# THE NEGATIVE PARTICIPATION PARADOX IN THREE-CANDIDATE INSTANT RUNOFF ELECTIONS

DAVID MCCUNE AND JENNIFER WILSON

Abstract. In this paper, we provide theoretical and empirical estimates for the likelihood of a negative participation paradox under instant runoff voting in three-candidate elections. We determine the probability of the paradox and related conditional probabilities based on the impartial anonymous culture and impartial culture models for both complete and partial ballots. We compare these results to the empirical likelihood of a negative participation paradox occurring under instant runoff voting in a large database of voter profiles which have been reduced to three candidates. Lastly, we analyze the relative likelihood of this paradox in comparison to other well-known paradoxes.

## 1. Introduction

The voting method of instant runoff voting (IRV) is famously susceptible to many classical paradoxes. For example, under IRV an election can demonstrate an upward or downward monotonicity paradox [4, 16, 18, 22], or a no-show paradox [3, 4, 20]. IRV is also susceptible to paradoxes of negative participation, a lesser-studied type of paradox. The purpose of this article is to examine the likelihood that a three-candidate IRV election demonstrates this paradox.

An election is said to demonstrate a negative participation paradox under a given voting method if the addition of voters who rank candidate $A$ last causes $A$ to go from losing to winning. This paradox is closely related to the notion of a positive participation paradox [11], also referred to in the language of "positive involvement" [25], where the addition of voters who rank candidate $A$ uniquely first causes $A$ to go from winning to losing. Such outcomes are sometimes referred to as strong no-show paradoxes [2, 24]. To the best of our knowledge, positive participation paradoxes have received more attention in prior literature than negative participation paradoxes, even though IRV is not susceptible to the positive type. Prior research on positive participation paradoxes focuses on other voting methods such as Condorcet methods [2, 8, 9, 24] (this literature tends to use the term "strong no-show paradox" instead of positive participation paradox). Much of the prior work concerning negative participation paradoxes also examines the susceptibility of Condorcet methods to this paradox [2, 24], while some focuses on scoring runoff rules like IRV [11, 13]. Much previous work on these kinds of voting paradoxes focuses on elections with a small number of candidates, often analyzing elections with only three candidates. Our work similarly focuses on the three-candidate case.

This paper makes one small and two substantive contributions regarding the likelihood that an election demonstrates a negative participation paradox under IRV in a three-candidate election. First, for our small contribution we provide new theoretical results under the impartial anonymous culture (IAC) model of voter behavior. Most of the interesting work in this vein has been done in [11, 13], and our results in this area fill some small gaps left from those papers. Second, we provide results under the impartial culture (IC) model of voter behavior. We are able to obtain these new results using geometric techniques developed in [10] and [26]. Third, we provide empirical results using a large database of IRV elections from American and Scottish political elections, as well as elections from the American Psychological Association. Ours is the first study which provides such empirical results; the only other previous work which searched for negative participation paradoxes in real-world ballot data is [15], which focuses on multiwinner single transferable vote elections from Scotland. Our general finding is that negative participation paradoxes occur at a much higher rate than other paradoxes.

Because we are interested in providing both theoretical and empirical analyses of the negative participation paradox we consider situations when all voters provide complete rankings and when voters are allowed to provide only partial ballots. This allows for a better comparison between theoretical and empirical results since partial ballots are extremely common in real-world election data. On the theoretical side, we provide estimates that the likelihood of a negative participation paradox occurs under IRV for 3 candidates under under both the IAC model and IC models. For the IAC model, we analyze the complete and partial ballot cases using the software package Normaliz and Monte Carlo simulations, respectively. For the IC model we use geometric techniques for both full and partial ballots.

The widespread use of the method of single transferable vote in Scotland and other locations, as well as the recent rise of IRV in the US, has made available a large collection of ranked data from real elections. When an election contained more than three candidates, we reduce each election to three candidates by running a number of initial rounds of IRV. Then we determine the number of elections susceptible to a negative participation paradox using the existing ballot data as well as data which we make more "complete" using numerical techniques which we describe in Section 4.

The paper is organized as follows. Section 2 provides relevant definitions as well as a motivating example. In Section 3 we provide theoretical results under the IAC and IC models of voter behavior. As mentioned above, our primary contribution is for the IC model, as the IAC model has been mostly analyzed previously. In Section 4 we give results using a large database of real-world elections. Section 5 provides a discussion and comparison of the likelihood of negative participation paradoxes to other well-known paradoxes. We conclude in Section 6.

## 2. Preliminaries

Instant-runoff voting (IRV), often colloquially referred to as "ranked-choice voting", has become increasingly popular in the United States as an alternative to the method of plurality when there are more than 2 candidates. In such elections voters cast preference ballots with a (possibly partial) linear ranking of the candidates. After an election, the ballots are aggregated into a preference profile, which shows the number of each type of ballot cast. For example, Table 1 shows a preference profile involving the three candidates $J$, $S$, and $W$. The number 1430 denotes that 1430 voters ranked $J$ first, $S$ second, and $W$ third. We use $\succ$ to denote when a voter ranks one candidate over another, so that 1430 voters cast the ballot $J \succ S \succ W$. An IRV election proceeds in rounds. In each round, the number of first-place votes for each candidate is calculated. If a candidate has a majority of the first-place

| *Num. Voters* | *1149* | *1430* | *1012* | *830* | *1989* | *821* | *149* | *303* | *216* |
|---|---|---|---|---|---|---|---|---|---|
| *1st choice* | $J$ | $J$ | $J$ | $S$ | $S$ | $S$ | $W$ | $W$ | $W$ |
| *2nd choice* | | $S$ | $W$ | | $J$ | $W$ | | $J$ | $S$ |
| *3rd choice* | | $W$ | $S$ | | $W$ | $J$ | | $S$ | $J$ |

Table 1. Vote totals for the top 3 candidates in the 2023 election for Minneapolis City Council Ward 8

votes, they are declared the winner. If no candidate receives a majority of votes, the candidate with the fewest first-place votes is eliminated and the names of the remaining candidates on the affected ballots are moved up. The process repeats until a candidate is declared the winner[1]. In the case when partial ballots are allowed, any ballot in which all candidates have been eliminated is considered "exhausted" and plays no further role in the determination of the winner.

When there are only three candidates, at most 2 rounds are required to determine a winner; in general, several rounds may be required. In our analysis of the real-world database in Section 4, we run a sufficient number of rounds for each set of voter preferences until the number of candidates is reduced to three and consider the resulting three-candidate election. We demonstrate IRV in such a case in the following example.

**Example 1.** *The 2003 election for Minneapolis City Council in Ward 8 involved four not write-in candidates: Andrea Jenkins, Soren Stevenson, Bob Sullentrop, and Terry White. Sullentrop earned the fewest first-place votes and is eliminated first, creating the three-candidate preference profile shown in Table 1. The resulting first-place vote totals for Jenkins, Stevenson, and White are 3591, 3640, and 668, respectively. White is eliminated in the next round, causing a transfer of 303 votes to Jenkins and 216 to Stevenson. As a result, Jenkins defeats Stevenson 3894 votes to 3856.*

It is well-known that IRV can produce strange outcomes when ballots are added or removed from an election [3, 4, 5, 6, 11, 13, 15, 16], or when candidates are added or removed from an election [7, 17, 19, 28]. The example below illustrates a type of paradoxical outcome which can occur when ballots are added to an election.

**Example 2.** *Suppose we add 3000 ballots of the form $W \succ J \succ S$ to the ballots from in Table 1. What effect should this have on the electoral outcome? Since these voters all prefer $W$, it would make sense for White to win after receiving this large boost of support. Since these voters prefer Jenkins to Stevenson, it would also make sense for these ballots to bolster Jenkins' victory. However, the addition of these ballots causes Jenkins to be eliminated before White and Stevenson edges out White in the final round.*

Thus, adding thousands of ballots in which Stevenson is ranked last can result in a good outcome for him. Note that this outcome is also true in the original four-candidate election including Sullentrop ($Su$): the same result is obtained if we add 3000 ballots of the form $W \succ J \succ Su \succ S$ or 3000 ballots of the form $W \succ Su \succ J \succ S$. That is, when we add ballots to create a negative participation paradox in

[1]The issue of ties does not arise in our work in this paper, and thus we ignore the possibility of multiple winners.

an IRV election which has been reduced to three candidates, only the first and last rankings matter. We can fill in the intermediate rankings in an arbitrary fashion; the original election with all candidates also demonstrates a negative participation paradox as long as sufficient votes are added so that the plurality loser remains until eliminated in the final round.

This example motivates the primary definition of the paper.

**Definition 1.** *An election with preference profile $P$ is said to demonstrate a* ***negative participation paradox*** *if there exists a losing candidate $A$ and a set of identical ballots $\mathcal{B}$ with $A$ ranked last such that if we add the ballots from $\mathcal{B}$ to the ballots from $P$, candidate $A$ wins the resulting election.*

That is, an election demonstrates this paradox when a losing candidate can be made into a winner by adding ballots on which the loser is ranked last. We use the language of "negative participation paradox" following [11], but there is no standard vocabulary around the topic. For example, much of the literature concerning these paradoxes and Condorcet methods [8, 9, 24] uses the the language of no-show paradoxes and negative and positive involvement. The previous example shows that IRV is susceptible to negative participation paradoxes; we note that other voting methods such as positional scoring rules cannot exhibit this type of paradox.

In the US, the most common voting method for political elections is the method of plurality, in which the candidate with the most first-place votes wins the election. Throughout the article, we refer to the candidate who receives the most first-place votes in a three-candidate election as the *plurality winner*. For example, Stevenson is the plurality winner of the preference profile in Table 1. The *plurality loser* is the candidate with the fewest first-place votes. In the Minneapolis Ward 8 election Sullentrop was the actual plurality loser; because we care only about the three-candidate case, we ignore Sullentrop and say that White is the plurality loser of the election in Table 1.

In a three-candidate election, a negative participation paradox can only occur when several conditions are satisfied. First, the added votes cannot result in the plurality loser being made into the IRV winner: if the added votes rank the plurality loser last, the candidate will remain the plurality loser and be eliminated in the first round. Second, the IRV winner must be different from the plurality winner. To see this, suppose that the IRV winner is $A$ and the candidates are ranked by first-place votes $A > B > C$. Since $A$ is the IRV winner, $B$ must be the losing candidate who can be made into a winner by the addition of votes ranking $B$ last. But $B$ cannot be the winner in such a fashion: if $B$ becomes the new plurality loser then they are eliminated in the first round, and if $C$ remains the plurality loser than $C$ is eliminated in the first round, leaving $B$ to be eliminated in the second round as they were in the original election (the added votes increase the margin of $A$ over $B$). Hence the IRV and plurality winners must be distinct, implying it is the plurality winner who can be made to win under IRV by the addition of votes ranking them last. Similar reasoning implies that the new votes must rank the plurality loser first so that the original IRV winner is eliminated in the first round of the new election.

We use this logic to determine the probability of a negative participation paradox when there are three candidates in the next section.

## 3. The Impartial Anonymous and Impartial Culture Models

In this section we provide probabilities that an election demonstrates a negative participation paradox under the impartial anonymous culture (IAC) and impartial culture (IC) models of voter behavior. Suppose the three candidates are labeled $A$, $B$ and $C$, and there are $V$ voters. If voters must provide complete rankings, there are six possible candidate rankings; if voters can provide partial rankings, this number increases to 9. Assume the preference profiles for complete and partial rankings are as shown in Tables 2 and 3 where in each case the numbers $a_i, b_i$, and $c_i$ indicate the number of voters with the corresponding ranking and that these numbers sum to $V$. Note in Table 3 we equate the rankings $A \succ B$ and $A \succ B \succ C$ and hence do not list them separately. For either model of voter behavior, we are concerned with the *limiting probability* that an election demonstrates a negative participation paradox, where $V \to \infty$.

Before providing the limiting probabilities, we examine the conditions (expressed as inequalities) under which an election demonstrates a negative participation paradox.

Table 2. Preference profile with complete preferences.

| $a_1$ | $a_2$ | $b_1$ | $b_2$ | $c_1$ | $c_2$ |
|---|---|---|---|---|---|
| A | A | B | B | C | C |
| B | C | A | C | A | B |
| C | B | C | A | B | A |

Table 3. Preference profile with some partial preferences.

| $a_0$ | $a_1$ | $a_2$ | $b_0$ | $b_1$ | $b_2$ | $c_0$ | $c_1$ | $c_2$ |
|---|---|---|---|---|---|---|---|---|
| A | A | A | B | B | B | C | C | C |
| | B | C | | A | C | | A | B |
| | C | B | | C | A | | B | A |

In the complete ballot case, we can assume WLOG that

$$a_1 + a_2 > b_1 + b_2 > c_1 + c_2; \tag{1}$$

similarly, in the partial ballot case we assume WLOG that

$$a_0 + a_1 + a_2 > b_0 + b_1 + b_2 > c_0 + c_1 + c_2. \tag{2}$$

That is, we assume $A$ is the plurality winner and $C$ is the plurality loser. As described in Section 2, a negative participation paradox can only occur when the plurality ($A$) and IRV ($B$) winners are distinct and when additional votes are added with the ranking $C \succ B \succ A$. Hence, initially

$$a_1 + a_2 + c_1 < b_1 + b_2 + c_2. \tag{3}$$

If $z$ votes are added, then we require $B$ to be eliminated in the first round followed by $C$ in the second round. Hence,

$$b_1 + b_2 < c_1 + c_2 + z, \quad a_1 + a_2 + b_1 > b_2 + c_1 + c_2 + z$$

Solving for $z$ in each of these yields for $z$ yields

$$b_1 + b_2 - (c_1 + c_2) < z < a_1 + a_2 + b_1 - (b_2 + c_1 + c_2)$$

which implies

$$b_1 + b_2 - (c_1 + c_2) < a_1 + a_2 + b_1 - (b_2 + c_1 + c_2) \quad \text{or} \quad a_1 + a_2 > 2b_2. \tag{4}$$

It is easily seen that (3) and (4) are also sufficient to ensure a negative participation paradox can occur given that $A$ is the plurality winner and $C$ is the plurality loser.

In the partial rankings case, (3) and (4) are replaced by

$$b_0 + b_1 + b_2 + c_2 > a_0 + a_1 + a_2 + c_1 \tag{5}$$

and we obtain

$$a_0 + a_1 + a_2 > b_0 + 2b_2. \tag{6}$$

Before defining and analyzing the two models, we note that they are widely used in the social choice literature to provide *a priori* estimates for the probabilities of many phenomena in social choice, and they tend to provide theoretical upper bounds for such probabilities. For example, in three-candidate elections the probability that an election fails to contain Condorcet winner is 6.25% under the IAC model (this is easily calculated using Normaliz); the probability falls essentially to 0 in real-world elections [6, 21, 27]. Thus, the models are useful for providing theoretical "worst-case" estimates for various election probabilities.

3.1. **Impartial Anonymous Culture.** Under the impartial anonymous culture (IAC) model, the vector of vote totals at the top of the preference profile is assumed to be chosen at random. For example, in the complete ballot case we choose at random a vector of the form $(a_1, a_2, b_1, b_2, c_1, c_2)$ subject to the constraints that each entry is non-negative and the six vote totals sum to the number of voters. Since we examine only limiting probabilities, we choose a vector of proportions of length six in the complete ballot case or a vector of proportions of length nine in the partial ballot case. Limiting probabilities involving IAC can be determined using the theory of Ehrhart polynomials, which has been implemented in the software package Normaliz [1]. The probabilities presented in Proposition 1 were calculated using Normaliz. The probability of a negative participation paradox using IRV assuming complete rankings appears in [11]. In the following proposition, we include this result along with related conditional probabilities for reference.

Assuming, as in the previous section, that first-place vote ranking is $A > B > C$, the probability that an election demonstrates a negative participation paradox is equal to the probability that (3) and (4) are satisfied. The results in (ii) and (iii) below are conditional probabilities. If an election demonstrates a negative participation paradox then there cannot be a majority candidate (since the plurality and IRV winners would coincide). Thus, in (ii), the probability that an election demonstrates a negative participation paradox winner assuming the election does not contain a majority candidate is equal to the the probability of a negative participation paradox divided by the probability that an election does not have a majority candidate. Likewise, in (iii), the probability that an election demonstrates a negative participation paradox assuming the IRV winner is not the plurality winner is equal to the probability that an election demonstrates a negative participation paradox divided by the probability the IRV winner is not the plurality winner.

**Proposition 1.** *In a 3-candidate election in which all voters provide complete preferences, under IAC* *i* *the limiting probability that an election demonstrates a negative participation paradox is* $7/96 = 7.29\%$ [11].

*ii* *the limiting probability that an election demonstrates a negative participation paradox is* $7/54 = 12.96\%$, *assuming the election does not contain a majority candidate.*

*iii* *the limiting probability that an election demonstrates a negative participation paradox is* $42/71 = 59.15\%$, *assuming the IRV winner is not the plurality winner.*

We are unaware of analytical methods which extend the Ehrhart polynomial techniques employed by Normaliz to the partial ballot setting where a preference profile has nine parameters. Thus, to estimate the corresponding IAC probabilities under partial ballots we use Monte Carlo simulation. For each run, we choose a percentage distribution of length nine at random (straightforward techniques exist for this task, see [12] for example) and check for a paradox. Using 100,000 runs, in the partial ballot setting we estimate the corresponding limiting probabilities from Proposition 1 to be approximately 7.24%, 16.74%, and 62.49%.

3.2. **Impartial Culture.** Under the impartial Culture (IC) model, each voter selects uniformly and independently among all possible rankings of the candidates. We can approximate the likelihood of a negative participation paradox for a large number of voters by calculating the limiting probability as $V \to \infty$ using the geometric methods introduced in [26] and subsequently elaborated on in [10]. Note that the probability that there is a majority candidate is zero in the limiting case, thus we provide counterpoints to only two of the probabilities identified in Proposition 1.

**Proposition 2.** *In a 3-candidate election in which all voters provide complete preferences, under IC*

*i* *the limiting probability that an election demonstrates a negative participation paradox is approximately* $0.151084 \approx 15.11\%$..

*ii* *the limiting probability that an election demonstrates a negative participation paradox is* $0.617525 \approx 61.75\%$, *assuming the IRV winner is not the plurality winner.*

**Proposition 3.** *In a 3-candidate election in which voters provide partial preferences, under IC*

*i* *the limiting probability that an election demonstrates a negative participation paradox is approximately* $0.143861 \approx 14.39\%$.

*ii* *the limiting probability that an election demonstrates a negative participation paradox is approximately* $0.672803 \approx 67.28\%$, *assuming the IRV winner is not the plurality winner.*

The proof of Proposition 3 is similar to that of Proposition 2. It can be found in the Appendix. We prove Proposition 2 below.

*Proof.* **Proof of (i)**

**Step 1** First, we find the probability that $A$ and $B$ are the plurality and IRV winners respectively. Following the arguments of [26], the probability that (1) and (3) are satisfied is equal to the area of the spherical simplex $S$ defined by these inequalities on the surface of the unit sphere in $\mathbb{R}^3$, divided by the area of this sphere. Let

$$\mathbf{v_1} = (1, 1, -1, -1, 0, 0)$$
$$\mathbf{v_2} = (0, 0, 1, 1, -1, -1)$$
$$\mathbf{v_3} = (-1, -1, 1, 1, -1, 1).$$

be the normal vectors of the three hyperplanes bounding $S$. By the Gauss-Bonet Theorem, $Vol_2(S) = \alpha_{12} + \alpha_{13} + \alpha_{23} - \pi$ where $\alpha_{ij}$ is the angle between vectors $\mathbf{v}_i$ and $\mathbf{v}_j$ and $Vol_n$ is the $n-$dimensional volume.

Using the law of cosines,

$$\alpha_{12} = \angle(\mathbf{v}_1, \mathbf{v}_2) = \cos^{-1}(\frac{-\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\|\|\mathbf{v}_2\|}) = \cos^{-1}(\frac{2}{2 \cdot 2}) = \cos^{-1}(\frac{1}{2})$$
$$\alpha_{13} = \angle(\mathbf{v}_1, \mathbf{v}_3) = \cos^{-1}(\frac{-\mathbf{v}_1 \cdot \mathbf{v}_3}{\|\mathbf{v}_1\|\|\mathbf{v}_3\|}) = \cos^{-1}(\frac{4}{2 \cdot \sqrt{6}}) = \cos^{-1}(\frac{1}{\sqrt{6}})$$
$$\alpha_{23} = \angle(\mathbf{v}_2, \mathbf{v}_3) = \cos^{-1}(\frac{-\mathbf{v}_2 \cdot \mathbf{v}_3}{\|\mathbf{v}_2\|\|\mathbf{v}_3\|}) = \cos^{-1}(\frac{-2}{2 \cdot \sqrt{6}}) = \cos^{-1}(\frac{-1}{\sqrt{6}}).$$

Thus the probability is equal to

$$\frac{1}{4\pi} Vol_2(S) = \frac{1}{4\pi}[\alpha_{12} + \alpha_{12} + \alpha_{12} - \pi] = 0.04077671.$$

**Step 2**: Next, we find the probability that additionally (4) is satisfied. Following [10], we introduce the modified inequality with parameter $t \in [0, 1]$,

$$x_1 + x_2 > x_3 + x_4 + t(x_4 - x_3). \tag{7}$$

Inequality (7) replaces (4). If $t = 0$, (7) reduces to the first inequality in (1); if $t = 1$, we recover (4). Similarly to Step 1, the probability that these conditions are met for a large number of voters $N$ is equal to the volume of the spherical simplex $S$ defined by these inequalities on the surface of the unit sphere in $\mathbb{R}^4$, divided by the volume of the surface of this sphere, $2\pi^2$. To find the $Vol_3(S)$, we use the Schäfli formula

$$dVol_n = \frac{1}{n-1} \sum_{1 \le j < k \le n} Vol_{n-2}(S_j \cap S_k) d\alpha_{jk}.$$

where $n = 3$ and $S_i$ are the hyperplanes bounding $S$. Let $S_i$ be the hyperplane with normal vector $\mathbf{v}_i$. Since (7) has normal vector $\mathbf{v}_4 = (1, 1, t-1, -t-1, 0, 0)$ and $\|\mathbf{v}_4\| = \sqrt{4 + 2t^2}$, we have,

$$\alpha_{14} = \angle(\mathbf{v}_1, \mathbf{v}_4) = \cos^{-1}(\frac{-4}{2 \cdot \sqrt{4 + 2t^2}}) = \cos^{-1}(\frac{-2}{\sqrt{4 + 2t^2}})$$
$$\alpha_{24} = \angle(\mathbf{v}_2, \mathbf{v}_4) = \cos^{-1}(\frac{2}{2 \cdot \sqrt{4 + 2t^2}}) = \cos^{-1}(\frac{1}{\sqrt{4 + 2t^2}})$$
$$\alpha_{34} = \angle(\mathbf{v}_3, \mathbf{v}_4) = \cos^{-1}(\frac{4}{\sqrt{6} \cdot \sqrt{4 + 2t^2}}).$$

Differentiating, gives

$$d\alpha_{14} = \frac{-\sqrt{2}}{2 + t^2} \quad d\alpha_{24} = \frac{t}{\sqrt{3 + 2t^2} \cdot (2 + t^2)} \quad \text{and} \quad d\alpha_{34} = \frac{2t}{\sqrt{2 + 3t^2} \cdot (2 + t^2)}.$$

To calculate $Vol_1(S_j \cap S_k)$, we must determine where the hyperplanes intersect. Let $P_{ijk}$ be the vertex lying at the intersection of $S_i$, $S_j$ and $S_k$. We identify a basis for the subspace orthogonal to that spanned by $\mathbf{v}_1 \ldots, \mathbf{v}_4$. One such basis is $\mathbf{v}_5 = (1,1,0,0,0,0)$ and $\mathbf{v}_6 = (1,1,1,1,1,1)$.

Next, we find points $P_{ijk}$ lying at the intersection of the hyperplanes with normals given by the vectors $\mathbf{v}_i$, $\mathbf{v}_j$ and $\mathbf{v}_k$. Vertex $P_{124}$, for instance, can then be found by solving the conditions

$$\begin{aligned} n_1 + n_2 - n_3 - n_4 &= 0 \\ n_3 + n_4 - n_5 - n_6 &= 0 \\ -n_1 - n_2 + n_3 + n_4 - n_5 + n_6 &> 0 \\ n_1 + n_2 + (t-1)n_3 + (-t-1)n_4 &= 0 \\ n_1 - n_2 &= 0 \\ n_1 + n_2 + n_3 + n_4 + n_5 + n_6 &= 0. \end{aligned}$$

This yields $P_{124} = (0,0,0,0,-1,1)$. Similarly, we have $P_{134} = (1,1,1,1,-2,-2)$ and $P_{234} = (1,1,\frac{-3-t}{2t},\frac{3-t}{2t},-2,1)$.

Since $\|P_{124}\| = \sqrt{2}$, $\|P_{134}\| = 2\sqrt{3}$ and $\|P_{234}\| = \frac{1}{t}\sqrt{\frac{3}{2}}\sqrt{5t^2+3}$, we have

$$\begin{aligned} Vol(S_1 \cap C_4) &= \angle(P_{124}, P_{134}) = \cos^{-1}\left(\frac{0}{\sqrt{2}\cdot 2\sqrt{3}}\right) = \frac{\pi}{2} \\ Vol(S_2 \cap C_4) &= \angle(P_{124}, P_{234}) = \cos^{-1}\left(\frac{\sqrt{3}t}{\sqrt{5t^2+3}}\right) \\ Vol(S_3 \cap C_4) &= \angle(P_{134}, P_{234}) = \cos^{-1}\left(\frac{\sqrt{2}t}{\sqrt{5t^2+3}}\right) \end{aligned}$$

By the Schäfli formula, $dVol_3 = \frac{1}{2}\sum_{1\leq j<k\leq 4} Vol_1(S_j \cap S_k)d\alpha_{jk}$, so

$$\begin{aligned} Vol_3(S_{t=1}) &= Vol_3(S_{t=0}) + \int_0^1 dVol_3S \\ &= Vol_3(S_{t=0}) + \frac{1}{2}\int_0^1 Vol_1(S_1 \cap S_4)d\alpha_{14} + Vol_1(S_2 \cap S_4)d\alpha_{24} + Vol_1(S_3 \cap S_4)d\alpha_{34}dt \\ &= Vol_3(S_{t=0}) + \frac{1}{2}\left[I_1 + I_2 + I_3\right] \end{aligned}$$

where $I_1 = \int_0^1 \frac{\pi}{2}\cdot\frac{-2}{4+2t^2}dt = -0.99679$ and similarly $I_2 = 0.11231$ and $I_3 = 0.268772$

By Step 1, $Vol_3(S_{t=0}) = Vol_2(S_{t=0})\frac{2\pi^2}{4\pi} = 0.04077671(2\pi^2)$. Hence the probability that all these conditions are met is equal to

$$\frac{1}{2\pi^2}\left[0.04077671(2\pi^2) + \frac{1}{2}(-0.99679 + 0.11231 + 0.268772)\right] = 0.02518064.$$

Since the choice of $A$ and $B$ as plurality and IRV winners was arbitrary, the probability of a negative participation paradox is $6 \times 0.02518064 \approx 0.151084$.

**Proof of (ii)** Since a negative participation paradox can only occur when the IRV and plurality winners are distinct, the probability that an election demonstrates a negative participation paradox given that these winners are different is equal to the probability of negative participation paradox divided by the probability that the IRV and plurality winners are distinct. The latter was determined in Step 1; hence this probability is equal to $0.02518386/0.04077671 \approx 0.617525$. □

3.3. **Analysis of IAC and IC Results.** Overall, a negative participation paradox is twice as likely under the IC than the IAC model, and somewhat more likely assuming the IRV and plurality winners are distinct. This is to be expected, as the IAC model frequently produces preference profiles containing a majority candidate while this is impossible under IC in the limiting case. Under both models, the probability of a negative participation paradox is slightly higher for complete voter preferences than for partial voter preferences; the relationship is reverse when conditioned on the elections having distinct IRV and plurality winners. This is possibly because the probability that these winners are distinct is somewhat lower for partial preferences than for complete preferences (in the IC case, 21% versus 24%), With partial ballots, it is less likely that a candidate ranking second in first-place votes will gain sufficient additional votes, after the plurality loser is eliminated, to overcome the vote total of the plurality winner. We compare these results with an empirical analysis in the next section.

## 4. Empirical Results

In this section we present empirical results using a large database of real-world IRV elections. Our data comes from three sources. First, we use single-winner political elections from the United States with at least three (not write-in) candidates. This data is available at the FairVote data repository [23], and contains ballot data from municipal elections in cities such as San Francisco, CA and Minneapolis, MN. The repository also contains ballot data from federal elections in Alaska and Maine. Second, we use single-winner elections from the American Psychological Association, which generally is willing to share ballot data for ranked-choice presidential elections and elections for the Board of Directors. Some of this data is available on preflib.org [14], and some was shared directly with the first author. Third, we use single-winner political elections from Scotland from a database of Scottish local government elections. The ballot data for these elections is available at `https://github.com/mggg/scot-elex`.

In total, we have access to the preference profiles for 361 single-winner IRV elections with at least three candidates. Many of these elections contain more than three candidates; because we focus on the three-candidate case, for each election we run the IRV algorithm until three candidates remain and use the resulting profile. As with Example 1, any paradox we find in the resulting three-candidate profile will also be demonstrated in the original election. The reason is that when we create the ballots which produce the paradox only the first and last rankings matter; the intermediate rankings can be filled in arbitrarily using the remaining candidates from the original candidate set. Of course, it is possible that an election with four or more candidates demonstrates a paradox which cannot be found when we reduce down to three candidates; our methodology would not find such elections.

In Section 3 we analyze the likelihood of negative participation paradoxes in two cases: when all voters provide complete preferences and when some voters cast partial ballots. To provide comparisons to those results, with our real data we also give results using fully complete preferences and using partial ballots. In real elections voters often provide partial rankings of the candidates, and thus our results from partial ballots represent the "actual" results, the results obtained from using the actual ballot data. To create complete preferences, we proportionally fill in partial ballots using the available complete ballots to create hypothetical results for the case that all voters submit complete preferences.

To see how we fill in the ballots, consider the preferences profile from Table 1. We need to take the bullet votes for $J$ and extend them to complete ballots over the three candidates. Since $1430 + 1012 = 2442$ voters rank $J$ first and provide complete rankings, we extend $1430/2442 * 100\%$ of the 1149 bullet votes for $J$ to ballots of the form $J \succ S \succ W$ and $1012/2442 * 100\%$ of the bullet votes for $J$ to ballots of the form $J \succ W \succ S$, rounding to the nearest integer. As a result, 673 of the 1149 bullet votes are extended to $J \succ S \succ W$ and 476 are extended to $J \succ W \succ S$. This methodology is used on every three candidate profile to create complete preferences.

| Jurisdiction | USA | APA | Scotland |
|---|---|---|---|
| Elections | 304 | 27 | 30 |
| No Majority Cand | 191 | 23 | 30 |
| IRV Winner $\neq$ plurality winner | 18, 30 | 3, 3 | 5, 8 |
| Paradoxes | 7, 17 | 3, 3 | 2, 3 |

TABLE 4. Results from real-world elections. In the final two rows, the first number is obtained using the actual ballots and the second is obtained from ballots which have been completed.

Our results are summarized in Table 4. The top row shows the number of elections from each jurisdiction and the second row shows the number of elections without a majority candidate. The third row shows the number of elections in which the IRV winner is different from the plurality winner. The first number is the number of such elections when using the actual data with partial ballots; the second number is the number of such elections when using complete preferences, created as described above. The fourth row shows the number of participation paradoxes, using the same format. For example, there are 191 American political elections without a majority candidate. Of these, the IRV winner is not the plurality winner in 18 (respectively 30) elections when using the actual ballot data (respectively, complete preference data). Seven American elections demonstrate a paradox when using the actual data, which rises to 17 when we complete the ballots.

Table 5 provides a complete list of elections which demonstrate a paradox when using the actual ballots. The table also provides the range of the number of ballots which can be added to produce the paradox. For example, the third row shows that we can add anywhere from 2924 to 3389 ballots of the form $W \succ J \succ S$ to create the paradox in Example 1. Note that the last election in the table just barely produces a paradox, as the paradox can be demonstrated using only 266 ballots exactly.

When using the actual ballot data, in total we find 12 elections which demonstrate this paradox. From Table 4, we obtain an empirical probability of $12/361 = 3.3\%$ that the paradox occurs. When we condition on the non-existence of a majority candidate we obtain a probability of $12/244 = 4.9\%$, which rises to $12/26 = 46.2\%$ when we restrict to elections when the IRV and plurality winners are different. The corresponding three probabilities for complete preferences are 6.4%, 9.4%, and 56.1%, respectively. As with the IAC and IC models we obtain higher probabilities for complete preferences for the two conditional probabilities.

| Election | Num. Voters | Ballot Range |
|---|---|---|
| 2020 San Francisco Board of Supervisors D7 | 38321 | 798-1606 |
| 2021 Minneapolis City Council Ward 2 | 8907 | 80-224 |
| 2023 Minneapolis City Council Ward 8 | 7899 | 2924-3389 |
| 2010 Oakland Mayor | 113217 | 2315-3951 |
| 2016 Oakland School Director D5 | 12950 | 2876-4325 |
| 2022 Oakland School Director D4 | 26432 | 38-598 |
| 2021 New York City Rep Primary D50 | 8182 | 773-830 |
| 2005 APA President | 14079 | 2275-2780 |
| 2007 APA President | 12925 | 2397-2718 |
| 2020 APA Board of Directors Race 2 | 6227 | 1093-1883 |
| 2021 Argyll Bute By-Election Isle of Bute Ward$^\dagger$ | 1804 | 26-38 |
| 2021 Highland By-Election Aird Ward$^\dagger$ | 3321 | 266-266 |

Table 5. Each election which demonstrates a negative participation paradox when using the actual ballot data.

Unlike those models, we obtain higher probabilities for complete preferences for the unconditioned probabilities as well. Overall, the frequency with which the paradox of negative participation occurs in the database is lower than that which is predicted by the IAC and IC models, as expected since these models tend to provide upper bounds for the probabilities of paradoxical behavior. It is interesting, though, that some of the empirical probabilities are not significantly lower than the theoretical probabilities. Often, an empirical estimate of the probability that an election demonstrates a particular voting paradox is much lower than what is predicted by a model like IAC or IC. When we assume the IRV and plurality winners are different, the empirical probability for negative participation paradoxes is surprisingly close to the probabilities reported in Section 3.

## 5. Comparison to other paradoxes

In this section we compare the conditions and likelihood that an IRV election demonstrates a negative participation paradox to other more well-known (and well-studied) paradoxes: upward and downward monotonicity paradoxes, and no-show paradoxes (see e.g. [4] and [18] for definitions and discussions of these paradoxes). In the three-candidate complete ballot setting, for an upward (resp. downward) monotonicity paradox to occur, the plurality loser must earn at least 25% (respectively $16\frac{2}{3}$%) of the first-place votes [18]. In either the complete or partial ballot setting, the plurality loser must earn at least $16\frac{2}{3}$% of the first-place votes for a no-show paradox to occur [5]. By contrast, for a large enough electorate there is no such positive lower bound for the plurality loser regarding negative participation paradoxes. That is, it is possible for the plurality lower to receive an arbitrarily small amount of first-place votes and yet the election demonstrates a negative participation paradox.

To see this, consider Table 6. The top row shows the percentage of the first-place vote controlled by each candidate. Suppose these percentages satisfy $p_C < p_B < p_A < 0.5$. Note that $B$ is the IRV winner, but if we add enough $C \succ B \succ A$ ballots so that the adjusted percentages satisfy $p'_B < p'_C < p'_A$ (which is possible for a large enough electorate), then $A$ wins. Thus, it is possible for the plurality loser to be arbitrarily weak and yet observe this paradox. In particular, as long as no candidate earns 50% or more of the first-place votes, it is possible to find a preference profile which demonstrates a negative participation paradox for any distribution of first-place votes.

| % Voters | $p_A$ | $p_B$ | $p_C$ |
|---|---|---|---|
| 1st choice | $A$ | $B$ | $C$ |
| 2nd choice | $B$ | $A$ | $B$ |
| 3rd choice | $C$ | $C$ | $A$ |

Table 6. A table illustrating that an election can demonstrate a negative participation paradox even if the plurality loser controls a small percentage of the first-place votes.

Furthermore, to create such a paradox we can control the type and number of ballots added to the original ballot data. Thus, we expect that such paradoxes are exhibited more frequently in real-world data than other paradoxes. Prior empirical research bears this out. Using essentially the same database of American elections that we use, [6] found four elections which demonstrate an upward monotonicity paradox, three which demonstrate a downward monotonicity paradox, and one which demonstrates a no-show paradox. The authors did not include the APA elections in that study, but they evaluated those elections using the same methodology as in [6] and found no paradoxes of any kind. The study in [16] included the 30 Scottish elections we evaluate as part of a much larger study of single transferable vote elections, and of these 30 elections found only one which demonstrated a paradox (this election demonstrated both an upward monotonicity paradox and a no-show paradox). In total, of the 361 elections in our database there are five documented elections which demonstrate an upward monotonicity paradox, three which demonstrate a downward monotonicity paradox, and two which demonstrate a no-show paradox. As in this paper, all of the previously documented paradoxes occurred at the "three candidate level" in the sense that the paradox can be demonstrated after eliminating all but the final three candidates.

Based on prior empirical work, IRV three-candidate elections demonstrate a negative participation paradox 2-6 times more frequently than other classical paradoxes, an unsurprising result given the above discussion.

## 6. Conclusion

We conclude with two final comments. First, IRV proponents sometimes claim that a strength of IRV is that the method can choose a different (and presumably more deserving) winner than the plurality winner. This is a natural point to make, since if the two winners always coincide then there is no reason to use IRV, and arguably the plurality winner sometimes does not deserve to win an election (e.g. if the plurality winner is a Condorcet loser). However, in some sense our work shows that this feature of IRV comes with a price: when the IRV and plurality winners differ, it is often the case that the election demonstrates a negative participation paradox. That is, when IRV "works properly" by choosing a more deserving candidate than the plurality winner, often the plurality winner could have been the IRV winner if only the turnout of some voters who don't like the plurality winner were increased.

Second, we make no claim regarding the normative status of negative participation paradoxes. Social choice theorists care about paradoxes such as monotonicity and participation paradoxes in part because such paradoxes show there is some mathematical irrationality built into the mechanics of IRV. Put simply, sometimes IRV does not behave rationally in response to changes in the ballot data, and this seems normatively undesirable. On the other hand, paradoxes in this vein are hypothetical–negative participation paradoxes especially so. In Example 2, it is extremely unlikely that there were thousands of potential voters in Minneapolis' eighth ward whose favorite candidate was White and whose least favorite was Stevenson, and these voters abstained from the actual election. Example 2 is built using thousands of people who probably do not exist, which perhaps ultimately says nothing about whether IRV is a "good" or "appropriate" voting method.

Regardless of one's normative stance on these paradoxes, the results in this paper contribute to the literature on voting paradoxes by providing limiting probabilities regarding negative participation paradoxes in the three-candidate case under the impartial culture model, empirical probabilities using a large database of real-world elections, and limiting probabilities in the three-candidate case for some cases left over from [11] under the impartial anonymous culture model.

## References


[1] Bruns, W., Ichim, B. Söger, C. & von der Ohe, U. (2022). *Normaliz. Algorithms for rational cones and affine monoids.*, *Normaliz.* `https://www.normaliz.uni-osnabrueck.de`. Accessed Mar. 8, 2024.

[2] C. Duddy. (2014). Condorcet's principle and the strong no-show paradoxes. *Theory and Decision* **77**(2): 275-285.

[3] D. Felsenthal and N. Tideman. (2013). Varieties of failure of monotonicity and participation under five voting methods. *Theory and Decision* **75**: 59-77. `https://doi.org/10.1007/s11238-012-9306-7`.

[4] P. Fishburn and S. Brams. (1983). Paradoxes of preferential voting. *Mathematics Magazine*, **56** (4): 207-214. `https://doi.org/10.2307/2689808`.

[5] A. Graham-Squire. (2024). Conditions for Fairness Anomalies in Instant-Runoff Voting. In M. Jones, D. McCune, J. Wilson (Eds), *Mathematical Analyses of Decisions, Voting and Games*. Contemporary Mathematics, Volume 795. American Mathematical Society.

[6] A. Graham-Squire and D. McCune. (2023). An Examination of Ranked-Choice Voting in the United States, 2004-2022. *Representaion*, `https://doi.org/10.1080/00344893.2023.2221689`.

[7] J. Green-Armytage, T. Tideman, and R. Cosman. (2016). Statistical evaluation of voting rules. *Social Choice Welfare*, 46: 183–212.

[8] W. Holliday. (2024). An impossibility theorem concerning positive participation in voting. *Economics Letters* **236**. `https://doi.org/10.1016/j.econlet.2024.111589`.

[9] W. Holliday and E. Pacuit. (2021). Measuring Violations of Positive participation in Voting. `arXiv:2106.11502`.

[10] E. Kamwa and V. Merlin. (2015). Scoring Rules over Subsets of Alternatives: Consistency and Paradoxes. *J. of Math. Econ.* **61**: 130-138. `https://doi.org/10.1016/j.jmateco.2015.08.008`.

[11] E. Kamwa, V. Merlin, and F.M. Top. (2023). Scoring Run-off Rules, Single-peaked Preferences and Paradoxes of Variable Electorate. Preprint: `https://hal.science/hal-03143741v2`.

[12] D. Lepelley, A. Louichi, and F. Valognes. (2000). Computer simulations of voting systems. *Adv Complex Syst* **3**(1):181–194. `https://doi.org/10.1142/S0219525900000145`.

[13] D. Lepelley and V. Merlin. (2001). Scoring runoff paradoxes for variable electorates. *Economic Theory* **14**(1): 53-80.

[14] N. Mattei and T. Walsh. (2013). Preflib: A library for preferences. In *ADT*, 259–270. Springer.

[15] D. McCune. (2024). Single Transferable Vote and Paradoxes of Negative and Positive participation. To appear in *Mathematics Magazine*.

[16] D. McCune and A. Graham-Squire. (2023). Monotonicity Anomalies in Scottish Local Government Elections. Preprint: `arxiv.org/abs/2305.17741`.
[17] D. McCune and J. Wilson. (2023). Ranked-Choice Voting and the Spoiler Effect. *Public Choice* **196**: 19-50.
[18] N. Miller. (2017). Closeness matters: monotonicity failure in IRV elections with three candidates. *Public Choice*, **173**: 91-108. DOI 10.1007/s11127-017-0465-5.
[19] N. Miller. (2019). Reflections on Arrow's theorem and voting rules. *Public Choice*, 179: 113–124.
[20] H. Moulin. (1988). Condorcet's principle implies the no show paradox. *Journal of Economic Theory*, **45** (1): 53-64. `https://doi.org/10.1016/0022-0531(88)90253-0`.
[21] A. Myers. The Frequency of Condorcet Winners in Real Non-Political Elections, preprint.
[22] J. Ornstein and R. Norman. (2014). Frequency of monotonicity failure under Instant Runoff Voting: estimates based on a spatial model of elections. *Public Choice*, **161** (1/2): 1-9. `https://doi.org/10.1007/s11127-013-0118-2`.
[23] D. Otis. (2022). Single winner ranked choice voting CVRs. Harvard Dataverse, `https://doi.org/10.7910/DVN/AMK8PJ`.
[24] J. Pérez. (2001). The strong no show paradoxes are a common flaw in Condorcet voting correspondences. *Social Choice and Welfare* **18** (3): 601-616. doi:10.1007/s003550000079.
[25] D. Saari. (1995). *Basic Geometry of Voting*. Springer, Berlin.
[26] D.Saari and M. Tataru. (1999). The likelihood of dubious election outcomes. *Economic Theory* **13**: 345-363.
[27] C.G. Song. (2023). What difference does a voting rule make?. *Constitutional Political Economy* **34**: 275-285. `https://doi.org/10.1007/s10602-022-09375-9`.
[28] T. Tideman. (1987). Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, **4** (3): 105–206.

## Appendix

The proof of Proposition 3 is similar to that of Proposition 2.

**Proof of Proposition 3 (i)**

**Step 1** First, we find the probability that $A$ and $B$ are the plurality and IRV winners respectively. The inequalities (2) and (5) have normal vectors

$$\begin{aligned}\mathbf{v}_1 &= (1,1,1,-1,-1,-1,0,0,0)\\ \mathbf{v}_2 &= (0,0,0,1,1,1,-1,-1,-1)\\ \mathbf{v}_3 &= (-1,-1,-1,1,1,1,0,-1,1).\end{aligned}$$

So

$$\begin{aligned}\alpha_{12} &= \angle(\mathbf{v}_1,\mathbf{v}_2) = \cos^{-1}(\frac{-\mathbf{v}_1\cdot\mathbf{v}_2}{\|\mathbf{v}_1\|\|\mathbf{v}_2\|}) = \cos^{-1}(\frac{3}{\sqrt{6}\cdot\sqrt{6}}) = \cos^{-1}(\frac{1}{2})\\ \alpha_{13} &= \angle(\mathbf{v}_1,\mathbf{v}_3) = \cos^{-1}(\frac{-\mathbf{v}_1\cdot\mathbf{v}_3}{\|\mathbf{v}_1\|\|\mathbf{v}_3\|}) = \cos^{-1}(\frac{6}{\sqrt{6}\cdot\sqrt{8}}) = \cos^{-1}(\frac{\sqrt{3}}{2})\\ \alpha_{23} &= \angle(\mathbf{v}_2,\mathbf{v}_3) = \cos^{-1}(\frac{-\mathbf{v}_2\cdot\mathbf{v}_3}{\|\mathbf{v}_2\|\|\mathbf{v}_3\|}) = \cos^{-1}(\frac{-3}{\sqrt{6}\cdot\sqrt{8}}) = \cos^{-1}(\frac{-\sqrt{3}}{4}).\end{aligned}$$

Thus, the probability is equal to

$$\frac{1}{4\pi}Vol_2(S) = \frac{1}{4\pi}[\alpha_{12}+\alpha_{12}+\alpha_{12}-\pi] = 0.03563737.$$

**Step 2**: Next, we find the probability that additionally (6) is satisfied. We introduce the inequality $x_1+x_2+x_3 > x_4+x_5+x_6+t(x_6-x_5)$. When $t=0$ this reduces to a (2); when $t=1$ we get the required extra condition.

Since this inequality has normal vector equal to $\mathbf{v}_4 = (1,1,1,-1,t-1,-t-1,0,0,0)$, we have

$$\begin{aligned}
\alpha_{14} &= \angle(\mathbf{v}_1,\mathbf{v}_4) = \cos^{-1}(\frac{-6}{\sqrt{6}\cdot\sqrt{6+2t^2}}) = \cos^{-1}(\frac{-\sqrt{3}}{\sqrt{3+t^2}})\\
\alpha_{24} &= \angle(\mathbf{v}_2,\mathbf{v}_4) = \cos^{-1}(\frac{3}{\sqrt{6}\cdot\sqrt{6+2t^2}}) = \cos^{-1}(\frac{\sqrt{3}}{2\sqrt{3+t^2}})\\
\alpha_{34} &= \angle(\mathbf{v}_3,\mathbf{v}_4) = \cos^{-1}(\frac{6}{\sqrt{8}\cdot\sqrt{6+2t^2}}) = \cos^{-1}(\frac{3}{2\sqrt{3+t^2}})
\end{aligned}$$

leading to

$$d\alpha_{14} = \frac{-\sqrt{3}}{(3+t^2)} \quad d\alpha_{24} = \frac{\sqrt{3}t}{\sqrt{9+4t^2}(3+t^2)} \quad \text{and} \quad d\alpha_{34} = \frac{3t}{\sqrt{3+4t^2}(3+t^2)}.$$

Next, we require a basis for the space in $\mathbb{R}^9$ that is orthogonal to $\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3$ and $\mathbf{v}_4$. One such basis is

$$\begin{aligned}
\mathbf{v}_5 &= (1,1,-2,0,0,0,0,0,0)\\
\mathbf{v}_6 &= (1,-1,0,0,0,0,0,0,0)\\
\mathbf{v}_7 &= (0,0,0,-2,1,1,0,0,0)\\
\mathbf{v}_8 &= (0,0,0,0,0,0,-2,1,1)\\
\mathbf{v}_9 &= (1,1,1,1,1,1,1,1,1).
\end{aligned}$$

We find $P_{124}$, $P_{134}$ and $P_{234}$ as before to get $P_{124} = (0,0,0,0,0,0,0,-1,1)$, $P_{134} = (1,1,1,1,1,1,-2,-2,-2)$ and $P_{234} = (4t,4t,4t,-2t,-9-2t,9-2t,-2t,-11t,7t)$. Since $\|P_{124}\| = \sqrt{2}$, $\|P_{134}\| = 3\sqrt{2}$ and $\|P_{234}\| = \sqrt{48t^2+8t^2+121t^2+49t^2+162+8t^2} = \sqrt{234t^2+162} = 3\sqrt{2(13t^2+9)}$, we have

$$\begin{aligned}
Vol(S_1\cap C_4) &= \angle(P_{124},P_{134}) = \cos^{-1}(\frac{0}{\sqrt{2}\cdot 2\sqrt{18}}) = \frac{\pi}{2}\\
Vol(S_2\cap C_4) &= \angle(P_{124},P_{234}) = \cos^{-1}(\frac{18t}{\sqrt{2}\cdot 3\sqrt{2(13t^2+9)}}) = \cos^{-1}(\frac{3t}{\sqrt{13t^2+9}})\\
Vol(S_3\cap C_4) &= \angle(P_{134},P_{234}) = \cos^{-1}(\frac{18t}{3\sqrt{2}\cdot 3\sqrt{2(13t^2+9)}}) = \cos^{-1}(\frac{t}{\sqrt{13t^2+9}})
\end{aligned}$$

So

$$\begin{aligned}
Vol_3(S_{t=1}) &= Vol_3(S_{t=0}) + \int_0^1 dVol_3S\\
&= Vol_3(S_{t=0}) + \frac{1}{2}\int_0^1 Vol_1(S_1\cap S_4)d\alpha_{14} + Vol_1(S_2\cap S_4)d\alpha_{24} + Vol_1(S_3\cap S_4)d\alpha_{34}dt\\
&= Vol_3(S_{t=0}) + \frac{1}{2}\left[I_1+I_2+I_3\right]
\end{aligned}$$

where $I_1 = \int_0^1 \frac{\pi}{2}\cdot\frac{-\sqrt{3}}{(3+t^2)}dt = -0.82247$ and similarly $I_2 = 0.0806918$ and $I_3 = 0.28144178$

By Step 1, $Vol_3(S_{t=0}) = Vol_2(S_{t=0})\frac{2\pi^2}{4\pi} = 0.03563737(2\pi^2)$. Hence the probability that all these conditions are met is equal to

$$\frac{1}{2\pi^2}\left[0.03563737(2\pi^2) + \frac{1}{2}(-0.82247 + 0.0806918 + 0.28144178)\right] = 0.023976912.$$

Since the choice of $A$ and $B$ was arbitrary, the probability of a negative participation paradox is is $6 \times 0.023976912 \approx 0.143861$

**Proof of (ii)** As with the case with complete ballots, the probability that an election demonstrates a negative participation paradox given that these winners are different is equal to the probability of negative participation paradox divided by the probability that the IRV and plurality winners are distinct. The latter was determined in Step 1; hence this probability is equal to $0.023976912/0.03563737 \approx 0.672803$.

David McCune, Department of Mathematics and Data Science, William Jewell College, 500 College Hill, Liberty, MO, 64068

*Email address*: mccuned@william.jewell.edu

Jennifer Wilson, The New School, Department of Natural Sciences and Mathematics, 66 West 12th Street New York, NY 10011

*Email address*: wilsonj@newschool.edu