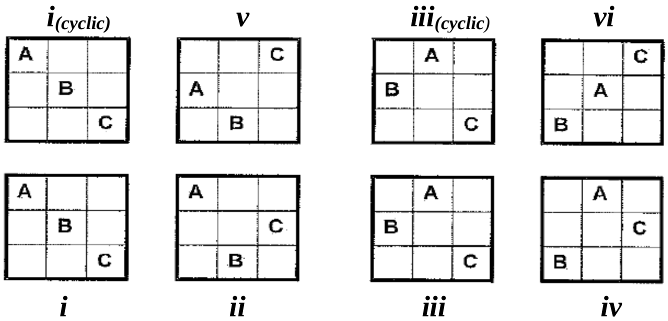

FIGURE (0.1) *Information not revealed in the standard IRV tally*

All eight diagrams show that C is eliminated (C in the right hand column) and that A wins over B in the final round (A in the higher row). The standard IRV-tally tells nothing about C in pairwise comparisons. By increasing y in (0.5), B-supporters may change *i* to *$i_{(cyclic)}$* and *iii* to *$iii_{(cyclic)}$*. By decreasing x, A-supporters may change *v* to *$i_{(cyclic)}$* and *vi* to *$iii_{(cyclic)}$*. Only *vi* is non-cyclic and gives an IRV-winner who is neither Condorcet- nor Plurality winner. Labelling candidates by top-rank order, *$i_{(cyclic)}$* has cycle 1231 and *$iii_{(cyclic)}$* has the reverse cycle 2132.

In Figure (0.1), notation (0.4) gives the candidates their rôles C, B, and A according to an IRV tally. In cyclic versions of *ii*, *iv*, *v*, and *vi*, cyclic row permutations show that

$$i_{(cyclic)} = iv_{(cyclic)} = v_{(cyclic)} \text{ and } iii_{(cyclic)} = ii_{(cyclic)} = vi_{(cyclic)}.$$

The finalists switch rôles with passage from *ii* to $ii_{(cyclic)}$ and from *iv* to $iv_{(cyclic)}$. With tiebreak rules in case of equality, every vote vector (0.5) belongs to one of the eight constellations.

(0.9) *Definition* The *constellation family* $\mathcal{A}$ consists of *i*, *ii*, *v*, and $i_{(cyclic)}$, where IRV-winner A is also Plurality winner. *Family* $\mathcal{B}$ consists of *iii*, *iv*, *vi*, and $iii_{(cyclic)}$, where runner-up B is Plurality winner. See (0.8) and Figure (0.1). The standard 3-candidate IRV tally (0.7) reveals what family the election belongs to, but all four constellations in the family are compatible with the tally.

Two facts, known in other formulations, are that only constellations *iii* and $iii_{(cyclic)}$

- may allow some supporters of another candidate to let their favorite snatch victory from A by applying the strategic (tactical) voting of *Pushover;*
- may let additional voters in one voting category cause a worse result according to their own ballot ranking through a *No-Show accident*.

There are different ways to establish these facts. However, constellation diagrams allow a hands-on reasoning, in close touch with the tally process. They also help to explain why the only voter actions that may cause these effects are the Pushover strategy and the No-Show accident.

THEOREM (0.1) Consider tallies with no tiebreaks. The preference distributions that allow supporters of B or C to make their favorite become IRV-winner with any kind of strategic voting, form a subset of all preference distributions in constellations *iii* and $iii_{(cyclic)}$. The only possibility is that suitably many supporters of B yield their top-rank to C, as in the Pushover strategy.

Proof: The voters who rank C on top cannot change their ballots in a way that prevents elimination of C. We must consider what may be possible for the supporters of runner-up B.

The voters who rank B on top cannot make B an IRV-winner in constellations *ii*, *iv*, *v* or *vi*, because no change in their ballots can change the fact that B is Condorcet loser and, if promoted to the final round, will lose whether the opponent is C or A.

The voters who rank B on top cannot make B an IRV-winner in constellations *i* or $i_{(cyclic)}$ either: No change in their ballots can prevent that A, as Plurality winner, thus with more than 1/3 of the top-ranks, qualifies for the final round. In order to win, they must ensure that B still qualifies for the final round, but no change in their ballots can prevent that A wins over B in pairwise comparison.

Only constellations *iii* and *iii*$_{(cyclic)}$ remain. The supporters of B cannot change the fact that a majority prefers A to B. The only possibility is to get rid of A; a suitable number of B-supporters yield top-rank to C, promote C to the final round and get A eliminated. ☐

If |A| < N/4, then |A|+|C| < N/2 < |B| and B wins. If N/3 < |A|, then A cannot be eliminated:

(0.10) With regard to Pushover and its reverse, the scope of this paper is N/4 < |A| < N/3.

REMARK (0.1) **By Theorem (0.1), moving from BAC or BCA into categories CBA or CAB is the only way for supporters of another candidate X to make X defeat A through strategic voting. Each actionist can move in steps between neighbor categories, as shown in Figure (0.2).**

BAC → BCA → CBA → CAB

FIGURE (0.2) ***A move seen as a sequence of ballot changes in neighbor pairs.***
Moves into CBA and CAB have the same effect after elimination of A, so it is enough to consider moves into CBA. Moves from BAC and BCA also have the same effect. Contributions from category BAC are required only if |BCA| is too small.

Thus, it is enough to concentrate on categories BCA and CBA, i.e. voters who rank IRV-winner A last; then the action is a case of Pushover (0.3). A Pushover attempt to help B may miss, but cause C to win, e.g. if 400 voters up-rank C from BCA to CBA in row 1 of (0.6). This is an improvement for both voter groups involved, BCA and CBA. C is then a *fallback security* for the actionists who start from BCA. In effect, the action becomes a case of Compromise strategy (0.3), common in Plurality elections. However, voters who start from BAC, run a risk to turn their bottom-ranked candidate C into a winner; this risk is an argument against joining a Pushover attempt.

THEOREM (0.2) **Consider tallies with no tiebreaks. The preference distributions that allow new voters to be added to one of the six voter categories and cause a candidate whom they rank after the IRV-winner A to become new IRV-winner, form a subset of all preference distributions in constellations *iii* and *iii*$_{(cyclic)}$. The only possibility is that the new voters have preference CAB.**

Proof: **We first establish that the new voters must have preference CAB. The new voters cannot give top-rank to IRV-winner A, because with higher margin than before, A would qualify for the final round, and there win against B with higher margin than before.**

They cannot rank A last either, because then there cannot be a new winner whom they rank *after* A. Thus the extra voters must give A second rank and vote either BAC or CAB.

If they vote BAC, then C still is last in top-ranks and gets eliminated: The new winner cannot be the one they rank after A. Therefore, the only possibility is that they vote CAB and make B new IRV-winner. However, A will still beat B in a final, so they must eliminate A.

In what constellations from Figure (0.1) may additional CAB-ballots cause elimination of the IRV-winner A and make B win? In constellations *i*, *i*$_{(cyclic)}$, *ii*, and *v*, A is ahead of B in top ranks, and cannot possibly be eliminated. In constellations *iv* and *vi*, B is already Condorcet loser and cannot possibly win the final tally round. Thus only *iii* and *iii*$_{(cyclic)}$ remain. ☐

REMARK (0.2) ***If Pushover is possible, then a No-Show accident is possible too.***

To see this, suppose that h voters switching from BCA to CBA will succeed in helping B to win with Pushover. With some good luck, B may become IRV-winner without Pushover: Suppose instead that extra top-ranks for C come from h *new voters* who vote CBA or CAB. A is eliminated, and B wins over C with higher margin than if the new top-ranks for C came in a Pushover action.

Every distribution of the h new ballots on CBA and CAB give exactly the same tally, but those new voters who vote CBA, have B as fallback security and may feel some happiness if they cause B to win instead of A. If they change mind and switch to CAB, all h new voters experience that their participation in the election caused their bottom-ranked B to win instead of their second-ranked A; thus they became "victims" of a No-Show accident, i.e. they would have been better off by not participating. ☐

"Abstention" **i.e. non-participation of some CAB-voters in order to avoid election of B is a very artificial remedy. If a suitable number of the unfortunate CAB-voters in Theorem (0.2) had switched from CAB to ACB, they would have helped A by Compromise (0.3). An adjustment in the IRV tally rules considered below is to use a Condorcet method when 3 candidates remain.**

EXAMPLE (0.2) **A table illustrates how a Pushover possibility is sufficient for the possibility of a No-Show accident, and proves that it is not necessary. In elections I, II, and III*, A wins IRV over Plurality winner B after C is eliminated; in II* and III, A is eliminated and B wins.**

	|ABC|	|ACB|	|CAB|	|CBA|	|BCA|	|BAC|	Voters	Const.
I	19	13	<u>16</u>	10	15	22	95	*iii*
II	19	13	21	10	15	22	100	*iii*
II*	19	13	21	<u>12</u>	<u>13</u>	22	100	*v*
III	19	13	<u>23</u>	10	15	22	102	*v*
III*	19	13	<u>23</u>	08	17	22	102	*iii*

In reference election II, two B-supporters make B snatch IRV-victory from A by changing ballot according to Theorem (0.1). In the table they move from BCA to CBA, change only the anti-A components in the vote vector, create election II*, and B wins by a Pushover "trick".

According to Remark (0.2), two new C-supporters have the same effect in election II, i.e. making B IRV-winner instead of A. In the table they vote CAB, create election III, and B wins. They break a *"No-Show quarantine"* and make all CAB-voters "victims" of a No-Show accident.

Election III is a non-monotonicity "trap": Two CBA-voters, maybe intending to consolidate winner B, move to BCA, create election III*, eliminate C, and A wins. (Obviously II* is also a trap, a reversed action leads back to II.)

***A No-Show accident with* |A| > N/3** **Removing five CAB-voters from election II leads to election I, where |A| is too large for Pushover: |A| = 32 > 95/3. However, in election I, seven new CAB-voters will create election III, and there is a No-Show accident.**

REMARK (0.3) **The effects described in Theorems (0.1) and (0.2), and their reverses, are often labelled "anomalous", but the sign of |A|−|C| is more likely to change in other ways:**

- **One anti-B move from CAB to ACB matches two anti-A moves from BCA to CBA.**
- **|A| and |C| have random components of importance in a close race to be challenger of B.**

Equal preference

Like in most Australian IRV-elections, the voters in our main example (Frome 2009) were obliged to express a complete and strict ranking of the candidates. If voters may declare equality, e.g. that j candidates share the ranks k+1, k+2, …, k+j, the tally may count them with *symmetrizing*:

(0.11) The tally is done *as if* the ballot is replaced by j! *"mini-ballots"* of weight 1/(j!),
one mini-ballot for each permutation of the j candidates.

Many IRV elections in other countries allow incomplete ballots: Voters may rank strictly k of n candidates from top, 1, 2, …, k, where k < n – 1, and leave out the remaining n–k. Emptied ballots are usually discarded; counting with symmetrizing would instead interpret them as expressing equal preference, but give the same winner. However, the report from the last IRV-tally round will show correctly the number of votes and the finalists' relative strengths.

1 *The standard IRV tally and Frome 2009*

Frome 2009 was a by-election for a single seat in the state assembly. It had six candidates. As usual in Australian IRV elections, ballots were required to specify one of the 6! (= 720) complete, strict, and transitive preferences. Three candidates remained for tally round 4:

(1.1) (A, B, C) = (Brock – independent, Boylan – liberal, Rohde – labor)

(1.2) Round 4 established |A| = 5562, |B| = 8215, and |C|= 5532.
Round 5 established |CAB| = 4425 and |CBA| = 1107. Thus,
A became IRV-winner with 5562 + 4425 = 9987, while B got 8215 + 1107 = 9322.

After the standard tally in (1.2), the *vote vector* (0.5) still keeps the number of subsidiary ranks given to C by the supporters of A and B as unknowns x and y:

(1.3) (|ABC|, |ACB|, |CAB|, |CBA|, |BCA|, |BAC|)
= (5562–x , x , 4425, 1107, y , 8215–y).
The unknowns are x = |ACB| $\in$ [0, 5562], y = |BCA| $\in$ [0, 8215].

The standard IRV-tally (0.7) is the same for all points (x, y) in Figure (1.1), but conditions for the Pushover strategy, the No-Show accident, or their reverses depend on (x, y). Figure (1.1) illustrates a discussion that uses both constellation diagrams (0.8) and *pictograms*.

Pictograms **A unique *pictogram* represents the 3-candidate vote vectorr (0.5) faithfully (Stensholt 1996). With specified (x, y), it consists of a circle and three chords that meet pairwise inside or on the circle. Distinct chords form a triangle T not corresponding to any voter group. In real elections, T is usually very small. The pictograms of Figure (1.2) illustrate three realistic choices of (x, y), but published extra information shows that the real (x, y) was on the stapled line in Figure (1.1); see Figure (1.3).**

"Ideal points" **A, B, and C, are the corners of a *"candidate triangle"* ΔABC, inserted so that its perpendicular bisectors almost coincide with the chords of the pictogram. Exact coincidence occurs in *"Perfect Pie-sharing"*; either the chords are not distinct or T has zero area. The other areas are proportional to the components of the vote vector. Perfect Pie-sharing is a 2D model of spatial voting; it generalizes the familiar 1D *"Single-Peak"*-model of Duncan Black (1948): The voters are distributed in a circle, and each voter ranks the candidates according to the Euclidean distance from the voter to the ideal points. [2]**

[2] The candidate triangle ΔABC is unique in shape, but not in size. One may see from Figure (1.2) that the size changes with homothetic transformations centered on the intersection point of the perpendicular bisectors.

Main features of Figure (1.1) **The 5563x8216 possible points (x, y) with integer coordinates form a grid in the big rectangle of Figure (1.1). One main feature is the partitioning into four sub-rectangles labelled SW, SE, NE, and NW, defined by two lines where supporters of A or B give δ_C second ranks to C, see (0.2). In Condorcet's pairwise comparison, with N = |A|+|B|+|C|,**

(1.4) A-supporters make C tie with B at x = δ_C = N/2 – |C| = 4122.5

(1.5) B-supporters make C tie with A at y = δ_C = N/2 – |C| = 4122.5

In each sub-rectangle is a constellation diagram from Figure (0.1) that changes when a line, (1.4) or (1.5), is crossed. Only pairwise comparisons change; A, B, and C stay in columns 2, 1, and 3, respectively. When x < δ_C and y < δ_C, Plurality loser C is also Condorcet loser. Thus, in SW, C occupies the lower right corner in the constellation diagram, and the constellation in SW is either *i* or *iii*; see Figure (0.1). According to the standard tally in (1.2), B is Plurality winner in Frome 2009; thus, the constellation in SW is *iii*, and the election is in family $\mathfrak{B}$.

Another main feature is a set of three curves. The middle curve connects two corners. In the SW corner, (x, y) = (0, 0), no voter ranks C as number 2; the pictogram degenerates, i.e. two chords coincide, and there is Perfect Pie-sharing. In the NE corner, (x, y) = (|A|, |B|), no voter ranks C last; the preference distribution is single-peaked, and there is again Perfect Pie-sharing.

Along the middle curve, (x, y) *maintains Perfect Pie-sharing*. Three choices of grid points close to the curve in SW, SE, and NE, give the pictograms in Figure (1.2). They illustrate that *the Condorcet relation is transitive along the middle curve*. The reason is that, with Perfect Pie-sharing, the Condorcet relation ranks candidates by their distance to the circle center, thus *transitively*.

Transitivity implies that the middle curve cannot pass through the NW sub-rectangle where all points (x, y) define a cyclic vote vector. The two other curves have endpoints on the edges of the rectangle and give pictograms where triangle T covers a fraction 0.001 of the circle area.

Why Condorcet cycles are rare **In real 3-candidate elections where this many voters have similar perceptions of the political landscape, a pictogram usually fits Perfect Pie-sharing visually well; (x, y) is often close to the middle curve and likely to be inside the "0.001-zone" between the upper and the lower curve. Empirical data on cycle occurrence due to Tideman are in (Gehrlein 2006, pp.47-48.)**

At the NE corner, i.e. (x, y) = (|A|, |B|), nobody ranks C last; it is the only point in Figure (1.1) satisfying the Single-Peak condition (Black 1948). Along the middle curve, Perfect Pie-sharing avoids cycles, thus also generalizing Black's sufficiency condition for Condorcet transitivity.

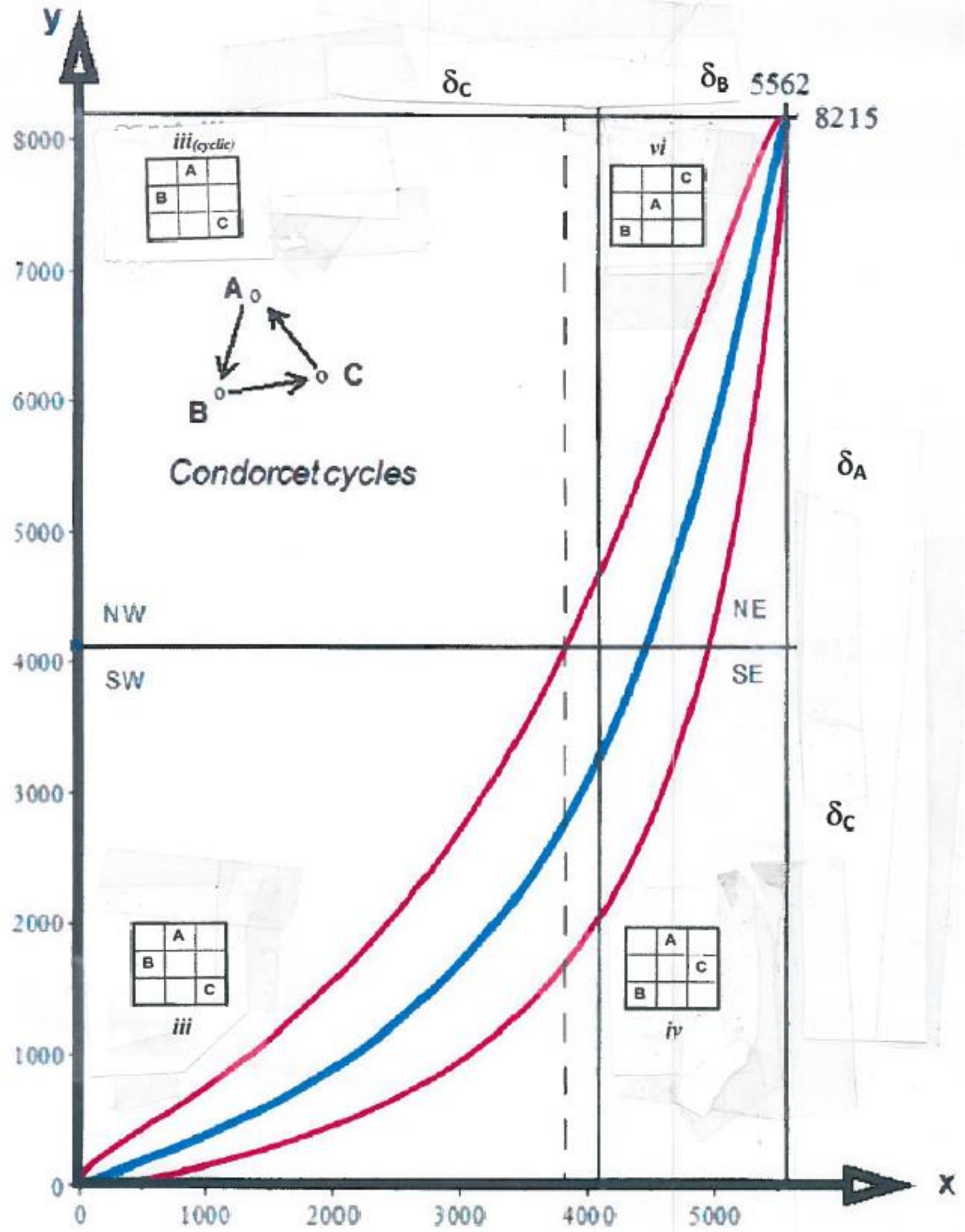

FIGURE (1.1) ***The last two rounds of the standard IRV tally in Frome 2009***

The quantitative information (|A|, |B|, |CAB|, CBA|) = (5562, 8215, 4425, 1107) is all that is found in the standard tally; both x = |ACB| and y = |BCA| are unknown.

The constellation changes within family $\mathcal{B}$ when x or y passes δ_C = 4122.5.

Along the middle curve are the (x, y) that represent vote vectors (1.3) with *Perfect Pie-sharing*, specializing to Black's *Single-Peak condition* at (x, y) = (5562, 8215): Pictograms for some (x, y) are in Figure (1.2). Along the other curves, T covers 0.001 of the circle.

***Additional information* from the Electoral Commission revealed that x = 3801. Thus, the real election corresponds to an unknown point on the stapled line. Pictograms for selected points on the line are in Figure (1.3), Figure (3.2), and Figure (3.4). Except for the stapled line, the figure builds on information from the standard tally (1.2).**

Each pictogram in Figure (1.2) corresponds to a grid-point (x, y) close to the middle curve:

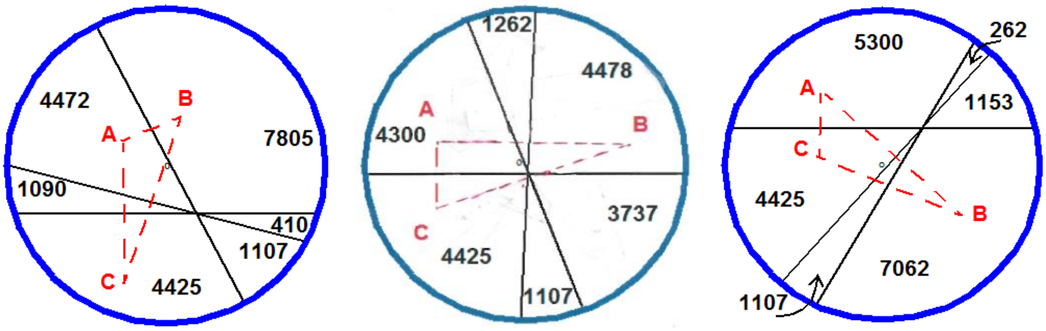

FIGURE (1.2) ***Almost Perfect Pie-sharing compatible with the standard tally***

Imagine that (x, y) moves along the middle curve in Figure (1.1), with snapshots taken at

(x, y) = (1090, 410) in SW; (4300, 3737) in SE, and (5300, 7062) in NE.

Moving closer to the corner in SW, all subsidiary support for C disappears; in the corner, two chords coincide, and the pictogram degenerates.

In NE, the constellation is *vi*, and the eliminated candidate C is Condorcet-winner. C has the smallest number of bottom-ranks: 262+1153. Moving closer to the corner, all bottom-ranks for C disappear; in the corner, the preference distribution is single-peaked.

The fate of Condorcet-winners **In 3-candidate IRV, it is only in NE, i.e. constellations *v* (family 𝒜) and *vi* (family ℬ), that a Condorcet winner is eliminated. In *vi*, B is Plurality winner and C Condorcet winner, but none of them wins the IRV-tally; the constellation diagram visualizes the double incentive that IRV gives to candidates and their parties, often adduced in favor of IRV: Campaign for primary support from enthusiastic followers (A beats C in top-ranks), and also for subsidiary support from political neighbors (A beats B in Condorcet's pairwise comparison).**

At (x, y) = (5300, 7062) in Figure (1.1), the pictogram in Figure (1.2) even shows that A beats C by (5300 – 4425) among the 9725 anti-B voters. However, it is not wise to keep a rule that always declares A as winner in constellations *v* and *vi*: The pictogram also shows two massive pairwise victories for C (12594 – 6715) and (10832 – 8477). What would the public reaction have been if this had been the real election with A as winner, and the full ballot data had been published? It is a matter of fairness, not just to Condorcet winner C, but particularly to the B-supporters. Their overwhelming subsidiary support of C (7062 – 1153) is ignored in IRV, despite the fact that this election method is promoted with its transfer of subsidiary votes as a trademark.

Figures like (1.1) illustrate that cycles are rare and likely to have a small T. and therefore all pairwise comparisons close to 50–50. The vast majority of 3-candidate IRV elections have a Condorcet-winner. Can it be justified to eliminate the Condorcet winner? In Frome 2009, both A and C had between 28% and 29% of the top-ranks. It seems reasonable to declare both as

eligible, but C as winner in constellation *vi*. An alternative is to stop the elimination ordeal based on top-ranks while three (or more) candidates still remain, and then use a Condorcet method. Almost all 3-candidate elections are non-cyclic, usually in constellation *i*, *ii*, *iii*, or *iv*, and IRV picks the Condorcet-winner; Baldwin's Condorcet method, considered below, eliminates the "Borda loser" instead of the "top-rank loser", and then picks the Condorcet-winner in constellations *v* and *vi* too.

Additional information in Frome 2009 **Fortunately, the reality in Frome 2009 was far from the scenario of the third pictogram in Figure (1.2). Since the winner A was an independent candidate, the electoral board published also the pairwise comparison between the candidates of the major parties. There,**

(1.6) in the pair (B, C}, B won with 9976 versus C with 9333. This reveals that in (1.3), x = |ACB|= 3801.

Thus, with only one unknown, the *vote vector* was

(1.7) (|ABC|, |ACB|, |CAB|, |CBA|, |BCA|, |BAC|) = (1761, 3801, 4425, 1107, y, 8215–y)

It is likely that the real (x, y) was in the 0.001-zone with 1662 < y < 4067. In the *pictograms* below, y = |BCA| = 1663; 2748; 4066. A natural estimate for the unknown y is 2748.

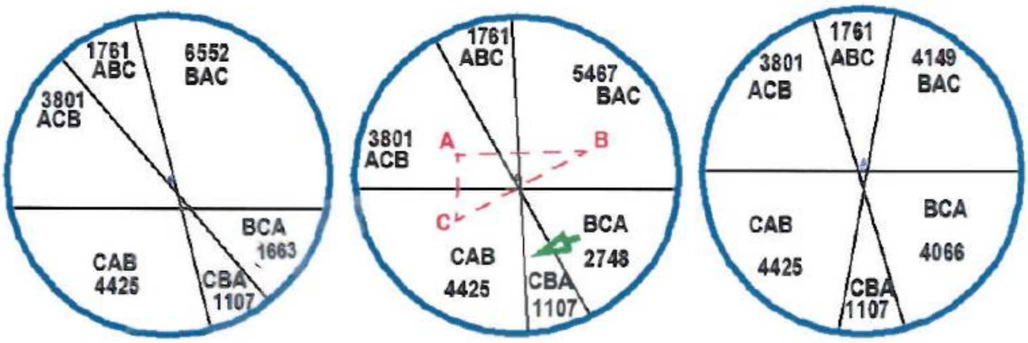

FIGURE (1.3) ***Pictograms on the stapled line in Figure (1.1)***

The grid points (x, y) = (3801, 1663); (3801, 2748); (3801, 4066) are close to the curves in Figure (1.1). For y = |BCA| = 2748, the triangle T defined by the chords covers a fraction $4\cdot10^{-11}$ of the circle area, a good approximation to Perfect Pie-sharing.

T changes continuously with y. For 1663 ≤ y ≤ 4066, T covers < 0.001 of the circle area. This illustrates the robustness of the Perfect Pie-sharing model. The arrow shows how, in all three cases, h voters perform Pushover by switching from BCA to CBA, 31 ≤ h ≤ 321. Figures (3.2) and (3.4) show pictograms of cyclic elections in NW, on the stapled line.

2 *How No-Show and Pushover work in 3-candidate IRV*

No-show effects **occur if *one component* of the vote vector passes a critical value: New voters joining [would-be voters skipping] cause the result to be worse [better] according to their own ballots, i.e. worse by *Accident* when joining [better by *Abstention* when skipping].**

No-Show; a gain by Abstention **In the election at (5300, 7062) in Figure (1.1), pictogram in Figure (1.2), let q BCA-voters skip the election. The new vote vector becomes**

$$(|ABC|, |ACB|, |CAB|, |CBA|, |BCA|, |BAC|) =$$

$$(262, 5300, 4425, 1107, 7062–q, 1153) \tag{*}$$

C wins if q ∈ [2684, 5878]. A No-Show effect occurs when increasing q passes 2683.

Ward Paradox I (No-Show case) **In (*), A wins if q < 2683 and if 5879 < q. Imagine a constituency where IRV is tallied in two wards, both with vote vector (*), but with different values of q. Choose q = q_1 in ward 1 and q = q_2 in ward 2 so that**

$$q_1 < 2683 < (q_1+q_2)/2 < 5879 < q_2$$

Candidate A wins in each ward separately; the vote vector *in the constituency* is twice the one at $(q_1+q_2)/2$; there C is IRV- and Condorcet-winner. However, one cannot expect that all three elections in a *"ward paradox"* constructed this way have pictogram with a realistically small T.

Abstention and Compromise **When increasing q passes 2683, C wins by Abstention. More efficient than Abstention is to exit from BCA, then *join* CBA, and get the vote vector**

$$(262, 5300, 4425, 1107+q, 7062–q, 1153);$$

If q ∈ [31, 7062], they eliminate B and make C win by *Compromise* (0.3). However, in the subinterval [2684, 5878], the exit part alone, i.e. Abstention, is enough. Being so close to Compromise, an Abstention win should not be called an anomaly.

Intuitively, a reduction of |BCA| is likely to reduce some advantage B has over C and some advantage C has over A. The net effect may help or it may harm C; in neither case one should be surprised. However, a theoretically possible win by Abstention is an alarm signal: The B-supporters' very strong subsidiary preference for C is ignored in the IRV-tally.

Strategic voting **does not change the total number of voters; it changes two accounts, and from tally data alone, one cannot prove it really happened. Only in some elections in constellations *iii* or $iii_{(cyclic)}$ from family $\mathfrak{B}$ is it, by Theorem (0.1), possible to make one's top-ranked candidate defeat an IRV-winner by means of strategic voting.**

A suitable number (h) of supporters of Plurality winner B in Figure (1.1) may then, for some values of x = |ACB|, apply Pushover (0.3) to snatch victory from the IRV-winner A. In constellation iii, the target A is also Condorcet winner. It is convenient to focus on the anti-A voters, categories BCA and CBA; see Remark (0.1). They turn A into top-rank loser. However, in the final, either B wins by Pushover or C wins by Compromise. Below, Figures (2.1) and (2.2) show what possibilities the anti-A voters have to avoid election of A.

Pushover; action space with pitfalls

In a Pushover trick, h actionists switch ballot ranking from BCA to CBA.

By Theorem (0.1), they start at h = 0 in constellation *iii* or $iii_{(cyclic)}$, in the SW or NW of Figure (1.1). Increasing h may cause B to snatch the IRV victory from A by creating a new constellation where B is winner. By Remark (0.1), every BAC-voter may join a Pushover action by first moving to BCA. The anti-A group has |CBA|+|BCA| members; thus the full *action space* is $-|CBA| \le h \le |BCA|$.

In notation from (0.2), ranking by top-rank changes when h passes α, β, and γ:

$$\alpha = |A| - |C| = \delta_C - \delta_A; \quad (2.1)$$
$$\gamma = |B| - |A| = \delta_A - \delta_B;$$
$$\beta = (|B| - |C|)/2 = (\alpha + \gamma)/2 = (\delta_C - \delta_B)/2$$

Necessary for elimination of A without tiebreak is |A| < N/3, see condition (0.10). Then

$$\alpha < \beta < \gamma \quad (2.2)$$

When an increasing h passes α; β; γ, then, respectively, in terms of top-ranks,

C passes A; C passes B; B passes A.

In the constellation diagram, the *columns* are switched for A and C; B and C; A and B.

The *pairwise comparison* changes once with h. This happens in {B, C}, but obviously not in {A, B} or {A, C}. After h ballot switches from BCA to CBA, the vote vector (0.5) is

$$(|ABC|, |ACB|, |CAB|, |CBA| + h, |BCA| - h, |BAC|) \quad (2.3)$$
$$= (|A| - x,\ x,\ |CAB|, |CBA| + h,\ y - h,\ |B| - y)$$

B and C are equal in pairwise comparison when

$$(y - h) + (|B| - y) + (|A| - x) = x + |CAB| + (|CBA| + h), \quad (2.4)$$
$$2\delta_C = N - 2|C| = |A| + |B| - |C| = 2(x + h)$$

Thus, the *rows* of the constellation diagram are permuted once, i.e. at h = ξ where

$$\xi + x = \delta_C = N/2 - |C| \quad (2.5)$$

By (1.4) and Figure (1.1), $0 \le x \le \delta_C$ in constellation *iii* and $iii_{(cyclic)}$; thus, $0 \le \xi \le \delta_C$. and

$$\xi = \alpha \text{ for } x = \delta_A ; \quad \xi = \beta \text{ for } x = (\delta_B + \delta_C)/2 ; \quad \xi = \gamma \text{ for } x = \delta_C - \delta_A + \delta_B \quad (2.6)$$

When α, β, γ, and ξ are distinct, x = |ACB| defines one of four different sequences:

(2.7)	sequence 1:	$\xi < \alpha < \beta < \gamma$	for	$\delta_A < x \le \delta_C$
	sequence 2:	$\alpha < \xi < \beta < \gamma$	for	$(\delta_B + \delta_C)/2 < x < \delta_A$
	sequence 3:	$\alpha < \beta < \xi < \gamma$	for	$\delta_C - \delta_A + \delta_B < x < (\delta_B + \delta_C)/2$
	sequence 4:	$\alpha < \beta < \gamma < \xi$	for	$0 \le x < \delta_C - \delta_A + \delta_B$

The action space, the interval [−|CBA|, |BCA|], may be small and not contain all transition points α, β, γ, and ξ in (2.7), but it may be extended as in Remark (0.1) and Figure (0.2). If |BCA| = |CBA| = 0, Black's Single-Peak condition is satisfied.

Cycles are very rare, so in the vast majority of cases, a Pushover action must start at h = 0 in constellation *iii*. Passing ξ and γ, the sequence ends in constellation *iii*, but with new rôles for B and C. In Figure (2.1), each sequence follows a unique path of arrows from the upper left to the lower right constellation diagram. Figure (2.2) shows the sequences after a start in *iii*

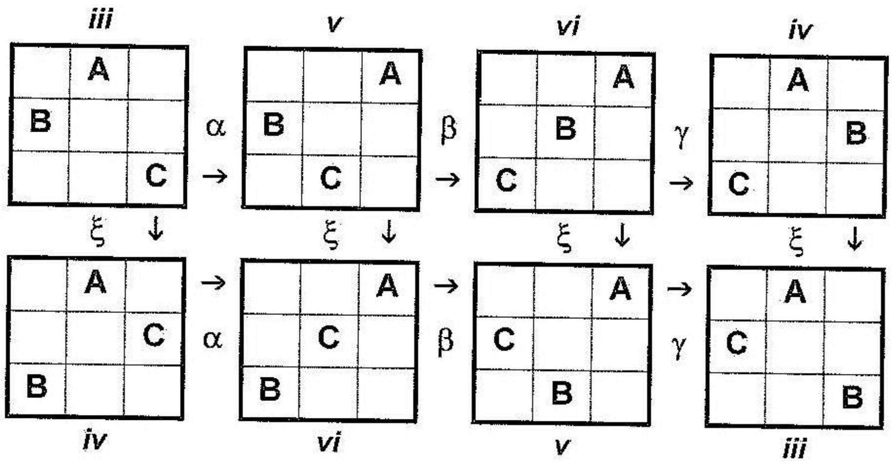

FIGURE (2.1) ***Action space: Pushover from constellation iii,*** **N/4 < |A| < N/3**

Starting at h = 0 in the upper left constellation *iii*, sequences 2, 3, and 4 lead to *v* or *vi*, and B wins by Pushover if the ballot changes stop in time, before the pitfalls at γ and ξ: Passing γ makes B Plurality loser, and passing ξ makes B Condorcet loser. Passing α and ξ but not γ makes C win by Compromise (0.3). At the other end of the action space, by symmetry, C-supporters may follow sequences 3, 2, and 1 in reverse, reduce h, move from constellation *iii* to *v* or *vi*, and win if they stop in time, before the pitfalls at h = α and h = ξ. The vertical pitfall satisfies $\xi - \beta = (|ABC| - |ACB|)/2$.

With h ϵ (α, γ), B or C wins instead of A for all ξ, by Pushover or Compromise after a successful anti-A action. In Frome 2009, (α, ξ, β, γ) = (30, 321.5, 1341.5, 2653).

A-supporters and the Pushover threat from B-supporters

Pushover is impossible if $\xi < \alpha$, i.e. in sequence 1. Then by (2.7), since |A| < N/3,

(2.8) $$N/6 < N/2 - |A| = \delta_A < x = |ACB|$$

(2.9) $$|ABC| < N/6 < |ACB|.$$

By (2.8) and (2.9, moves from ABC to ACB may prevent B-supporters from winning by Pushover, but a closer look at Figure (2.1) reveals that this would be a bad deal for those who move:

> **By making $x > \delta_A$, i.e. |ACB| > N/2 – |A|, A-supporters set a track switch so that Pushover actionists move along sequence 1 and cannot reach a constellation where B snatches the IRV-victory from A. Thus, ABC-voters moving to ACB may, by (2.5), set $\xi < \alpha$, intending it as a "prophylactic strategy". However, they are wiser leaving ξ as it is: ABC-voters cannot help A anyway if a Pushover attempt passes α, and it is better to let the actionists pass from *iii* into *v* so B wins in the final if $\alpha < h < \xi$, than into iv and thence *vi* where C wins.**

The B-C symmetry of Figure (2.1)

In the end constellation *iii*, B and C switch rôles: C may win by Pushover if a suitable number of voters move from CBA to BCA and reduce h. This symmetry illustrates the non-monotonic events "tricks" and "traps", and also another appearance of the *"Ward Paradox"*.

Tricks and traps **Anti-A voters prevent election of Condorcet-winner A by making $h \in (\alpha, \gamma)$. They enter the interval at α by moving from BCA to CBA and make**

B win with Pushover ($\alpha < h < \min(\xi,\gamma)$), or C win with Compromise ($\max(\xi,\alpha) < h < \gamma$).

Symmetrically (α, γ) may be entered at γ by voters moving from CBA to BCA. A wins if actionists leave (α, γ). Tricks and traps, being reverses of each other, are the two kinds of *non-monotonic events*. In sequence 1 and 4, only one end-point crossing may be non-monotonic (trick or trap).

Ward Paradox II (Non-monotonicity case) **In (2.3), A wins if $h < \alpha$ and if $\gamma < h$; but B or C wins if $\alpha < h < \gamma$; see Figure (2.1). Imagine there are two wards, with the same vote vector (2.3), except for the h-value.**

Choose $h = h_1$ in ward 1 and $h = h_2$ in ward 2 so that

$$h_1 < \alpha < (h_1+h_2)/2 < \gamma < h_2$$

Candidate A is Condorcet-winner in all of Figure (2.1) and wins in both wards, but not in the constituency. In general, if X wins both wards, but Y wins at the midpoint, the set of vote vectors won by X is not convex. Such non-convexity can always be used to construct a ward paradox.

Ward Paradox III (early example) **Fishburn and Brams (1983, p.209) gave an example of a ward paradox and called it the *"Multiple districts paradox":***

(|BHW|, |BWH|, |WBH|, |WHB|, |HWB|, |HBW|) =

(160, 0, 0, 285, 0, 143) in ward 1 (single-peaked);

(257, 82, 285, 39, 357, 0) in ward 2 (cyclic).

In the entertaining story, both wards are won by B (Mrs. Bitt), but the constituency is won by its Condorcet winner H (Mr. Huff). The constituency has constellation *iii* and |W| > |H| > |B|. The construction idea in *Ward Paradox I* gives constellation *v* in the constituency (where Condorcet-winner C is saved from elimination). In *Ward Paradox II* it gives constellation *v* or *vi* in the constituency (where Condorcet-winner A is eliminated but wins both wards). Can all ward paradoxes in 3-candidate IRV be obtained by continuous deformation of a few simple cases?

Pushover/Compromise in Cycle-land

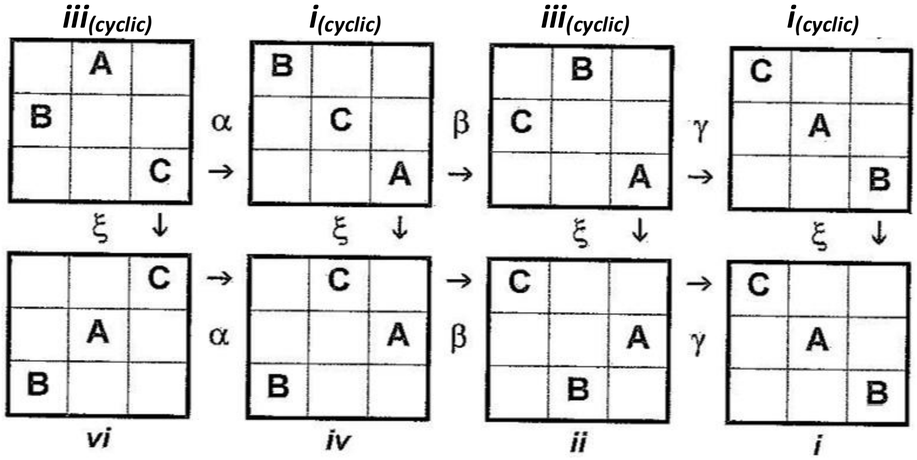

FIGURE (2.2) ***Action space: Pushover from*** $iii_{(cyclic)}$, **N/4 < |A| < N/3**

In Cycle-land the action starts start in $iii_{(cyclic)}$, **by Theorem (0.1). The cycle is ABCA in all four cyclic constellations; α, β, γ, ξ are as before. The pitfall at ξ leads out from Cycle-land.**

Baldwin's Condorcet method – based on the Borda Count

Figure (2.1) illustrates how the Pushover trick or its reverse (a trap) may decide the fate of a Condorcet-winner in constellations *iii*, *v* and *vi*. IRV may be modified by switching to a Condorcet method when three (or more) candidates are still not eliminated. The election method of Baldwin (1926) is a Condorcet method, but also an elimination method like IRV. The Baldwin tally ends with a two-candidate final; thus also a Baldwin winner beats a unique runner-up. Baldwin is a "cousin" of IRV, related to the Borda Count as IRV is related to Plurality. This use of Baldwin is based on (2.10):

(*) For X ≠ Y, define Bo(X, v, Y) = 1 [0] if voter v strictly prefers X to Y [Y to X].

The Borda Count aggregates (*) over all Y≠X and all v:

(**) X receives Σ_Y Bo(X, v, Y) *Borda-points* from v, and total *Borda-score* $\Sigma_v \Sigma_Y$ Bo(X, v, Y).

By (**), Condorcet's relation of pairwise comparisons is linked to the Borda Count:

(***) X is at least as good as Y when Σ_v Bo(X, v, Y) ≥ Σ_v Bo(Y, v, X); i.e.

Σ_v Bo(X, v, Y) ≥ N/2 in an election with N voters.

Thus, if X is *Condorcet-winner*, then (***) holds with > for every other candidate Y:

(2.10) $$\Sigma_Y \Sigma_v \mathrm{Bo}(X, v, Y) \;>\; N\cdot(M-1)/2;$$

Notice that (M–1)/2 is the average score received from one voter; thus (2.10) implies that a Condorcet-winner X has total Borda-score above average. Thus, *Baldwin cannot eliminate* X.[3]

Burying and Pushover (0.3) are strategic voting methods and, with their reverses, often seen as severe anomalies. The following observations on IRV and Baldwin elections start with the "parent" of each, respectively Plurality and its "sibling", the Borda Count. Both of these are positional election methods, i.e. they rank candidates by points obtained from ballot rankings.

Nobody minds coarseness, but one must draw the line at cruelty[4]

In Plurality elections, Compromise (0.3) is the only kind of strategic voting, but there it is used routinely. The voter casts an *instrumental* vote for the candidate deemed best among those with a reasonable chance to win, instead of an *expressive* vote for the one deemed best for the office. In both cases, Plurality elections expose voters to mental stress and coarse scolding. Either: *You commit favorite betrayal!* Or: *You waste your vote!*

There is no insincerity behind instrumental voting; voters simply include *electability* as one of their ranking criteria. With substantial friction, Plurality works in open political landscapes: Voters find information about the electorate's preference distribution; that is essential to those who will avoid wasting their vote. A voter without access to information on the size of the supporter alliances behind the candidates, is at a disadvantage with no basis for a choice between expressive and instrumental voting.

Compromise, like all other strategic voting, leaves no trace in tally accounts, see Example (0.1). Still, there is abundant evidence that Compromise is important in political Plurality elections,

---

[3] **I.e. X cannot be Condorcet winner and Borda loser. Reversal of all ballots gives an equivalent, common but less useful formulation: The Borda Count satisfies the "Condorcet loser criterion"; i.e. *a Borda winner cannot lose against every other candidate in pairwise comparison.***

The connection (2.10) of the Borda Count to Condorcet's relation was observed by Nanson (1882), a mathematician at the University of Melbourne. Nanson eliminates *all candidates* not above average Borda sum. Baldwin gives other cycle-breaks, but is with its one-by-one eliminations closer to IRV/STV, by 1926 in common use in Australia. Nanson's method was used at the University of Melbourne until 1983, and then repealed: *"The reason for abandoning the Nanson system was that it was perceived to advantage inoffensive but not outstanding candidates as against those who attracted strong support."* (McLean 2002). However, with IRV-eliminations until only three candidates remain, those who are too inoffensive run the risk of an earlier elimination for weak primary support.

[4] **Said Dorothy L. Sayers' character Peter Wimsey on mindless sophistication (*Lord Peter Views the Body*, 1928).**

but comparison of opinion polls with an election result reflects only voters who in the meantime decided to vote instrumentally. This may be only the tip of an iceberg: Those who always, without qualms, vote instrumentally in order not to waste their vote, remain invisible.

For protagonists of the Plurality method it is an obvious argument that Compromise is a *sincere manipulation* (Dowding and Van Hees 2008). With only single-seat constituencies, Plurality contributes to shape or preserve a political landscape with two major parties (Duverger's law). Many see this as a good thing. Electoral reformists propose various preferential election methods to replace Plurality (or the 2-day election with Plurality ballots), and they often see both Plurality and unnecessary use of single-seat constituencies as polarizing a society.

The Borda Count is a remedy intended to alleviate the coarseness of Plurality, but is worse than the ailment it should cure. In Borda elections, there is an overwhelming urge to apply "Burying" (0.3): In a race with front-runners P and Q, every voter who switches from PQRSTUVW to PRSTUVWQ, makes up for seven others who rank Q first, P second. Attempting to trade sincerity for power, they rank *insincerely* in {Q, R}, {Q, S}, {Q, T}, {Q, U}, {Q, V}, and {Q, W}. Fear that a political enemy will use Burying to make Q defeat P motivates more voters to join the action. They act under pressure from a dilemma *cruelly* imposed on them by the Borda Count.

Mutually stimulated insincerity may even hide that a sincere majority would have ranked both P and Q before the declared winner, and is a step towards anarchy.

When confronted with *Burying* (0.3), Borda is reported as saying *"My scheme is only intended for honest men"* (Black 1958). However, the Borda Count has a serious flaw even when all voters are both honest and well informed:

> ***Sports journalists use the Borda Count to elect their country's "athlete of the year".* Comparison of candidates in the same sport is objectively based on results. Comparison of candidates from different sports is much more demanding and cannot be equally fine-tuned. Candidate A in sport I has 60% of the top-ranks, and B in sport II has 40%. However, in everybody's mind, candidate C in sport II is very close to B, just slightly behind B according to results. Every ballot is ABC… or BCA… . Every B-supporter gets a double vote in {A, B}, and B becomes Borda-winner.**

A set of k candidates are *"clones"* if they occupy k consecutive ranks in every ballot (Tideman 1987). The Borda protagonist Dummett (1998) was concerned about the *(dis)similarity effect*,

where A loses to B because of dissimilarity between a small group of candidates similar to A and a larger group similar to B. Clones make a special case suited for theoretical study.

Electing the candidate with highest Borda score is mindless sophistication; eliminating the one with lowest Borda score makes sense. Condorcet methods are recognized improvements. Dummett suggested an alternative remedy; a description and assessment is in Schulze (2002).

On Borda's suggestion, the French Academy of Sciences for a while used the already controversial Borda Count to elect new members. The combined efforts of two other prominent members caused its repeal: P-S de Laplace and his younger associate N. Bonaparte (Szpiro 2010).

Each derived method, IRV and Baldwin, is remarkably better than its parent, but not completely free from some of the same unwanted possibilities, and even comes with some new ones.

Problems with IRV

Opportunity and motivation **The pictograms of Figure (1.3) show a Plurality winner B with all the "right wing" support, while a "center/left" majority includes |ACB|+|CAB| anti-B voters. A high value of x = |ACB| means that most A-supporters are closer to C than to B and gives a small ξ, see (2.5): If possible at all, a Pushover win for B cannot be easy; see Figure (2.1).**

Technically, an anti-A action may be realistic if the interval (α, γ) is not too short. However, actionists must be B-supporters moving from BCA to CBA, and most likely |BCA| < |BAC|, see Figure (1.3). When BAC-voters resent an anti-A action favoring C, motivation will be weak among BCA-voters. When A is Condorcet winner, an anti-A campaign is not likely to be popular.

EXAMPLE (2.1*)* ***Motivation for strategic voting in multi-candidate IRV*** **In round 1 of Frome 2009, the six candidates B, C, A, L, M, N got 7576, 5041, 4557, 1267, 734, 134 top-ranks, respectively. A pre-election opinion poll might have detected the gap from A to L and been used to predict elimination of L, M, and N. But an anti-A action is less realistic than in an election with {A, B, C} only: B-supporters will hardly antecipate eliminations of {L, M, N}, see A as a threat, and move C ahead of B, e.g. from BLMNCA to CBLMNA, in order to eliminate A.**

Voters want their ballot to express their true opinion, and it is even harder to misrepresent their preference in {N, C}, {M, C}, {L, C}, and {B, C) than just in {B, C}. The situation must be common in IRV-elections with many candidates: In Frome, 2135 voters had L, M or N on top, but every ballot, also the 7576 with B on top, had to include them; it seems likely that more ballots had some of them ahead of the "leftists", C and A.

For fair comparison: The Borda Count allows *manipulation by Burying,* which even lets one "manipulated ballot" neutralize *several* ballots from sincere supporters of an opponent.

Anyway, voters mentally ready to join an anti-A action and familiar with its logic, probably saw no sign that it could make a difference in Frome: For an independent, A did remarkably well in round 1. Surprisingly, A also got 5562–4557 = 1005 of the 2135 votes transferred from {L, M, N}, becoming Condorcet winner and a serious contender for the vacant seat; see Figure (1.3).[5]

Counter-strategy from anti-B voters **If B-supporters still should try an anti-A action and move from BCA to CBA, each actionist would *decrease* |A|−|C| by 1. Anti-B voters may fear the Plurality winner; this is an incentive to switch from CAB to ACB and protect Condorcet-winner A. Many of the 4425 CAB-voters would have been motivated to compromise, each of them *increasing* |A|−|C| by 2.**

Random changes **More realistically, unavoidable random fluctuations in participation could have caused |A|−|C| to change sign but this cannot be classified as a "non-monotonic" event: Random participation among the 8226 anti-B voters would have influenced |A|−|C| much more than random participation among a much smaller set of |CBA| = 1107 anti-A voters.**

Below, in section 3, the conditions for non-monotonic events in 3-candidate IRV and the number of actionists are determined for a given standard tally (0.7) in terms of x or y. However, it is *the possibility* of non-monotonic effects which is behind many attacks on IRV. What is possible is, in general, unlikely; see Example (0.1) on creative accounting.

Legitimacy of IRV **Only *v* and *vi* have an IRV-winner who is not also Condorcet-winner. The third pictogram of Figure (1.2) shows how C, in constellation *vi*, is eliminated despite two massive pairwise victories and makes it easy to see the associated Abstention mechanism. The possibility that a weak No-Show accident has occurred is a strong signal that structural features of the preference distribution should get attention. Both non-monotonic and No-Show events are linked to the fact that IRV ignores the subsidiary rankings of the supporters of runner-up B. The B-supporters experience that their subsidiary rankings are ignored. A Baldwin/Borda elimination when three candidates remain should improve the method's legitimacy.**

If a Condorcet method is used instead of IRV, the outcome is the same in constellations *i*, *ii*, *iii*, and *iv*, where IRV-winner A also is Condorcet-winner. However, in every 3-candidate Condorcet

[5] As incumbent, but still independent, A (Geoff Brock) kept the seat in ordinary elections 2010, 2014, 2018.

tally, a given cycle XYZX may have been reached from a Condorcet ranking XYZ [or YZX or ZXY] in an attempt to win with Burying (0.3): Supporters of the runner-up Y [or Z or X] change ballots from YXZ to YZX, making Z beat X in Condorcet's relation [*mutatis mutandis*]. In retrospect, it is *theoretically possible* that the supporter group who won the cycle-break did it. Burying in a Condorcet election is visualized with an "Election Box".

REMARK (2.1) *The "Election Box"* extends the rectangles of Figure (1.1) to 3D modelling. Three cuts make eight small boxes with edge lengths from a (unique) set of three numbers:

$$|A| \text{ x } |B| \text{ x } |C| = (\delta_B+\delta_C) \text{ x } (\delta_C+\delta_A) \text{ x } (\delta_A+\delta_B)$$

There is one *block* (small box) for each constellation in Figure (0.1). Family $\mathcal{B}$ is top layer with blocks of size $\delta_C\delta_C\delta_B$, $\delta_B\delta_C\delta_B$, $\delta_B\delta_A\delta_B$, $\delta_C\delta_A\delta_B$ for constellations *iii*, *iv*, *vi*, $iii_{(cyclic)}$. Respectively, blocks of size $\delta_C\delta_C\delta_A$, $\delta_B\delta_C\delta_A$, $\delta_B\delta_A\delta_A$, $\delta_C\delta_A\delta_A$ for constellations *i*, $i_{(cyclic)}$, *v*, *ii* in family $\mathcal{A}$ fit underneath. Since $\delta_B < \delta_A < \delta_C$, the block for *vi* [i] is always the smallest [largest]. One may check that the surface for Perfect Pie-sharing includes the six outer edges (of lengths |A|, |B|, and |C|) not touching any of the two cycle blocks, and also includes the point where the blocks meet.

A matter of rôles C-supporters have the power to decide which family the election belongs to, see (0.9): With $|CAB| > \delta_A$, they make A defeat B in Condorcet's relation (election in family $\mathcal{B}$), and Plurality winner B is runner-up in IRV, as in Figure (1.1) which is a cross section through the top layer of the Election Box. With $|CAB| < \delta_A$, they make B defeat A in Condorcet's relation; then Plurality winner B is also IRV-winner (election in family $\mathcal{A}$).
However, for a study of non-monotonic and No-Show events, it is convenient to relabel according to the rôles of winner and runner-up in IRV, A and B respectively, and let the candidates switch rôle labels when the election drops to the lower layer.
Moreover, it is convenient to let (x, y) = (|ACB|, |BCA|), as in Figure (1.1) also with IRV-winner A as Plurality winner. The big rectangle which in Figure (1.1) has "portrait shape" in family $\mathcal{B}$, then gets "landscape shape" in family $\mathcal{A}$. Thereby, the cycles are still in NW.

A niche for Baldwin

In IRV the No-Show possibilities are connected to the elimination of a Condorcet-winner and that the subsidiary ranking from B-supporters are not noticed. Below, in section 3, the possibilities for No-Show events in 3-candidate IRV and the number of actionists are determined for a given standard tally (0.7) in terms of x or y. However, in Condorcet methods the No-Show events are technically very different. A Baldwin elimination towards the end is a natural adjustment. Baldwin has also flaws to be considered, e.g. its own version of the "Abstention *strategy*", which violates the Participation criterion. [6]

Weak or Strong No-Show. The Participation criterion has this definition:

"Adding one or more ballots that vote X over Y should never change the winner from X to Y." (Electowiki: *Participation criterion*)

If X is *top-ranked* in added ballots which change winner from X to Y, this is the *strong No-Show accident*, which must be seen as an anomaly, but by Theorem (0.2) it cannot occur in IRV. In IRV Abstention starts in constellation *v* or *vi*; see the construction of *Ward Paradox* I above. The criterion is obviously violated, but this is only a *weak No-Show*. Abstention is "good news and bad news" for C, who may be helped first, then harmed.

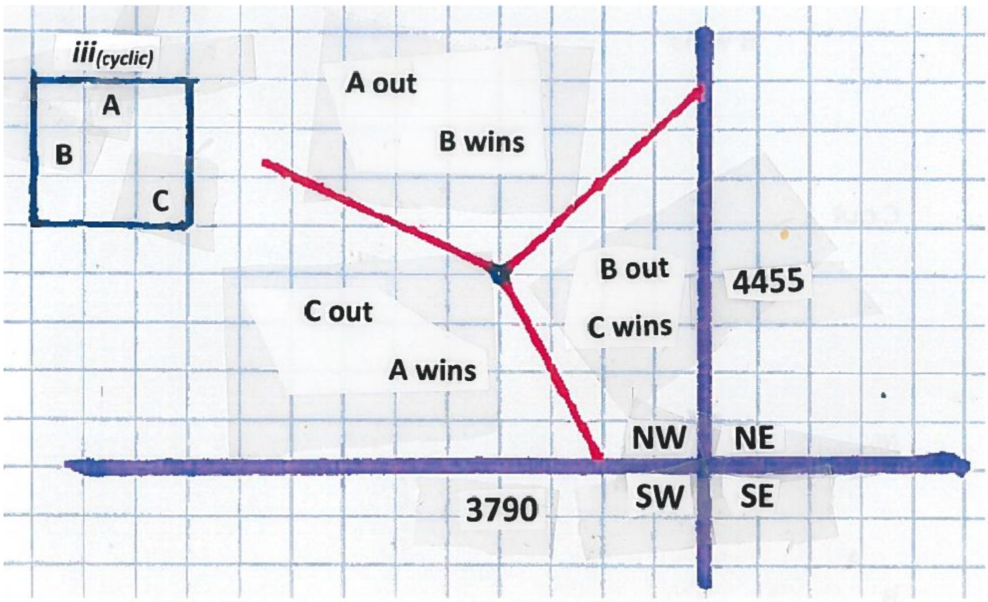

FIGURE (2.3) *Baldwin's treatment of the Condorcet cycles in Figure (1.1)*
As in Figure (1.1), IRV-winner A is also Baldwin winner in SW and SE; C is Baldwin winner in NE. Baldwin's cycle-breaks in NW are shown. Along three rays from (x,y) = (3790,4455) there is a tie for second and third in Borda-score. In non-cyclic elections, the Burying strategy must be performed by supporters of the candidate who is second in the Condorcet-ranking. A is Baldwin winner in SW and SE, but also in a part of Cycle-land NW which makes A's territory not convex. Similarly, C's territory, NE and part of NW, is not convex. Only with positional methods (i.e. point-awarding, like Plurality or Borda Count) is the set V(X) of vote vectors won by candidate X convex (Young 1975).

[6] Abstention is not a strategy in the usual sense, as it changes only one account; see e.g. (0.6).

Baldwin's cycle-break **With eliminations based on Borda-score, the Baldwin tally performs its cycle-breaks invisibly. Possibilities for strategic voting or No-Show events are illustrated in Figure (2.3), from a part of the xy-plane in Figure (1.1). It shows how**

- **A-supporters let A snatch Baldwin victory from C: Burying from constellation *vi* in the NE, they switch from ACB to ABC, reduce x, bury C, and move to NW, provided that Baldwin's cycle-break is favorable for A;**
- **B-supporters let B snatch Baldwin victory from A: starting from *iii* in SW, they switch from BAC to BCA, increase y, and bury A. Similarly by Remark (2.1), in the Election Box**
- **C-supporters may start from constellation *ii* below and switch from CBA to CAB.**

Ward Paradox IV (Baldwin's method) **If candidate X is *Condorcet-winner* in both wards, then X wins the pairwise comparisons also in the constituency. Thus, a case of the Ward Paradox requires a win by cycle-break in at least one ward. Figure (2.3) visualizes the Ward Paradox structure with all three vote vectors in the xy-plane: One paradox is visualized by a line through "B out, C wins" connecting a point in "C out, A wins" with a suitable point in SW or SE: A wins both wards and C wins the constituency. For another case, draw a line through "A out, B wins" connecting a point in "B out, C wins" with a suitable point in NE.**

Condorcet methods and Burying ***All Condorcet methods allow Burying,*** **most importantly with start from a transitive Condorcet relation: Starting from Condorcet ranking XYZ, Z-supporters moving between ballots ZXY and ZYX cannot change the fact that Z loses in two pairwise comparisons. Thus, only Y-supporters can win with Burying. Y-supporters have the power to decide exactly one pairwise comparison, i.e. in {X,Z}. With |YZX| > δ_Z, Y-supporters create a cycle XYZX. The same cycle could have been created from Condorcet rankings YZX and ZXY, respectively by Z-supporters and by X-supporters.**

For Baldwin, Figure (2.3) illustrates this, e.g. with an election in "B out, C wins": It could have been created, even with start from Perfect Pie-sharing, possibly from SW by B-supporters or from NE by A-supporters. However, only C-supporters from below in the Election Box may be "suspected" of a victory with the Burying strategy (*cui bono*). In "C out, A wins", see Figure (3.2), A-supporters are "suspects". In Figure (3.4), B-supporters are "suspects".

An action against a Condorcet-winner will hardly be a popular project anyway, but just in case, actionists should know that the result may attract attention, especially if the election lands deep inside Cycle-land and far outside the 0.001-zone. Then the pictogram shows an unusually large

T; it will seem that many ballots did not reflect a sincere preference. This prospect is more curious than serious: Burying is a severe flaw in the Borda Count, but it is hard to see why the mere possibility should influence voter behavior in its offspring Baldwin.

The Condorcet winner in Baldwin and IRV **These methods have in common that a Condorcet winner X may be attacked with strategic voting, but in very different ways.**

In Baldwin, with Condorcet ranking XYZ, actionists are Y-supporters who apply Burying; they switch from YXZ to YZX and create a cycle won by Y as shown in Figure (2.3), where X=A in SW and SE, X=C in NE. In IRV, actionists are supporters of B (runner-up and Plurality winner) who switch from BCA to CBA in an anti-A strategy as shown in Figure (2.1), with A=X.

A thorough technical comparison of the anti-X strategy in IRV and the pro-Y strategy (targeting X) in Baldwin is unlikely to change a perception of both as being too impractical. They are at most minor nuisances when a post-election analysis discloses a missed opportunity.

In political elections, a Condorcet-winner X is likely to be accepted as a legitimate representative of the constituency. The fact that IRV in constellations *v* and *vi* eliminates Condorcet-winner C is a consequence of another fact, i.e. that IRV systematically ignores the subsidiary rankings of a large voter group. It is this second fact that damages IRV's legitimacy.

Condorcet methods and No-Show **In a *strong* No-Show accident, new ballots *with winner* X *on top* let another candidate snatch the victory from X. In Condorcet methods, this is impossible if winner X is also Condorcet winner, since the new ballots then make X win with higher margins in each pair. Thus,**

(2.11) a *strong* No-Show accident must *start* from a cycle.

A No-Show accident starting from a Condorcet ranking XYZ must, by (2.11), be weak. If the new ballots are YXZ, Z is still Condorcet loser, and there is no accident. Thus, the new ballots must be ZXY, but X still beats Y in pairwise comparison. To give victory to Y, whom they rank after X, they must create a cycle ZXYZ where Y wins the cycle-break. Thus,

(2.12) a No-Show accident starting from transitivity is weak and leads to a cycle.

With a combinatorial approach, Moulin (1988) showed that every Condorcet method for ≥4 candidates and ≥25 voters allows No-Show.[7] Because of (2.11) and (2.12) one cannot expect No-Show to be of much practical importance in Condorcet elections.

Strong No-Show in Baldwin Baldwin allows strong No-Show events already for 3 candidates. This is harmless since it only happens in Cycle-land, but it is still a genuine anomaly since it is anti-intuitive that new ballots with X on top can take victory away from X. Figure (2.3) illustrates how it may work in Baldwin: On the ray,

$$(x, y) = (3790 + r, 4455 + r),\ r > 0, \qquad (2.13)$$

at the border between B-territory and C-territory, the vote vector is

$$(|ABC|, |ACB|, |CAB|, |CBA|, |BCA|, |BAC|)$$
$$= (1772 - r, 3790 + r, 4425, 1107, 4455 + r, 3760 - r)$$

For $r < \delta_C - 3790 = 332.5$, the vote vector is still in $iii_{(cyclic)}$. With k new CBA-ballots, the Borda-scores become

$$A: 19309 - r,\quad B: 19309 - r + k,\quad C: 19309 + 2r + 2k$$

With k in some interval (μ, ν), $\mu < 0 < \nu$, the vote vector remains in Cycle-land, but is no more in the xy-plane of Figures (1.1) and (2.3) if $k \neq 0$.

Let $\mu < k < \nu$. When increasing k passes 0, there is a strong No-Show accident: The new CBA-voters eliminate A instead of B, and B wins instead of C. In reverse, C wins by Abstention.

[7] For recent work in a quest to find the Condorcet methods that avoid the strong No-Show, see (Duddy, 2014).

3 *Non-monotonic and No-Show events in IRV; conditions and number of actionists*

Each actionist in non-monotonic or No-Show events changes one or two of |A|, |B|, and |C| and therefore also changes Figure (1.1) gradually.

Non-monotonicity **In elections with a transitive Condorcet relation, a Pushover trick starts in constellation *iii* (family $\mathfrak{B}$), according to Theorem (0.1). Its reverse is a trap effect starting in constellation *v* or *vi* according to Figure (2.1).**

The rôles for A, B, and C get permuted in Pushover ↑ or a trap effect ↓; see Figure (3.1)

	C	B	A	action
I *vi*	run.up	winner	elim	
	↑	↑	↑	trick; h moves from BCA to CBA
II *iii*	elim	run.up	winner	
II *vi*	elim	run.up	winner	
	↓	↓	↓	trap; g moves from BAC to ABC
III *iii*	winner	elim	run.up	

FIGURE (3.1) ***Rôle changes in non-monotonic effects: tricks and traps***
The rôles for A, B, C get permuted in a trick effect ↑; see Figure (2.1). h *ballots* change
(*) from 1) runner-up, 2) eliminated, 3) winner to 1) eliminated, 2) runner-up, 3) winner.
To check that the trap effect ↓ is as shown here, it is enough to control that the *opposite ballot change,* from ABC to BAC, is described in (*), and that it changes the rôles the same way as ↑ does. The argument works for any start position in constellation *v* as well as for *vi* in the figure.

The non-monotonicity trick (Pushover) **The vote vector in (2.3) shows when the move of h voters from BCA to CBA is a Pushover. In top-ranks, C must pass A, and similarly, B must stay ahead of A, as shown in Figure (2.1):**

$$(|A| - x) + x \;<\; |CAB| + (|CBA| + h)$$
$$|A| - |C| < h < |B| - |A|$$

Next, B must stay ahead of C in pairwise comparison:

$$x + |CAB| + (|CBA| + h) \;<\; (y - h) + (|B| - y) + (|A| - x); \quad \text{thus,}$$
$$h < N/2 - |C| - 1 = \delta_A - 1$$

Figures (2.1) and (2.2) show in detail how h voters, moving from BCA to CBA, h ∊ (α, min (γ, ξ)), make B snatch the IRV-victory from A with Pushover. Clearly, $|A| - |C| < N/2 - |C| - x$. Since h and x are integers:

(3.1) $$|A| - |C| + 1 \le h < \min(|B| - |A|, N/2 - |C| - x)$$
(3.2) $$0 \le x < N/2 - |A| - 1$$

The non-monotonicity trap **By Figure (2.1), g actionists change ballot from BAC to ABC, and the new vote vector becomes**

$$(|ABC| + g,\ |ACB|,\ |CAB|,\ |CBA|,\ |BCA|,\ |BAC| - g)$$
$$= \ (|A| - x + g,\quad x,\quad |CAB|,\ |CBA|,\quad y,\quad |B| - y - g)$$

Clearly $g \ \leq \ |B| - y$. The trap requires that B is eliminated, so that C meets A in the final.

This is obtained without tiebreak if C gets more top-ranks than B:

$$y + (|B| - y - g) \ < \ |CAB| + |CBA|, \text{ i.e. } |B| - |C| \ < \ g, \text{ and so}$$

(3.3) $$1 + |B| - |C| \ \leq \ g \ \leq \ |B| - y$$

The trap requires C to win the final pairwise comparison with A. Without tiebreaks, this means

$$(|B| - y - g) + (|A| - x + g) + x \ < \ |CAB| + |CBA| + y; \quad \text{i.e} \quad N - 2|C| + 1 \ \leq \ 2y$$

From (3.3) follows $1 - |C| \ \leq \ - y$. Thus, a non-monotonicity trap requires

(3.4) $$N/2 \ - \ |C| + ½ \ \leq \ y \ \leq \ \ |C| - 1.$$

This implies that

(3.5) **a necessary condition for a non-monotonicity trap in a 3-candidate IRV election is that $N/4 < |C|$**

In Frome 2009, the condition (3.2) is $0 \leq x \leq 4191$ and is satisfied; (3.4) and (3.3) become

(3.6) $$4123 \ \leq \ y \ \leq \ 5531$$
(3.7) $$2684 \ \leq \ g \ \leq \ 8215 - y$$

EXAMPLE (3.1) **The point (x, y) = (3801, 4200) is on the stapled line in Figure (1.1), in the NW, just outside the 0.001-zone. The pictogram is in Figure (3.2).**

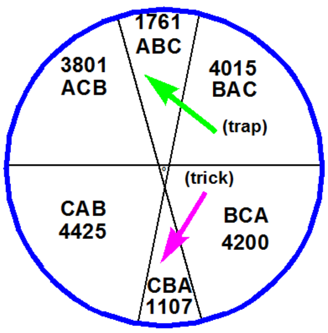

FIGURE (3.2) ***A trick (Pushover) and a trap on the stapled line in Figure (1.1)*** **[8]:**
With (x, y) = (3801, 4200), on the stapled line in Figure (1.1), the election is cyclic; thus, triangle T covers the circle center. T also covers a fraction 0.001213 of the circle area. The constellation is $iii_{(cyclic)}$.
For a trick, h voters move from BCA to CBA; by (3.1), $31 \leq h \leq 321$; getting B instead of A.
For a trap, g voters move from BAC to ABC; by (3.3), $2684 \leq g \leq 4015$, getting C instead of A.

---

[8] **As indicated in Figure (1.1), cycles are very rare in political elections, but are useful to illustrate the two non-monotonic events, trick and trap, in one pictogram. Figure (3.4) has a similar purpose.**

No-Show "paradox" In elections with a transitive result, a No-Show accident starts in constellation *iii* (family $\mathfrak{B}$), according to Theorem (0.2). Its reverse is a "win by Abstention". By (3.15) below, this is possible only in constellations *v* and *vi*.

The rôles for A, B, and C get permuted in an accident ↑ or an abstention win ↓; see Figure (3.3):

	C	B	A	action
I *vi*	run.up	winner	elim	
↑	↑	↑	↑	accident; k CAB-voters enter into election
II *iii*	elim	run.up	winner	
II *vi*	elim	run.up	winner	
↓	↓	↓	↓	abstention; q BCA-voters exit from election
III *iii*	winner	elim	run.up	

FIGURE (3.3) *Rôle changes in No-Show effects: accident and abstention*

In the ↑ action (*an accident*), k actionists *join the election* and vote CAB, Theorem (0.2). They are unfortunate: C passes winner A in top-ranks, and B becomes winner instead of A. To check that the abstention ↓ is as shown in Figure (3.3), it is enough to control that the *opposite effect* is as in ↑: B (1^{st} in the *ballot*) goes from eliminated to runner-up; C (2^{nd} in the *ballot*) goes from winner to eliminated; A (3^{rd} in the *ballot*) from runner-up to winner.

The No-Show accident k new voters change the vote vector from (1.3) to:

$$(|ABC|, |ACB|, |CAB| + k, |CBA|, |BCA|, |BAC|)$$
$$= (|A| - x, x, |CAB| + k, |CBA|, y, |B| - y).$$

One condition is that C passes A in top-ranks. This happens without tiebreak if

$$(|A| - x) + x < (|CAB| + k) + |CBA|,$$

(3.8) $$|A| - |C| < k$$

Next, B still wins over C in pairwise comparison. This happens without tiebreak if

$$x + (|CAB| + k) + |CBA| < y + (|B| - y) + (|A| - x)$$

(3.9) $$k < N - 2|C| - 2x$$

with $N = |A| + |B| + |C|$. Since the numbers are integers, (3.8) and (3.9) are equivalent to

(3.10) $$|A| - |C| + 1 \le k \le N - 2|C| - 2x - 1$$

The maximal value for x follows from (3.10):

(3.11) $$x \le |B|/2 - 1$$

The No-Show Abstention In row II of Figure (3.3), q BCA ballots are removed:

$$(|ABC|, |ACB|, |CAB|, |CBA|, |BCA| - q, |BAC|)$$
$$= (|A| - x, x, |CAB|, |CBA|, y - q, |B| - y).$$

B loses to C in top-ranks and gets eliminated. This happens without tiebreak if

$$(y - q) + (|B| - y) < |CAB| + |CBA|$$

(3.12) $$|B| - |C| < q$$

A must lose to C in pairwise comparison. This happens without tiebreak if

$$(|B| - y) + (|A| - x) + x < |CAB| + |CBA| + (y - q)$$

(3.13) $$q < 2|C| - N + 2y$$

Since the numbers are integers, (3.12) and (3.13) imply

(3.14) $$|B| - |C| + 1 \leq q \leq 2|C| - N + 2y - 1$$

Clearly $y \leq |B|$; another condition follows from (3.14):

(3.15) $$|B| - |C| + |A|/2 + 1 \leq y \leq |B|,$$

$$\text{here } \delta_C = N/2 - |C| < |B| - |C| + |A|/2 + 1$$

(The second line simplifies to $|C|/2 < |B|/2 + 1$.) Figure (2.1) shows that, in family $\mathcal{B}$, the constellation is *vi* or *iii(cyclic)*; similarly it is *v* or *i(cyclic)* *in family* $\mathcal{A}$. From (3.15) follows also a necessary condition for an election to be the result of a No-Show accident:

(3.16) $$|A|/2 \leq |C| - 1$$

EXAMPLE (3.2) In Frome 2009, the condition (3.11) becomes $x \leq 4191$ and is satisfied. The point $(x, y) = (3801, 5600)$ is on the stapled line in Figure (1.1), in the NW, and far from the 0.001-zone. The pictogram is in Figure (3.4).

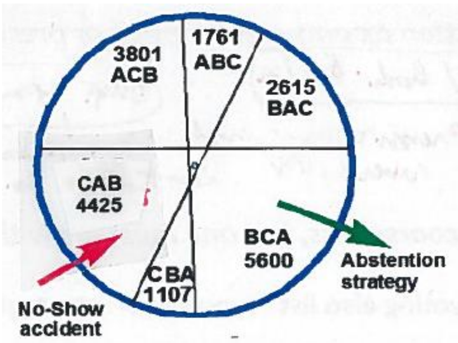

FIGURE (3.4) *No-Show: accident and abstention on stapled line in Figure (1.1)*

With $(x, y) = (3801, 5600)$, on the stapled line in Figure (1.1), the election is cyclic; thus, triangle T covers the circle center. T also covers a fraction 0.005049 of the circle area. The constellation is *iii(cyclic)*.

By (3.10), k voters *join* CAB, $31 \leq k \leq 642$, eliminate A, B wins, but C is too strong at 643.

By (3.11), q voters *leave* BCA, $2684 \leq q \leq 2954$, the influence of B-supporters drops in two steps: B gets eliminated at $q = 2684$, letting C win, and at $q = 2955$, the remaining BCA ballots get too few to help C against A.

A vote vector with Perfect Pie-sharing that allows a win by abstention is obtained by following the straight line $y = 5600$, until intersection with the middle curve in Figure (1.1).

4 *IRV in various contexts*

IRV is the single-seat version of STV (Single Transferable Vote). Elimination and vote transfer are ideas developed in the 19th century, in the wider context of STV which was intended for proportional representation in multi-seat constituencies. This part of IRV's historical context is described by Tideman (1995). Glimpses of older historical roots, in work of Borda and Condorcet, appear above, and some of their basic ideas have in modern times been found in medieval writings. Below, we consider the importance of Black's work two generations ago.

IRV also has a special legal/mathematical framework. Claimed anomalies in IRV, e.g. the Pushover strategy, have got a conspicuous place in the public and political debate, but motivation and feasibility seem inadequate for strategic voting.

IRV, Black and Condorcet's paradox

Duncan Black (1948) had the idea that a vote vector has a structure reflecting a common perception among voters of their political landscape: Different voters see themselves at different points in the landscape and rank the candidates (alternatives) accordingly.

For n candidates, Black's 1D proximity model implies that voters choose from 2^{n-1} "single-peaked" ballots and that Condorcet's social relation of pairwise comparison then cannot contain cycles. One should imagine that voters in the Single-Peak model agree on the candidate sequence on the line, *but have different perceptions of the distances in between.*[9]

In the 2D model of Perfect Pie-sharing, 3 candidates are represented by ideal points that are corners of a "candidate triangle". The area of T in the pictogram measures how well this proximity model fits a real election. Some caution is required: The curves of Figure (1.1) illustrate that at the SW and NE corners, the area of T as function of (x,y) has cusps and is highly volatile.

In real political elections, the Single-Peak condition may, at best, be approximately satisfied for certain candidate triples like "far left, center, far right". Nevertheless, Condorcet cycles are remarkably rare: Near the two corners, T may be large but is far away from the circle center. Near $(x,y)=(\delta_C,\delta_C)$ T gets small and is unlikely to cover the center (which is necessary for a cycle).

---

[9] Already 4 candidates A, B, C, D from left to right make it difficult to assign "ideal points" to voters on the same line as the candidates: Ballots BCDA and CBAD are both single-peaked. Rankings in {B,C} show the first voter as most leftist, but rankings in {A,D} show the second voter as most leftist. Moving from left of A to right of D, a voter changes ballot 6 times, choosing 7 of the 8 single-peaked ballots along the way.

A mathematical framework **General results show that designing good preferential election methods must be difficult. The message of Arrow (1950) is the incompatibility of two axioms:**

● SO (Social Ordering), i.e. the tally gives a complete and transitive ordering of the candidates;
● IIA (Independence of Irrelevant Alternatives), i.e. the social ordering inside {X, Y} depends only on the voters' rankings inside {X, Y}.

However, Arrow also included another axiom which blurred the message. Wilson (1972) found the Weak Pareto axiom *essentially* irrelevant, removed it, and added to Arrow's *dictatorship* just two other undemocratic tally results: *Either* there is an *anti-dictator* (whose ballot reversed is the social relation), *or* the social relation is *trivial* (all candidates are in one indifference class).

The strategies in (0.3) violate IIA. The G–S theorem (Gibbard 1973, Satterthwaite 1975) concerns all forms of strategic voting (i.e. a ballot with X before Y is changed so that X wins instead of Y; the purpose with the original ballot is better served by the changed ballot). The G–S axioms are:

● The tally always picks a unique winner;
● every candidate will win with a suitable preference distribution.

G–S concludes that there is a *"G–S-dictator"* (i.e. whose top-ranked candidate always wins).

However, Black (1948) had already made the picture less gloomy: His Single-Peak condition describes voter behavior that satisfies SO and IIA. So does the Perfect Pie-sharing condition. Figure (1.1) illustrates some robustness to a cyclic result when a vote vector with Perfect Pie-sharing is exposed to perturbations of the real world.

Normative axioms **SO, IIA and "no dictator" are among many natural desiderata, useful as benchmarks, but often mutually inconsistent. Enforcing one of them *by law* makes it harder to design useful preferential election methods. There have been attempts to require Monotonicity or Participation criteria by law, but this only narrows down the field of useful methods. There is no way around assessing how frequent and how annoying violations of a desired criterion will be, given how real voters behave.**

EXAMPLE (4.1) **In 2008, the German federal constitutional court found the rules then used to distribute party seats in the national assembly to be *unconstitutional* because, under certain conditions, a suitable decrease [increase] in the number of "*Zweitstimme*" (votes supporting a party) would have increased [decreased] the party's total number of seats (usually by 1).[10] In an**

[10] Ballots contain an "Erststimme" in a single-seat constituency and a "Zweitstimme" in a nationwide tally

introductory "Leitsatz", the court states that equality and legal immediacy is harmed if such events are possible.

Such events are analogues of the strong No-Show in Baldwin's method. There it exists only in Cycle-land, where real vote vectors almost never land. Only there is it theoretically possible that extra XYZ-ballots may let Y win instead of X. To prevent use of Baldwin's method by law for such reason, would be absurd. To prohibit the Borda Count in public elections because it makes voters fight in a "chicken game" with Burying as a weapon, would be a different story.

In general, comparison of two proposed election methods should build on empirical facts and realistic simulations that reflect a structure in real vote vectors.[11] Common voter behavior gives very few 3-candidate finals in Cycle-land, and if it happens, a small T means that pairwise comparisons are not so far from 50–50. The remote possibility of a cycle is unlikely to harm legitimacy.[12]

Baldwin for legitimacy. **IRV's immunity to Burying comes from counting top-ranks only. Although it is a principle cherished by IRV-advocates, it should be weighed against other principles. Elimination of Condorcet-winner C in constellation *v* and *vi* may even ignore the clear signal from a violation of the Participation criterion.**

For all (x, y), the vote vector (1.3) is compatible with the standard tally (0.7) in Frome 2009. This includes a realistic possibility (Perfect Pie-sharing) in Figure (1.2), where C is Condorcet-winner with overwhelming pairwise victories and 28.6% of the top-ranks, but is eliminated in IRV since the massive subsidiary preference of B-supporters (BCA 7062 v BAC 1153) is also ignored.

In IRV, those who have supported the runner up from round one experience that their subsidiary rankings are completely wasted, not even noticed. With one Baldwin/Borda elimination, this cannot happen to a voter with a complete ballot.

[11] Smith (2010; revised 2016) gives tables of the frequencies of various combinations of anomalies in IRV with 3 candidates, one of them for a stochastic selection of ballots satisfying Black's Single-Peak model.

[12] In a political assembly where members from the same party coordinate their preferences, the vote vector is likely to be further away from Perfect Pie-sharing, and more often be cyclic; see Figure (1.1). However, with n≥3 alternatives, the voting procedure is usually a sequence of dichotomous steps, either an "amendment" process of n-1 steps, or a "successive" process of j steps, 1 ≤ j ≤ n (Rasch 1995). Extra information beyond the voting procedure is then needed to conclude if there was an underlying cycle in the distribution of coordinated preferences.

Continuing one-by-one eliminations, also a Baldwin finish leads to a winner with >50% support, but now against a runner-up with supporters whose subsidiary votes have not been ignored. That should be good for legitimacy. Still, the sequence of IRV-eliminations is important, making it impossible for a candidate to reach the penultimate 3-candidate round by means of Burying. That should also be good for legitimacy.

Perceived anomalies

Fishburn and Brams (1983) listed four "paradoxes" in IRV:

1) a Condorcet winner fails to win IRV; 2) No-Show; 3) non-monotonicity; 4) the Ward Paradox. Of these, 2) and 3) get most public attention. In Figure (1.1), constellations *vi* and *iii*$_{(cyclic)}$,

y ϵ [4123, 5531] (up from bottom) gives a monotonicity failure, by (3.4) allowing a trap;
y ϵ [5465, 8215] (down from top) allows a No-Show effect, an Abstention gain by (3.15).

The two intervals occasionally intersect like here. In constellations *v* and *vi*, both effects are linked to paradox 1), the elimination of Condorcet winner C. The actionists are

either unfortunate trap "victims" moving from BAC to ABC as in Figure (3.2)
or BCA voters abstaining to promote C in order to defeat A, see Figure (3.4).

With x = 3801, Figures (3.2) and (3.4) show two elections, and one of them was *possibly* the true but unknown Frome 2009. If they had been observed with other x-values, near the middle curve of Figure (1.1) in constellation *vi*, they should not have surprised anyone. The trap and Abstention effects should generally be considered as paper events anyway, but as theoretic possibilities they are signals that should be observed and understood.

Among the many thousand Australian standard IRV-tallies (0.7), there must be many in constellations *v* and *vi*, with elimination of a Condorcet winner. However, they could not be observed from the published standard tally and did not become evidence in a debate on election rules. Unfairness to an eliminated Condorcet-winner is one issue, particularly if the election is in the upper y-interval. The treatment of the runner-up's supporters is likely to damage IRV's legitimacy even more.

Paradox 4) of Fishburn and Brams concerns non-convexity through a concomitant, the Ward Paradox. For every *positional* method, the set V(X) of vote vectors where candidate X wins, is convex. They also have the *scaling property*, i.e. vote vectors v and λv give the same result for λ>0.[13] Assume that we study another preferential election method with this property.

[13] One may well imagine election methods without the scaling property: Let IRV eliminations stop when all remaining candidates have reached a certain fixed number of top-ranks, e.g. 1000, and then switch to Baldwin/Borda eliminations.

If a preferential method with the scaling property makes V(X) non-convex, it is possible to pick two vote vectors, v_1 and v_2 in V(X) such that $(v_1+v_2)/2 \in V(Y)$ with $Y \neq X$. Then $(v_1+v_2) \in V(Y)$, and non-convexity is didactically converted to the Ward Paradox, with v_1 in ward 1 and v_2 in ward 2. The result of Peyton Young (1975) is often formulated accordingly: Only positional methods avoid the Ward Paradox.

In *Ward Paradox I &II* above, the No-Show and the anti-A strategy seen in Figure (2.1) are behind the construction of cases with *non-cyclic vote vectors* in both wards and in the constituency. The Ward Paradox is odd enough to attract attention, but it is not easy to see how often it might appear in IRV-elections tallied in real wards established before the election.[14]

The set of vote vectors which make candidate X Condorcet winner is also defined by linear inequalities and is therefore convex. Needing a cyclic election, as in *Ward Paradox IV* above, the Ward Paradox cannot have any practical significance in Condorcet methods.

In a Condorcet method, a Condorcet winner may

- **only lose to one strategy, i.e. Burying, which must start in or lead to a cycle;**
- **only lose in a No-Show accident, which must start in or lead to a cycle;**
- **not be Condorcet winner in both wards of a Ward Paradox case.**

Paradox 1) of Fishburn and Brams occurs if the constellation is *v* or *vi*, as in Figure (1.1). Counterfactual but realistic cases are hidden behind the same standard tally (0.7). Figure (1.2) visualizes the elimination of a Condorcet winner C, with opponents A and B on two sides, despite C's massive pairwise victories over both.

Strategic voting: motivation and feasibility

The most common form of strategic voting is Compromise in Plurality elections: A voter with preference ranking XYZ switches to YXZ, and uses a ballot for Y instead of one for X; it is an anti-Z strategy. In IRV, an anti-A strategy might attract B-supporters because also B may win, as in Figure (2.1). However, there is a simple counter-strategy and a natural lack of motivation.

The vote vector in Frome 2009 remains unknown, but the three pictograms of Figure (1.3) are possibilities in the 0.001-zone. There is a close race between A and C to be promoted to the final and become challenger of B. The result depends on the sign of the difference |A|−|C|.

[14] Irish presidents are elected by IRV in 43 wards, but these can only count top-ranks and report numbers to a central where numbers are aggregated to a national total and the candidate to be eliminated is identified. The identity is reported back to the wards before a new round starts.

With strategic voting (Pushover/Compromise) anti-A voters may make the difference negative by moving from BCA to CBA, and prevent election of Condorcet winner A. More obviously, anti-B voters could make or keep it positive by moving from CAB to ACB, and avoid election of Plurality winner B.

In Frome 2009, 31 moves from BCA to CBA would have avoided election of A, but if the strong performance of independent A had been expected, the importance of moves by anti-B voters from CAB to ACB would have become obvious; see Figure (1.3). The political distance from C and A to B is motivating, and the recruitment reservoir of CAB-voters (for an anti-B Compromise) is larger than the recruitment reservoir of BCA-voters for an anti-A action (trick or Compromise). Moreover, one move from CAB to ACB neutralizes two moves from BCA to CBA. The anti-B voters' dominance over anti-A voters is clear if strategic voting seems relevant.

Also pure randomness in participation may influence the sign of $|A|-|C|$, but random participation means more in a set of $|A|+|C|$ voters than in a subset of $|CBA|$anti-A voters.

Burlington 2009 **Motivation and feasibility are key factors when voters consider strategic voting. When IRV allows incomplete ballots, one gets a glimpse of the motivational background. In the mayoral election in Burlington (Vermont) 2009, the political structure had one similarity to Frome 2009: In both, the Plurality winner lost in the final round.**

However, Burlington 2009 was in constellation *vi*. Wright (republican), Kiss (progressive), and Montroll (democrat) got top-ranks as follows:

$$(|W|, |K|, |M|) = (3297, 2982, 2554)$$

The constellation was *vi* since pairwise comparison gave the opposite ordering, with margins

590 (M over K), 250 (K over W), 929 (M over W)

With three candidates left, 1289 incomplete ballots supported W, i.e. expressed no preference in {K, M}. One explanation is indifference, another is understanding that with W in the final (obviously to be expected), a subsidiary preference in {K, M} would be ignored anyway. Could M still have been the target of a successful anti-M action? If so, before the action,

***either* K had < 2554 top-ranks, and got > 428 top-ranks from W-supporters,**

***or* W had < 2554 top-ranks, and got > 743 top-ranks from K-supporters.**

Thus, Burlington 2009 was well inside the interval (α, γ) of Figure (2.1). However, c*ounting only complete ballots*, published election data shows smaller numbers:

$$(|KWM|, |WKM|) = (371, 495).$$

It takes some motivation to move from XYZ to YXZ in an anti-Z action, a particularly clear preference gap from X and Y to Z. For safety and to show some respect for X, it is natural that an actionist expresses a complete preference YXZ. If an anti-M action (trick or Compromise) caused M's elimination, many actionists omitted this (428 > 371; 743 > 495).

The symmetrized vote vector in Burlington 2009 is a trap in the theoretical sense, i.e. it made a non-monotonicity event *possible*, but it was utterly unrealistic that enough "victims" should walk into one of the traps, leaving the interval at α or γ, see Figure (2.1).

Burlington aftermath **Burlington 2009 was followed by a heated debate. The elimination of Condorcet-winner M was an obvious fact and according to the rules. However, two other facts had strong impact:**

1) (|WMK|, |WKM|) = (2157.5, 1139.5), and 2) this was ignored according to the rules. Subsidiary voting among the runner-up's supporters made M a clear Condorcet winner, and they were ignored. Enraged W-supporters were active in the political fire that followed. The often repeated statement, *"Voters can vote for their favorite candidate without worrying about wasting their vote"*, was not likely to stop the fire. In 2010, Burlington quashed IRV in a referendum. A Baldwin final will provide foundation for the statement.

It is a legitimate opinion that the large |K|−|M| justified the elimination of M. Moreover, in the symmetrized vote vector (|KMW|,|MKW|) = (2327,1559.5); thus the anti-W voters preferred K to challenge W in the final. However, it is an at least equally legitimate opinion that ballots from the supporters of runner-up W should not be ignored and wasted.

Familiar anomalies were also adduced: A No-Show label was (mistakenly) attached by anti-IRV campaigners, e.g (Gierzynski 2009, 2011), but the election came close with y=|BCA|=|WMK| just outside the interval of (3.15). The vote vector was a *trap* in constellation *vi*, in the interval (α, γ) of Figure (2.1), but as shown above, it was very far from both end points.

Conclusion **A Baldwin/Borda elimination in the penultimate round is suggested to avoid wasting up to one half of the ballots and to improve the legitimacy. With an elimination ordeal by top-ranks until then, Burying does not work. Also the anomaly mechanisms remain until then, but the exploration above indicates that they should not damage the method's legitimacy.**

References

K. ARROW (1950) *A difficulty in the concept of social welfare*, Journal of Political Economy, pp. 328-346.

J. M. BALDWIN (1926), *The technique of the Nanson preferential majority system of election*, Proc. Roy. Soc. Victoria n. s. 39, pp. 42-52.

D. BLACK (1948), *On the rationale of group decision making*, Journal of Political Economy, pp. 23-34.

D. BLACK (1958), *The theory of committees and elections*, Cambridge University Press.

K. DOWDING AND M. VAN HEES (2008), *In praise of manipulation*, British Journal of Political Science, pp. 1-15.

C. DUDDY (2014), *Condorcet's principle and the strong no-show paradoxes*, Theory and Decision, 77, pp. 275 – 285.

M. DUMMETT (1998), *The Borda Count and agenda manipulation*, Social Choice and Welfare, pp. 289-296.

P. C. FISHBURN AND S. J. BRAMS (1983), *Paradoxes of preferential voting*, Math. Mag., 56, pp. 207-214.

W. V. GEHRLEIN (2006), *Condorcet's paradox*, Springer.

A. GIBBARD (1973), *Manipulations of voting schemes: A general result*, Econometrica, 41, pp.587-601.

A. GIERZYNSKI (2009), *Instant runoff voting*, Vermont Legislative Research Shop, University of Vermont.

A. GIERZYNSKI (2011), *Saving American elections: A diagnosis and prescription for a healthier democracy*, Cambria Press.

I. MCLEAN (2002), *Australian electoral reform and two concepts of representation*. Paper for APSA Jubilee Conference, Canberra, October 2002.

N. R. MILLER (2017), *Closeness matters: monotonicity failure in IRV elections with three candidates*, Public Choice 173, pp. 91-108.

H. MOULIN (1988), *Condorcet's principle implies the No-Show paradox*, Journal of Economic Theory 45, pp.53-64.

E. J. NANSON (1882), *Methods of election*, Trans. Proc. Roy. Soc. Victoria, 18, pp. 197-240.

J. T. ORNSTEIN AND R. Z. NORMAN (2014), *Frequency of monotonicity failure under instant runoff voting: estimates based on a spatial model of elections*, Public Choice 161: 1-9.

B. E. RASCH (1995), Parliamentary Voting Procedures, ch.15 in H. Döring (ed.) *Parliaments and Majority Rule in Western Europe*, University of Mannheim.

M. A. SATTERTHWAITE (1975), *Strategy proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions*, J. Econ. Theory, 10, pp. 187-217.

M. SCHULZE (2002), On Dummett's 'Quota Borda System', Voting Matters, Issue 15, 10-13.

W. D. SMITH (2010 – 2016), *3-candidate instant runoff voting: Master list of paradoxes and their probabilities.* RangeVoting.org.

E. STENSHOLT (1996), *Circle pictograms for vote vectors*, SIAM Review 38, pp. 96-119.

G. G. SZPIRO (2010), *Number rule. The vexing mathematics of democracy from Plato to the present*, Princeton University Press.

N. TIDEMAN (1987), *Independence of clones as a criterion for voting rules*, Social choice and welfare 4, pp. 185-206.

N. TIDEMAN (1995), *The Single Transferable Vote,* Journal of Economic Perspective 9, pp. 27-38.

R. WILSON (1972), *Social choice theory without the Pareto principle*, Journal of Economic Theory pp. 478-486.

H. P. YOUNG (1975), *Social Choice Scoring Functions*, SIAM J. Appl. Math. 28 no. 4, pp. 824-838.

NHH

NORGES HANDELSHØYSKOLE
Norwegian School of Economics

Helleveien 30
NO-5045 Bergen
Norway

T +47 55 95 90 00
E nhh.postmottak@nhh.no
W www.nhh.no