European Journal of Political Economy 76 (2023) 102228

Contents lists available at ScienceDirect

European Journal of Political Economy

journal homepage: www.elsevier.com/locate/ejpe

The likelihood of the referendum paradox for a given referendum result

Pavlo Blavatskyy [1]

Montpellier Business School, 2300, Avenue des Moulins, 34185, Montpellier, Cedex 4, France

ARTICLE INFO

JEL classificationcodes
D71

Keywords
Social choice
Voting
Referendum paradox
Compound majority paradox
Impartial anonymous culture

ABSTRACT

The referendum (or compound majority) paradox occurs when the majority of voters in the majority of districts supports an issue/candidate but the majority of voters across all districts opposes the same issue/candidate (or vice versa). We calculate the likelihood of this social choice anomaly for any (possibly large) odd number of districts and any (possibly large) odd number of voters per district. The likelihood of the paradox is close to 50% when the issue/candidate is divisive (voters across all districts are split almost 50%–50%). The paradox virtually disappears when the issue/candidate is supported/opposed by at least two-thirds of all voters.

1. Introduction

The referendum (or compound majority) paradox is also known as election (or majority) inversion. It occurs when the majority of voters in the majority of districts supports an issue/candidate but the majority of voters across all districts opposes the same issue/candidate or vice versa (*cf*. Nurmi, 1998, Section 2.4, p. 336). For example, in the 2016 United States presidential election the Republican Party candidate Donald Trump received 2.8 million votes less than the Democratic Party candidate Hillary Clinton but Donald Trump won the election by receiving 304 electoral votes (while Hillary Clinton received only 227 electoral votes). Another example is the 2000 United States presidential election when the Republican Party candidate George W. Bush received half a million votes less than the Democratic Party candidate Al Gore but won the election by receiving 271 electoral votes (Al Gore received only 266 electoral votes). Lahrach and Merlin (2009) provide further examples of the referendum paradox in the French local elections.

The result of a referendum or a direct presidential election depends only on the popular majority. The result of an indirect (representative) election additionally depends on the distribution of votes across districts. For example, if all but one state have only a tiny majority for a Republican Party candidate and the remaining state (say California) has a huge majority in favor of a Democratic Party candidate, then the latter wins the popular vote but the former wins the electoral votes (*cf*. Nurmi, 1998, pp. 344–345). Such referendum (or compound majority) paradox is closely related to the Ostrogorski (1903) paradox.

May (1948) was the first to calculate the likelihood of the referendum paradox for an odd (possibly infinite) number of districts each with an odd (possibly infinite) number of voters under the assumptions that each voting situation is equally likely within districts and independently distributed across districts. For three districts May (1948, Section 3, pp. 204–205) used a combinatorial technique. In general, when the number of districts is not necessarily three, May (1948, Section 4, pp. 205–206) employed the method of Euler

E-mail address: p.blavatskyy@montpellier-bs.com.

[1] Pavlo Blavatskyy is a member of the Entrepreneurship and Innovation Chair, which is part of LabEx Entrepreneurship (University of Montpellier, France) and funded by the French government (Labex Entreprendre, ANR-10-Labex-11-01). I am indebted to the editor Toke Aidt and two anonymous referees for their very helpful comments.

https://doi.org/10.1016/j.ejpoleco.2022.102228
Received 12 August 2021; Received in revised form 3 April 2022; Accepted 6 April 2022
Available online 15 April 2022

based on collecting terms of a polynomial generating function.

However, May's contribution (published in a mathematical journal) received a relatively limited acknowledgement in the economic literature. Feix et al. (2004, Table 4, p. 236) calculated the likelihood of the referendum paradox when there are only three, four and five districts each with the same (up to 21) number of voters as well as asymptotic results for an infinitely large number of voters. Feix et al. (2004, Section 3.3) also provided computer simulations for a larger number of states and voters per state.

Wilson and Pritchard (2007, Section 3.5, p. 254) calculated the likelihood of the referendum paradox for up to nine districts by counting integer points in convex polytopes with Ehrhart polynomials (*cf*. Lepelley et al., 2008). However, Wilson and Pritchard (2007) ignored the constraint that the number of votes in each district cannot exceed the number of voters in the district. Lepelley et al. (2011, Box I, p. 30) corrected this omission and calculated the likelihood of the referendum paradox when there are five, seven and nine districts. Lepelley et al. (2011) also calculated the asymptotic results for an infinitely large number of voters and approximative results for a large number of districts and a large number of voters. Geruso et al. (2022) numerically estimate the likelihood of the referendum paradox in the U.S. presidential elections in 1836–2016 as a function of the national popular vote. They find this conditional probability to be relatively stable for close elections.

Several studies of the referendum paradox employ more distinct voting models. Hinich et al. (1972) study the likelihood of the referendum paradox using a voting model where the proportion of voters supporting an issue/candidate is independently distributed across equally-sized districts and drawn from a beta distribution within districts. de Mouzon et al. (2020) estimate numerically the probability of the referendum paradox for three equally-sized districts and any (possibly large) number of voters per district. de Mouzon et al. (2020) employ a voting model where each voting situation within the population is equally likely. Kaniovski and Zaigraev (2018) study the probability of the referendum paradox for three possibly unequal districts and various voting rules. Le Breton et al. (2014) estimate the likelihood of the referendum paradox for mixed voting systems. Kikuchi (2017) explores the impact of voting rules on the likelihood of the referendum paradox.

The main contribution of this paper is to derive an explicit formula for the probability of the referendum paradox conditional on the referendum result (voter majority). For example, the United States presidential election in 1984, when a Republican candidate Ronald Reagan enjoyed a solid majority, was qualitatively different from elections in 2000 or 2016, when American voters were much more divided. Aggregating these elections together overstates the likelihood of the referendum paradox in the former example and understates in the latter example. The formula of conditional probability of the referendum paradox derived in this paper is arguably more appropriate.

The second contribution of this paper is to calculate the overall likelihood of the referendum paradox by aggregating conditional probabilities. This gives us an alternative method for calculating the overall probability of the referendum paradox by means of a relatively elementary combinatorial technique. As mentioned above, May (1948, Section 4, pp. 205–206) employed a relatively complex method of Euler, which is more difficult to follow.

We find that the likelihood of the referendum paradox is close to 50% when the issue/candidate is divisive (voters across all districts are split almost 50%–50%).[2] Yet, it quickly decreases for larger margins in the popular vote. For example, with 19 districts the likelihood of the referendum paradox is less than 5% when at least 56% of all voters support/oppose an issue/candidate. The paradox virtually disappears when the majority in the popular vote is two thirds or more. Nurmi (1998, p. 345) already showed that the referendum paradox cannot exist when the majority in the popular vote is greater than or equal to 75%. This paper illustrates that this theoretical upper bound is rather excessive.

The remainder of the paper is organized as follows. Section 2 presents our mathematical notation and modelling assumptions. Section 3 presents main results. Section 4 concludes with policy recommendations on how to limit the occurrence of the referendum paradox.

2. Mathematical notation and modelling assumptions

For simplicity, we consider an odd number of districts $2m+1$ for some $m\in\{1,2,\ldots\}$ (so that ties are not possible when district representatives vote on a binary issue). Each district has an odd number of voters $2n+1$ for some $n\in\{1,2,\ldots\}$ (so that ties are not possible when district representatives are elected). We assume that all district representatives are faithful to the majority of voters in their respective districts (if the majority of voters in a district supports/opposes an issue then so does its district representative). There is no abstention in any voting. We consider a binary issue so that each voter either supports or opposes this issue.

The total number of voters across all districts is $(2m+1)(2n+1)$. Thus, the majority in the popular vote for a binary issue can be written as $2mn+m+n+k$ for some $k\in\{1,2,\ldots\}$. When $k=1$ there is a minimal majority of one vote (voters across all districts are split almost 50%–50%). When k is large there is a solid majority in the popular vote.

The referendum paradox occurs when there is a popular majority but at least $m+1$ districts do not have a majority for this issue (or vice versa). Notice that when k is very large then there cannot be any referendum paradox. In fact, the largest majority in the popular vote that can be opposed by $m+1$ districts is when each of these $m+1$ districts has exactly n voters for the issue and all voters in the remaining m districts support this issue. This largest majority is then $n(m+1)+(2n+1)m$ or when $k=mn$. Therefore, to study the likelihood of the referendum paradox, we can limit our attention only to $k\in\{1,\ldots,mn\}$. As a percentage of all voters, the largest "reversible" majority is given by

[2] This higher incidence of the referendum paradox for closely contested elections appears to be consistent with empirical findings of Lahrach and Merlin (2009) for French local elections.

$$\frac{n(m+1)+(2n+1)m}{(2m+1)(2n+1)}=1-\left(\frac{m+1}{2m+1}\right)\left(\frac{n+1}{2n+1}\right)$$

We assume that each voting situation is equally likely within districts, which is known as the impartial anonymous culture in the social choice literature (*cf*. Kuga and Nagatani, 1974; Gehrlein and Fishburn, 1976). Voting situations are independent across districts. Under these assumptions, the likelihood of any voting paradox is then given by the number of integer solutions to a system of linear inequalities (*cf*. Huang and Chua, 2000; Gehrlein, 2002; Lepelley et al., 2008). Calculating the likelihood of the referendum paradox for a given majority in the popular vote is a combinatorial problem. We need to count in how many ways a number can be decomposed into a sum of nonnegative integers. For a compact notation, let $\binom{a}{b}$ be zero when a<b and a standard binomial coefficient when $a \geq b \geq 0$.

3. Results

We consider voting situations when there is a popular majority for some $k \in \{1, \ldots, mn\}$ but at least $m+1$ districts do not have a district majority. Voting situations when there is no popular majority but at least $m+1$ districts have a district majority are symmetric.

Proposition 1. The number of voting situations when there is a popular majority for some $k \in \{1, \ldots, mn\}$ but at least $m+1$ districts do not have a district majority is given by

$$P(n,m,k)=\sum_{j=1}^{m}\binom{2m+1}{j}\left[\sum_{i=0}^{j-1}(-1)^i\binom{2m+1}{i}\binom{m-k+(n+1)(j-i)}{2m}\right]$$

The Proof is presented in the appendix.

Proposition 1 counts the number of voting situations when the referendum paradox occurs. To assess its likelihood, we also need to count the number of voting situations when there is popular majority $2mn+m+n+k$ but no referendum paradox is observed. This is done by the following Proposition 2.

Proposition 2. The number of voting situations when there is a popular majority for some $k \in \{1, \ldots, mn\}$ and at least $m+1$ districts have a district majority is given by

$$Q(n,m,k)=\sum_{j=0}^{m}\binom{2m+1}{j}\left[\sum_{i=0}^{m+j-1}(-1)^i\binom{2m+1}{i}\binom{m+k-1+(n+1)(j-i)}{2m}\right]$$

The Proof is presented in the appendix.

The likelihood of the referendum paradox conditional on a given voter majority is then $P(n,m,k)/[P(n,m,k)+Q(n,m,k)]$. Fig. 1 plots this likelihood as a function of the referendum result (the fraction of the total population supporting the issue, given by $0.5+\frac{k-0.5}{(2m+1)(2n+1)}$) for 7, 9, 19, 39 and 99 districts with 99 and 999 voters per district. The likelihood of the referendum paradox is close to 50% when the referendum result is close to 50%. This is intuitive. If exactly 50% of the population support an issue then there is an equal chance that the majority of same-size districts supports or opposes the issue (by the symmetry of the problem). Thus, if a referendum results in a very slim margin, the chances of the referendum paradox are close to 50%. In other words, there is always a danger of this paradox for very divisive issues/candidates.

Fig. 1 show that the impact of the number of voters per district on the likelihood of the referendum paradox is rather limited—there is almost no difference between 99 and 999 voters per district. The number of districts, however, has a more important effect on the likelihood of the referendum paradox. *Ceteris paribus*, the higher is the number of districts the less likely is the referendum paradox. In the limit when there are as many districts as voters (so that each district has only one voter) the paradox disappears completely.

Fig. 1 shows that the theoretical upper limit of $1-\left(\frac{m+1}{2m+1}\right)\left(\frac{n+1}{2n+1}\right)$ for the existence of the referendum paradox is rather excessive. The likelihood of the referendum paradox is less than 5% when the referendum result is 59.3% for seven districts with 99 voters, 56% for nineteen districts, and 52.8% for 99 districts (the corresponding theoretical upper bounds are 71.1%, 73.4%, and 74.5%). In other words, even though the referendum paradox is inevitable for very divisive issues/candidates, its likelihood rapidly diminishes when there are many districts and a solid majority in the popular vote.

The overall likelihood of the referendum paradox $N(n,m)$ for $2m+1$ districts each with $2n+1$ voters is the sum of $P(n,m,k)$ from Proposition 1 over all possible values of k, divided by the total number of voting situations, which is $(2n+2)^{2m+1}$, and multiplied by two.[3]

[3] One possibility is when the majority of the population votes in favor of an issue, but the majority of districts votes against the issue. The second possibility is the opposite when the majority of the population votes against the issue but the majority of districts votes in favor of this issue.

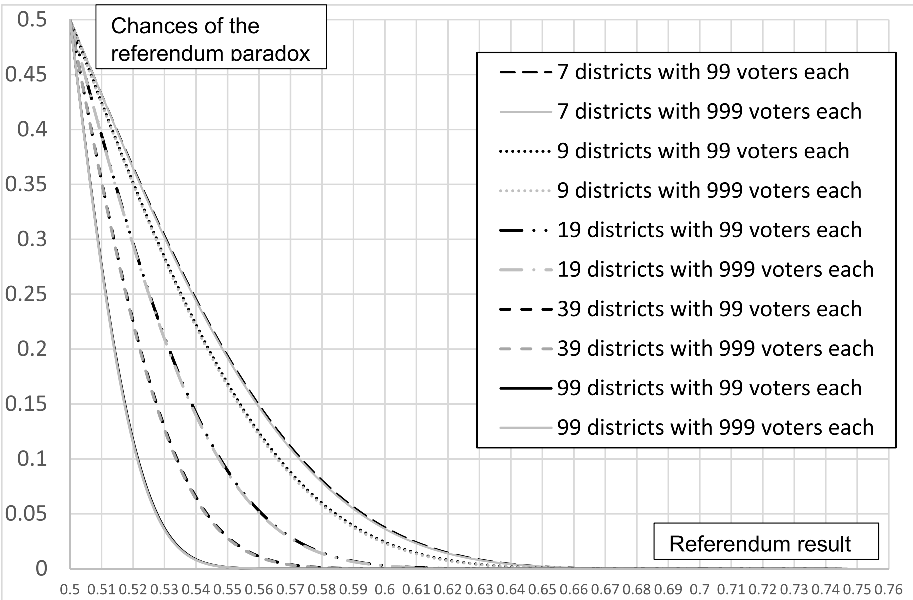

Fig. 1. The likelihood of the referendum paradox as a function of the referendum result.

$$N(n,m) = \frac{1}{2^{2m}(n+1)^{2m+1}} \sum_{k=1}^{mn} P(n,m,k)$$

This formula is equivalent to equation (9) in May (1948, p. 206). Table 1 shows overall likelihood of the referendum paradox $N(n,m)$ for three, five, seven, nine, 49 and 99 districts and various numbers of voters per district. Python 3.5 script for computing Table 1 is presented in the online supplementary appendix. The second and the third columns of Table 1 (for three and five districts) coincide with the results computed in Table 4 of Feix et al. (2004, p. 236). The third and the fourth columns of Table 1 (for five and seven districts) coincide with the results computed in *Box* I in Lepelley et al. (2011, p. 30). The fifth column (for nine districts) does not coincide with the results computed in *Box* I in Lepelley et al. (2011, p. 30). It seems that there is a minor typo in *Box* I in Lepelley et al. (2011, p. 30) for nine districts: one of the multipliers in the nominator should be $n+3$ instead of n-3. If we correct this typo in Lepelley et al. (2011, p. 30) then our results for nine districts correspond perfectly to theirs. Results from Table 1 for three, five, seven and nine districts with 999 voters coincide with asymptotic probabilities in May (1948, table on p.208).

Results from Table 1 for five, seven and nine districts with 999 voters coincide with the limiting probabilities 55/384, 577/3840 and 1589879/10321920 reported in Lepelley et al. (2011, p. 29). For 99 districts with 999 voters the probability of the referendum paradox is 0.16555. It is close to the result of a computer simulation ("around 16.5%") reported in Feix et al. (2004, section 3.3, p. 236) and the asymptotical limit of 1/6 derived by May (1948).

The last two rows of Table 1 also show that the overall likelihood of the referendum paradox lies between 15% and 16.6% for different *m*-*n* combinations considered in Fig. 1. Thus, if we look only at the unconditional probability of the referendum paradox, we have an impression that these scenarios are quite similar. However, as Fig. 1 shows, if we look at the conditional probability, these scenarios are rather heterogeneous (in particular, in terms of their variation with respect to the number of districts).

4. Conclusion

Direct democracy is often viewed as a true reflection of voting preferences but, with the exception of Switzerland (Frey, 1994), it is often replaced with representative democracy on practical grounds (*cf*. Coffman, 2016). This may produce conflicting results such as the referendum paradox when the majority of population supports an issue/candidate but only a minority of district representatives do the same (or vice versa). The Brexit impasse in the Parliament of the United Kingdom in 2019 shows that such conflicts may result in high macroeconomic costs for the society. The existing literature largely focusses on the overall likelihood of the referendum paradox. In contrast, the question how its likelihood depends on the referendum result, to the best of our knowledge, has not been researched before.

We find that the likelihood of the referendum paradox is close to 50% when total population is nearly evenly split between supporting and opposing an issue or a candidate. However, this likelihood rapidly diminishes when referendum results move away from the 50%–50% split. The decline is particularly strong when there are many districts. For example, when at least 55% of the total population support an issue, the likelihood that only a minority of 99 (19) district representatives do the same is 0.1% (8.7%). The referendum paradox virtually disappears (its likelihood becomes less than 0.05%) when at least two thirds of the total population votes for or against an issue/candidate.

Table 1
Overall likelihood of the referendum paradox.

Voters per district	Number of districts					
	3	5	7	9	49	99
3	0,09375	0,117188	0,126465	0,131378	0,144705	0,146164
5	0,111111	0,132459	0,140432	0,144646	0,156167	0,157436
7	0,117188	0,137329	0,144863	0,14887	0,159862	0,161074
9	0,12	0,1395	0,146843	0,150761	0,161523	0,162711
11	0,121528	0,140657	0,147901	0,151772	0,162413	0,163587
13	0,122449	0,141347	0,148533	0,152376	0,162945	0,164112
15	0,123047	0,141792	0,148941	0,152766	0,163289	0,164451
17	0,123457	0,142096	0,149219	0,153033	0,163524	0,164683
19	0,12375	0,142313	0,149418	0,153223	0,163692	0,164848
21	0,123967	0,142472	0,149565	0,153363	0,163816	0,16497
99	0,12495	0,143193	0,150227	0,153997	0,164375	0,165522
999	0,125	0,143229	0,15026	0,154029	0,164404	0,16555

Representative majority voting is widely used in everyday politics. It is only occasionally complemented by direct voting such as referenda. Thus, the practical solution for avoiding the conflict between the two could be an introduction of a minimum threshold in the popular vote. The exact probability value of such threshold is ultimately a political decision. Yet, our results indicate that setting this threshold anywhere between 55% and 60% dramatically reduces the likelihood of the referendum paradox.

Data availability

No data was used for the research described in the article.

Appendix

Proof of Proposition 1

Step 1 Let us first consider the simplest voting situation, which is the strongest possible district opposition, when $2m$ districts do not have a majority and only one district has a majority. Let $x_i \geq 0$ denote the number of votes in the i-th district. Suppose that the first district has a majority so that $x_1 \geq n+1$ and $x_i \leq n$ for all $i \in \{2, \ldots, 2m+1\}$. Since the popular majority is $2mn + m + n + k$ the following equation must hold:

$$\sum_{i=1}^{2m+1} x_i = 2mn + m + n + k$$

This equation can be rearranged as follows:

$$(2n + 1 - x_1) + \sum_{i=2}^{2m+1} (n - x_i) = n + 1 - m - k$$

Each bracket on the left-hand-side of this equation is a non-negative integer less than or equal to n. Thus, the right-hand-side must be a non-negative integer as well, so that $k \leq n+1-m$. This is quite intuitive—the strongest district opposition is only possible when the popular majority is very thin.

The right-hand-side of the above equation is at most n-m (when $k = 1$). Therefore, the constraints that each bracket on the left-hand-side must be less than or equal to n are not binding and can be ignored. We have then a standard combinatorial problem: what is the number of $(2m + 1)$-tuples of non-negative integer numbers whose sum is $n+1-m-k$? By stars-and-bars method, this number is given by

$$\binom{n + 1 + m - k}{2m}$$

Since the district that has a majority can be not only the first district but also—any other district, the total number of voting situations when $2m$ districts do not have a majority but the popular majority is $2mn + m + n + k$ is equal to

$$N_1 = \binom{2m+1}{1}\binom{n + 1 + m - k}{2m}$$

Step 2 Let us now consider the voting situation when $2m$-1 districts do not have a majority and two districts have a majority. Note that this step is only relevant for the referendum paradox when $m > 1$. Suppose that the first and the second districts have a majority so that $x_1 \geq n+1$, $x_2 \geq n+1$ and $x_i \leq n$ for all $i \in \{3, \ldots, 2m+1\}$. Since the popular majority is $2mn + m + n + k$ we have:

$$\sum_{i=1}^{2m+1} x_i = 2mn + m + n + k$$

This equation can be rearranged as follows:

$$\sum_{i=1}^{2}(2n+1-x_i) + \sum_{i=3}^{2m+1}(n-x_i) = 2n+2-m-k$$

Each bracket on the left-hand-side of this equation is a non-negative integer less than or equal to n. Thus, the right-hand-side must be a non-negative integer as well, so that this voting situation is only possible when $k \leq 2n+2-m$.

The right-hand-side of the above equation is at most $2n$-1 (when $k = 1$ and $m = 2$). This means that constraints that each bracket on the left-hand-side must be less than or equal to n could be binding, but only for one district at most. We first calculate the number r_2 of decompositions ignoring this constraint. Second, we calculate the number s_2 of decompositions when one district violates this constraint. Third, we subtract the latter from the former to find the number of decompositions when no district violates the constraint.

By stars-and-bars method, the number of $(2m + 1)$-tuples of non-negative integer numbers whose sum is $2n+2-m-k$ is given by

$$r_2 = \binom{2n+2+m-k}{2m}$$

The number of $(2m + 1)$-tuples of non-negative integer numbers, one of which is greater than or equal to $n+1$, whose total sum is $2n+2-m-k$ can be calculated as follows. Suppose that the first number is greater than or equal to $n+1$. The number of $(2m + 1)$-tuples of non-negative integer numbers, the first being greater than or equal to $n+1$, whose sum is $2n+2-m-k$ is the same as the number of $(2m + 1)$-tuples of non-negative integer numbers whose sum is $n+1-m-k$. By stars-and-bars method, it is

$$\binom{n+1+m-k}{2m}$$

Since the number that is greater than or equal to $n+1$ could be not only the first but also—any other number, we have

$$s_2 = \binom{2m+1}{1}\binom{n+1+m-k}{2m}$$

Thus, the number of $(2m + 1)$-tuples of non-negative integer numbers, each less than or equal to n, whose total sum is $2n+2-m-k$ is given by

$$r_2 - s_2 = \binom{2n+2+m-k}{2m} - \binom{2m+1}{1}\binom{n+1+m-k}{2m}$$

Since the pair of districts that have a majority can be not only the first pair of districts but also—any other two districts, the total number of voting situations when $2m$-1 districts do not have a majority but the popular majority is $2mn + m + n + k$ is equal to

$$N_2 = \binom{2m+1}{2}\left[\binom{2n+2+m-k}{2m} - \binom{2m+1}{1}\binom{n+1+m-k}{2m}\right]$$

Step 3 Let us now consider the voting situation when $2m$-2 districts do not have a majority and three districts have a majority. To observe the referendum paradox, the number of districts that do not have a majority must be at least $m+1$ so that $2m-2 \geq m+1$, or $m \geq 3$.

Suppose that the first three districts have a majority so that $x_1 \geq n+1$, $x_2 \geq n+1$, $x_3 \geq n+1$ and $x_i \leq n$ for all $i \in \{4, \ldots, 2m+1\}$. Since the popular majority is $2mn + m + n + k$ we must have:

$$\sum_{i=1}^{2m+1} x_i = 2mn + m + n + k$$

This equation can be rearranged as follows:

$$\sum_{i=1}^{3}(2n+1-x_i) + \sum_{i=4}^{2m+1}(n-x_i) = 3n+3-m-k$$

Each bracket on the left-hand-side of this equation is a non-negative integer less than or equal to n. Thus, the right-hand-side must be a non-negative integer as well, so that this voting situation is only possible when $k \leq 3n+3-m$.

The right-hand-side of the above equation is at most $3n$-1 (when $k = 1$ and $m = 3$). This means that constraints that each bracket on the left-hand-side must be less than or equal to n could be binding, but only for two districts at most. We first calculate the number of decompositions ignoring this constraint. Second, we calculate the number of decompositions when one or two districts violate this constraint. Third, we subtract the latter from the former to find the number of decompositions when no district violates the constraint.

By stars-and-bars method, the number of $(2m + 1)$-tuples of non-negative integer numbers whose sum is $3n+3-m-k$ is given by

$$r_3 = \binom{3n+3+m-k}{2m}$$

The number of $(2m+1)$-tuples of non-negative integer numbers, at least one of which is greater than n, whose total sum is $3n+3-m-k$ is given by

$$s_3 = \binom{2m+1}{1}\binom{2n+2+m-k}{2m} - \binom{2m+1}{2}\binom{n+1+m-k}{2m}$$

Note that we subtract the second term to avoid double-counting of instances when two numbers are greater than n.

Thus, the number of $(2m+1)$-tuples of non-negative integer numbers, each less than or equal to n, whose total sum is $3n+3-m-k$ is given by

$$r_3 - s_3 = \sum_{i=0}^{2}(-1)^i\binom{2m+1}{i}\binom{m-k+(n+1)(3-i)}{2m}$$

Since the triple of districts that have a majority can be not only the first three districts but also—any other triple of districts, the total number of voting situations when $2m$-2 districts do not have a majority but the popular majority is $2mn+m+n+k$ is equal to

$$N_3 = \binom{2m+1}{3}\sum_{i=0}^{2}(-1)^i\binom{2m+1}{i}\binom{m-k+(n+1)(3-i)}{2m}$$

By repeating analogous steps m times we obtain that the number of voting situations when there is a popular majority but at least $m+1$ districts do not have a district majority is given by

$$P(n,m,k) = \sum_{j=1}^{m} N_j = \sum_{j=1}^{m}\binom{2m+1}{j}\left[\sum_{i=0}^{j-1}(-1)^i\binom{2m+1}{i}\binom{m-k+(n+1)(j-i)}{2m}\right]$$

Q.E.D.

Proof of Proposition 2.

Let us first consider the strongest voting situation of harmony between the popular and representative voting: the case when all districts have a majority. Let $x_i \in \{n+1, \ldots, 2n+1\}$ denote the number of votes in the i-th district. Since the popular majority is $2mn+m+n+k$ we have:

$$\sum_{i=1}^{2m+1} x_i = 2mn+m+n+k$$

This equation can be rearranged as follows:

$$\sum_{i=1}^{2m+1}(x_i - n) = m+k$$

Each bracket on the left-hand-side is a positive integer less than or equal to $n+1$. Thus, the right-hand-side must be at least $2m+1$, so that this voting situation is only possible when $k \geq m+1$.

By stars-and-bars method, the number of $(2m+1)$-tuples of positive integers whose sum is $m+k$ is given by $\binom{m+k-1}{2m}$. This is not quite what we need as some of these tuples may include integers greater than $n+1$. Since we consider only $k \in \{1, \ldots, mn\}$, at most m-1 integers can be greater than $n+1$. The number of $(2m+1)$-tuples of positive integers, at least one of which is greater than $n+1$, and whose total sum is $m+k$ is given by

$$\binom{2m+1}{1}\binom{m+k-n-2}{2m} - \binom{2m+1}{2}\binom{m+k-2n-3}{2m} + \ldots \binom{2m+1}{m-1}\binom{k+n-mn}{2m}$$

Thus, the total number of voting situations when the popular majority is $2mn+m+n+k$ and every district has a district majority as well is

$$H_1 = \binom{m+k-1}{2m} - \binom{2m+1}{1}\binom{m+k-n-2}{2m} + \ldots + (-1)^{m-1}\binom{2m+1}{m-1}\binom{k+n-mn}{2m}$$

Let us now consider a slightly less harmonic voting situation when $2m$ districts have a majority and one district does not have a majority. Suppose that the first district does not have a majority so that $x_1 \in \{0, \ldots, n\}$ and, as before, $x_i \in \{n+1, \ldots, 2n+1\}$ for all $i \in \{2, \ldots, 2m+1\}$. The number of votes across all districts should sum up to the popular majority $2mn+m+n+k$. This can be written as the following equation:

$$(x_1+1) + \sum_{i=2}^{2m+1}(x_i - n) = m+k+n+1$$

Each bracket on the left-hand-side is a positive integer less than or equal to $n+1$. By stars-and-bars method, the number of $(2m+1)$-tuples of positive integers whose sum is $m+k+n+1$ is given by $\binom{m+k+n}{2m}$. From this number we need to subtract the number of tuples that include integers greater than $n+1$. Since we consider only $k\in\{1, \ldots, mn\}$, at most m integers can be greater than $n+1$. The number of $(2m+1)$-tuples of positive integers, at least one of which is greater than $n+1$, and whose total sum is $m+k+n+1$ is given by

$$\binom{2m+1}{1}\binom{m+k-1}{2m}-\binom{2m+1}{2}\binom{m+k-n-2}{2m}+\ldots\binom{2m+1}{m}\binom{k+n-mn}{2m}$$

Thus, the number of $(2m+1)$-tuples of positive integers less than or equal to $n+1$ and whose total sum is $m+k+n+1$ is given by

$$\binom{m+k+n}{2m}-\binom{2m+1}{1}\binom{m+k-1}{2m}+\ldots+(-1)^m\binom{2m+1}{m}\binom{k+n-mn}{2m}$$

Since the district that does not have a majority can be not only the first district but also—any other district, the total number of voting situations when the popular majority is $2mn+m+n+k$ and $2m$ districts have a district majority as well is given by

$$H_2=\binom{2m+1}{1}\left[\binom{m+k+n}{2m}-\binom{2m+1}{1}\binom{m+k-1}{2m}+\ldots+(-1)^m\binom{2m+1}{m}\binom{k+n-mn}{2m}\right]$$

By repeating analogous steps $m+1$ times we obtain that the number of voting situations when there is a popular majority $2mn+m+n+k$ and the majority of districts have a district majority as well is given by

$$Q(n,m,k)=\sum_{j=1}^{m+1}H_j=\sum_{j=0}^{m}\binom{2m+1}{j}\left[\sum_{i=0}^{m+j-1}(-1)^i\binom{2m+1}{i}\binom{m+k-1+(n+1)(j-i)}{2m}\right]$$

Q.E.D.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejpoleco.2022.102228.

References

Coffman, Katherine, 2016. Representative democracy and the implementation of majority-preferred alternatives. Soc. Choice Welfare 46, 477–494.
de Mouzon, Olivier, Thibault Laurent, Le Breton, Michel, Lepelley, Dominique, 2020. The theoretical Shapley–Shubik probability of an election inversion in a toy symmetric version of the US presidential electoral system. Soc. Choice Welfare 54, 363–395.
Feix, Marc, Lepelley, Dominique, Merlin, Vincent, Rouet, Jean-Louis, 2004. The probability of conflicts in a U.S. presidential type election. Econ. Theor. 23, 227–257.
Frey, Bruno, 1994. Direct democracy: politico-economic lessons from the Swiss experience. Am. Econ. Rev. 84 (2), 338–342.
Gehrlein, W., 2002. Obtaining representations for probabilities of voting outcomes with effectively unlimited precision integer arithmetic. Soc. Choice Welfare 19, 503–512.
Gehrlein, W., Fishburn, P., 1976. Condorcet's paradox and anonymous preference profiles. Publ. Choice 26, 1–18.
Geruso, M., Spears, D., Talesara, I., 2022. Inversions in US presidential elections: 1836–2016. Am. Econ. J. Appl. Econ. 14 (1), 327–357.
Hinich, M.J., Mickelsen, R., Ordeshook, P.C., 1972. The electoral college versus a direct vote: policy bias, reversals, and indeterminate outcomes. J. Math. Sociol. 4, 3–35.
Huang, H., Chua, V., 2000. Analytical representation of probabilities under the IAC condition. Soc. Choice Welfare 17, 143–155.
Kaniovski, Serguei, Alexander, Zaigraev, 2018. The probability of majority inversion in a two-stage voting system with three states. Theor. Decis. 84 (4), 525–546.
Kikuchi, K., 2017. The likelihood of majority inversion in an indirect voting system. Mimeo.
Kuga, K., Nagatani, H., 1974. Voter antagonism and the paradox of voting. Econometrica 42, 1045–1067.
Lahrach, Rahhal, Merlin, Vincent, 2009. Assessing the probability of the referendum paradox: the French local election case. Mimeo.
Lepelley, Dominique, Ahmed, Louichi, Smaoui, Hatem, 2008. On Ehrhart polynomials and probability calculations in voting theory. Soc. Choice Welfare 30, 363–383.
Lepelley, Dominique, Merlin, Vincent, Rouet, Jean-Louis, 2011. Three ways to compute accurately the probability of the referendum paradox. Math. Soc. Sci. 62, 28–33.
Le Breton, Michel, Dominique, Lepelley, Merlin, Vincent, 2014. Evaluating the likelihood of the referendum paradox for mixed voting systems. Mimeo.
May, Kenneth, 1948. Probability of Certain Election Results. Am. Math. Monthly 55 (4), 203–209.
Nurmi, Hannu, 1998. Voting paradoxes and referenda. Soc. Choice Welfare 15, 333–350.
Ostrogorski, Moise, 1903. La démocratie et l'organisation des partis politiques" Paris. Calmann-Levy.
Wilson, Mark, Pritchard, Geoffrey, 2007. Probability calculations under the IAC hypothesis. Math. Soc. Sci. 54, 244–256.