The Probability of Majority Inversion in a Two-stage Voting System with Three States

Serguei Kaniovski *　　Alexander Zaigraev †

December 23, 2017

Abstract

Two-stage voting is prone to majority inversions, a situation in which the outcome of an election is not backed by a majority of popular votes. We study the probability of majority inversion in a model with two candidates, three states and uniformly distributed fractions of supporters for each candidate. The model encompasses equal or distinct population sizes, with equal, population-based or arbitrary voting weights in the second stage. We prove that, when no state can dictate the outcome of the election by commanding a voting weight in excess of one half, the probability of majority inversion increases with the size disparity among the states.

JEL-Codes: D72

Key Words: majority inversion, two-stage voting, population weights, voting weights

*Austrian Institute of Economic Research (WIFO), Arsenal, Object 20, 1030 Vienna, Austria.
E-mail: serguei.kaniovski@wifo.ac.at

†Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Chopin str. 12/18, 87-100 Toruń, Poland.
Email: alzaig@mat.uni.torun.pl

1 Introduction

Two-way elections conducted using two-stage voting procedures are prone to majority inversions, a situation in which the outcome of an election does not represent the will of a majority of voters. We study a model of two-stage voting with three states (constituencies, districts), derive the probability of majority inversion as a function of the population weights, and relate this probability to the inequality in the distribution of population among the states.

Two-stage voting requires the voters to be grouped into states. In the first stage, each voter casts a single vote for one of the two candidates (parties, referendum options). The outcome of the first stage is called the *popular vote*. In the absence of ties, the simple majority rule picks a winner in each state and in the country. The hallmark of a two-stage voting system is that the popular vote does not determine the outcome of the election. It is decided by the *electoral vote* in the second stage. Electoral votes are cast by the electors, each representing one state in an assembly of states. The electors may command several votes. We assume that each elector casts all of his votes as a bloc for the candidate who obtained a majority in the state the elector represents, thus sidestepping the issue of 'faithless electors'. The election is awarded to the candidate with a majority of electoral votes, who may not be the candidate that won the popular vote. Majority inversion takes place if the outcome of the election does not coincide with the popular vote.

The U.S. presidential election is not a two-way election.[1] If we held the U.S. presidential election for what it essentially always was: a contest between the Democrats and Republicans, we would identify four majority inversions in the past. Since more populous states command more electoral votes, the second-stage votes in the Electoral College system are weighted. By contrast, the second-stage votes in legislative elections are not weighted. In a single-member-district majority system, each district elects one member of the parliament in the first stage. The member then casts a single vote on a series of bills over the course of a legislature term, which can be viewed as the second stage of a two-stage voting process. Majority inversion occurs if the party that obtained a majority of seats in parliament is not the party that won a majority of votes.[2] This variety suggests that a comprehensive model of two-stage voting should admit constituencies of different sizes and weighted voting in the second stage. We shall refer to the system with population-weighted second-tier votes as Electoral College (**EC**)[3] and the single-member-district majority system as Westminster (**WM**). The third model, denoted **EP**, is the baseline model of equipopulous states that is commonly studied in the theoretical literature on majority inversions. The forth model, denoted **GM**, is the most general one. It involves states of different population sizes and weighted electoral votes.

May (1948) appears to have taken the first step towards computing the probability of majority inversion in a two-stage model with an odd number of equally-sized states. May workes with a discrete uniform distribution for the number of supporters of a certain candidate in each state, which became a continuous uniform distribution in the limiting case of infinitely many voters. The assumption of a uniform distribution implies that all levels of support for a given candidate,

[1] For example, ten candidates ran for office in each of the three recent elections: 2008, 2012, 2016.

[2] Miller (2012) documents inversions in legislative elections in the British-type or Westminster parliamentary systems. For parliamentary elections contested by more than two parties, one can ensure that the voters face a choice between two alternatives by using the typical left-right ideological dichotomy to define the outcome.

[3] We are aware of the fact that the actual weights used by the U.S. Electoral College do not accurately reflect the relative difference in the population among the states. Nevertheless, weighting electoral votes by the population is the intended design. The current allocation of votes to the states is based on the 2010 Census and applies to the 2012, 2016 and 2020 presidential elections.

whether expressed in absolute terms in persons or in relative terms as a percentage, are equally likely. The recent wave of theoretical results placed Mays work in the context of the stochastic models of voting behavior used in the contemporary voting theory. Feix, Lepelley, Merlin and Rouet (2004) and Lepelley, Merlin and Rouet (2011) extend the calculations to an even number of states and provide numerical simulations when the number of states is large. May assumes that the levels of popular support in any two states are independent random variables. This assumption is relaxed in De Mouzon, Laurent, Le Breton and Lepelley (2017), who perform a comprehensive analysis of the rate of converge of the probability of majority inversion as the number of states increases, under the assumption that any two votes in the electorate correlate.

The baseline model (**EP**) maintains equally-sized states and a 'one person, one vote' principle at both stages of the two-stage voting procedure. The 'one person, one vote' principle is a natural assumption for the first stage, where it embodies the democratic principle of equal suffrage among the voters. It is also consistent with the assumption of equally-sized states, since the number of electoral votes awarded to a state usually depends on its size. The next level of complexity involves states of different population sizes (**WM**), followed by electoral votes weighted either with the population weights (**EC**), or with arbitrary weights (**GM**).

Studies of the weighted model can be found in Lepelley, Merlin, Rouet and Vidu (2014) for the case of three states, and in Kikuchi (2016) as the number of states tends to infinity. Lepelley et al. (2014) showed that weighting the second-tier votes proportionally to the ratio of the square-roots of population sizes, a method famously known to equalize the indirect voting power in two-stage voting systems, does not minimize the probability of majority inversion.[4] Consistent with the literature on the Square-Root Rule, Lepelley et al. (2014) used the Impartial Culture (IC) model. While the probability of majority inversion can be computed under IC, the companion Impartial Anonymous Culture (IAC) model is preferable due to it being extendable to infinite populations. The continuous uniform model used in this paper arises as a limiting case of the IAC model applied at the state level, while maintaining May's independence assumption between the states. De Mouzon et al. (2017) refer to this model as IAC*.

The present paper is motivated by Kikuchi (2016), who provided several far-reaching results for weighted votes. He established that equal weights minimize the probability of majority inversion and that this probability increases monotonically with the dispersion of population weights to an upper bound of one half. Kikuchi's results are asymptotic in the number of states and, therefore, do not apply when the number of states is small. In particular, letting the number of states tend to infinity implies that the weight of each state tends to zero. Indeed, we show that, for three states, increasing the inequality of voting weights increases the probability only if the population weight of the largest state does not exceed one half, and that the distribution of the population among the three states which attains the bounds on this probability depends on the weighting scheme at the second stage. The largest voting weight exceeding one half is equivalent to the largest state being a 'dictator'. Our principal tool is the theory of majorization and Schur-convexity (Marshall, Olkin and Arnold 2011), which are closely related to the theory of inequality measurement (Yitzhaki and Schechtman 2013).

2 The model

Models must make assumptions to stay amenable to analytical treatment. Three such assumptions stand out in the literature. The first is a near-universal equation of turnout and population.

[4]For a discussion of the Penrose Square-Root Rule, see Felsenthal and Machover (1998).

This simplification overlooks the fact that not every resident of a state may be eligible to vote, and that some of those who are eligible may abstain. The equation of turnout and population is admissible when the turnout is treated as exogenous. The second assumption limits the number of states. Confining to a minimum of three states may be the only path to a tractable model, as will be the case in this paper. The third assumption commonly found in the literature imposes the equality of state sizes. We seek to relax this assumption, while equating turnout and population. This allows us to equate the turnout share with the population share, both of which will be referred to as the *population weight*. The second weight of the state is given by the share of the electoral votes it commands. This is the *voting weight*.

The voters face two alternatives, A and B. Let the number of states be n, where n is odd, and $n > 1$. The variations of the standard two-stage voting models studied in this paper rest on three assumptions. For each state $i = 1, \ldots, n$:

A.1 $w_i > 0$ the population weights, such that $\sum_{i=1}^{n} w_i = 1$;

A.2 $v_i > 0$ the voting weights, such that $\sum_{i=1}^{n} v_i = 1$;

A.3 $a_i \sim U(0,1)$ the share of voters in state i who support A, where the uniformly distributed random variables $a_1, \ldots, a_n$ are independent.

Three types of asymptotic analysis can be found in the literature: 1) the number of voters in each state tends to infinity, 2) the number of states tends to infinity, or 3) both quantities tend to infinity, a typology that goes back to the pioneering work by May (1948). The second and third kind of asymptotic analysis should be particularly susceptible to producing distorted results in comparison with the exact calculations. The assumption [A.3] forebodes an asymptotic analysis of the first kind. The analysis presented in this paper indeed assumes a 'continuum' of voters in each state, which is an approximation to what happens in large electorates.

Instead of specifying the behavior of an individual voter, the stochastic element of the model describes the share of supporters of A in each state. The uniform distribution for the share of supporters arises as a limiting case of the beta-binomial model of individual voting behavior. In the beta-binomial model, the probability of a voter supporting the alternative A, denoted p_a, is drawn from a uniform distribution $U(0,1)$, which is also a $Beta(1,1)$ distribution (Section 4.4 in Casella and Berger (2002)). This probability is then assigned as the probability of success to a binomial random variable S_a, denoting the absolute number of supporters of A. Let m be the number of voters. The random variable S_a thus follows a discrete uniform distribution on $[0, 1, \ldots, m-1, m]$ and the share of voters who support A, or S_a/m, follows a discrete uniform distribution on $[0, 1/m, \ldots, (m-1)/m, 1]$, which becomes continuous in the limit as $m \to \infty$. This is how May (1948) handles large electorates.

The beta-binomial model of individual voting behavior has the binomial model of voting as its simpler companion. The binomial model assumes that each vote has an equal probability of being cast for or against an alternative, and that all votes are independent. The two models are well-known to the voting theorist under the names of the Impartial Anonymous Culture (IAC) and Impartial Culture (IC). They are standard in much of the literature on voting power and procedures. The beta-binomial model, or IAC, is consistent with the Shapley-Shubik index of voting power, whereas the binomial model, or IC, underlies the Penrose-Banzhaf measure of power (Straffin 1977). Table 1 summarizes the implications of the two models for the joint probability distribution of the combinations of votes and their sums (number of supporters). With m voters, there will be 2^m combinations of binary votes called *voting profiles*, with $m+1$

sums ranging from 0 to m. The number of profiles containing k supporters is given by the binomial coefficient $C_m^k = m!/[k!(m-k)!]$ for $m, k \in \mathbb{N}$, where $C_m^k = 0$ for $m < k$. It is evident that the uniform model for a share of supporters is the limiting case of the IAC. Under this model of individual voting behavior, any attainable level of support is equally likely. This paper adopts the limiting case of the IAC model at the state level, while maintaining May's independence assumption between the states.

Table 1: Probabilities for IC and IAC models.

of	Binomial (IC)	Beta-binomial (IAC)
one of 2^m voting profiles	$\frac{1}{2^m}$	$\frac{1}{(m+1)C_m^k}$
any of C_m^k voting profiles	$\frac{C_m^k}{2^m}$	$\frac{1}{m+1}$

Coming back to the assumptions A.1 to A.3, if the turnout is equal to the population, then the share of state i in the total population of the county, or w_i, is also the share of i in total turnout. If the share of supporters of A in state i is a uniformly distributed random variable, then the share of supporters of B in state i will also follow a uniform distribution. Although the analysis in this paper will be conducted in terms of shares, all voting outcomes can be expressed in absolute terms. If the total population equals the total turnout m, then the total number of votes cast in state i equals $w_i m$, of which $a_i w_i m$ favored A and the remaining $(1-a_i)w_i m$ favored B. The total number of voters who voted for A thus becomes $m \sum_{i=1}^n a_i w_i$.

A majority inversion occurs when A loses the popular vote but wins the electoral vote, or when A wins the popular vote but loses the electoral vote. The probabilities of these two compound events are, respectively,

$$P\left(\sum_{i=1}^n w_i a_i < \frac{1}{2}\,,\ \sum_{j=1}^n v_j \mathbf{1}_{\{a_j > \frac{1}{2}\}} > \frac{1}{2}\right) \quad \text{and} \quad P\left(\sum_{i=1}^n w_i a_i > \frac{1}{2}\,,\ \sum_{j=1}^n v_j \mathbf{1}_{\{a_j > \frac{1}{2}\}} < \frac{1}{2}\right).$$

Since both compound events are equally probable and mutually exclusive, the probability of majority inversion can be written more compactly as

$$P(w_1, \ldots, w_n, v_1, \ldots, v_n) = 2P\left\{\sum_{i=1}^n w_i a_i < \frac{1}{2}\,,\ \sum_{j=1}^n v_j \mathbf{1}_{\{a_j > \frac{1}{2}\}} > \frac{1}{2}\right\}.$$

The following analysis will be conducted in terms of the random variables $x_1, \ldots, x_n$, where $x_i = a_i - \frac{1}{2} \sim U(-\frac{1}{2}, \frac{1}{2})$ for $i = 1, \ldots, n$. The probability of majority inversion then becomes:

$$P(w_1, \ldots, w_n, v_1, \ldots, v_n) = 2P\left\{\sum_{i=1}^n w_i x_i < 0\,,\ \sum_{j=1}^n v_j \mathbf{1}_{\{x_j > 0\}} > \frac{1}{2}\right\}. \tag{1}$$

The above transformation simplifies the computations without altering the sought probability.

Several additional remarks concern the notation. Since each set of weights sums to unity, formula (1) has $2(n-1)$ free variables. Nonetheless, we shall use the entire vector of weights as an argument, keeping in mind that one of its elements is determined implicitly. Secondly, with few exceptions that will be noted, all results in this paper apply to $n = 3$ only. Thirdly, we assume, without any loss of generality, that the population weights are sorted in a *descending* order. Finally, we derive the probability of majority inversion for the following variations of the two-stage voting model:

Equal Population (**EP**): $w_i = v_i = 1/n$ for all $i = 1, 2, \ldots, n$;

Westminster (**WM**): $v_i = 1/n$ for all $i = 1, 2, \ldots, n$;

Electoral College (**EC**): $v_i = w_i$ for all $i = 1, 2, \ldots, n$;

General Model (**GM**) with weights as in the assumptions A.1 and A.2.

In general, the probability of majority inversion depends on the population weights $w_1, \ldots, w_n$ and the voting weights $v_1, \ldots, v_n$. In the particular models **EP**, **WM** and **EC**, however, the voting weights are either fixed or tied to the population weights. Since in three of the models the probability depends on the population weights only, and in view of our focus on the effect of the inequality in population weights on the inversion probability, we shall always write $P(w_1, \ldots, w_n)$ instead of $P(w_1, \ldots, w_n; v_1, \ldots, v_n)$, separately listing all those combinations of voting weights, for which the probability takes distinct values.

The final remark pertains to ties. The possibility of ties significantly encumbers the analysis of voting models, often without producing new insights. We need to ensure that neither the popular vote nor the electoral vote can result in a tie. The assumption of a continuous probability distribution for the share of the electorate supporting a certain alternative technically precludes ties in the popular vote. The parity of n rules out ties in the electoral vote for models **EP** and **WM**, but not for **GM** and **EC**. To exclude ties in the electoral vote for any n, we assume that none of 2^n combinations of the voting weights add to 1/2. Whenever such cases implicitly appear in calculations, they should be viewed as limiting cases.

3 The case of three states

For any $x \in \mathbb{R}^n$, let $x_{[1]} \geq \cdots \geq x_{[n]}$ denote the elements of x in a *descending* order. The first result is a formula for the probability of majority inversion for an arbitrary distribution of population and voting weights sorted in a descending order.

Theorem 1. *For* ***GM****, the probability of majority inversion as a function of the population weights assumes the following expressions.*
i). If the largest voting weight $v_{[1]} \leq \frac{1}{2}$*, then*

$$P(w_1, w_2, w_3) = \begin{cases} \frac{w_{[1]}-w_{[2]}}{8w_{[3]}} + \frac{w_{[1]}-w_{[3]}}{8w_{[2]}} + \frac{w_{[2]}-w_{[3]}}{8w_{[1]}} - \frac{w_{[1]}^2}{24w_{[2]}w_{[3]}} + \frac{w_{[2]}^2}{24w_{[1]}w_{[3]}} + \frac{3w_{[3]}^2}{24w_{[1]}w_{[2]}} & \textit{if } w_{[1]} \leq \frac{1}{2}; \\ \frac{1}{4} - \frac{w_{[3]}}{4w_{[1]}} + \frac{w_{[3]}^2}{12w_{[1]}w_{[2]}} & \textit{if } w_{[1]} > \frac{1}{2}. \end{cases}$$

ii). If $v_{[1]} > \frac{1}{2}$*, and the indices of* $v_{[1]}$ *and* $w_{[1]}$ *coincide, then*

$$P(w_1, w_2, w_3) = \begin{cases} \frac{1}{4} - \frac{w_{[1]}-w_{[2]}}{8w_{[3]}} - \frac{w_{[1]}-w_{[3]}}{8w_{[2]}} + \frac{w_{[2]}-w_{[3]}}{8w_{[1]}} + \frac{w_{[1]}^2}{24w_{[2]}w_{[3]}} - \frac{w_{[2]}^2}{24w_{[1]}w_{[3]}} + \frac{w_{[3]}^2}{24w_{[1]}w_{[2]}} & \textit{if } w_{[1]} \leq \frac{1}{2}; \\ \frac{w_{[2]}}{4w_{[1]}} + \frac{w_{[3]}^2}{12w_{[1]}w_{[2]}} & \textit{if } w_{[1]} > \frac{1}{2}. \end{cases}$$

iii). If $v_{[1]} > \frac{1}{2}$, and the indices of $v_{[1]}$ and $w_{[2]}$ coincide, then

$$P(w_1, w_2, w_3) = \begin{cases} \frac{1}{4} + \frac{w_{[1]}-w_{[2]}}{8w_{[3]}} + \frac{w_{[1]}-w_{[3]}}{8w_{[2]}} - \frac{w_{[2]}-w_{[3]}}{8w_{[1]}} - \frac{w_{[1]}^2}{24w_{[2]}w_{[3]}} + \frac{w_{[2]}^2}{24w_{[1]}w_{[3]}} + \frac{w_{[3]}^2}{24w_{[1]}w_{[2]}} & \textit{if } w_{[1]} \leq \frac{1}{2}; \\ \frac{1}{2} - \frac{w_{[2]}}{4w_{[1]}} & \textit{if } w_{[1]} > \frac{1}{2}. \end{cases}$$

iv). If $v_{[1]} > \frac{1}{2}$, and the indices of $v_{[1]}$ and $w_{[3]}$ coincide, then

$$P(w_1, w_2, w_3) = \begin{cases} \frac{1}{4} + \frac{w_{[1]}-w_{[2]}}{8w_{[3]}} + \frac{w_{[1]}-w_{[3]}}{8w_{[2]}} + \frac{w_{[2]}-w_{[3]}}{8w_{[1]}} - \frac{w_{[1]}^2}{24w_{[2]}w_{[3]}} + \frac{w_{[2]}^2}{24w_{[1]}w_{[3]}} + \frac{w_{[3]}^2}{24w_{[1]}w_{[2]}} & \textit{if } w_{[1]} \leq \frac{1}{2}; \\ \frac{1}{2} - \frac{w_{[3]}}{4w_{[1]}} & \textit{if } w_{[1]} > \frac{1}{2}. \end{cases}$$

A proof of Theorem 1 can be found in Appendix A.

The first case in Theorem 1 is the main case. In the other three cases, the largest state can dictate the outcome of the election. The 'dictatorship' cases are more peculiar. They differ amongst themselves and from the main case in the population shares that minimize or maximize the probability of majority inversion, and in the way this probability relates to the inequality of population weights. Both of these aspects will be explored below.

The simplest of the three models, **EP**, corresponds to the limiting case of the model in May (1948), as the electorate size tends to infinity, but the number of states remains fixed. We can adapt Theorem 1 to **WM**, which is a two-stage voting model with three states of unequal size commanding an equal number of electoral votes, or **EC**, in which the electoral votes are weighted based on the population.

Corollary 1. *For **WM**, the probability of majority inversion is:*

$$P(w_1, w_2, w_3) = \begin{cases} \frac{w_{[1]}-w_{[2]}}{8w_{[3]}} + \frac{w_{[1]}-w_{[3]}}{8w_{[2]}} + \frac{w_{[2]}-w_{[3]}}{8w_{[1]}} - \frac{w_{[1]}^2}{24w_{[2]}w_{[3]}} + \frac{w_{[2]}^2}{24w_{[1]}w_{[3]}} + \frac{3w_{[3]}^2}{24w_{[1]}w_{[2]}} & \textit{if } w_{[1]} \leq \frac{1}{2}; \\ \frac{1}{4} - \frac{w_{[3]}}{4w_{[1]}} + \frac{w_{[3]}^2}{12w_{[1]}w_{[2]}} & \textit{if } w_{[1]} > \frac{1}{2}. \end{cases}$$

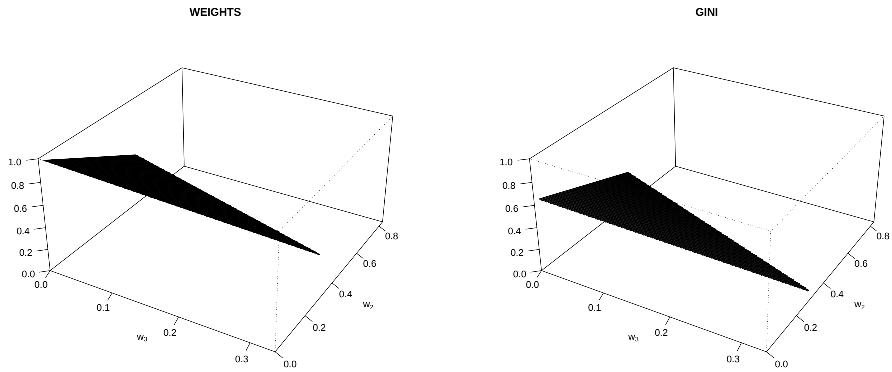
The top formula for the probability holds for the case when the 'one person, one vote' rule applies at each stage, as **EP** is subsumed in **WM** for $w_{[1]} \leq \frac{1}{2}$. The top formula also holds for **EC**.

Corollary 2. *For **EC**, the probability of majority inversion is:*

$$P(w_1, w_2, w_3) = \begin{cases} \frac{w_{[1]}-w_{[2]}}{8w_{[3]}} + \frac{w_{[1]}-w_{[3]}}{8w_{[2]}} + \frac{w_{[2]}-w_{[3]}}{8w_{[1]}} - \frac{w_{[1]}^2}{24w_{[2]}w_{[3]}} + \frac{w_{[2]}^2}{24w_{[1]}w_{[3]}} + \frac{3w_{[3]}^2}{24w_{[1]}w_{[2]}} & \textit{if } w_{[1]} \leq \frac{1}{2}; \\ \frac{w_{[2]}}{4w_{[1]}} + \frac{w_{[3]}^2}{12w_{[1]}w_{[2]}} & \textit{if } w_{[1]} > \frac{1}{2}. \end{cases}$$

The proof of Theorem 1 shows that the formula for **WM** turns out to be identical to the formula for **EC** for any given odd n, provided that the largest population weight does not exceed the sum of the two smallest weights, $w_{[n-1]} + w_{[n]} \geq w_{[1]}$. The population weights must be sufficiently distinct for them to have an effect on the probability of majority inversion when used as voting weights in the second stage. With equipopulous states in **EP**, this probability equals 1/8. It increases to 7/24 in the presence of a 'dictator' at the second stage of the voting procedure.

Figure 1: Population weights and the Gini coefficient for $n = 3$.



4 Inequality of weights

To study how the probability of majority inversion varies with inequality of population weights, we first need to express the degree of inequality as a function of weights. Despite the existence of several inequality metrics, the Gini coefficient is certainly among the most popular metrics, if not the most popular one (Yitzhaki and Schechtman 2013). Figure 1 shows the admissible population weights and the Gini coefficient as a measure of their inequality, when $w_1 \geq w_2 \geq w_3$. The assumed order of the population weights entails the following upper bounds on the smallest two weights: $w_3 \leq 1/3$ and $w_2 \leq 1/2$, whereas the implicit largest weight $w_1 = 1 - w_2 - w_3$ can vary freely in $[1/3, 1]$. For a general input vector of positive elements, the Gini coefficient takes values between zero and one. The value of zero indicates the equality, the value of one the maximal inequality. This range may narrow to a subset of the unit interval if the elements of the input vector are constrained. For an input vector of n shares that add to one, the maximum equals $(n-1)/n$. The main property still holds: the larger the coefficient, the more unequal the distribution. The Gini coefficient is defined as

$$G(w_1, w_2, \ldots, w_n) = \frac{2\sum_{i=1}^{n}(n-i+1)w_i}{n\sum_{i=1}^{n} w_i} - \frac{n+1}{n} \text{ for } w_1 \geq w_2 \geq \cdots \geq w_{n-1} \geq w_n > 0.$$

In the case of three weights, such that $w_1 + w_2 + w_3 = 1$, it assumes a particularly simple form:

$$G(w_1, w_2, w_3) = \frac{2(w_1 - w_3)}{3},$$

taking the minimal value of zero for $w_1 = w_2 = w_3 = 1/3$, and the maximal value of $2/3$ for $w_1 = 1$ and $w_2 = w_3 = 0$.

The essential property of the Gini coefficient is strict *Schur-convexity*, a natural property for an inequality metric. This property ensures that taking from the poor and giving to the rich strictly increases inequality. A Schur-convex function preserves the order under *majorization* as a relational property between a pair of vectors x and y. To define majorization, we first need to compare the partial sums of the sorted elements of each vector. The following two definitions can be found in Marshall et al. (2011) (Ch. 1, Definition A.1; Ch. 3, Definition A.1):

Definition 1. *For $x, y \in \mathbb{R}^n$,*

$$x \prec y \;\; if \;\; \begin{cases} \sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]}, \;\; k = 1, \ldots, n-1, \\ \sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}. \end{cases}$$

When $x \prec y$, x is said to be majorized by y (y majorizes x).

The vector y majorizes x if their elements sum to the same value, but the elements of y are more dispersed than the elements of x. Higher dispersion causes the partial sum of the sorted y to lead the partial sum of the sorted x for any k. Marshall et al. (2011) provide numerous applications in which majorization arises and discuss several tests of this property.

Definition 2. *A real-valued function ϕ defined on a set $\mathcal{A} \subset \mathbb{R}^n$ is said to be Schur-convex on $\mathcal{A}$ if*

$$x \prec y \;\; on \;\; \mathcal{A} \Rightarrow \phi(x) \leq \phi(y).$$

If in addition, $\phi(x) < \phi(y)$ whenever $x \prec y$ but x is not a permutation of y, then ϕ is said to be strictly Schur-convex on $\mathcal{A}$.

Being isotonic relative to the preordering $\prec$, a Schur-convex function increases with dispersion in its arguments. As a strictly Schur-convex function, the Gini coefficient strictly increases with dispersion, expressing the fact that higher concentration implies more inequality.

The following characterization quoting Marshall et al. (2011) (Ch. 3, Lemma A.2) is particularly intuitive in the context of inequality measurement.

Lemma 1. *Let ϕ be a continuous real-valued function defined on*

$$\mathcal{D} = \{(x_1, x_2, \ldots, x_n) : x_1 \geq x_2 \geq \ldots \geq x_n\}.$$

Then ϕ is Schur-convex on $\mathcal{D}$, if and only if for all $z \in \mathcal{D}$ and $k = 1, \ldots, n-1$,

$$\phi(z_1, \ldots, z_{k-1}, z_k + \varepsilon, z_{k+1} - \varepsilon, z_{k+2}, \ldots, z_n)$$

is increasing in ε over the region

$$\begin{aligned} &0 \leq \varepsilon \leq z_2 - z_3, \quad k = 1, \\ &0 \leq \varepsilon \leq \min\{z_{k-1} - z_k, z_{k+1} - z_{k+2}\}, \quad k = 2, \ldots, n-2, \\ &0 \leq \varepsilon \leq z_{n-2} - z_{n-1}, \quad k = n-1. \end{aligned}$$

The last set of inequalities ensures that transferring a positive ε from z_{k+1} to z_k increases the dispersion among the arguments. Consequently, the lemma says that increasing the dispersion cannot decrease the value of the function. There exist alternative definitions of Schur-convexity for narrower classes of functions, such as symmetric functions with or without partial derivatives.

The second theorem establishes that the probability of majority inversion is a Schur-convex function of the population weights.

Theorem 2. *For* ***GM****, the probability of majority inversion is Schur-convex if the largest voting weight $v_{[1]} \leq \frac{1}{2}$.*

A proof of Theorem 2 is given in Appendix B.

Theorem 2 implies that the probability of majority inversion is a Schur-convex function of the population weights for **EC** for $w_{[1]} \leq \frac{1}{2}$ and also for **WM**. For these common types of indirect voting systems, we can formulate the following rule: the more unequal the states, the more likely inversion is to occur.

The superposition rules for the Schur-convex function summarized in Ch. 3, Table 1 in Marshall et al. (2011) imply a monotonic relationship between the probability of majority inversion and the inequality of population weights. The Schur-convexity of the two functions ensures that if a third coupling function h exists, such that $P(w_1, w_2, w_3) = h(G(w_1, w_2, w_3))$, then h must be increasing. Putting the two results together allows us to conclude that with three states the probability of majority inversion increases with inequality in state sizes, provided that the share of the largest state does not exceed one half.

4.1 Bounds on the probability for $n = 3$

Figure 2 shows the probability of majority inversion for **WM** and **EC** when $w_1 \geq w_2 \geq w_3$. A comparison of the right panel of Figure 1 with Figure 2 illustrates the property established in Theorem 2. The probability of majority inversion increases with the inequality in population weights as measured by the Gini coefficient, provided that no state can dictate the outcome of an election.

Turning to the extremal values of the probability of majority inversion for different distribution of weights, let us again first consider **WM** and **EC**. The left panel shows that for **WM** the global maximum takes the value 1/4 for $w_3 = 0$. The global maximum is not unique. The global minimum appears to lie somewhere between 0.1 and 0.15; the actual value being 1/8 at $w_1 = w_2 = w_3 = 1/3$. The limiting behavior of the probability for **EC** is different from that for **WM**. The right panel shows a crest of local maxima for **EC**. They occur for all population weights such as $(1/2, 1/2 - w_3, w_3)$, with the global maximum of 1/4 for $w_3 = 0$. This fact is proven in Appendix C. For **EC**, substituting $n = 3$ and $(1, 0, 0)$ in the general formula (1) yields $P(1, 0, 0) = 0$, which is the global minimum. Arguments similar to those used in Appendix C yield the following complete characterization of the extrema, which we state without a proof.

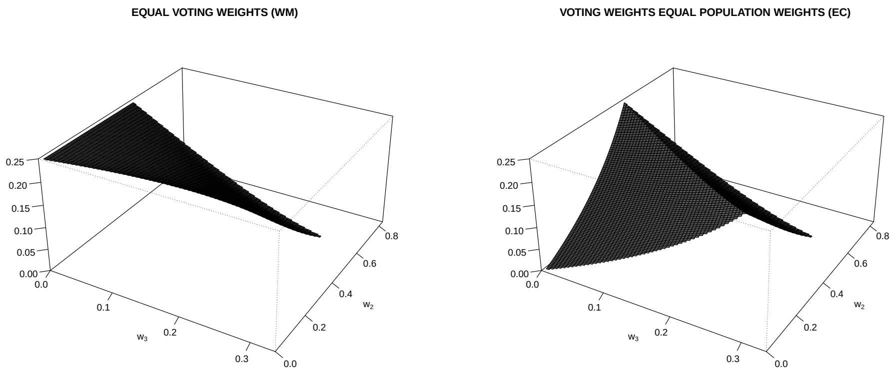
Theorem 3. *For **GM**, the probability of majority inversion as a function of the population weights assumes the following expressions.*
i). If the largest voting weight $v_{[1]} \leq \frac{1}{2}$, then

$$\min P(w_1, w_2, w_3) = \frac{1}{8} \text{ for } (w_1^*, w_2^*, w_3^*) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$
$$\max P(w_1, w_2, w_3) = \frac{1}{4} \text{ for } (w_1^*, w_2^*, w_3^*) = (w_1^*, 1 - w_1^*, 0) \text{ for any } w_1^* \geq \frac{1}{2}.$$

ii). If $v_{[1]} > \frac{1}{2}$, and the indices of $v_{[1]}$ and $w_{[1]}$ coincide, then

$$\min P(w_1, w_2, w_3) = 0 \text{ for } (w_1^*, w_2^*, w_3^*) = (1, 0, 0)$$
$$\max P(w_1, w_2, w_3) = \frac{7}{24} \text{ for } (w_1^*, w_2^*, w_3^*) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

Figure 2: The probability of majority inversion for $n = 3$.



iii). If $v_{[1]} > \frac{1}{2}$*, and the indices of* $v_{[1]}$ *and* $w_{[2]}$ *coincide, then*

$$\min P(w_1, w_2, w_3) = \frac{1}{4} \text{ for } (w_1^*, w_2^*, w_3^*) = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$$
$$\max P(w_1, w_2, w_3) = \frac{1}{2} \text{ for } (w_1^*, w_2^*, w_3^*) = (1, 0, 0).$$

iv). If $v_{[1]} > \frac{1}{2}$*, and the indices of* $v_{[1]}$ *and* $w_{[3]}$ *coincide, then*

$$\min P(w_1, w_2, w_3) = \frac{7}{24} \text{ for } (w_1^*, w_2^*, w_3^*) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$
$$\max P(w_1, w_2, w_3) = \frac{1}{2} \text{ for } (w_1^*, w_2^*, w_3^*) = (w_1^*, 1 - w_1^*, 0) \text{ for any } w_1^* \geq \frac{1}{2}.$$

The models **WM** and **EC** are covered by i) and i)-ii), respectively.

5 Discussion

How do the above results relate to the existing literature? Let us start with the simplest model **EP**, in which all states have equal population and voting weights. May (1948) offered a complete analysis of the discrete case, as well as various limiting cases, one of which corresponds to **EP** for any odd $n \geq 3$. His formula for the probability of majority inversion for **EP** reads:

$$P(\infty, n) = \frac{1}{n! 2^{n-1}} \sum_{r=1}^{\frac{n-1}{2}} \sum_{i=0}^{\frac{n-1}{2}-r} (-1)^i C_n^i C_n^{i+r} r^n. \tag{2}$$

The notation $P(\infty, n)$ indicates a 'continuum' of voters populating n states. This probability is best viewed as an approximation, which May calls the 'large number of voters in each district' model, because assuming equally sized states populated by infinitely many voters is a simplification characteristic of a continuous population model. These qualifications equally apply to the three models studied in this paper. May has also established that, as the number of states tends to infinity, the probability of majority inversion tends to 1/6, concluding with

'On the basis of the simple case here considered, the best we can do is to suggest that for the values of m and n usually encountered the probability is of the order of 1/6'.

May used his formula to compute the probability of majority inversion for $n = 3, 5, 7, 9$. We verify May's first four values using the software `LattE` by Baldoni, Berline, De Loera, Dutra, Köppe, Moreinis, Pinto, Vergne and Wu (2014).[5] Our results

$$\frac{1}{8}, \frac{55}{384}, \frac{577}{3840}, \frac{1589879}{10321920}$$

agree with May's $(1/8, 0.143, 0.150, 0.154)$ (Table on p. 208 in May 1948). We switch to numerical simulation starting from $n = 11$. The bottom line in Figure 3 shows the probability of majority inversion for **EP**. The first four exact values are connected by a solid line. The simulated values from $n = 11$ till $n = 101$ are connected with a dashed line. It is evident that the numerically simulated probability tends to 1/6, as $n \to \infty$.

Let us now turn to the more general model **WM**. The model **EP** is a special case of **WM** with equally-sized states. For **WM** and $n = 3$, Theorem 2 shows that the probability is a Schur-convex function of population weights. Consequently, for $n = 3$, the lower bound on the probability is attained for $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, whereas the upper bound is attained for $(w_1, 1 - w_1, 0)$ for any $w_1 \geq \frac{1}{2}$, if the population weights are sorted in a descending order (Theorem 3). Moreover for **WM**, we can use formula (1) to obtain the following probability

$$P(1, \underbrace{0, \ldots, 0}_{n-1}) = \sum_{k=1}^{(n-1)/2} C_{n-1}^{(n-1)/2+k} \frac{1}{2^{n-1}} = \frac{1}{2^{n-1}} \sum_{k=0}^{(n-3)/2} C_{n-1}^{k} = \frac{1}{2} - \frac{C_{n-1}^{(n-1)/2}}{2^n}. \tag{3}$$

We show the probability (3) in Figure 3 as the solid line. The probability (3) increases from 1/4 to 1/2, when n increases from three to infinity.

The maximum inequality among the states is achieved by concentrating all voters in a single state. The cases of 'empty' states are purely hypothetical, because the absence of voters at the first stage makes voting in the second stage meaningless. Note also that the formulae in Theorem 1 and its corollaries require strictly positive weights. Nevertheless, the general formula for the probability of majority inversion (1) remains valid. We have already noted that substituting the vector $(1, 0, 0)$ in (1) yields zero probability of majority inversion for **EC**. This result generalizes to any odd $n \geq 3$, as

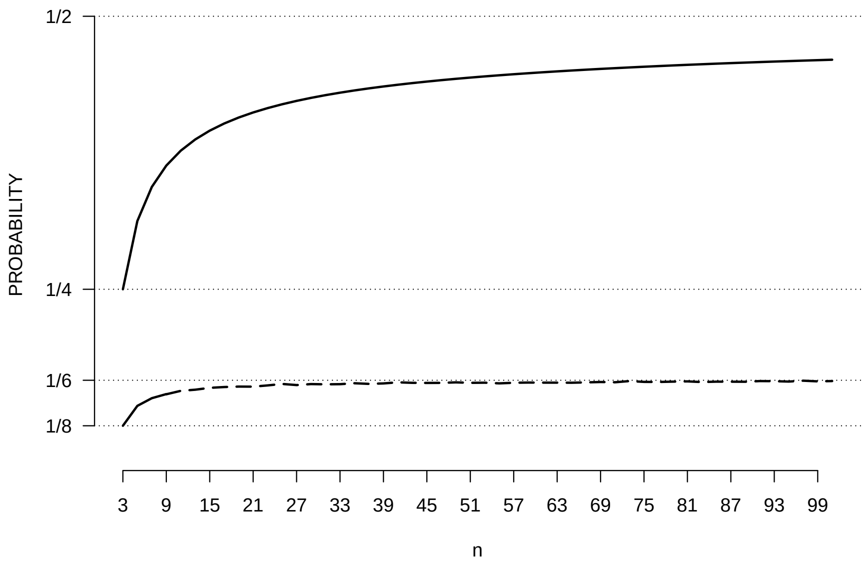
$$P(1, \underbrace{0, \ldots, 0}_{n-1}) = 0.$$

The overall lower and upper bounds for the probability of majority inversion of 0 and 1/2 agree with the results of the asymptotic analysis in Kikuchi (2016).

We shall conclude the paper by revisiting the relationship between weight inequality and the probability of majority inversion and formulating a conjecture. We established that for $n = 3$, the probability of majority inversion for **WM** is bounded between May's formula (2) and (3). For $n = 3$, the bounding values are 1/8 and 1/4. We conjecture that these bounding functions apply for any odd $n \geq 3$, or that May's formula (2) remains the lower bound and (3) remains the upper bound on the probability of majority inversion for **WM** for any odd $n \geq 3$. This

[5] The theoretical background for the methods used in `LattE` is elaborated in De Loera, Dutra, Köppe, Moreinis, Pinto and Wu (2013).

Figure 3: Conjectured bounds on the probability of majority inversion for **WM**.



conjecture would be proven if one could establish Schur-convexity of the probability of majority inversion for **WM** for any odd number of states, a challenge that we leave to future work. Although the asymptotic analysis in Kikuchi (2016) lends optimism in this matter, we surmise that proving this conjecture may require methods more advanced than those used in this paper.

Acknowledgments: We would like to thank Michel Le Breton for constructive criticism.

References

Baldoni, V., Berline, N., De Loera, J. A., Dutra, B., Köppe, M., Moreinis, S., Pinto, G., Vergne, M. and Wu, J.: 2014, *A user's guide for LattE integrale v1.7.2.* *https://www.math.ucdavis.edu/~latte

Casella, G. and Berger, R. L.: 2002, *Statistical inference*, Duxbury Press.

De Loera, J. A., Dutra, B., Köppe, M., Moreinis, S., Pinto, G. and Wu, J.: 2013, Software for exact integration of polynomials over polyhedra, *Computational geometry: Theory and applications* **46**, 232–252.

De Mouzon, O., Laurent, T., Le Breton, M. and Lepelley, D.: 2017, The theoretical Shapley-Shubik probability of an election inversion in a toy symmetric version of the U.S. presidential electoral system. Working Paper.

Feix, M. R., Lepelley, D., Merlin, V. R. and Rouet, J.-L.: 2004, The probability of conflicts in a U.S. presidential type election, *Economic Theory* **23**, 227–257.

Felsenthal, D. S. and Machover, M.: 1998, *The measurement of voting power: Theory and practice, problems and paradoxes*, Edward Elgar.

Kikuchi, K.: 2016, The likelihood of majority inversion in an indirect voting system. Waseda University - School of Political Science and Economics, Working Paper.

Lepelley, D., Merlin, V. R. and Rouet, J.-L.: 2011, , *Mathematical Social Sciences* **62**, 28–33.

Lepelley, D., Merlin, V. R., Rouet, J.-L. and Vidu, L.: 2014, Referendum paradox in a federal union with unequal populations: the three state case, *Economics Bulletin* **34**, 2201–2207.

Marshall, A. W., Olkin, I. and Arnold, B. C.: 2011, *Inequalities: Theory of majorization and its applications*, Springer-Verlag New York.

May, K.: 1948, Probabilities of certain election results, *The American Mathematical Monthly* **55**, 203–209.

Miller, N. R.: 2012, Election inversions by the U.S. Electoral College, *in* D. S. Felsenthal and M. Machover (eds), *Electoral Systems. Studies in Choice and Welfare*, Springer-Verlag Berlin Heidelberg, pp. 93–127.

Straffin, P. D.: 1977, Homogeneity, independence and power indices, *Public Choice* **30**, 107–118.

Yitzhaki, S. and Schechtman, E.: 2013, *The Gini methodology: A primer on a statistical methodology. Springer series in statistics, vol 272*, Springer-Verlag New York.

A Proof of Theorem 1 (Probability of Inversion)

If $v_{[1]} \leq \frac{1}{2}$, then the probability of majority inversion

$$\begin{aligned} P(w_1, w_2, w_3) = & \, 2P\left(w_1x_1 + w_2x_2 + w_3x_3 < 0, x_1 \in (-\tfrac{1}{2}, 0), x_2 \in (0, \tfrac{1}{2}), x_3 \in (0, \tfrac{1}{2})\right) && (P_1) \\ & + 2P\left(w_1x_1 + w_2x_2 + w_3x_3 < 0, x_1 \in (0, \tfrac{1}{2}), x_2 \in (-\tfrac{1}{2}, 0), x_3 \in (0, \tfrac{1}{2})\right) && (P_2) \\ & + 2P\left(w_1x_1 + w_2x_2 + w_3x_3 < 0, x_1 \in (0, \tfrac{1}{2}), x_2 \in (0, \tfrac{1}{2}), x_3 \in (-\tfrac{1}{2}, 0)\right). && (P_3) \end{aligned}$$

Each summand can be expressed as a multiple integral

$$P_i = \frac{2}{w_1w_2w_3} \int\!\!\int\limits_{A_i}\!\!\int dxdydz \quad \text{for} \quad i = 1, 2, 3,$$

evaluated over the following convex polytopes

$$\begin{aligned} A_1 &= \{(x, y, z) : x \in (0, \tfrac{w_1}{2}), y \in (0, \tfrac{w_2}{2}), z \in (0, \tfrac{w_3}{2}), x > y + z\}, \\ A_2 &= \{(x, y, z) : x \in (0, \tfrac{w_1}{2}), y \in (0, \tfrac{w_2}{2}), z \in (0, \tfrac{w_3}{2}), y > x + z\}, \\ A_3 &= \{(x, y, z) : x \in (0, \tfrac{w_1}{2}), y \in (0, \tfrac{w_2}{2}), z \in (0, \tfrac{w_3}{2}), z > x + y\}. \end{aligned}$$

Here, we use the fact that if $x_i \sim U(-\frac{1}{2}, \frac{1}{2})$, then also $-x_i \sim U(-\frac{1}{2}, \frac{1}{2})$ for $i = 1, 2, 3$.

Assume $w_1 \geq w_2 \geq w_3$. Eliminating y yields

$$\begin{aligned} P(w_1, w_2, w_3) = & \frac{2}{w_1w_2w_3} \int\!\!\int\limits_{A_1'} (x - z) dxdz + \frac{1}{w_1w_3} \int\!\!\int\limits_{A_1''} dzdx && (P_1) \\ & + \frac{2}{w_1w_2w_3} \int\!\!\int\limits_{A_2'} \left(\frac{w_2}{2} - x - z\right) dxdz && (P_2) \\ & + \frac{2}{w_1w_2w_3} \int\!\!\int\limits_{A_3'} (z - x) dzdx, && (P_3) \end{aligned}$$

where the transformed integration regions are given by

$$\begin{aligned} A_1' &= \{(x, z) : x \in (0, \tfrac{w_1}{2}), z \in (0, \tfrac{w_3}{2}), 0 < x - z < \tfrac{w_2}{2}\}, \\ A_1'' &= \{(x, z) : x \in (0, \tfrac{w_1}{2}), z \in (0, \tfrac{w_3}{2}), x - z \geq \tfrac{w_2}{2}\}, \\ A_2' &= \{(x, z) : x \in (0, \tfrac{w_1}{2}), z \in (0, \tfrac{w_3}{2}), x + z < \tfrac{w_2}{2}\}, \\ A_3' &= \{(x, z) : x \in (0, \tfrac{w_1}{2}), z \in (0, \tfrac{w_3}{2}), z - x > 0\}. \end{aligned}$$

The integrals corresponding to P_2 and P_3 evaluate as

$$\begin{aligned} P_2 &= \frac{2}{w_1w_2w_3} \int_0^{\frac{w_3}{2}} dz \int_0^{\frac{w_2}{2} - z} \left(\frac{w_2}{2} - x - z\right) dx = \frac{1}{w_1w_2w_3} \int_0^{\frac{w_3}{2}} \left(\frac{w_2}{2} - z\right)^2 dz = \frac{3w_2^2 - 3w_2w_3 + w_3^2}{24w_1w_2}, \\ P_3 &= \frac{2}{w_1w_2w_3} \int_0^{\frac{w_3}{2}} dz \int_0^z (z - x) dx = \frac{1}{w_1w_2w_3} \int_0^{\frac{w_3}{2}} z^2 dz = \frac{w_3^2}{24w_1w_2}. \end{aligned}$$

The value of P_1 depends on whether the largest weight exceeds $\frac{1}{2}$.

If $w_1 > \frac{1}{2}$, then

$$P_1 = \frac{2}{w_1 w_2 w_3} \int_0^{\frac{w_3}{2}} dz \int_z^{z+\frac{w_2}{2}} (x-z)dx + \frac{1}{w_1 w_3} \int_0^{\frac{w_3}{2}} dz \int_{z+\frac{w_2}{2}}^{\frac{w_1}{2}} dx$$
$$= \frac{w_2}{4w_1 w_3} \int_0^{\frac{w_3}{2}} dz + \frac{1}{w_1 w_3} \int_0^{\frac{w_3}{2}} \left(\frac{w_1 - w_2}{2} - z\right) dz = \frac{2w_1 - w_2 - w_3}{8w_1}.$$

If $w_1 \le \frac{1}{2}$, then

$$P_1 = \frac{2}{w_1 w_2 w_3} \int_0^{\frac{w_1-w_2}{2}} dz \int_z^{z+\frac{w_2}{2}} (x-z)dx + \frac{2}{w_1 w_2 w_3} \int_{\frac{w_1-w_2}{2}}^{\frac{w_3}{2}} dz \int_z^{\frac{w_1}{2}} (x-z)dx$$
$$+ \frac{1}{w_1 w_3} \int_0^{\frac{w_1-w_2}{2}} dz \int_{z+\frac{w_2}{2}}^{\frac{w_1}{2}} dx = \frac{w_2}{4w_1 w_3} \int_0^{\frac{w_1-w_2}{2}} dz + \frac{1}{w_1 w_2 w_3} \int_{\frac{w_1-w_2}{2}}^{\frac{w_3}{2}} \left(\frac{w_1}{2} - z\right)^2 dz$$
$$+ \frac{1}{w_1 w_3} \int_0^{\frac{w_1-w_2}{2}} \left(\frac{w_1 - w_2}{2} - z\right) dz = \frac{w_1 - w_2}{8w_3} + \frac{w_2^2}{24w_1 w_3} - \frac{w_1^2}{24 w_2 w_3} + \frac{w_1 - w_3}{8w_2} + \frac{w_3^2}{24w_1 w_2}.$$

Therefore, if $w_1 > \frac{1}{2}$, then

$$P(w_1, w_2, w_3) = \frac{3w_2^2 - 3w_2 w_3 + 2w_3^2}{24 w_1 w_2} + \frac{2w_1 - w_2 - w_3}{8w_1} = \frac{1}{4} - \frac{w_3}{4w_1} + \frac{w_3^2}{12 w_1 w_2},$$

else

$$P(w_1, w_2, w_3) = \frac{3w_2^2 - 3w_2 w_3 + 2w_3^2}{24 w_1 w_2} + \frac{w_1 - w_2}{8w_3} + \frac{w_2^2}{24 w_1 w_3} - \frac{w_1^2}{24 w_2 w_3} + \frac{w_1 - w_3}{8w_2} + \frac{w_3^2}{24 w_1 w_2}$$
$$= \frac{w_1 - w_2}{8w_3} + \frac{w_1 - w_3}{8w_2} + \frac{w_2 - w_3}{8w_1} - \frac{w_1^2}{24 w_2 w_3} + \frac{w_2^2}{24 w_1 w_3} + \frac{3w_3^2}{24 w_1 w_2}.$$

Consider the case $v_{[1]} > \frac{1}{2}$. As before, let $w_1 \ge w_2 \ge w_3$.
i) If $v_{[1]} = v_1$, then the probability of majority inversion

$$P(w_1, w_2, w_3) = 2P\left(w_1 x_1 + w_2 x_2 + w_3 x_3 < 0, x_1 \in (0, \tfrac{1}{2}), x_2 \in (-\tfrac{1}{2}, 0), x_3 \in (-\tfrac{1}{2}, 0)\right) \quad (P_1')$$
$$+ 2P\left(w_1 x_1 + w_2 x_2 + w_3 x_3 < 0, x_1 \in (0, \tfrac{1}{2}), x_2 \in (-\tfrac{1}{2}, 0), x_3 \in (0, \tfrac{1}{2})\right) \quad (P_2')$$
$$+ 2P\left(w_1 x_1 + w_2 x_2 + w_3 x_3 < 0, x_1 \in (0, \tfrac{1}{2}), x_2 \in (0, \tfrac{1}{2}), x_3 \in (-\tfrac{1}{2}, 0)\right). \quad (P_3')$$

Thus, $P_2' = P_2, P_3' = P_3$. For P_1', we have

$$P_1' = \frac{2}{w_1 w_2 w_3} \int\int_A\int dxdydz,$$

where $A = \{(x,y,z) : x \in (0, \frac{w_1}{2}), y \in (0, \frac{w_2}{2}), z \in (0, \frac{w_3}{2}), x < y + z\}$. Since $P_1' = \frac{1}{4} - P_1$,

$$P(w_1, w_2, w_3) = \frac{w_2}{4w_1} + \frac{w_3^2}{12 w_1 w_2} \text{ for } w_1 > \frac{1}{2},$$
$$P(w_1, w_2, w_3) = \frac{1}{4} - \frac{w_1 - w_2}{8w_3} - \frac{w_1 - w_3}{8w_2} + \frac{w_2 - w_3}{8w_1} + \frac{w_1^2}{24 w_2 w_3} - \frac{w_2^2}{24 w_1 w_3} + \frac{w_3^2}{24 w_1 w_2} \text{ for } w_1 \le \frac{1}{2}.$$

ii) If $v_{[1]} = v_2$, then the probability of majority inversion

$$\begin{aligned} P(w_1,w_2,w_3) &= 2P\left(w_1x_1+w_2x_2+w_3x_3<0, x_1\in(-\tfrac{1}{2},0), x_2\in(0,\tfrac{1}{2}), x_3\in(0,\tfrac{1}{2})\right) && (P_1'') \\ &+ 2P\left(w_1x_1+w_2x_2+w_3x_3<0, x_1\in(-\tfrac{1}{2},0), x_2\in(0,\tfrac{1}{2}), x_3\in(-\tfrac{1}{2},0)\right) && (P_2'') \\ &+ 2P\left(w_1x_1+w_2x_2+w_3x_3<0, x_1\in(0,\tfrac{1}{2}), x_2\in(0,\tfrac{1}{2}), x_3\in(-\tfrac{1}{2},0)\right). && (P_3'') \end{aligned}$$

Here $P_1'' = P_1, P_2'' = \frac{1}{4} - P_2, P_3'' = P_3$. Therefore,

$$P(w_1,w_2,w_3) = \frac{1}{2} - \frac{w_2}{4w_1} \text{ for } w_1 > \frac{1}{2},$$
$$P(w_1,w_2,w_3) = \frac{1}{4} + \frac{w_1-w_2}{8w_3} + \frac{w_1-w_3}{8w_2} - \frac{w_2-w_3}{8w_1} - \frac{w_1^2}{24w_2w_3} + \frac{w_2^2}{24w_1w_3} + \frac{w_3^2}{24w_1w_2} \text{ for } w_1 \le \frac{1}{2}.$$

iii) If $v_{[1]} = v_3$, then the probability of majority inversion

$$\begin{aligned} P(w_1,w_2,w_3) &= 2P\left(w_1x_1+w_2x_2+w_3x_3<0, x_1\in(-\tfrac{1}{2},0), x_2\in(0,\tfrac{1}{2}), x_3\in(0,\tfrac{1}{2})\right) && (P_1''') \\ &+ 2P\left(w_1x_1+w_2x_2+w_3x_3<0, x_1\in(0,\tfrac{1}{2}), x_2\in(-\tfrac{1}{2},0), x_3\in(0,\tfrac{1}{2})\right) && (P_2''') \\ &+ 2P\left(w_1x_1+w_2x_2+w_3x_3<0, x_1\in(-\tfrac{1}{2},0), x_2\in(-\tfrac{1}{2},0), x_3\in(0,\tfrac{1}{2})\right). && (P_3''') \end{aligned}$$

Here $P_1''' = P_1, P_2''' = P_2, P_3''' = \frac{1}{4} - P_3$. Thus,

$$P(w_1,w_2,w_3) = \frac{1}{2} - \frac{w_3}{4w_1} \text{ for } w_1 > \frac{1}{2},$$
$$P(w_1,w_2,w_3) = \frac{1}{4} + \frac{w_1-w_2}{8w_3} + \frac{w_1-w_3}{8w_2} + \frac{w_2-w_3}{8w_1} - \frac{w_1^2}{24w_2w_3} + \frac{w_2^2}{24w_1w_3} + \frac{w_3^2}{24w_1w_2} \text{ for } w_1 \le \frac{1}{2}.$$

B Proof of Theorem 2 (Schur-convexity)

Assume for simplicity that $w_1 \ge w_2 \ge w_3$. To prove that $P(w_1,w_2,w_3)$ is Schur-convex, we use Lemma 1 to show that

a). $P(w_1, w_2+\varepsilon, w_3-\varepsilon)$ is an increasing function of ε in $0 \le \varepsilon \le \min\{w_1-w_2, w_3\}$;

b). $P(w_1+\varepsilon, w_2-\varepsilon, w_3)$ is an increasing function of ε in $0 \le \varepsilon \le w_2 - w_3$.

Part a).
If $w_1 > w_2 + w_3$, then

$$P(w_1, w_2+\varepsilon, w_3-\varepsilon) = \frac{1}{4} - \frac{w_3-\varepsilon}{4w_1} + \frac{(w_3-\varepsilon)^2}{12w_1(w_2+\varepsilon)},$$

and the derivative

$$\frac{dP(w_1, w_2+\varepsilon, w_3-\varepsilon)}{d\varepsilon} = \frac{1}{12w_1}\left(4 - \frac{(w_2+w_3)^2}{(w_2+\varepsilon)^2}\right) > 0.$$

The inequality follows since $\frac{w_2+w_3}{w_2+\varepsilon} < 2$ due to $w_2 - w_3 + 2\varepsilon > 0$.

If $w_1 \le w_2 + w_3$, then from

$$P(w_1, w_2+\varepsilon, w_3-\varepsilon) = \frac{w_1 - w_2 - \varepsilon}{8(w_3-\varepsilon)} + \frac{w_1 - w_3 + \varepsilon}{8(w_2+\varepsilon)} + \frac{w_2 - w_3 + 2\varepsilon}{8w_1} - \frac{w_1^2}{24(w_2+\varepsilon)(w_3-\varepsilon)} + \frac{(w_2+\varepsilon)^2}{24w_1(w_3-\varepsilon)} + \frac{3(w_3-\varepsilon)^2}{24w_1(w_2+\varepsilon)}$$

we obtain

$$24w_1 P(w_1, w_2+\varepsilon, w_3-\varepsilon) = 6w_1 - 2w_2 - 10w_3 + 8\varepsilon + \frac{3w_1(w_1-w_2-w_3) + (w_2+w_3)^2 - \frac{w_1^3}{w_2+w_3}}{w_3-\varepsilon} + \frac{3w_1(w_1-w_2-w_3) + 3(w_2+w_3)^2 - \frac{w_1^3}{w_2+w_3}}{w_2+\varepsilon},$$

and the derivative

$$24w_1 \frac{dP(w_1, w_2+\varepsilon, w_3-\varepsilon)}{d\varepsilon} = 8 + \frac{3w_1(w_1-w_2-w_3) + (w_2+w_3)^2 - \frac{w_1^3}{w_2+w_3}}{(w_3-\varepsilon)^2} - \frac{3w_1(w_1-w_2-w_3) + 3(w_2+w_3)^2 - \frac{w_1^3}{w_2+w_3}}{(w_2+\varepsilon)^2}.$$

The expression in the numerator of the first fraction is non-negative:

$$3w_1(2w_1-1) + (1-w_1)^2 - \frac{w_1^3}{1-w_1} = \frac{(1-2w_1)^3}{1-w_1} \ge 0.$$

Therefore,

$$24w_1 \frac{dP(w_1, w_2+\varepsilon, w_3-\varepsilon)}{d\varepsilon} \ge 8 + \frac{3w_1(w_1-w_2-w_3) + (w_2+w_3)^2 - \frac{w_1^3}{w_2+w_3}}{(w_2+\varepsilon)^2} - \frac{3w_1(w_1-w_2-w_3) + 3(w_2+w_3)^2 - \frac{w_1^3}{w_2+w_3}}{(w_2+\varepsilon)^2} = 8 - \frac{2(w_2+w_3)^2}{(w_2+\varepsilon)^2} > 0.$$

Part b).
If $w_1 + 2\varepsilon > w_2 + w_3$, then

$$P(w_1+\varepsilon, w_2-\varepsilon, w_3) = \frac{1}{4} - \frac{w_3}{4(w_1+\varepsilon)} + \frac{w_3^2}{12(w_1+\varepsilon)(w_2-\varepsilon)},$$

and

$$\frac{dP(w_1+\varepsilon, w_2-\varepsilon, w_3)}{d\varepsilon} = \frac{w_3}{4(w_1+\varepsilon)^2} + \frac{w_3^2(w_1-w_2+2\varepsilon)}{12(w_1+\varepsilon)^2(w_2-\varepsilon)^2} \ge 0.$$

If $w_1 + 2\varepsilon \le w_2 + w_3$, then

$$P(w_1+\varepsilon, w_2-\varepsilon, w_3) = \frac{w_1 - w_2 + 2\varepsilon}{8w_3} + \frac{w_1 - w_3 + \varepsilon}{8(w_2-\varepsilon)} + \frac{w_2 - w_3 - \varepsilon}{8(w_1+\varepsilon)} - \frac{(w_1+\varepsilon)^2}{24(w_2-\varepsilon)w_3} + \frac{(w_2-\varepsilon)^2}{24(w_1+\varepsilon)w_3} + \frac{3w_3^2}{24(w_1+\varepsilon)(w_2-\varepsilon)}.$$

Consequently,

$$24w_3P(w_1+\varepsilon,w_2-\varepsilon,w_3)=4w_1-4w_2-6w_3+8\varepsilon+\frac{3w_3(w_1+w_2-w_3)+\frac{3w_3^3}{w_1+w_2}-(w_1+w_2)^2}{w_2-\varepsilon}$$
$$+\frac{3w_3(w_1+w_2-w_3)+\frac{3w_3^3}{w_1+w_2}+(w_1+w_2)^2}{w_1+\varepsilon}.$$

The derivative

$$24w_3\frac{dP(w_1+\varepsilon,w_2-\varepsilon,w_3)}{d\varepsilon}=8+\frac{3w_3(w_1+w_2-w_3)+\frac{3w_3^3}{w_1+w_2}-(w_1+w_2)^2}{(w_2-\varepsilon)^2}$$
$$-\frac{3w_3(w_1+w_2-w_3)+\frac{3w_3^3}{w_1+w_2}+(w_1+w_2)^2}{(w_1+\varepsilon)^2}$$
$$=8-\frac{2(w_1+w_2)^2}{(w_2-\varepsilon)^2}+A\left(\frac{1}{(w_2-\varepsilon)^2}-\frac{1}{(w_1+\varepsilon)^2}\right),$$

where $A=3w_3(1-2w_3)+\frac{3w_3^3}{1-w_3}+(1-w_3)^2$. We show that

$$A\frac{(w_1-w_2+2\varepsilon)(w_1+w_2)}{(w_1+\varepsilon)^2(w_2-\varepsilon)^2}\geq 2\left[\left(\frac{w_1+\varepsilon}{w_2-\varepsilon}+1\right)^2-2^2\right]=\frac{2(w_1-w_2+2\varepsilon)(w_1+3w_2-2\varepsilon)}{(w_2-\varepsilon)^2},$$

or, equivalently,

$$2(w_2-\varepsilon)^2(3(1-w_3)-2(w_2-\varepsilon))\geq 1-6w_3+12w_3^2-10w_3^3.$$

The expression on the left-hand side is not smaller than $4(\frac{1}{2}-w_3)^2(1-w_3)$ since $w_2-\varepsilon\geq\frac{1}{2}-w_3$, $2(w_2-\varepsilon)\leq 1-w_3$. But $4(\frac{1}{2}-w_3)^2(1-w_3)=1-5w_3+8w_3^2-4w_3^3\geq 1-6w_3+12w_3^2-10w_3^3$, as $(1-5w_3+8w_3^2-4w_3^3)-(1-6w_3+12w_3^2-10w_3^3)=w_3(1-4w_3+6w_3^2)=w_3((1-2w_3)^2+2w_3^2)\geq 0$.

C The upper bound on the probability for EC

Again, assume $w_1\geq w_2\geq w_3$. In part b) of Appendix B we have established that the function $P(w_1+\varepsilon,w_2-\varepsilon,w_3)$ is increasing of ε in $0\leq\varepsilon\leq w_2-w_3$ for **EC** if $w_1+2\varepsilon\leq w_2+w_3$. For **EC**, we show that if $w_1+2\varepsilon>w_2+w_3$, the function $P(w_1+\varepsilon,w_2-\varepsilon,w_3)$ is decreasing in ε for $0\leq\varepsilon\leq w_2-w_3$. Indeed,

$$P(w_1+\varepsilon,w_2-\varepsilon,w_3)=\frac{w_2-\varepsilon}{4(w_1+\varepsilon)}+\frac{w_3^2}{12(w_1+\varepsilon)(w_2-\varepsilon)}$$

implies

$$\frac{dP(w_1+\varepsilon,w_2-\varepsilon,w_3)}{d\varepsilon}=\frac{(w_1-w_2+2\varepsilon)w_3^2-3(w_1+w_2)(w_2-\varepsilon)^2}{12(w_1+\varepsilon)^2(w_2-\varepsilon)^2}.$$

Denote $X=w_2-\varepsilon$. Let us show that $3(w_1+w_2)X^2+2w_3^2X-(w_1+w_2)w_3^2>0$ for $X\in[w_3,w_2]$. The roots $X_{1,2}$ of the quadratic are

$$X_{1,2}=\frac{-w_3^2\pm\sqrt{w_3^4+3(w_1+w_2)^2w_3^2}}{3(w_1+w_2)}$$

with the largest root being smaller than w_3. It follows that $w_1^* = \frac{1}{2}$ maximizes $P(w_1, w_2, w_3)$.

The function

$$P\left(\frac{1}{2}, w_2, \frac{1}{2} - w_2\right) = \frac{w_2}{2} + \frac{(\frac{1}{2} - w_2)^2}{6w_2}, \quad \frac{1}{4} \leq w_2 \leq \frac{1}{2}$$

is increasing in w_2, and so it is maximized for $w_2^* = \frac{1}{2}$, because

$$\frac{dP(\frac{1}{2}, w_2, \frac{1}{2} - w_2)}{dw_2} = \frac{2}{3} - \frac{1}{24w_2^2} = \frac{2}{3}\left(1 - \frac{1}{16w_2^2}\right) > 0 \text{ for } w_2 > \frac{1}{4}.$$

Consequently, $\underset{(w_1,w_2,w_3)}{\arg\max}\ P(w_1, w_2, w_3) = (\frac{1}{2}, \frac{1}{2}, 0)$ and $\underset{(w_1,w_2,w_3)}{\max}\ P(w_1, w_2, w_3) = \frac{1}{4}$.