Asymptotic probability of majority inversion under a general binomial voting model

Serguei Kaniovski * Alexander Zaigraev †‡

September 17, 2021

Abstract

This paper characterizes the limit values of the probability of majority inversion when the number of voters tends to infinity, assuming a binomial model specific to each state, states of different population sizes and arbitrary voting quotas in both stages of the voting procedure. The main asymptotic theorem provides the limit values for most parametrizations of the model. A prominent special case in which the limit cannot be determined using the theorem is the classical binomial model. To address this well-known special case, we provide an exact expression for the inversion probability that applies to any parametrizations of the model.

Key Words: two-stage voting, probability of majority inversion, binomial voting model

*✉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.umk.pl

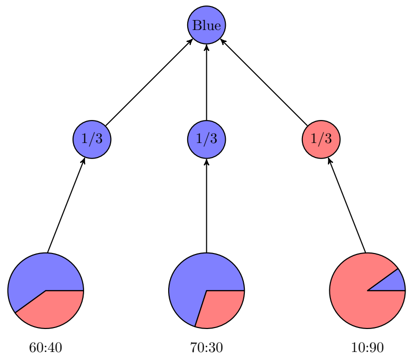
‡**Acknowledgment**: We thank two anonymous reviewers for their careful examination of the manuscript and their critical comments. All other errors are our own responsibility.

1 Introduction

In a two-stage election, voters apportioned among constituencies (states) determine the overall outcome of the election only indirectly. The following analysis is limited to the case of two-way or binary voting, where each voter casts his or her vote in favor of one of the two alternatives. In the first stage, a majority in each state is determined. In the second stage, the alternative supported by a majority of the states wins the election. A *majority inversion* occurs when the majority of the states in the second stage contradicts a nationwide majority of the votes in the first stage. The phenomenon of majority inversion attracts considerable attention in the political science literature and in the theory of voting, as the prospect of majority inversion undermines the democratic legitimacy of an election outcome.

To illustrate the phenomenon of majority inversion, consider the example of three states in Figure 1.[1] The votes are unweighted and the collective decision rule is simple majority. A total turnout of 300 voters is evenly distributed among the three states. Blue wins the election with a majority of states, even though Red would win a hypothetical direct nationwide vote, as the Blue faction totaling 140 votes is smaller then the Red faction totaling 160 votes.

Figure 1: Majority inversion with three states



Blue wins the election by a 2/3 majority of states, while receiving only 140 of 300 votes. This leads to a majority inversion, because Blue is not backed by a nationwide majority.

The probability of majority inversion in an idealized voting model depends on the voting procedure and the stochastic model that determines the voters' choice between the two alternatives. The two stochastic settings dominate the theoretical literature: the Impartial Culture (IC) and the Impartial Anonymous Culture (IAC). The IC setting relies on what we call the *classical binomial model* of voting. In this model, each vote is equally likely to support or oppose an alternative on the ballot, and all votes are independent. The model implies the same level of support for the two alternatives. The probability of majority inversion under the classical binomial model has been studied, for example, in Feix, Lepelley, Merlin and Rouet (2004) and Lepelley, Merlin and Rouet (2011). The main conceptual difference between the IC and the IAC is that the latter admits different levels of support, provided that all conceivable levels of support are equally likely. The level of support is expressed by an absolute number of supporters, which becomes a fraction of supporters asymptotically, as in the pioneering asymptotic analysis by May (1948) and, more recently, in Feix et al. (2004) and Lepelley et al. (2011). The underlying stochastic

[1]This figure is borrowed from Zaigraev and Kaniovski (2020).

model assumes that each vote has an equal probability of being in favor of an alternative, and that this probability is drawn from the uniform distribution on a unit interval. Having assigned the common probability of an affirmative vote, voters cast their votes independently. The model behind the IAC is a special case of a more general compound beta-binomial model, because the continuous uniform distribution from which the common probability parameter is drawn is a special case of a more general beta distribution. Drawing a common probability parameter for the uniform distribution introduces a positive correlation of 1/3 between the affirmative votes, even though the votes are independent, conditioned on the common parameter. The recent paper by De Mouzon, Laurent, Le Breton and Lepelley (2020) shows that the two settings can be combined to obtain an even richer setting in which the IAC assumption holds within each state, but any two votes from different states remain independent, and investigate the bounds on the probability of inversion in this setting. Three simplifying assumptions common to all of the above studies are i) states of equal population size, ii) simple majority rule in the first stage of the two-stage voting procedure, and iii) unweighted votes in the second stage. Each of these assumptions greatly simplifies the combinatorial aspect of the probability calculations.

The IC and IAC dichotomy is rooted in the literature on voting power, where it serves as a probabilistic foundation for the Penrose-Banzhaf and the Shapley-Shubik measures of power (Straffin 1977). According to Morriss (2002) and Felsenthal and Machover (2004), a power measure should reflect *a priori* voting power, or power granted by a set of decision-making rules, rather than the preferences or the behavior of the voters.[2] This perspective favors simple stochastic models that should be agnostic and neutral with respect to implied voting behavior. The stochastic model considered in this paper introduces a minimal departure from an overly restrictive interpretation of the *a priori* perspective embodied by the classical binomial model. It does so by allowing voters from different states to have different probabilities of supporting the alternative or candidate on the ballot.

In the recent theoretical literature on the probability of majority inversions, exact calculations and estimates of the inversion probability under the two alternative settings often go hand in hand, not least due to the similarity of the analytical methods used to perform them. Yet the IC setting is often considered too contrived and consequently less interesting than the IAC setting because the underlying classical binomial model is extremely rigid and essentially assumes that every voter in every state tosses a fair coin. In this paper we show that asymptotic estimates of the inversion probability can be performed for a richer binomial setting in which different binomial models are assumed for different states. This brings the proposed binomial setting closer to the IAC setting, the difference being that in the former different probability parameters are assumed for each state, whereas in the IAC these parameters are drawn from a uniform distribution. Other generalizations pursued in this paper include relaxing the assumption that states have the same population size and that the decision rule is a simple majority. The present framework thus assumes a binomial model specific to each state, where states can differ in size and the collective decision rule allows for an arbitrary voting quota in both stages of the two-stage voting procedure.

In the next section, we outline the scope of this paper in terms of the voting procedure and the stochastic voting model considered in it. Before proceeding with the main asymptotic result, we pause to state some nonasymptotic results for the inversion probability in Section 3, which

[2]For a debate on the role of preferences in the measurement of voting power, the reader is referred to Napel and Widgrén (2004) and the critique by Braham and Holler (2005a), as well as the reply by Napel and Widgrén (2005) and the rejoinder by Braham and Holler (2005b).

will help to frame the asymptotic analysis in Section 4. A detailed proof of the main theorem can be found in Appendix A. The final section briefly summarizes the main findings.

2 The model

2.1 The voting procedure

The voting procedure is defined by the number of states, the number of voters in each state, the voting weights and voting quotas. We make the following assumptions:

A.1 the number of states s is fixed, the number of voters in each state is $n_1, \ldots, n_s$ and their overall number is $N = \sum_{i=1}^{s} n_i$;

A.2 in the asymptotic analysis, the total number of voters tends to infinity, $N \to \infty$, such that the population shares of the states converge to constant values: $n_1/N \to c_1, \ldots, n_s/N \to c_s$, where $c_i \in [0, 1]$ for all i and $\sum_{i=1}^{s} c_i = 1$;

A.3 the votes in both stages of the voting procedure are not weighted;

A.4 the voting quota in the first stage is given by $\alpha \in (0, 1)$, whereas the voting quota in the second stage is given by $\beta \in (0, 1)$;

A.5 $s\beta$, $N\alpha$ and $n_i\alpha$ for all $i = 1, \ldots, s$ are not integers;

A.6 we do not distinguish between turnout and population, thus ignoring the possibility that not every resident of a state is eligible to vote and that some eligible voters may abstain.

Let us briefly discuss each of the above assumptions. Assumptions [A.1] and [A.2] convey the asymptotic nature of our analysis. The paper investigates the limiting behavior of the inversion probability, which will be defined shortly, when the population distributed among a fixed number of states tends to infinity so that the proportions of states in the total population stabilize. The limit population shares are not necessarily equal. Assumptions [A.3] and [A.6] are fairly common to the existing theoretical studies of two-stage voting models. Assumption [A.6] equates turnout and population. This simplification is permissible if turnout is treated as exogenous, allowing us to use the terms "number of voters", "voter turnout" and "population" interchangeably. Assumption [A.4] allows us to see the effect of arbitrary quotas on the inversion probability. We differentiate between the quotas in the first and the second stages of the voting procedure, making the analysis fully general with respect to the voting rule. The value $\alpha = 0.5$ ($\beta = 0.5$) implies a *simple majority* and $\alpha > 0.5$ ($\beta > 0.5$) implies a *qualified majority* in the first (second) stage. Most existing theoretical studies assume simple majority rule. In the case of a qualified majority, the quota can be interpreted as the level of support needed to bring about a change in the status quo, such as the adoption of a policy or a constitutional amendment in a nationwide referendum. Assumption [A.5] is a common technical assumption that rules out ties in both stages of the voting procedure. This assumption is not particularly restrictive in an asymptotic analysis, but it removes the need to provide a tie-breaking rule.

Note that U.S. presidential elections do not exactly follow the above model. First, they are not two-way elections because there are usually more than two candidates running for office and not all candidates run in all states. Nonetheless, most theoretical and empirical models hold U.S. presidential elections to be what they have essentially always been: a contest between Democrats and Republicans, i.e., a binary election. Regarding the assumption of unweighted votes [A.3], note that real two-stage procedures can feature weighted votes in the second stage, with the weights typically reflecting the size of the population in each state, as is the case with the U.S. Electoral College. The assumption of weighted votes precludes an analytical analysis of the voting model for a large number of states, due to the resulting combinatorial complexity. This is the reason why estimates of the inversion probability and calculations of voting power for general weighted voting models rely on numerical methods. Explicit nonasymptotic results for several weighting schemes in the case of three states can be found in Lepelley, Merlin, Rouet and Vidu (2014), Kaniovski and Zaigraev (2018) and Zaigraev and Kaniovski (2020). A recent example of a numerical analysis of different weighting schemes in the context of inversion probability can be found in Feix, Lepelley, Merlin, Rouet and Vidu (2021). Their numerical analysis includes several stochastic voting models and weighting schemes, while maintaining simple majority rule as the universal decision rule. Empirical models of the U.S. presidential elections can be found in Katz, Gelman and King (2004), Miller (2012), who also documents inversions elsewhere around the world, and Geruso, Spears and Talesara (2019). Despite its apparent simplifications, the idealized formal model of two-stage voting, such as the one described above, has attracted considerable attention in the theory of voting because it allows for the study of the phenomenon of majority inversion in an analytically tractable framework.

2.2 The binomial model

The stochastic model assumes that each voter either supports or opposes the alternative on the ballot. Let $n_i^+ \in [0, n_i]$ be the number of *proponents* and $n_i^- = n_i - n_i^+$ the number of *opponents* in state i, then

A.7 all votes in the country are *stochastically independent* binary random variables;

A.8 the number of proponents n_i^+ follows a *binomial distribution* $\mathcal{B}(n_i, p_i)$, whose probability $p_i \in (0, 1)$ is specific to the state i.

Assumption [A.8] specifies the probability distribution of the sizes of factions $n_i^+ \sim \mathcal{B}(n_i, p_i)$ and $n_i^- \sim \mathcal{B}(n_i, 1 - p_i)$ as factions who support the same alternative in a given state. The total sizes of the two factions $N^+ = \sum_{i=1}^s n_i^+$ and $N^- = \sum_{i=1}^s n_i^-$ comprise the total population $N = N^- + N^+$. At the level of individual votes, [A.7] and [A.8] are equivalent to all affirmative votes in state i having the same probability and all votes in the country being independent. The assumption of states having different probabilities p_i generalizes the classical binomial model of voting with $p = 0.5$ assumed for each state.

2.3 Probability of majority inversion

The probability of majority inversion is the focus of this paper. Majority inversions occur when the nationwide majority in the first stage does not match the majority of states in the second stage. Let P_1 be the probability that proponents have a nationwide qualified majority of votes but not a qualified majority of states, and P_2 be the probability that proponents do not have a nationwide qualified majority of votes but have a qualified majority of states. Since the two scenarios are mutually exclusive events, the inversion probability is equal to the sum $P = P_1 + P_2$, where

$$P_1 = P(N^+ > N\alpha, \text{ at least } s - \lfloor s\beta \rfloor \text{ of } \{n_i^+ - n_i\alpha\} \text{ are negative}),$$
$$P_2 = P(N^+ < N\alpha, \text{ at least } \lfloor s\beta \rfloor + 1 \text{ of } \{n_i^+ - n_i\alpha\} \text{ are positive}).$$

Here, $\lfloor \cdot \rfloor$ denotes the integer part of the argument (floor function).

The main result allows us to determine the limit value of the inversion probability for most, but not all, parametrizations. A parametrization of an election scenario comprises: a vector of limit population shares of the states $\mathbf{c} = (c_1, \ldots, c_s)$, a vector of binomial probability parameters $\mathbf{p} = (p_1, \ldots, p_s)$ and a pair of voting quotas $\alpha, \beta \in (0, 1)$. The main discriminatory quantity with respect to the limit value

$$\lim_{N \to \infty} P(\mathbf{p}, \alpha, \beta)$$

is the scalar product $\langle \mathbf{c}, \mathbf{p} \rangle$, leading to two principal cases $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$ and $\langle \mathbf{c}, \mathbf{p} \rangle \neq \alpha$. The first case is an *knife-edge* case corresponding to a perfect split of popular support for the two alternatives on the ballot, while the second case is an *generic* case covering a richer variety of voting scenarios and model parametrizations. The asymptotic theorem provides a complete characterization of the limit values for the generic case $\langle \mathbf{c}, \mathbf{p} \rangle \neq \alpha$.

For the special case $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$, the asymptotic theorem can only confirm that the limit value is positive, provided that $p_i = \alpha$ for all i, or all c_i are smaller than 1. Exact limit values of the inversion probability for $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$ have so far only been established for the classical binomial model ($p_1, \ldots, p_s = 0.5$), states of equal population size ($n_1, \ldots, n_s = n$) and simple majority rules ($\alpha, \beta = 0.5$). For this setting, Feix et al. (2004) provide the exact limit values for $s = 3, 4, 5$, as well as a numerically simulated limit value for $s > 5$. In a follow-up paper, Lepelley et al. (2011) present an approximate expression for the limit value of the inversion probability when n and s are large, retaining the assumption of equal population size and simple majority rules. The main idea behind the approximation in Lepelley et al. (2011) is that for independently and identically distributed votes the distribution of a normalized margin of victory in each state tends to a normal law. The limit value for the case of infinitely many states approximated by simulation in Feix et al. (2004) has been confirmed analytically by Kikuchi (2016), who also provides the expression:

$$\frac{1}{2} - \frac{1}{\pi} \cdot \arctan\left(\sqrt{\frac{2}{\pi - 2}}\right) \approx 0.2059524.$$

The above limit value is again confirmed numerically in Feix et al. (2021) for up to $s = 50$, who also relax the assumption of equally-sized states and study the effect of different weighting schemes on the probability of majority inversion, while preserving simple majority rules. The bottom line from these studies, however, is that the knife-edge case $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$ does not appear to be amenable to analytic methods when the states differ in population size, even for unweighted voting and simple majority rules. For a small number of voters and states the inversion probability for unweighted votes can be calculated using an exact formula provided in the next section.

3 Nonasymptotic probability of majority inversion

For a pair of integers a, b, let $\binom{a}{b} = a!/[b!(a-b)!]$ be the binomial coefficient, if $a \geqslant b \geqslant 0$, and $\binom{a}{b} = 0$, otherwise. The following additive decomposition of the inversion probability $P(\mathbf{p}, \alpha, \beta)$

can be useful for obtaining theoretical results and in computations:

$$P(\mathbf{p},\alpha,\beta) = P_1(\mathbf{p},\alpha,\beta) + P_2(\mathbf{p},\alpha,\beta) = \sum_{k=s-\lfloor s\beta\rfloor}^{s-1} \sum P_1(i_1,\ldots,i_k) + \sum_{k=\lfloor s\beta\rfloor+1}^{s-1} \sum P_2(i_1,\ldots,i_k), \tag{1}$$

where for a fixed k each inner sum has $\binom{s}{k}$ summands and

$$\begin{aligned} P_1(i_1,\ldots,i_k) &= P\Big(N^+ > N\alpha, n^+_{i_1} < n_{i_1}\alpha, \ldots, n^+_{i_k} < n_{i_k}\alpha, n^+_{i_{k+1}} > n_{i_{k+1}}\alpha, \ldots, n^+_{i_s} > n_{i_s}\alpha\Big), \\ P_2(i_1,\ldots,i_k) &= P\Big(N^+ < N\alpha, n^+_{i_1} > n_{i_1}\alpha, \ldots, n^+_{i_k} > n_{i_k}\alpha, n^+_{i_{k+1}} < n_{i_{k+1}}\alpha, \ldots, n^+_{i_s} < n_{i_s}\alpha\Big). \end{aligned}$$

In P_1, $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant s$ indicate the states in which opponents win, while in P_2 they indicate the states in which proponents win. The number of summands in P_1 does not decrease with β, while the number of summands in P_2 does not increase with β. Since all summands are non-negative, $P_1(\mathbf{p},\alpha,\beta)$ is non-decreasing in β, whereas $P_2(\mathbf{p},\alpha,\beta)$ is non-increasing in β and vanishes for $\beta > \frac{s-1}{s}$. For $\beta > \frac{s-1}{s}$, we thus have

$$P(\mathbf{p},\alpha,\beta) = P\left(N^+ > N\alpha\right) - \prod_{i=1}^{s} P\left(n_i^+ > n_i\alpha\right).$$

The above formula states that if β is sufficiently high, the probability P_2 vanishes, while P_1 and thus the inverse probability is equal to the probability that proponents win in the first stage of the election, minus the probability that proponents win in each state.

The next expression for $P(\mathbf{p},\alpha,\beta)$ follows directly from [A.7] and [A.8]:

Exact Formula. *The probability of majority inversion $P(\mathbf{p},\alpha,\beta)$ equals the sum of two terms:*

$$\begin{aligned} P_1(\mathbf{p},\alpha,\beta) &= \sum_{k=s-\lfloor s\beta\rfloor}^{s-1} \sum_{t=\lfloor N\alpha\rfloor+1}^{t^*(k)} \sum_{\{k_i\}\in A_{t,k}} \prod_{j=1}^{s} \binom{n_j}{k_j} p_j^{k_j}(1-p_j)^{n_j-k_j}, \\ P_2(\mathbf{p},\alpha,\beta) &= \sum_{k=\lfloor s\beta\rfloor+1}^{s-1} \sum_{t=t_*(k)}^{\lfloor N\alpha\rfloor} \sum_{\{k_i\}\in A'_{t,k}} \prod_{j=1}^{s} \binom{n_j}{k_j} p_j^{k_j}(1-p_j)^{n_j-k_j}, \end{aligned}$$

where

$$\begin{aligned} A_{t,k} &= \{k_1+k_2+\cdots+k_s = t, k \textit{ integers in } [0,n_i\alpha), s-k \textit{ integers in } (n_i\alpha, n_i]\}, \\ A'_{t,k} &= \{k_1+k_2+\cdots+k_s = t, k \textit{ integers in } (n_i\alpha, n_i], s-k \textit{ integers in } [0,n_i\alpha)\}, \end{aligned}$$

and, for $n_1 \leqslant n_2 \leqslant \cdots \leqslant n_s$,

$$\begin{aligned} t^*(k) &= \lfloor n_1\alpha\rfloor + \cdots + \lfloor n_k\alpha\rfloor + n_{k+1} + \cdots + n_s, \\ t_*(k) &= \lfloor n_1\alpha\rfloor + \cdots + \lfloor n_k\alpha\rfloor + k. \end{aligned}$$

The complexity of computing the inversion probability using the above formula results from the need to determine all elements of the sets $A_{t,k}$ and $A'_{t,k}$ as the integers lying on the hyperplanes defined by $k_1 + k_2 + \cdots + k_s = t$. The summation bounds $t_*(k)$ and $t^*(k)$ reduce the number of summands in $P_1(\mathbf{p},\alpha,\beta)$ and $P_2(\mathbf{p},\alpha,\beta)$ by omitting those values of $\{k_i\}$ for which inversions cannot occur.

For $\beta = 0.5$, the expression for the inversion probability possesses the following symmetry:

Corollary 1. *$P_1(\mathbf{p}, \alpha, 0.5) = P_2(\mathbf{1_s} - \mathbf{p}, 1-\alpha, 0.5)$ and $P_2(\mathbf{p}, \alpha, 0.5) = P_1(\mathbf{1_s} - \mathbf{p}, 1-\alpha, 0.5)$ and, consequently, $P(\mathbf{p}, \alpha, 0.5) = P(\mathbf{1_s} - \mathbf{p}, 1-\alpha, 0.5)$, where $\mathbf{1_s}$ is a unit vector of length s.*

Indeed, if n_i^+ has the distribution $\mathcal{B}(n_i, p_i)$, then $n_i - n_i^+$ has the distribution $\mathcal{B}(n_i, 1-p_i)$. To prove $P_2(\mathbf{p}, \alpha, 0.5) = P_1(\mathbf{1_s} - \mathbf{p}, 1-\alpha, 0.5)$, note that

$$\begin{aligned}
P_2(\mathbf{p}, \alpha, 0.5) &= P(N^+ < N\alpha,\ \text{at least } (s+1)/2 \text{ of } \{n_i^+ - n_i\alpha\} \text{ are positive}),\\
&= P(N^+ - N < N\alpha - N,\ \text{at least } (s+1)/2 \text{ of } \{n_i^+ - n_i + n_i(1-\alpha)\} \text{ are positive}),\\
&= P(N - N^+ > N(1-\alpha),\ \text{at least } (s+1)/2 \text{ of } \{(n_i - n_i^+) - n_i(1-\alpha)\} \text{ are negative}),\\
&= P_1(\mathbf{1_s} - \mathbf{p}, 1-\alpha, 0.5).
\end{aligned}$$

The equality $P_1(\mathbf{p}, \alpha, 0.5) = P_2(\mathbf{1_s} - \mathbf{p}, 1-\alpha, 0.5)$ can be proven in a similar manner.

The above symmetry can reduce the computational complexity of the inversion probability, especially in the case of the classical binomial model, where $P(0.5 \cdot \mathbf{1_s}, 0.5, 0.5) = 2P_1(0.5 \cdot \mathbf{1_s}, 0.5, 0.5) = 2P_2(0.5 \cdot \mathbf{1_s}, 0.5, 0.5)$. It can be further reduced by assuming that all states have the same number of voters.

Corollary 2. *Let n be the population size of each state. For $n_i^+ \sim \mathcal{B}(n, 0.5)$ for all $i = 1, \dots, s$ and $\alpha, \beta = 0.5$, we have*

$$P(0.5 \cdot \mathbf{1_s}, 0.5, 0.5) = 2^{1-sn} \sum_{k=(s+1)/2}^{s-1} \binom{s}{k} \sum_{t=(sn+1)/2}^{sn-k(n+1)/2} \sum_{\{k_i\} \in A_{t,k}} \binom{n}{k_1}\binom{n}{k_2} \cdot \dots \cdot \binom{n}{k_s},$$

where

$$A_{t,k} = \{k_1 + \dots + k_s = t,\ \textit{the first } k \textit{ integers in } [0, (n-1)/2], s-k \textit{ integers in } [(n+1)/2, n]\}.$$

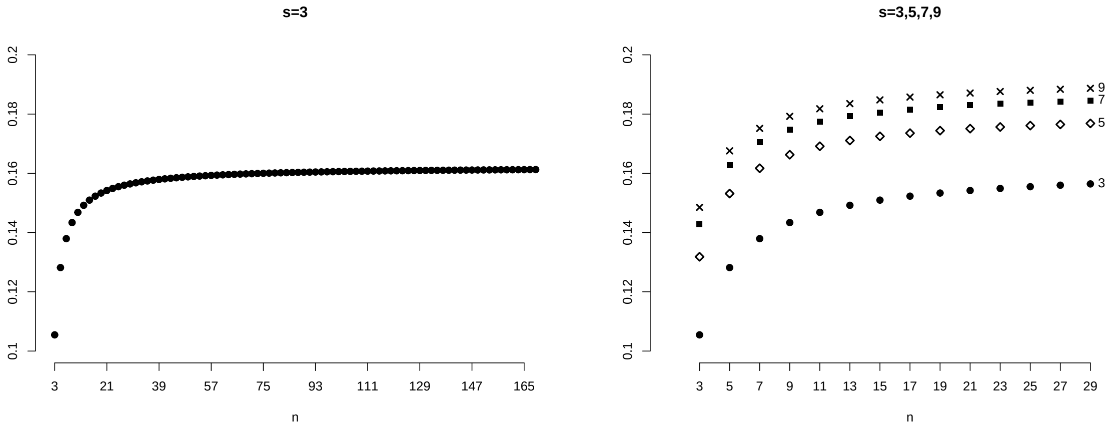
Figure 2 illustrates the inversion probability given in Corollary 2. The probability appears to converge as $n \to \infty$, where $n = N/s$, for a fixed s (left panel), for $s \to \infty$ for a fixed n, and for $n \to \infty$ and $s \to \infty$ (right panel). The scenario $n \to \infty$ for a fixed s is the most relevant of the three asymptotic scenarios because it can serve as an approximation of real-world electoral systems with a large electorate residing in a given number of states. The computational complexity precludes the practical application of the exact formula unless the number of states and, especially, the number of voters are unrealistically small. This motivates the following asymptotic analysis, the goal of which is to derive practical approximations to the inversion probability for a wide range of model parametrizations. In the next section we analyze the set of feasible limit values for the inversion probability under a more general binomial model in which p is specific to each state and the states can differ in population size.

4 Asymptotic probability of majority inversion

Let c_i be the limit population share of state i in the total population, $i = 1, \dots, s$, such that $\sum_{i=1}^s c_i = 1$. The limit shares need not be equal and can include the borderline case of a single state absorbing the entire population. Next, we introduce $\mathbf{m}$, the vector of *election margins* in each state:

Definition. *For a vector $\mathbf{m}$ with elements $m_i = \alpha - p_i, i = 1, \dots, s$, denote the number of negative, zero and positive elements in $\mathbf{m}$ by m^-, m^0 and m^+, respectively.*

Figure 2: Inversion probability for equally-sized states



The figure illustrates the inversion probability under the classical binomial model, simple majority rule and states of equal population size n given in Corollary 2. The probability appears to converge as $n \to \infty$ for a fixed number of states s, for $s \to \infty$ for a fixed n, as well as for $n \to \infty$ and $s \to \infty$.

Since we are dealing with a partition of $\mathbf{m}$, $m^- + m^0 + m^+ = s$, where s is the number of elements in $\mathbf{m}$, or the number of states. The vector of margins will play a prominent role in the following analysis. Following the logic of the model, we will say that the election outcome in state i is close if the i'th element of $\mathbf{m}$ is equal to 0, that is $p_i = \alpha$. We will call such states *close* states and the other states *partisan* states.

Asymptotic Theorem. *For the limit of the probability of inversion it holds that:*

a) if $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$, *then*

$$\lim_{N\to\infty} P(\mathbf{p}, \alpha, \beta) = \begin{cases} 2^{-m^0} \sum_{k=0}^{s-\lfloor s\beta \rfloor -1-m^+} \binom{m^0}{k}, \; \textit{for } 1 \leqslant m^+ \leqslant s - \lfloor s\beta \rfloor - 1 \\ 0, \; \textit{for } s - \lfloor s\beta \rfloor \leqslant m^+ \leqslant s; \end{cases} \tag{2}$$

b) if $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$, *then*

$$\lim_{N\to\infty} P(\mathbf{p}, \alpha, \beta) = \begin{cases} 2^{-m^0} \sum_{k=0}^{\lfloor s\beta \rfloor -m^-} \binom{m^0}{k}, \; \textit{for } 1 \leqslant m^- \leqslant \lfloor s\beta \rfloor \\ 0, \; \textit{for } \lfloor s\beta \rfloor + 1 \leqslant m^- \leqslant s; \end{cases} \tag{3}$$

c) if $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$, *then the limit is positive, unless* $\mathbf{m} \neq \mathbf{0_s}$, *where* $\mathbf{0_s}$ *denotes a zero vector of length* s, *and* $c_i = 1$ *for some* i, *in which case the limit can be 0.*

A proof of the theorem is given in Appendix A.

An immediate consequence of the above theorem is the following result:

Corollary 3. $\lim_{N\to\infty} P(\mathbf{p}, \alpha, \beta) = 0$ *for* $\alpha < \min\{p_i\}$ *or* $\alpha > \max\{p_i\}$.

Choosing a first-stage quota α outside of the range of p_i's leads to an asymptotic impossibility of majority inversion. Consequently, the inversion probability tends to zero if $\alpha \to 1$.

4.1 Limit values for $\langle \mathbf{c}, \mathbf{p} \rangle \neq \alpha$

The Asymptotic Theorem gives the limit of the inversion probability for all cases, except for the knife-edge case of $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$. Let us discuss the generic case $\langle \mathbf{c}, \mathbf{p} \rangle \neq \alpha$, which is the focus of this paper. Determining the limit in this case requires checking several conditions. The first step checks if the inversion probability asymptotically vanishes. Depending on whether $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$ or $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$, this occurs if $m^+ \geqslant s - \lfloor s\beta \rfloor$ or $m^- \geqslant \lfloor s\beta \rfloor + 1$. In view of the definition of $\mathbf{m}$ as a vector with elements $\alpha - p_i$, this happens because the nationwide majority and the majority of states agree. For all other parametrizations, the inversion probability converges to a positive value given by either (2) or (3). This value can be found by plugging m^0 and, depending on whether $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$ or $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$, either m^+ or m^-. For $\beta = 0.5$, plugging m^+ or m^- yields the same set of feasible limit values.

Table 1: Feasible limit values for $s = 3, 5$, $\beta = 0.5$ and $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$.

$s = 3$				$s = 5$					
$m^0 \backslash m^+$	1	2	3	$m^0 \backslash m^+$	1	2	3	4	5
0	1	0	0	0	1	1	0	0	0
1	1/2	0		1	1	1/2	0	0	
2	1/4			2	3/4	1/4	0		
				3	1/2	1/8			
				4	5/16				

The limit values can be collected in a table of size $s \times s$, whose entries depend on the number of states s and the voting quota β in the second stage of the voting procedure. Table 1 shows the limit values for $s = 3, 5$ and $\beta = 0.5$. Similarly, Table 2 shows the limit values for $s = 9$ and $\beta = 0.5$. We compute the entries in Tables 1 and 2 by plugging the values of m^0 and m^+ into the asymptotic formula (2), which corresponds to the case $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$. The tables for the case $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$ can be obtained by plugging the values of m^0 and m^- into the asymptotic formula (3), yielding the same entries since $\beta = 0.5$. Table 3 shows the effect of setting $\beta = 0.75$ on the set of feasible limit values for $s = 9$. Since $\beta \neq 0.5$, the set of feasible limit values will differ depending on whether $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$ or $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$, leading to two sub-tables.

The fact that the asymptotic theorem can be applied to any parametrization for any number of states makes it difficult to systematize all possibilities. To gain insight into the dispersion of feasible limit values under a variety of parametrizations, we created a complete set of tables for an odd number of states ranging from $s = 3$ to $s = 101$ and simple majority rule in the second stage of the voting procedure $\beta = 0.5$, collecting the absolute frequencies of the limit values in ten intervals.

The resulting absolute frequencies are presented in Appendix B. A comparison of the values found in Tables 1-2 with the corresponding rows of the frequency table in the appendix shows that the frequencies double. For example, the right side of Table 1 for $s = 5$ contains 1/8 as the sole limit belonging to the range $(0.1, 0.2]$, but the corresponding entry in the second (data) row of the frequency table reports two limits in this range. This happens due to the existence of two symmetric cases when $\beta = 0.5$, only one of which ($\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$) is shown in Tables 1-2. We added the cases in which the asymptotic inversion probability takes the value of 0 and 1 to the first and the last column of the frequency table, respectively.

Table 2: Feasible limit values for $s = 9$, $\beta = 0.5$ and $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$.

m^0 \ m^+	1	2	3	4	5	6	7	8	9
0	1	1	1	1	0	0	0	0	0
1	1	1	1	1/2	0	0	0	0	
2	1	1	3/4	1/4	0	0	0		
3	1	7/8	1/2	1/8	0	0			
4	15/16	11/16	5/16	1/16	0				
5	13/16	1/2	3/16	1/32					
6	21/32	11/32	7/64						
7	1/2	29/128							
8	93/256								

Table 3: Feasible limit values for $s = 9$ and $\beta = 0.75$.

$\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$									
m^0 \ m^+	1	2	3	4	5	6	7	8	9
0	1	1	0	0	0	0	0	0	0
1	1	1/2	0	0	0	0	0	0	
2	3/4	1/4	0	0	0	0	0		
3	1/2	1/8	0	0	0	0			
4	5/16	1/16	0	0	0				
5	3/16	1/32	0	0					
6	7/64	1/64	0						
7	1/16	1/128							
8	9/256								
$\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$									
m^0 \ m^-	1	2	3	4	5	6	7	8	9
0	1	1	1	1	1	1	0	0	0
1	1	1	1	1	1	1/2	0	0	
2	1	1	1	1	3/4	1/4	0		
3	1	1	1	7/8	1/2	1/8			
4	1	1	15/16	11/16	5/16				
5	1	31/32	13/16	1/2					
6	63/64	57/64	21/32						
7	15/16	99/128							
8	219/256								

The general formula for the frequency of boundary and interior cases of the asymptotic inversion probability is provided by the following corollary. Let L be the set of all feasible limits, including their multiples, for a given s and β. In other words, L is the set of all entries in a table. For example, in the case $s = 3$ and $\beta = 0.5$, the set $L = \{0, 0, 0, 0, 0, 0, 1/4, 1/4, 1/2, 1/2, 1, 1\}$.

Half of the elements of L can be found in Table 1, the other half being identical for $\beta = 0.5$. We obtain the following corollary of Asymptotic Theorem:

Corollary 4. *For $\langle \mathbf{c}, \mathbf{p} \rangle \neq \alpha$, the frequency of feasible limit values l equals*

$$\begin{aligned}
\#\{l \in L,\ \textit{such that } l = 0\} &= \left((s - \lfloor s\beta \rfloor)^2 + (\lfloor s\beta \rfloor + 1)^2 + s + 1\right)/2, \\
\#\{l \in L,\ \textit{such that } 0 < l < 1\} &= 2(\lfloor s\beta \rfloor + 1)(s - \lfloor s\beta \rfloor) - s - 1, \\
\#\{l \in L,\ \textit{such that } l = 1\} &= \left((s - \lfloor s\beta \rfloor)^2 + (\lfloor s\beta \rfloor + 1)^2 - s - 1\right)/2,
\end{aligned}$$

where $\#$ denotes the cardinality of a set.

For $\beta = 0.5$, the three frequencies simplify to $(s+1)(s+3)/4$, $(s^2-1)/2$ and $(s-1)(s+1)/4$, respectively. The validity of the above corollary for $\beta = 0.5$ can be checked using Tables 1-2 or, more directly, using the frequency table supplied in Appendix B. For $\beta = 0.75$, the resulting frequencies should be compared to those in Tables 3.

The absolute frequencies in Appendix B show that the number of feasible limit values increases with the number of states s in all ranges. The frequencies of the boundary limit values of 0 and 1 tend to accumulate fast with s, followed by the cases of limit values close to the boundaries. To a lesser extent, this is also true for limit values close to 0.5. We conclude that, at least for the baseline case of simple majority rule $\beta = 0.5$ in the second stage of the voting procedure, nontrivial limit values remain a relevant phenomenon as s increases. For $\beta > (s-1)/s$, it holds that

$$\begin{aligned}
\#\{l \in L,\ \text{such that } l = 0\} &= s(s+1)/2 + 1, \\
\#\{l \in L,\ \text{such that } 0 < l < 1\} &= s - 1, \\
\#\{l \in L,\ \text{such that } l = 1\} &= s(s-1)/2.
\end{aligned}$$

Since the total number of possible limit values is equal to $s(s+1)$, we can consider the relative frequencies of each of the three types of limit values by dividing the corresponding absolute frequencies in Corollary 4 by $s(s+1)$. Simple calculations show that for $\beta = 0.5$ and $s \to \infty$, the relative frequency of each of the two boundary limit values tends to $1/4$, while the relative frequency of the intermediate limit values tends to $1/2$. For $\beta > (s-1)/s$ and $s \to \infty$, the relative frequency of each boundary limit value tends to $1/2$, but the relative frequency of the intermediate limit values vanishes. In general, for a fixed β and $s \to \infty$, the relative frequency of each boundary limit value becomes $1/2 - \beta(1-\beta)$ and the relative frequency of the intermediate limit values become $2\beta(1-\beta)$.

To round up the discussion, note that in the general case $\langle \mathbf{c}, \mathbf{p} \rangle \neq \alpha$, the absence of close states is a *sufficient* condition for a limit value of 0 or 1. A *necessary* and *sufficient* condition for the limit value to be *different* from 0 or 1 is furnished by a pair of inequalities $m^- \leqslant \lfloor s\beta \rfloor$ and $m^+ \leqslant s - \lfloor s\beta \rfloor - 1$. If all states have a common p, the inversion probability vanishes asymptotically. This result does not depend on the shares, since the condition $\langle \mathbf{c}, \mathbf{p} \rangle \neq \alpha$ simplifies to $p \neq \alpha$ in view of $\sum_{i=1}^{s} c_i = 1$.

4.2 Example of a parametrization for nine states

Let us apply Table 2 to a particular parametrization in the case of nine states ($s = 9$) and simple majority rules $\alpha, \beta = 0.5$ in both stages of the voting procedure. The results are collected in Table 4.

Table 4: Example of a parametrization for $s = 9$.

State	1	2	3	4	5	6	7	8	9
Asymptotic shares $\mathbf{c}$	0.05	0.05	0.05	0.1	0.1	0.15	0.15	0.15	0.2
Probabilities $\mathbf{p}$	0.5	0.2	0.5	0.3	0.4	0.9	0.3	0.5	0.7
Margins $\mathbf{m}$	0.0	0.3	0.0	0.2	0.1	-0.4	0.2	0.0	-0.2
$m^- = 2$						✓			✓
$m^0 = 3$	✓		✓					✓	
$m^+ = 4$		✓		✓	✓		✓		

In this example, $\langle \mathbf{c}, \mathbf{p} \rangle = 0.525$, so that $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$ ($\alpha = 0.5$). The inversion probability is positive, as $m^- = 2$ is smaller than $(s+1)/2 = 5$. The number of close states is given by $m^0 = 3$. Equation (3) yields the limit value of 7/8 (see Table 2). Let us now slightly change the parametrization by lowering the probability for the fifth state from $p_5 = 0.4$ to $p_5 = 0.1$. In this case $\langle \mathbf{c}, \mathbf{p} \rangle = 0.495$, so that $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$. To verify whether the inversion probability is positive, we now have to check if m^+ is smaller than $(s+1)/2$, as is the case because $m^+ = 4$ and $(s+1)/2 = 5$. The number of close states, $m^0 = 3$, is the same as in the previous example. Equation (2) yields a new limit value of 1/8.

The above example illustrates a pattern that can be formulated in the following corollary:

Corollary 5. *Fix the numbers m^-, m^0, m^+ and let $\beta = 0.5$. If the inversion probability for $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$ equals to P, then the inversion probability for $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$ equals to $1 - P$.*

This fact can be easily verified by adding the sums in Equations (2) and (3) of Asymptotic Theorem under the above conditions. Changing from $\langle \mathbf{c}, \mathbf{p} \rangle < \alpha$ to $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$, without changing the number of close states m^0 and the number of partisan states of either type, m^- or m^+, results in the inversion probability changing from P to $1 - P$.

5 Summary

This paper provides the set of feasible limit values of the probability of majority inversion for generic parametrizations of the type $\langle \mathbf{c}, \mathbf{p} \rangle \neq \alpha$. Such parametrizations involve different levels of support embodied in a state-specific binomial voting model, different population sizes, and different voting quotas in the first and second stages of the voting procedure. In the asymptotic setting, the number of voters tends to infinity but the number of states is fixed. This setting is intended to serve as an approximation to real two-stage electoral systems with a large electorate residing in a small number of states.

The application of Asymptotic Theorem is straightforward and requires as input only the number of close and (the two kinds of) partisan states relative to the voting quota in the first stage of the voting procedure and the quota in the second stage. The set of feasible limit values obtained using the theorem can be conveniently presented in a table.

The Asymptotic Theorem does not provide any new tangible results for the knife-edge case of a close election nationwide, where $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$. The most prominent special case in which the limit cannot be determined is the classical binomial model of voting. Here, the theorem only confirms the positivity of the limit probability for all cases where the entire population is not asymptotically concentrated in a single state, or all states are close. Fortunately, the inversion probability under the classical binomial model with equally-sized states has been extensively studied in the existing literature. Our contribution to this important special case includes a new exact expression for the inversion probability that holds for any parametrization of the model. For all other parametrizations such that $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$, the Exact Formula offers a means of approximating the limit value numerically. But applying these nonasymptotic results is computationally intensive.

The combination of Exact Formula and Asymptotic Theorem thus fills an important gap in the literature. By considering a binomial model specific to each state with different-sized states and an arbitrary voting quota in either stage of the two-stage voting procedure, the results provided in this paper complement the exact and approximate results available in the theoretical literature on the probability of majority inversions for the classical binomial model with equally-sized states and simple majority rules. We believe that a comprehensive characterization of the limits leads to a better understanding of the implications of a binomial setting in the context of majority inversions and makes the binomial setting a more attractive analytical assumption than the special case of the classical binomial model may have suggested so far.

References

Braham, M. and Holler, M.: 2005a, Power and preferences again. A reply to Napel and Widgrén, *Journal of Theoretical Politics* **17**, 389–395.

Braham, M. and Holler, M.: 2005b, The impossibility of a preference-based power index, *Journal of Theoretical Politics* **17**, 137–157.

De Mouzon, O., Laurent, T., Le Breton, M. and Lepelley, D.: 2020, The theoretical Shapley–Shubik probability of an election inversion in a toy symmetric version of the US presidential electoral system, *Social Choice and Welfare* **54**, 363–395.

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.

Feix, M. R., Lepelley, D., Merlin, V. R., Rouet, J.-L. and Vidu, L.: 2021, Majority efficient representation of the citizens in a federal union, *in* M. Diss and V. R. Merlin (eds), *Evaluating Voting Systems with Probability Models. Essays by and in Honor of William Gehrlein and Dominique Lepelley. Studies in Choice and Welfare*, Springer International Publishing, pp. 163–187.

Felsenthal, D. S. and Machover, M.: 2004, A priori voting power: What is it all about?, *Political Studies Review* **2**, 1–23.

Geruso, M., Spears, D. and Talesara, I.: 2019, Inversions in US presidential elections, 1836-2016. Working Paper, The University of Texas at Austin.

Kaniovski, S. and Zaigraev, A.: 2018, The probability of majority inversion in a two-stage voting system with three states, *Theory and Decision* **84**, 525–546.

Katz, J. N., Gelman, A. and King, G.: 2004, Empirically evaluating the electoral college, *in* A. N. Crigler, M. R. Just and E. J. McCaffery (eds), *In Rethinking the Vote: The Politics and Prospects of American Election Reform*, Oxford University Press, pp. 75–88.

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

Lepelley, D., Merlin, V. R. and Rouet, J.-L.: 2011, Three ways to compute accurately the probability of the referendum paradox, *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.

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.

Morriss, P.: 2002, *Power: A philosophical analysis*, Manchester University Press.

Napel, S. and Widgrén, M.: 2004, Power measurement as sensitivity analysis: a unified approach, *Journal of Theoretical Politics* **16**, 517–538.

Napel, S. and Widgrén, M.: 2005, The possibility of a preference-based power index, *Journal of Theoretical Politics* **17**, 377–387.

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

Zaigraev, A. and Kaniovski, S.: 2020, Inequality in population weights, majority threshold and the inversion probability in the case of three states, *Operations Research Letters* **48**, 732–736.

A Proof of Asymptotic Theorem

Consider a decomposition of the probability of majority inversion (1), whose summands can be expressed in terms of $\eta_i = \frac{n_i^+ - n_i p_i}{\sqrt{n_i p_i (1-p_i)}}$ and $\mu_i = \frac{m_i}{\sqrt{p_i(1-p_i)}}$ for $i = 1, \ldots, s$ as

$$\begin{aligned}
P_1(i_1, \ldots, i_k) = P\Big(&\sum_{i=1}^{s} \sqrt{n_i p_i (1-p_i)}\eta_i > N\alpha - \sum_{i=1}^{s} n_i p_i, \\
&\eta_{i_1} < \sqrt{n_{i_1}}\mu_{i_1}, \ldots, \eta_{i_k} < \sqrt{n_{i_k}}\mu_{i_k}, \eta_{i_{k+1}} > \sqrt{n_{i_{k+1}}}\mu_{i_{k+1}}, \ldots, \eta_{i_s} > \sqrt{n_{i_s}}\mu_{i_s}\Big), \\
P_2(i_1, \ldots, i_k) = P\Big(&\sum_{i=1}^{s} \sqrt{n_i p_i (1-p_i)}\eta_i < N\alpha - \sum_{i=1}^{s} n_i p_i, \\
&\eta_{i_1} > \sqrt{n_{i_1}}\mu_{i_1}, \ldots, \eta_{i_k} > \sqrt{n_{i_k}}\mu_{i_k}, \eta_{i_{k+1}} < \sqrt{n_{i_{k+1}}}\mu_{i_{k+1}}, \ldots, \eta_{i_s} < \sqrt{n_{i_s}}\mu_{i_s}\Big).
\end{aligned}$$

Assume that $n_1/N \to c_1, \ldots, n_s/N \to c_s$ as $N \to \infty$. Letting the shares of the states in the total population converge as the number of voters tends to infinity, allows us to approximate the above probabilities as

$$\begin{aligned}
P_1(i_1, \ldots, i_k) \approx P\Big(&\sum_{i=1}^{s} \sqrt{c_i p_i (1-p_i)}\eta_i > \sqrt{N}(\alpha - \langle \mathbf{c}, \mathbf{p}\rangle), \\
&\eta_{i_1} < \sqrt{n_{i_1}}\mu_{i_1}, \ldots, \eta_{i_k} < \sqrt{n_{i_k}}\mu_{i_k}, \eta_{i_{k+1}} > \sqrt{n_{i_{k+1}}}\mu_{i_{k+1}}, \ldots, \eta_{i_s} > \sqrt{n_{i_s}}\mu_{i_s}\Big), \\
P_2(i_1, \ldots, i_k) \approx P\Big(&\sum_{i=1}^{s} \sqrt{c_i p_i (1-p_i)}\eta_i < \sqrt{N}(\alpha - \langle \mathbf{c}, \mathbf{p}\rangle), \\
&\eta_{i_1} > \sqrt{n_{i_1}}\mu_{i_1}, \ldots, \eta_{i_k} > \sqrt{n_{i_k}}\mu_{i_k}, \eta_{i_{k+1}} < \sqrt{n_{i_{k+1}}}\mu_{i_{k+1}}, \ldots, \eta_{i_s} < \sqrt{n_{i_s}}\mu_{i_s}\Big).
\end{aligned}$$

a) Case $\langle \mathbf{c}, \mathbf{p}\rangle < \alpha$

For $\langle \mathbf{c}, \mathbf{p}\rangle < \alpha$, the inequality $\sum_{i=1}^{s} \sqrt{c_i p_i(1-p_i)}\eta_i > \sqrt{N}(\alpha - \langle \mathbf{c}, \mathbf{p}\rangle)$ implies $P_1(i_1, \ldots, i_k) \to 0$ for any $s - \lfloor s\beta \rfloor \leqslant k \leqslant s-1$ and any $(i_1, \ldots, i_k)$. Therefore, $\sum_{k=s-\lfloor s\beta\rfloor}^{s-1} \sum P_1(i_1, \ldots, i_k) \to 0$, but

$$\begin{aligned}
P_2(i_1, \ldots, i_k) &\approx P\Big(\eta_{i_1} > \sqrt{n_{i_1}}\mu_{i_1}, \ldots, \eta_{i_k} > \sqrt{n_{i_k}}\mu_{i_k}, \eta_{i_{k+1}} < \sqrt{n_{i_{k+1}}}\mu_{i_{k+1}}, \ldots, \eta_{i_s} < \sqrt{n_{i_s}}\mu_{i_s}\Big) \\
&\approx \prod_{j=1}^{k}\Big(1 - \Phi(\sqrt{n_{i_j}}\mu_{i_j})\Big) \prod_{r=1}^{s-k} \Phi\Big(\sqrt{n_{i_{k+r}}}\mu_{i_{k+r}}\Big). \qquad (4)
\end{aligned}$$

Note that $m^+ \geqslant 1$. Since $m^+ \geqslant s - \lfloor s\beta \rfloor$ implies $P_2(i_1, \ldots, i_k) \to 0$ for any $\lfloor s\beta \rfloor + 1 \leqslant k \leqslant s-1$ and any $(i_1, \ldots, i_k)$, the inversion probability (1) tends to 0.

Let $1 \leqslant m^+ \leqslant s - \lfloor s\beta \rfloor - 1$ and assume, without any loss of generality, that $m_1, \ldots, m_s$ are sorted in an ascending order.

For $\mu \in \mathbb{R}$, as $n \to \infty$,

$$\Phi(\sqrt{n}\mu) \to \begin{cases} 1, & \text{for } \mu > 0 \\ 0, & \text{for } \mu < 0 \end{cases} \quad \text{and } \Phi(\sqrt{n}\mu) = 1/2 \text{ for } \mu = 0.$$

Therefore, any probability $P_2(i_1, \ldots, i_k)$, which does not satisfy the condition:

$$\{1, \ldots, m^-\} \subset \{i_1, \ldots, i_k\} \text{ and } \{s - m^+ + 1, \ldots, s\} \subset \{i_{k+1}, \ldots, i_s\} \tag{5}$$

tends to 0, whereas any probability $P_2(i_1, \ldots, i_k)$, which satisfies the condition (5), tends to 2^{-m^0}. It remains to calculate the number of probabilities satisfying (5).

First, consider the case $k = \lfloor s\beta \rfloor + 1$. For probabilities satisfying (5), there are $s - \lfloor s\beta \rfloor - 1 - m^+$ free indexes among $\{i_{k+1}, \ldots, i_s\}$, so the number of such probabilities is $\binom{m^0}{s-\lfloor s\beta \rfloor -1-m^+}$. For $k = \lfloor s\beta \rfloor + 2$ there are $s - \lfloor s\beta \rfloor - 2 - m^+$ free indexes among $\{i_{k+1}, \ldots, i_s\}$, so the number of such probabilities is $\binom{m^0}{s-\lfloor s\beta \rfloor -2-m^+}$, and so on. Therefore, the total number of such probabilities is

$$\sum_{k=0}^{s-\lfloor s\beta \rfloor -1-m^+} \binom{m^0}{k},$$

which furnishes the proof of (2).

b) Case $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$

For $\langle \mathbf{c}, \mathbf{p} \rangle > \alpha$, the inequality $\sum_{i=1}^{s} \sqrt{c_i p_i (1-p_i)} \eta_i < \sqrt{N}(\alpha - \langle \mathbf{c}, \mathbf{p} \rangle)$ implies $P_2(i_1, \ldots, i_k) \to 0$ for any $\lfloor s\beta \rfloor + 1 \leqslant k \leqslant s - 1$ and any $(i_1, \ldots, i_k)$. Therefore, $\sum_{k=\lfloor s\beta \rfloor+1}^{s-1} \sum P_2(i_1, \ldots, i_k) \to 0$, but

$$\begin{aligned} P_1(i_1, \ldots, i_k) &\approx P\Big(\eta_{i_1} < \sqrt{n_{i_1}}\mu_{i_1}, \ldots, \eta_{i_k} < \sqrt{n_{i_k}}\mu_{i_k}, \eta_{i_{k+1}} > \sqrt{n_{i_{k+1}}}\mu_{i_{k+1}}, \ldots, \eta_{i_s} > \sqrt{n_{i_s}}\mu_{i_s}\Big) \\ &\approx \prod_{j=1}^{k} \Phi(\sqrt{n_{i_j}}\mu_{i_j}) \prod_{r=1}^{s-k} \Big(1 - \Phi(\sqrt{n_{i_{k+r}}}\mu_{i_{k+r}})\Big). \end{aligned} \tag{6}$$

Since the rest of the proof in case b) follows the same logic as in case a), we omit the details and proceed with the proof of case c).

c) Case $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$

For $\langle \mathbf{c}, \mathbf{p} \rangle = \alpha$, we have

$$\begin{aligned} P_1(i_1, \ldots, i_k) &\approx P\Big(\sum_{i=1}^{s} \sqrt{c_i p_i (1-p_i)} \eta_i > 0, \\ &\qquad \eta_{i_1} < \sqrt{n_{i_1}}\mu_{i_1}, \ldots, \eta_{i_k} < \sqrt{n_{i_k}}\mu_{i_k}, \eta_{i_{k+1}} > \sqrt{n_{i_{k+1}}}\mu_{i_{k+1}}, \ldots, \eta_{i_s} > \sqrt{n_{i_s}}\mu_{i_s}\Big), \\ P_2(i_1, \ldots, i_k) &\approx P\Big(\sum_{i=1}^{s} \sqrt{c_i p_i (1-p_i)} \eta_i < 0, \\ &\qquad \eta_{i_1} > \sqrt{n_{i_1}}\mu_{i_1}, \ldots, \eta_{i_k} > \sqrt{n_{i_k}}\mu_{i_k}, \eta_{i_{k+1}} < \sqrt{n_{i_{k+1}}}\mu_{i_{k+1}}, \ldots, \eta_{i_s} < \sqrt{n_{i_s}}\mu_{i_s}\Big). \end{aligned}$$

From the additional first inequality in the probabilities it follows that each $P_1(i_1, \ldots, i_k)$ is not larger than the corresponding probability (6), and that each $P_2(i_1, \ldots, i_k)$ is not larger than the corresponding probability (4).

Each probability $P_1(i_1, \ldots, i_k)$, where $s - \lfloor s\beta \rfloor \leqslant k \leqslant s - 1$, tends to 0 for $m^- \geqslant \lfloor s\beta \rfloor + 1$ and each probability $P_2(i_1, \ldots, i_k)$, where $\lfloor s\beta \rfloor + 1 \leqslant k \leqslant s - 1$, tends to 0 for $m^+ \geqslant s - \lfloor s\beta \rfloor$.

Therefore, if $m^- \geqslant \lfloor s\beta \rfloor + 1$ and $m^+ \geqslant s - \lfloor s\beta \rfloor$ simultaneously, then $P_1 \to 0$, $P_2 \to 0$ and $P = P_1 + P_2 \to 0$. But the two inequalities cannot hold simultaneously, as $m^- + m^+ \leqslant s$.

If one of the population shares $\{c_i\}$ is equal to 1 and not all states are close ($\mathbf{m} \neq \mathbf{0}_s$), then zero limit can occur due to a contradiction between the first and the others inequalities in the probability. For example, the inversion probability vanishes in the limit for $\beta = 0.5$ and the population concentrated in a single close state, $m^- = m^+ = (s-1)/2$ and $m^0 = 1$.

B Frequencies of limit values for $s = 3, \ldots, 101$ and $\beta = 0.5$

Table 5: Frequencies of limit values for $s = 3, \ldots, 101$ and $\beta = 0.5$.

s	0	(0,0.1]	(0.1,0.2]	(0.2,0.3]	(0.3,0.4]	(0.4,0.5]	(0.5,0.6]	(0.6,0.7]	(0.7,0.8]	(0.8,0.9]	(0.9,1)	1
3	6	0	0	2	0	2	0	0	0	0	0	2
5	12	0	2	2	2	4	0	0	2	0	0	6
7	20	2	4	2	4	6	0	2	2	2	0	12
9	30	4	6	4	6	8	0	4	2	4	2	20
11	42	8	8	6	8	10	0	6	4	6	4	30
13	56	14	10	8	10	12	0	8	6	8	8	42
15	72	20	14	10	12	14	0	10	8	10	14	56
17	90	28	16	12	14	18	0	12	10	14	20	72
19	110	38	18	14	16	22	2	14	12	16	28	90
21	132	48	22	16	18	26	4	16	14	18	38	110
23	156	60	26	18	20	30	6	18	16	22	48	132
25	182	74	30	20	22	34	8	20	18	26	60	156
27	210	90	32	24	24	38	10	22	20	30	74	182
29	240	106	36	28	26	42	12	24	24	32	90	210
31	272	124	40	32	28	46	14	26	28	36	106	240
33	306	144	44	36	30	50	16	28	32	40	124	272
35	342	166	48	40	32	54	18	30	36	44	144	306
37	380	188	54	42	36	58	20	32	40	48	166	342
39	420	212	60	44	40	62	22	36	42	54	188	380
41	462	238	64	48	44	66	24	40	44	60	212	420
43	506	266	68	52	48	70	26	44	48	64	238	462
45	552	296	72	56	52	74	28	48	52	68	266	506
47	600	328	76	60	56	78	30	52	56	72	296	552
49	650	360	82	64	60	82	32	56	60	76	328	600
51	702	394	88	68	64	86	34	60	64	82	360	650
53	756	430	94	72	68	90	36	64	68	88	394	702
55	812	468	100	76	72	94	38	68	72	94	430	756
57	870	508	104	82	76	98	40	72	76	100	468	812
59	930	548	110	88	80	102	42	76	82	104	508	870
61	992	590	116	94	84	106	44	80	88	110	548	930
63	1056	634	122	98	90	110	46	84	94	116	590	992
65	1122	680	128	102	94	114	50	90	98	122	634	1056
67	1190	728	134	106	98	120	54	94	102	128	680	1122
69	1260	778	140	110	102	124	60	98	106	134	728	1190
71	1332	828	148	114	106	130	64	102	110	140	778	1260
73	1406	880	156	118	110	136	68	106	114	148	828	1332
75	1482	934	164	122	114	142	72	110	118	156	880	1406
77	1560	990	170	128	118	146	78	114	122	164	934	1482
79	1640	1048	176	134	122	150	84	118	128	170	990	1560
81	1722	1108	182	140	126	154	90	122	134	176	1048	1640
83	1806	1170	188	146	130	160	94	126	140	182	1108	1722
85	1892	1232	196	152	134	164	100	130	146	188	1170	1806
87	1980	1296	204	158	138	170	104	134	152	196	1232	1892
89	2070	1362	212	164	142	176	108	138	158	204	1296	1980
91	2162	1430	220	170	146	182	112	142	164	212	1362	2070
93	2256	1500	228	176	150	186	118	146	170	220	1430	2162
95	2352	1572	236	182	154	190	124	150	176	228	1500	2256
97	2450	1646	244	186	160	196	128	154	182	236	1572	2352
99	2550	1722	250	192	166	202	132	160	186	244	1646	2450
101	2652	1798	258	198	172	206	138	166	192	250	1722	2550