Majority Rule and Utilitarian Welfare

Vijay Krishna* John Morgan†

March 15, 2012
Preliminary and Incomplete‡

Abstract

Majority rule is known to be at odds with utilitarianism—majority rule follows the preferences of the median voter whereas a utilitarian planner would follow the preferences of the mean voter. In this paper, we show that when voting is costly and voluntary, turnout endogenously adjusts so that the two are completely reconciled: In large elections, majority rule is utilitarian. We also show that majority rule in unique in this respect: Among all supermajority rules, only majority rule is utilitarian. Finally, we show that majority rule is utilitarian even in the presence of aggregate uncertainty, a robustness not shared by other voting models.

1 Introduction

This paper is concerned with the following classic conundrum. Suppose 51% of the populace favors candidate A over B but only very slightly, while 49% favors B over A but very strongly. Majority voting will elect A but a utilitarian planner would choose B since the near-indifference of A voters would be weighed against the strong preferences of B voters. The reason, of course, is that majority voting does not take intensity of preferences into account. Put another way, majority voting follows the preferences of the *median* voter while a utilitarian planner would follow the preferences of the *mean* voter.

The argument given above assumes implicitly that all members of the population vote. In this paper we show that the conundrum is completely resolved if voting is costly and participation is voluntary. In this setting, participation rates will adjust to reflect intensity of preferences and to such an extent that the outcome of majority voting will, in large elections, coincide with the utilitarian outcome. This is true even in extreme versions of the example above. Suppose now that 90% of the populace favors A over B but the preferences the other 10% are at least 9 times as strong as those of A voters. Our main result concludes that, in large elections, B voters will

*Penn State University, E-mail: vkrishna@psu.edu
†University of California, Berkeley, E-mail: rjmorgan@haas.berkeley.edu
‡Please do not quote or use without permission.

show up to vote at a rate that is at least 9 times as large as the show-up rate of A voters. Thus B will be elected nevertheless.

This paper establishes the following results:

1. When voting is costly and voluntary, in large elections, majority rule is utilitarian (Theorem 1).

2. Among all supermajority rules, *only* majority rule is utilitarian in large elections (Theorem 2).

It is then natural to conjecture that supermajority rules maximize a *weighted* utilitarian welfare function. This is true but supermajority rules exaggerate the inherent bias. For instance, the 2/3 supermajority rule, under which A must obtain *twice* as many votes as B in order to win, maximizes a weighted welfare function in which the preferences of B voters are given *four* times as much weight as those of A voters.

3. The utilitarian property of majority rule is robust to the introduction of aggregate uncertainty (Theorem 3).

The third result is noteworthy because the introduction of aggregate uncertainty is known to erode other welfare properties of majority rule. For instance, it substantially weakens the information aggregation properties that form the basis of the celebrated Condorcet Jury Theorem (see Mandler, 2012).

2 The Model

There are two candidates, named A and B, who are competing in an election decided by majority voting with ties resolved by the toss of a fair coin.[1]

The size of the electorate is a random variable N which is distributed on $\mathbb{Z}_+ = \{0, 1, 2, ...\}$ according to the probability distribution function π^*. Thus the probability that there are exactly m eligible voters (or *citizens*) is $\pi^*(m)$. We suppose that N has a finite expectation, say n.

Voter types are determined as follows. First, with probability $\lambda \in (0, 1)$ a voter is determined to be an A supporter and with probability $1-\lambda$, a B supporter. Next, each A supporter draws a value v from the distribution G_A over $[0, 1]$ which measures the intensity of preference—the value of electing A over B. Similarly, each B supporter draws a value v from the distribution G_B over $[0, 1]$ which is the value of electing B over A. The combination of the direction of a voter's preferences and its intensity will be referred to as his *type*. Types are distributed independently across voters and independently of the number of voters.[2]

Types are private information. A citizen knows his own type and that the types of the others are distributed according to λ, G_A and G_B. Unless π^* is degenerate,

[1] Supermajority rules are considered in Section 7.

[2] In Section 8, we extend the basic model to allow for aggregate uncertainty about λ, the ex ante proportion of A supporters.

a voter does not know the realized number of potential voters, but rather only that N is distributed according to π^*. Of greater interest to an individual voter is the distribution of the number of *other* voters given that at least one voter (herself) is present. Let $\pi(m) = \Pr[N-1 = m \mid N \geq 1]$ denote the probability that a given voter assigns to the event that there are exactly m other voters present.

2.1 Utilitarianism

The expected welfare of A supporters from electing A over B is

$$v_A = \int_0^1 v dG_A(v)$$

and similarly, the expected welfare of B supporters from electing B over A is

$$v_B = \int_0^1 v dG_B(v)$$

Since the probability that a voter is an A supporter is λ and that she is a B supporter is $1-\lambda$, (ex ante) *utilitarian* welfare is higher from electing A rather than B if and only if

$$\lambda v_A > (1-\lambda) v_B$$

We will say that a voting rule is *utilitarian* if the outcome maximizes utilitarian welfare.

2.2 Pivotal Events

Since a voter's payoff depends on which candidate is elected, a key determinant of voting decisions is the chance that the votes of others lead to a situation where her individual vote is decisive. In making this determination, a voter needs to consider the set of decisive events. An *event* is a pair of vote totals (j,k) such that there are j votes for A and k votes for B. An event is *pivotal* for A if a single additional vote for A will affect the outcome of the election. We denote the set of such events by Piv_A. One additional vote for A makes a difference only if either (i) there is a tie; or (ii) A has one vote less than B. Let $T = \{(k,k) : k \geq 0\}$ denote the set of ties and let $T_{-1} = \{(k-1,k) : k \geq 1\}$ denote the set of events in which A is one vote short of a tie. Similarly, Piv_B is defined to be the set of events which are pivotal for B. This set consists of the set T of ties together with events in which A has one vote more than B. Let $T_{+1} = \{(k,k-1) : k \geq 1\}$ denote the set of events in which A is ahead by one vote.

Suppose that voting behavior is such that, *ex ante*, each voter casts a vote for A with probability q_A and a vote for B with probability q_B. Then $q_0 = 1 - q_A - q_B$ is the probability that a voter abstains. Fix a voter, say 1. Consider an event where the number of *other* voters is exactly m and from these, there are k votes in favor of

A and l votes in favor of B. The remaining $m-k-l$ voters abstain. If voters make decisions independently, the probability of this event is

$$\mathbb{P}\left[(k,l)\mid m\right]=\binom{m}{k,l}\left(q_A\right)^k\left(q_B\right)^l\left(q_0\right)^{m-k-l}$$

where

$$\binom{m}{k,l}=\binom{m}{k+l}\binom{k+l}{k}$$

denotes the trinomial coefficient.[3] For a fixed m, the probability of a tie—that is, the probability of events of the form (k,k)

$$\mathbb{P}\left[T\mid m\right]=\sum_{k=0}^{m}\binom{m}{k,k}\left(q_A\right)^k\left(q_B\right)^k\left(q_0\right)^{m-2k} \tag{1}$$

This probability can be rewritten as

$$\mathbb{P}\left[T\mid m\right]=\left[\frac{1}{m}\sum_{r=0}^{m-1}\left(\omega^r q_A+\omega^{-r}q_B+q_0\right)^m\right]-\left(q_A\right)^m-\left(q_B\right)^m \tag{2}$$

where $\omega=\exp\left(2\pi i/m\right)$ is a primitive mth root of unity, an expression that is very useful when we study the asymptotic properties of these probabilities (see Appendix A for a derivation). The probability of a tie from the perspective of a single voter facing an electorate of uncertain size is then simply

$$\Pr\left[T\right]=\sum_{m=0}^{\infty}\pi\left(m\right)\mathbb{P}\left[T\mid m\right]$$

Similarly, for fixed m, the probability that A falls one vote short is

$$\mathbb{P}\left[T_{-1}\mid m\right]=\sum_{k=1}^{m}\binom{m}{k-1,k}\left(q_A\right)^{k-1}\left(q_B\right)^k\left(q_0\right)^{m-2k+1} \tag{3}$$

which can be rewritten as

$$\mathbb{P}\left[T_{-1}\mid m\right]=\left[\frac{1}{m}\sum_{r=0}^{m-1}\omega^r\left(\omega^r q_A+\omega^{-r}q_B+q_0\right)^m\right]-m\left(q_A\right)^{m-1}q_0 \tag{4}$$

(again, see Appendix A) and, from the perspective of a single voter, the overall probability that A falls one vote short is

$$\Pr\left[T_{-1}\right]=\sum_{m=0}^{\infty}\pi\left(m\right)\mathbb{P}\left[T_{-1}\mid m\right]$$

The probabilities $\mathbb{P}\left[T_{+1}\mid m\right]$ and $\Pr\left[T_{+1}\right]$ are analogously defined.

[3] We follow the convention that if $m<k+l$, then $\binom{m}{k+l}=0$ and so $\binom{m}{k,l}=0$, as well.

Let Piv_A be the set of events where one additional vote for A is decisive. Then,

$$\Pr\left[Piv_A\right] = \tfrac{1}{2}\Pr\left[T\right] + \tfrac{1}{2}\Pr\left[T_{-1}\right]$$

where the coefficient $\frac{1}{2}$ arises since, in the first case, the additional vote for A breaks a tie while, in the second, it leads to a tie. Likewise, define Piv_B to be the set of events where one additional vote for B is decisive. Hence,

$$\Pr\left[Piv_B\right] = \tfrac{1}{2}\Pr\left[T\right] + \tfrac{1}{2}\Pr\left[T_{+1}\right]$$

Our first, rather simple, proposition establishes that, when other voters are more likely to choose B than A, then casting an A vote is more likely to be decisive. Conversely, when an A vote is more likely to be decisive, then it must be that other voters are more likely to vote for B than for A. Since the main benefit from voting occurs in casting a decisive vote for the preferred candidate, the proposition embodies a kind of "underdog effect"—a vote for a candidate who is behind is more valuable than a vote for the candidate who is ahead.

Proposition 1 $\Pr\left[Piv_A\right] > \Pr\left[Piv_B\right]$ *if and only if* $q_A < q_B$.

Proof. Note that

$$\Pr\left[Piv_A\right] - \Pr\left[Piv_B\right] = \tfrac{1}{2}\left(\Pr\left[T_{-1}\right] - \Pr\left[T_{+1}\right]\right)$$

and since

$$\begin{aligned} q_A \Pr\left[T_{-1}\right] &= \sum_{m=0}^{\infty} \pi\left(m\right) \sum_{k=0}^{m} \binom{m}{k, k+1} \left(q_A\right)^{k+1} \left(q_B\right)^{k+1} \left(q_0\right)^{m-2k-1} \\ &= q_B \Pr\left[T_{+1}\right] \end{aligned}$$

$\Pr\left[T_{-1}\right] > \Pr\left[T_{+1}\right]$ if and only if $q_A < q_B$. ■

3 Costly Voting

We suppose that voting is costly and voters can abstain if they wish. A citizen's cost of voting is private information and determined by an independent realization from a continuous probability distribution F with strictly positive density over the support $[0, 1]$. Finally, we assume that voting costs are independent of the type and the number of voters. Thus, prior to the voting decision, each citizen has two pieces of private information—his type and his cost of voting.

In this section, we will show that there exists an equilibrium to the voting game where both A and B supporters participate at strictly positive rates and vote for their preferred candidate.

Among those who show up at the polls, voting behavior is very simple—A supporters vote for A and B supporters for B. For both, voting for their preferred candidate is a weakly dominant strategy. Thus it only remains to consider the participation behavior of voters.

We will study *type-symmetric* equilibria. In these equilibria, all voters of the same type and same realized cost follow the same strategy. Myerson (1998b) has shown that in voting games with population uncertainty, *all* equilibria are type-symmetric.[4] Thus, when we refer to equilibrium, we have in mind type-symmetric equilibrium.

Formally, an equilibrium consists of two functions $c_A(v)$ and $c_B(v)$ such that (i) an A supporter (resp. B supporter) with cost c votes if and only if $c < c_A$ (resp. $c < c_B$); (ii) the participation rates $p_A(v) = F(c_A(v))$ and $p_B(v) = F(c_B(v))$ are such that the resulting pivotal probabilities make an A supporter (resp. B supporter) with value v and costs $c_A(v)$ (resp. $c_B(v)$), indifferent between voting and abstaining. An equilibrium is thus defined by the equations: for all v,

$$\begin{aligned} c_A(v) &= v \Pr[Piv_A] \\ c_B(v) &= v \Pr[Piv_B] \end{aligned}$$

If we denote by $p_A(v)$ (resp. $p_B(v)$) the probability that an A supporter (resp. B supporter) with value v will vote, then the equilibrium conditions become

$$\begin{aligned} F^{-1}(p_A(v)) &= v \Pr[Piv_A] \\ F^{-1}(p_B(v)) &= v \Pr[Piv_B] \end{aligned}$$

Equivalently,

$$\begin{aligned} p_A(v) &= F(v \Pr[Piv_A]) \\ p_B(v) &= F(v \Pr[Piv_B]) \end{aligned}$$

Integrating the function $p_A(v)$ over $[0,1]$ determines p_A, the ex ante probability that a given voter will vote for A. Similarly, integrating $p_B(v)$ over $[0,1]$ determines p_B, the ex ante probability that a given voter will vote for B. Thus, we have that in a costly voting equilibrium

$$\begin{aligned} p_A &= \int_0^1 F(v \Pr[Piv_A])\, dG_A(v) \\ p_B &= \int_0^1 F(v \Pr[Piv_B])\, dG_B(v) \end{aligned}$$

It is useful to formulate these in terms of the voting propensities—the ex ante probability of a vote for a particular candidate, that is, $q_A = \lambda p_A$ and $q_B = (1-\lambda) p_B$. In terms of voting propensities, the equilibrium conditions are

$$q_A = \lambda \int_0^1 F(v \Pr[Piv_A])\, dG_A(v) \tag{5}$$

$$q_B = (1-\lambda) \int_0^1 F(v \Pr[Piv_B])\, dG_B(v) \tag{6}$$

As in Ledyard (1984), it is now straightforward to establish:

[4] For the degenerate case where the number of eligible voters is fixed and commonly known, type asymmetric equilibria may arise; however, such equilibria are not robust to the introduction of even a small degree of uncertainty about the number of eligible voters.

Proposition 2 *With costly voting, there exists an equilibrium. In every equilibrium, all types of voters participate with a probability strictly between zero and one.*

Proof. Since both $\Pr[Piv_A]$ and $\Pr[Piv_B]$ are continuous functions of q_A and q_B, Brouwer's Theorem ensures that there is a solution $(q_A, q_B) \in [0,1]^2$ to (5) and (6). Let p_A and p_B be the corresponding expected participation rates.

First, note that neither p_A nor p_B can equal 1. If $p_A = 1$, say, then it must be that for all v, $p_A(v) = 1$ and hence for all v, $c_A(v) = 1$ as well. But the benefits from voting for A for a voter with value v cannot exceed v and so this is impossible. Second, neither p_A nor p_B can equal 0. Suppose to the contrary that $p_A = 0$, say. Then from the perspective of an A supporter there is a strictly positive probability that no one else shows up. To see this note, that if the realized number of other voters is m then there is a probability λ^m that all of these are A supporters. Thus $\Pr[Piv_A] > 0$ and hence so for all v, $c_A(v) > 0$ and, in turn, $p_A(v) > 0$ as well. ■

3.1 Uniform Costs

It is useful to first consider the case where the costs are *uniformly* distributed, that is, $F(c) = c$. In this case, the equilibrium conditions (5) and (6) can be rewritten as

$$\begin{aligned} q_A &= \lambda \Pr[Piv_A] \int_0^1 v dG_A(v) = \lambda v_A \Pr[Piv_A] \\ q_B &= (1-\lambda) \Pr[Piv_B] \int_0^1 v dG_B(v) = (1-\lambda) v_B \Pr[Piv_B] \end{aligned}$$

where v_A is the expected welfare of an A supporter from electing A rather than B and v_B is the expected welfare of a B supporter from electing B rather than A.

Notice that if we rewrite these expressions as a ratio and multiply each side by n, then we have

$$\frac{nq_A}{nq_B} = \frac{\lambda v_A}{(1-\lambda) v_B} \times \frac{\Pr[Piv_A]}{\Pr[Piv_B]} \tag{7}$$

The left-hand side of this expression is simply the ratio of the expected number of A versus B votes. The first term on the right-hand side is the ratio of the welfare from choosing A versus B. When electing A maximizes utilitarian welfare, this expression is more than 1 whereas it is fractional when B is the utilitarian choice. The last term is the ratio of the pivot probabilities. Our next proposition shows that the vote ratio and the welfare ratio must both lie on the same side of 1, which implies that A receives more votes than B (in expectation) when it is the utilitarian choice and vice-versa when it is not. The key to the result comes from Lemma 1, which highlighted that a vote for the trailing candidate is more likely to be pivotal. Thus, if A were trailing when it is the utilitarian choice, both expressions on the right-hand side would exceed 1, which is a contradiction. We formalize this idea in the following proposition:

Proposition 3 *Suppose voting costs are uniformly distributed. The expected number of votes for A exceeds the expected number of votes for B if and only if electing A maximizes utilitarian welfare. Precisely, $q_A > q_B$ if and only if $\lambda v_A > (1-\lambda) v_B$.*

Proof. Suppose to the contrary that $\lambda v_A > (1-\lambda) v_B$ but $q_A \leq q_B$. Proposition 1 implies that $\Pr[Piv_A] \geq \Pr[Piv_B]$ and so we have that the left-hand side of (7) is no greater than one, whereas the right-hand side is greater than one. This is a contradiction and so we must have $q_A > q_B$. A similar argument implies that the converse holds as well. ■

4 Large Elections

Proposition 3 shows that the vote ratio mirrors the welfare ratio; that is, A is more likely to win than B when it maximizes utilitarian welfare and vice-versa. In this section, we will show that, in large elections,the probability that A wins when it maximizes utilitarian welfare goes to 1, and likewise for B when it maximizes utilitarian welfare. While Proposition 3 held when costs are uniformly distributed, the limit results hold regardless of the distribution of costs, the distribution of the number of potential voters, and the distribution of types.

Before proceeding, it helps to define precisely what we mean by *large* elections. Formally, consider a family of population distributions π_n^* such that for each n (i) the expected size of the population according to π_n^* is finite and equals n; (ii) for all K,

$$\lim_{n\to\infty} \sum_{m=K}^{\infty} \pi_n^*(m) = 1$$

The second property requires that, for large n, the distribution π^* places almost all the weight on large populations. In what follows, we will consider a sequence of such π_n^* distributions and when we speak of a "large election" we mean that n is large.

The requirements above are very weak. Indeed, most commonly used families of distributions satisfy these properties. Obviously, if the number of potential voters is a fixed size, n, the above properties hold. Likewise, voting populations drawn from Poisson, Negative Binomial, or Geometric also have the required properties.

Consider a sequence of equilibria, one for each n. Let $p_A(n)$ and $p_B(n)$ be the sequence of equilibrium participation rates of A supporters and B supporters, respectively. The following proposition says that in large elections, these participation rates tend to zero but at a rate slower than $1/n$ so that the expected number of voters of each type is unbounded.

Proposition 4 *In any sequence of equilibria, the participation rates $p_A(n)$ and $p_B(n)$ tend to zero, while the expected number of voters $np_A(n)$ and $np_B(n)$ tend to infinity.*

Proof. See Lemmas A.7 and A.8. ■

The intuition for the first part of the result seems straightforward. If the participation rates did not go to zero so that even in the limit, voters participate with positive probability, then no single voter would be pivotal. Thus, there would be no incentive to vote, contradicting the hypothesis that there was positive participation in the limit.

But how do we know that the probability of being pivotal goes to zero in the limit? As m increases, the number of ways a tie can occur increases and, because of the combinatorics, a term-by-term comparison of the sums in (1) and (3) is inconclusive. Instead, we use the formulae in (2) and (4) to show that the participation rates indeed converge to zero.

The second part of the proposition asserts that, despite the fact the participation *rates* go to zero, the expected number of A and B voters is unbounded. The basic intuition is that, if there were a finite number of expected voters, then the probability that a voter is pivotal (and hence the benefit from voting) would be strictly positive. But this would imply strictly positive participation rates in the limit. More care is needed in circumstances where there are an unbounded number of A voters (say) and a bounded number of B voters. The key insight here is that the benefits from voting are greater for B voters than for A voters since they are more likely to be pivotal. This then implies that B voters participate at higher rates, which is a contradiction.

Our next proposition shows that, given the limiting turnouts and participation rates, in large elections the candidate with the higher vote propensity wins with probability approaching one. In particular, if a random voter is more likely to vote for A than to vote for B, then A will be elected with near certainty in large elections. Were the vote propensities fixed, the result would follow simply as a consequence of the law of large numbers. The subtlety is that the vote propensities change with the expected number of voters and go to zero in the limit. However, since there are an infinite number of A and B voters (of the same order of magnitude), then differing vote propensities imply that the expected vote difference becomes unbounded in the limit. But this is not enough to argue that, in fact, the leading candidate will win with probability one. To make this claim, one needs to show that the variability in the vote difference is small relative to the expected vote difference. We show that this is indeed the case in the next result. Formally,

Proposition 5 *If for large n, $q_A(n) > q_B(n)$ then* $\lim_{n\to\infty} \Pr[A \text{ wins}] = 1$.

Proof. If we denote by T_{-k} the event that the number of votes for B less the number of votes for A is k, then

$$\begin{aligned}\mathbb{P}[B \text{ wins} \mid m] &= \tfrac{1}{2}\mathbb{P}[T \mid m] + \sum_{k=1}^{m} \mathbb{P}[T_{-k} \mid m] \\ &\leq \sum_{k=0}^{m} \mathbb{P}[T_{-k} \mid m]\end{aligned}$$

If the population were distributed according to a Poisson distribution with mean m, then the probability of T_{-k} is

$$\mathcal{P}[T_{-k} \mid m] = \sum_{l=0}^{\infty} e^{-m(q_A+q_B)} \frac{(mq_A)^l}{l!} \frac{(mq_B)^{l+k}}{(l+k)!}$$

When m is large, we know that[5]

$$\mathcal{P}\left[T_{-k} \mid m\right] \approx \frac{e^{-m\left(\sqrt{q_A}-\sqrt{q_B}\right)^2}}{\sqrt{4\pi m\sqrt{q_A q_B}}}\left(\sqrt{\frac{q_B}{q_A}}\right)^k$$

and so the probability that B wins calculated in the Poisson model when m is large is

$$\mathcal{P}\left[B \text{ wins} \mid m\right] \leq \sum_{k=0}^{\infty} \mathcal{P}\left[T_{-k} \mid m\right] = \frac{e^{-m\left(\sqrt{q_A}-\sqrt{q_B}\right)^2}}{\sqrt{4\pi m\sqrt{q_A q_B}}}\frac{1}{1-\sqrt{\frac{q_B}{q_A}}}$$

We know from Roos (1999) that the probability $\mathbb{P}\left[S \mid m\right]$ of *any* event $S \subset \mathbb{Z}_+^2$ in the multinomial model with population m is well-approximated by the corresponding probability $\mathcal{P}\left[S \mid m\right]$ in the Poisson model with expected population m (see Appendix D). In particular,

$$\left|\mathbb{P}\left[B \text{ wins} \mid m\right] - \mathcal{P}\left[B \text{ wins} \mid m\right]\right| \leq q_A + q_B$$

As $m \to \infty$,

$$\mathcal{P}\left[B \text{ wins} \mid m\right] \leq \frac{e^{-m\left(\sqrt{q_A}-\sqrt{q_B}\right)^2}}{\sqrt{4\pi m\sqrt{q_A q_B}}}\frac{1}{1-\sqrt{\frac{q_B}{q_A}}} \to 0$$

and since in large elections,

$$\lim_{n\to\infty} \sum_{m=K}^{\infty} \pi_n\left(m\right) = 1$$

we have that

$$\Pr\left[B \text{ wins}\right] \to 0$$

This completes the proof. ■

5 Main Result

We can now state our main result.

Theorem 1 *In large elections with costly voting, majority rule produces utilitarian outcomes with probability one.*

Notice that the theorem does not make any demands on the distribution of voter types. For instance, in circumstances where 90% of voters favor B but where the 10% favoring A feel much more strongly about their candidate than B supporters do about their candidate—precisely, at least 9 times as much—majority voting will produce sufficient enthusiasm among A voters, and sufficient apathy among B voters, that candidate A will prevail. Given the ordinal nature of majority rule, this is quite remarkable. Of course, the key is voluntary participation—voters vote with their "feet" as well as with their ballots, thereby registering, not just the direction, but the intensity of their preferences as well. This produces the (socially) correct outcome.

[5] We write $x_n \approx y_n$ to denote that $\lim_{n\to\infty}\left(x_n/y_n\right) = 1$.

6 Uniqueness of Equilibrium

Ledyard (1984) conjectured that if the population of voters were fixed and commonly known, then there was a unique solution to equilibrium conditions. Our next result shows that uniqueness certainly obtains in large elections, provided that the population follows a Poisson distribution.

Proposition 6 *Suppose that the population is distributed according to a Poisson distribution. Then, in large elections, there is a unique equilibrium.*

Proof. See Appendix . ■

7 Supermajority Rules

We have shown that in large elections with costly voting, majority rule produces the same outcomes as demanded by utilitarianism. The mechanism driving this result is, of course, turnout. Even if one candidate enjoys the support of a majority of the population, turnout always favors the candidate whose election would maximize utilitarian welfare and to a sufficient extent that, in large elections, the welfare-maximizing candidate is elected with probability close to one. But do other voting rules also share this property? For instance, suppose that B is the default candidate (if this is a referendum, then the default policy). Suppose that in order for A to win, 2/3 of those voting must favor A . Is it the case that, in large elections, A will win if and only if it is the welfare-maximizing candidate?

In this section, we will consider *supermajority* voting rules, defined as follows. Candidate B is the default alternative and A needs a fraction $\mu \geq \frac{1}{2}$ of the votes cast in order to be elected. We will assume that μ is a rational number and so will write $\mu = a/(a+b)$, where a and b are positive integers which are relatively prime (have no common factors) and such that $a \geq b$. Note that for majority rule, $a = b = 1$ whereas, say for the two-thirds supermajority rule, $a = 2$ and $b = 1$.

In the event of a "tie"—that is, a situation in which A obtains exactly μ proportion of the votes—the winner is chosen at random. In particular, A is chosen with probability $t \in [0,1]$ and B is chosen with probability $1-t$. A case of special interest is one where $t = 0$. In this case, A must obtain one vote more than a μ-majority to win.

The main analytic result of this section is the following:

Proposition 7 *Suppose that the population is distributed according to a Poisson distribution. In a large $\frac{\alpha}{a+b}$ supermajority election, if*

$$\frac{\lambda v_A}{(1-\lambda)\, v_B} > \left(\frac{a}{b}\right)^2$$

then A is elected with probability close to one. If the reverse inequality holds strictly, then B is elected with probability close to one.

Proof. See Appendix B. ■

When $a > b$, supermajority rules of course bias outcomes in favor of B. Proposition 7 quantifies the extent of the bias. It says that the $a/(a+b)$ supermajority rule maximizes a welfare function in which the utilities of B supporters are given a weight $(a/b)^2$ relative to the utilities of A supporters. For instance, the 2/3 supermajority rule—which requires A to obtain *twice* as many votes as B—is appropriate only if the weight placed by society on the welfare of B supporters is *four* times that placed on the welfare of A supporters. Why do supermajority rules have the "squaring property?" The key is Lemma B.1 which is a generalization, for large Poisson populations, of Proposition 1. This lemma shows again that in large Poisson elections, if A is on the losing side, that is, the ratio of voting propensities, q_A/q_B falls short of the a/b ratio that is required by the voting rule, then the ratio of the pivotal probabilities $\Pr[Piv_A]/\Pr[Piv_B]$ exceeds b/a. For instance, if A falls short of the $2:1$ ratio needed under the 2/3rds rule, then the relative likelihood of A being pivotal is greater than 1/2. The converse also holds.

To see that this implies the squaring property in utility weights, recall that, when costs are uniformly distributed, then the equilibrium conditions imply:

$$\frac{\lambda v_A}{(1-\lambda)v_B}\frac{\Pr[Piv_A]}{\Pr[Piv_B]} = \frac{q_A}{q_B}$$

Consider the 2/3 supermajority rule. Suppose that the welfare ratio, $\lambda v_A/(1-\lambda)v_B$, exceeds 4. Were A to lose this election then the ratio of the expected number of votes $q_A/q_B < 2$. From Lemma B.1, the pivot ratio then exceeds 1/2 and since the welfare ratio exceeds 4, this is a contradiction. Therefore, A must win in these circumstances. Similarly, suppose the welfare ratio is smaller than 4 but still favors A on utilitarian grounds. Were A to win, then the pivot ratio would be less than 1/2 and hence again we have a contradiction since the product of the pivot and welfare ratios would be smaller than 2 whereas A would need a vote share ratio exceeding 2 to secure a victory. Thus A must lose in these circumstances.

An immediate consequence of Proposition 7 is:

Theorem 2 *Among all supermajority rules, only majority rule produces utilitarian outcomes in large elections.*

When voters have pure common values and voting is compulsory, the main result, due to Feddersen and Pesendorfer (1998) is that the voting rule is irrelevant. Voters will simply adjust their voting decisions to "undo" the bias of the rule and come to the correct choice. One might have conjectured that such an adjustment process would also occur with voluntary and costly voting—voters adjust their turnout decisions to undo the particular voting rule and hence obtain the utilitarian outcome. Theorem 2 shows that this is not the case. While it is true that turnouts adjust depending on the voting rule, permitting voters to express their intensity of preferences by voting with their feet is not enough to guarantee good outcomes. The voting rule itself plays a crucial role. In particular, supermajority rules, in effect, place "too much" weight on the status quo.

8 Aggregate Uncertainty

There are many random elements in the model considered so far. The size of the electorate (N distributed as π^*) itself is random. For each potential voter, the direction of his or her preference (A with probability λ and B with probability $1-\lambda$) is random, as is the intensity of preference (v distributed as G_A or G_B). Finally, the costs of voting (c distributed as F) are random as well. In equilibrium, these elements together serve to determine the key factor in individual voters' decisions—the pivot probabilities $\Pr[Piv_A]$ and $\Pr[Piv_B]$.

In this section, we amend the model to allow for aggregate uncertainty about the parameter λ, the fraction of A supporters. In other voting models, Mandler (2012) has shown that aggregate uncertainty about some parameter of interest can overturn welfare results concerning majority rule. In particular, as discussed in more detail below, when there is aggregate uncertainty, the Condorcet Jury Theorem is weakened by the presence of equilibria in which information does not aggregate.

In our model, aggregate uncertainty concerning λ is of a fundamentally different nature than uncertainty about the other elements of the model. The other random elements are all independently distributed and as a result, when λ is fixed and commonly known, voters' beliefs about the aggregate behavior of the population are identical. Specifically, the voting propensities q_A, q_B and q_0 that an A voter uses to calculate $\Pr[Piv_A]$ are the same as those that a B supporter uses to calculate $\Pr[Piv_B]$. But when there is uncertainty about λ, an A supporter will hold different beliefs about λ than those that a B supporter will hold. Precisely, suppose that λ is distributed according to a continuous density h on $[0,1]$ with mean $\overline{\lambda}$. The posterior density of λ assigned by an A supporter is

$$h_A(\lambda) = \frac{h(\lambda)\lambda}{\int_0^1 h(\theta)\theta d\theta} = h(\lambda)\frac{\lambda}{\overline{\lambda}} \tag{8}$$

while the posterior density assigned by a B supporter is

$$h_B(\lambda) = \frac{h(\lambda)(1-\lambda)}{\int_0^1 h(\theta)(1-\theta)d\theta} = h(\lambda)\frac{1-\lambda}{1-\overline{\lambda}} \tag{9}$$

Thus, as is natural, the posterior beliefs of A supporters put more weight on higher values of λ while those of B supporters put more weight on smaller values of λ. This in turn means that their posterior distributions over the voting propensities $q_A = \lambda p_A$, $q_B = (1-\lambda)p_B$ and $q_0 = 1 - q_A - q_B$ will differ as well.

Once there is uncertainty about λ, voters base their participation decisions on the *expected* pivot probabilities, calculated according to their posterior beliefs h_A and h_B. Our goal here is to explore how majority rule fares under these circumstances. Once again, we are interested in large elections.

Two facts are key to our analysis.[6] First, when n is large, the Poisson pivot probabilities $\mathcal{P}[Piv_A \mid \lambda]$ and $\mathcal{P}[Piv_B \mid \lambda]$, viewed as functions of λ, are single-

[6] In what follows, the analysis is simpler in the Poisson setting and so in this section, we assume that the population follows a Poisson distribution with mean n.

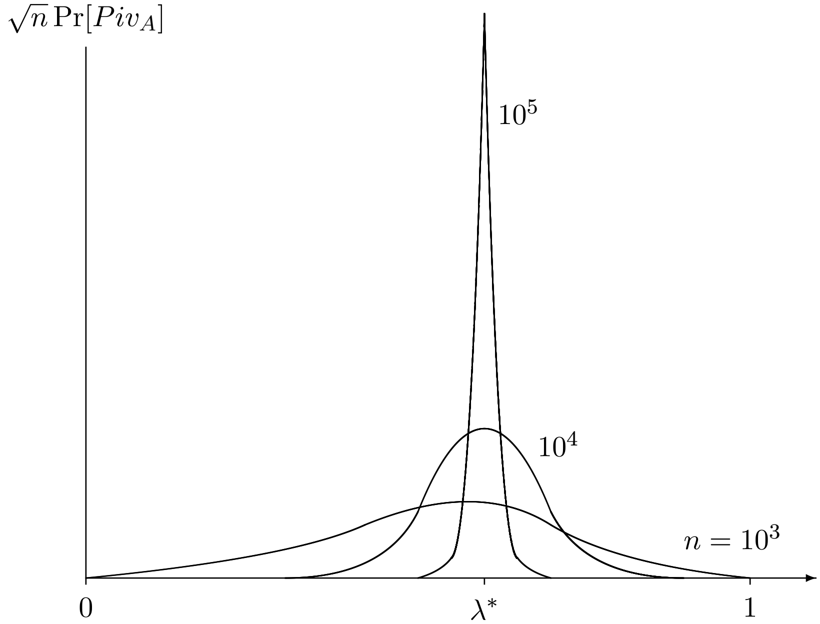

Figure 1: Asymptotic Behavior of Pivot Probabilities

peaked. Second, for large n, *both* pivot probabilities are maximized close to a *critical* value λ^* (determined by the asymptotic behavior of the participation rates) and "spike" in a neighborhood around this critical value—for all $\lambda \neq \lambda^*$, the ratio $\mathcal{P}\left[Piv_A \mid \lambda\right] / \mathcal{P}\left[Piv_A \mid \lambda^*\right]$ goes to zero as n increases. The accompanying Figure 1 depicts $\sqrt{n}\mathcal{P}\left[Piv_A \mid \lambda\right]$ as a function of λ for a particular sequence of participation rates, p_A and p_B, and varying electorate sizes. Notice that, when the population of potential voters reaches 10^5, the pivot probabilities close to the critical value, λ^*, overwhelm the pivot probabilities elsewhere.

The fact that $\mathcal{P}\left[Piv_A \mid \lambda\right]$ spikes at λ^* means that for large n, the expected pivot probability $E_\lambda\left[\mathcal{P}\left[Piv_A \mid \lambda\right]\right]$ is determined solely by the values of the pivot probability in a small neighborhood of λ^*. The same is true for $E_\lambda\left[\mathcal{P}\left[Piv_B \mid \lambda\right]\right]$. The study of the asymptotic behavior of such expected probabilities goes back to Bayes himself in the context of the following problem. Suppose a coin with an unknown probability of heads is tossed $2m$ times. Bayes (1763) showed that if the probability of heads, q, had a uniform prior, then the expected probability of a "tie"—that is, m heads and m tails, is

$$\int_0^1 \binom{2m}{m} q^m (1-q)^m \, dq = \frac{1}{2m+1}$$

When m is large, the function $\binom{2m}{m} q^m (1-q)^m$ has a spike at $q^* = \frac{1}{2}$. Using this fact, Good and Mayer (1975) and Chamberlain and Rothschild (1981) showed that if q

had a prior distribution with a continuous density h that is positive on $(0,1)$, then

$$\lim_{m\to\infty} 2m \int_0^1 \binom{2m}{m} q^m (1-q)^m h(q)\, dq = h\left(\tfrac{1}{2}\right) \tag{10}$$

The integral above is, of course, the expected probability of a tie in a majority election with $2m$ voters with full participation (compulsory voting). In this case, q represents the propensity to vote for A and is assumed to have a prior density of h. Notice that the voting propensities do not depend on the number of voters.

Unlike the case of compulsory voting, where participation rates are fixed, when voting is voluntary and costly, participation is determined endogenously. Moreover, the voting propensities $q_A = \lambda p_A$, $q_B = (1-\lambda) p_B$ and $q_0 = 1 - q_A - q_B$ vary with n, the expected size of the electorate. Hence, we require an analogous result to the one above that takes account of this dependence. Proposition 8 generalizes, in the Poisson model, equation (10) to situations where participation is endogenous. Note that compulsory voting, i.e., $p_A = p_B = 1$, falls out as a special case of Proposition 8 since in that case, $\lambda^* = \frac{1}{2}$.

Proposition 8 *Suppose that there is a sequence of elections such that* $\frac{p_B}{p_A+p_B} \to \lambda^* \in (0,1)$. *Then for any continuous density* h *that is positive on* $(0,1)$,

$$\lim_{n\to\infty} n(p_A+p_B) \int_0^1 \mathcal{P}[Piv_A \mid n, \lambda]\, h(\lambda)\, d\lambda = h(\lambda^*)$$

and

$$\lim_{n\to\infty} n(p_A+p_B) \int_0^1 \mathcal{P}[Piv_B \mid n, \lambda]\, h(\lambda)\, d\lambda = h(\lambda^*)$$

as well.

Proof. See Appendix C. ■

Proposition 8 highlights the dominant role played by the critical value, λ^*, in terms of the chance of being pivotal. In particular, suppose that the density h has the feature that it puts almost all the probability mass in a small neighborhood of the mean $\overline{\lambda}$ and that λ^* is outside this neighborhood. The proposition says that even though almost all the mass is close to $\overline{\lambda}$, the expected pivotal probability is, in the limit, exclusively determined not by the most likely value $\overline{\lambda}$ but rather by the very unlikely but critical value λ^*.

This disconnect between the "true" value of λ and the value voters use in determining expected pivot probabilities has the potential to create problems in terms of the welfare properties of majority voting. Mandler (2012), for instance, has shown that in the classic Condorcet jury set-up—compulsory voting, majority rule, and pure common values—the presence of aggregate uncertainty about signal precisions can lead to equilibria in which information does not aggregate. The reason is that there is, as above, a disconnect between "true" signal precision and the "critical" signal precision. Because of this disconnect, the Condorcet Jury Theorem is weakened when there is aggregate uncertainty.

One might suspect that a similar possibility would arise in our setting as well. After all, the crucial factor is the connection between the pivot probabilities and the utilitarian evaluation of the candidates. If, as in the example above, the utilitarian outcome obviously depends crucially on the mean and most likely value of λ, that is $\overline{\lambda}$, while the pivot probabilities do not appear to depend on this at all. Below we show that, despite this disconnect, the critical value λ^*, which is determined endogenously by the participation rates, still contains the information required for the utilitarian outcome to prevail.

Before proceeding with the analysis of equilibrium, it is useful to extend the utilitarian benchmark to the case where there is aggregate uncertainty about the fraction of voters favoring each side. As usual, we study this from an ex ante perspective. A utilitarian social planner will choose candidate A if and only if the expected welfare of A supporters is higher than that of B supporters. In expectation, the fraction of A supporters is $\overline{\lambda}$. Thus, the *expected* welfare from selecting A is higher than the expected welfare from selecting B if and only if

$$\overline{\lambda} v_A > \left(1 - \overline{\lambda}\right) v_B$$

How does majority rule perform in these circumstances? The equilibrium conditions when there is aggregate uncertainty (assuming uniform costs) generalize in the usual way: Voters participate so long as voting costs fall below the benefits from voting. Specifically, the equilibrium cost thresholds (which are equivalent to the participation rates in the uniform-cost case) are:

$$p_A = v_A \int_0^1 \mathcal{P}\left[Piv_A \mid n, \lambda\right] h_A\left(\lambda\right) d\lambda \tag{11}$$

$$p_B = v_B \int_0^1 \mathcal{P}\left[Piv_B \mid n, \lambda\right] h_B\left(\lambda\right) d\lambda \tag{12}$$

When n is large, we can use Proposition 8 to write

$$\frac{p_A}{p_B} \approx \frac{v_A}{v_B} \frac{h_A\left(\lambda^*\right)}{h_B\left(\lambda^*\right)}$$

or equivalently, from (8) and (9),

$$\frac{p_A}{p_B} \approx \frac{v_A}{v_B} \frac{\lambda^*}{1-\lambda^*} \frac{1-\overline{\lambda}}{\overline{\lambda}}$$

and since, in large elections, $\frac{p_A}{p_B} \approx \frac{1-\lambda^*}{\lambda^*}$, we have

$$\frac{\overline{\lambda} p_A}{\left(1-\overline{\lambda}\right) p_B} \approx \sqrt{\frac{\overline{\lambda} v_A}{\left(1-\overline{\lambda}\right) v_B}} \tag{13}$$

Equation (13) shows that the expected vote ratio favors candidate A if and only if the expected welfare ratio also favors A. In other words, even in the presence of aggregate uncertainty, majority rule with voluntary and costly voting implements the utilitarian outcome.

Proposition 9 *In large elections with aggregate uncertainty, the expected number of votes for A exceeds the expected number of votes for B if and only if electing A maximizes utilitarian welfare. Precisely,* $\overline{\lambda} p_A > \left(1-\overline{\lambda}\right) p_B$ *if and only if* $\overline{\lambda} v_A > \left(1-\overline{\lambda}\right) v_B$.

The workings of the proposition may be seen in the following example.

Example 1 *Suppose* $v_A = 1$, $v_B = 4$ *and that the fraction of A supporters,* λ, *is uniformly distributed, that is,* $h(\lambda) = 1$.

Note that for the example, from (8) and (9), the posterior beliefs of A and B voters are $h_A(\lambda) = 2\lambda$ and $h_B(\lambda) = 2(1-\lambda)$, respectively. Now, since $\overline{\lambda} = \frac{1}{2}$, the equilibrium condition (13) implies that for large n,

$$\frac{p_A}{p_B} = \frac{\overline{\lambda} p_A}{\left(1-\overline{\lambda}\right) p_B} \approx \sqrt{\frac{\overline{\lambda} v_A}{\left(1-\overline{\lambda}\right) v_B}} = \frac{1}{2}$$

and so, $p_B/(p_A + p_B) \approx \frac{2}{3}$ and $\lambda^* = \frac{2}{3}$ as well. Proposition 8 then implies that in large elections,

$$\rho \equiv \frac{\int_0^1 \mathcal{P}\left[Piv_B \mid n, \lambda\right] h_B(\lambda)\, d\lambda}{\int_0^1 \mathcal{P}\left[Piv_A \mid n, \lambda\right] h_A(\lambda)\, d\lambda} \approx \frac{h_B(\lambda^*)}{h_A(\lambda^*)} = \frac{1}{2}$$

Figure 2 depicts the equilibrium values of $p_B/(p_A + p_B)$ and ρ as functions of the expected number of voters, n. The rapid convergence of these quantities to their limits is particularly noteworthy.

Theorem 3 *Suppose* H_r *is a sequence of distributions (with continuous densities* h_r*) over* $[0,1]$ *that converges weakly to the distribution* H_0 *which is degenerate at* $\lambda_0 \in (0,1)$. *Then*

$$\lim_{r\to\infty} \lim_{n\to\infty} \Pr\left[A \text{ } wins\right] = 1$$

if and only if $\lambda_0 v_A > (1-\lambda_0) v_B$.

Proof. Suppose $\lambda_0 v_A > (1-\lambda_0) v_B$ and consider some distribution H^r. The equilibrium conditions (11) and (12) imply that the equilibrium participation rates depend on the distribution H^r and not on any particular realization of λ. To make this dependence explicit, let $p_A(n,r)$ and $p_B(n,r)$ denote the participation rates when the expected size of the electorate is n and λ is distributed according to H^r. If $\overline{\lambda}_r$ is the expected value of λ under the distribution H^r, then from (13) we know that for all r,

$$\lim_{n\to\infty} \frac{\overline{\lambda}_r p_A(n,r)}{\left(1-\overline{\lambda}_r\right) p_B(n,r)} = \sqrt{\frac{\overline{\lambda}_r v_A}{\left(1-\overline{\lambda}_r\right) v_B}}$$

Since $\lambda_0 v_A > (1-\lambda_0) v_B$ and $\overline{\lambda}_r \to \lambda_0$, when r is large, the right-hand side of the equality above exceeds one. Thus, there exists an R such that for all $r \geq R$,

$$\lim_{n\to\infty} \frac{\overline{\lambda}_r p_A(n,r)}{\left(1-\overline{\lambda}_r\right) p_B(n,r)} > 1 \tag{14}$$

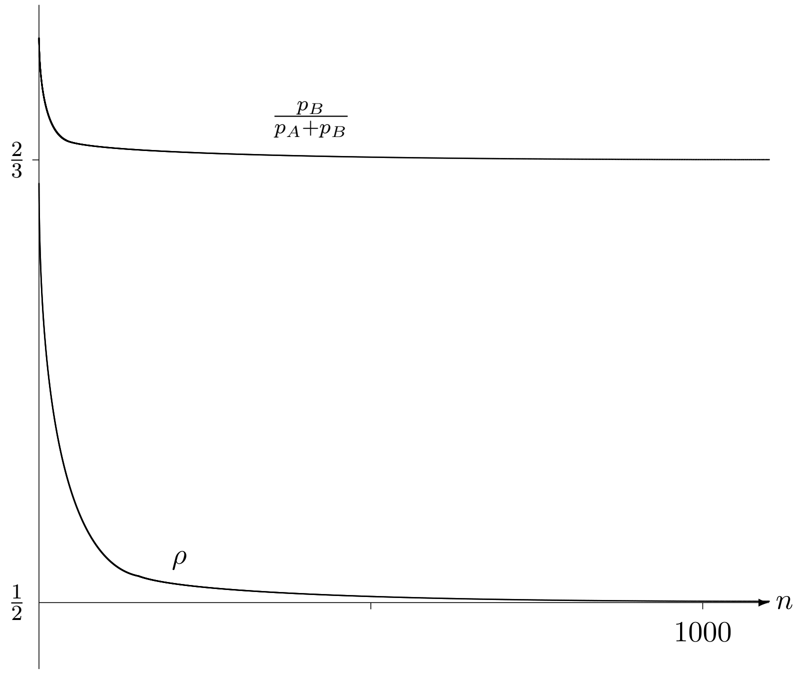

Figure 2: Ratios of Votes and Expected Pivot Probabilities in Example 1

Now, for a particular realization of λ, let $\Pr[A \text{ wins} \mid n, r, \lambda]$ denote the probability that A wins calculated using the voting propensities $q_A = \lambda p_A(n,r)$ and $q_B = (1-\lambda) p_B(n,r)$. Define

$$S_r = \left\{ \lambda \in [0,1] : \lim_{n\to\infty} \Pr[A \text{ wins} \mid n, r, \lambda] = 1 \right\}$$

and

$$S_r' = \left\{ \lambda \in [0,1] : \lim_{n\to\infty} \frac{\lambda p_A(n,r)}{(1-\lambda) p_B(n,r)} > 1 \right\}$$

Proposition 5 then implies that for all $r \geq R$, $S_r' \subseteq S_r$.

Finally, since $H_r \to H_0$ which is degenerate at λ_0, for all ε, there exists an $R' \geq R$, such that for all $r \geq R'$, $\Pr[S_r'] > 1-\varepsilon$ and so $\Pr[S_r] > 1-\varepsilon$, as well. This completes the proof. ■

A Asymptotics

The purpose of this appendix is to provide a proof of Proposition 4. This is done via Lemmas A.7 and A.8 below.

When studying the asymptotic behavior of the pivotal probabilities, it is useful to rewrite these in the form given in (2) and (4). We begin by establishing these "roots of unity" formulae.

A.1 Roots of Unity Formulae

For $m > 1$, let

$$\omega = \exp\left(2\pi i \frac{1}{m}\right)$$

Since $\omega^m = e^{2\pi i} = 1$, ω is an mth (complex) *root of unity.* Note that $\sum_{r=0}^{m-1} \omega^r = (1-\omega^m)\,/\,(1-\omega) = 0$.

Lemma A.1 *For x, y, z positive,*

$$\sum_{k=0}^{m-1} \binom{m}{k,k} x^k y^k z^{m-2k} = \left(\frac{1}{m}\sum_{r=0}^{m-1}\left(\omega^r x + \omega^{-r} y + z\right)^m\right) - (x^m + y^m)$$

Proof. Using the trinomial formula, for $r < m$,[7]

$$\left(\omega^r x + \omega^{-r} y + z\right)^m = \sum_{k=0}^{m}\sum_{l=0}^{m}\binom{m}{k,l}\omega^{rk}x^k\omega^{-rl}y^l z^{m-k-l}$$

and so, averaging over $r = 0, 1, ..., m-1$,

$$\begin{aligned}
\frac{1}{m}\sum_{r=0}^{m-1}\left(\omega^r x + \omega^{-r} y + z\right)^m &= \frac{1}{m}\sum_{r=0}^{m-1}\sum_{k=0}^{m}\sum_{l=0}^{m}\binom{m}{k,l}\omega^{r(k-l)}x^k y^l z^{m-k-l} \\
&= \frac{1}{m}\sum_{k=0}^{m}\sum_{l=0}^{m}\binom{m}{k,l}\left(\sum_{r=0}^{m-1}\omega^{r(k-l)}\right)x^k y^l z^{m-k-l} \\
&= \frac{1}{m}x^m\left(\sum_{r=0}^{m-1}\omega^{rm}\right) + \frac{1}{m}y^m\left(\sum_{r=0}^{m-1}\omega^{-rm}\right) \\
&\quad + \frac{1}{m}\sum_{k=0}^{m-1}\sum_{l=0}^{m-1}\binom{m}{k,l}\left(\sum_{r=0}^{m-1}\omega^{r(k-l)}\right)x^k y^l z^{m-k-l}
\end{aligned}$$

Now observe that since $\omega^m = 1$,

$$\sum_{r=0}^{m-1}\omega^{r(k-l)} = \begin{cases} m & \text{if } k = m \text{ or } l = m \\ m & \text{if } k = l \\ \frac{1-\omega^{m(k-l)}}{1-\omega^{(k-l)}} = 0 & \text{otherwise} \end{cases}$$

Thus,

$$\begin{aligned}
\frac{1}{m}\sum_{r=0}^{m-1}\left(\omega^r x + \omega^{-r} y + z\right)^m &= x^m + y^m + \frac{1}{m}\sum_{k=0}^{m-1}\binom{m}{k,k}\left(\sum_{r=0}^{m-1}1\right)x^k y^k z^{m-2k} \\
&= x^m + y^m + \sum_{k=0}^{m-1}\binom{m}{k,k}x^k y^k z^{m-2k}
\end{aligned}$$

■

[7] Recall the convention that if $m < k + l$, then $\binom{m}{k,l} = 0$.

Lemma A.2

$$\sum_{k=0}^{m-1} \binom{m}{k, k+1} x^k y^{k+1} z^{m-2k-1} = \left(\frac{1}{m} \sum_{r=0}^{m-1} \omega^r \left(\omega^r x + \omega^{-r} y + z \right)^m \right) - m x^{m-1} z$$

Proof. The proof is almost the same as that of Lemma A.1 and is omitted. ∎

The following lemma studies the asymptotic properties of the tie probability when the propensities to vote and abstain remain fixed as n increases.

Lemma A.3 *For x, y, z positive, satisfying $x + y + z = 1$,*

$$\lim_m \frac{1}{m} \sum_{r=0}^{m-1} \left(\omega^r x + \omega^{-r} y + z \right)^m = 0$$

Proof. First, note that since $|\omega^r| = 1 = |\omega^{-r}|$,

$$\begin{aligned} \left| \omega^r x + \omega^{-r} y + z \right| &\leq |\omega^r| \, x + \left| \omega^{-r} \right| y + z \\ &= 1 \end{aligned}$$

As a result, for all K and for all $m \geq K$

$$\begin{aligned} \frac{1}{m} \sum_{r=0}^{m-1} \left(\omega^r x + \omega^{-r} y + z \right)^m &\leq \frac{1}{m} \sum_{r=0}^{m-1} \left| \omega^r x + \omega^{-r} y + z \right|^m \\ &\leq \frac{1}{m} \sum_{r=0}^{m-1} \left| \omega^r x + \omega^{-r} y + z \right|^K \end{aligned}$$

and thus for all K,

$$\begin{aligned} \lim_{m \to \infty} \frac{1}{m} \sum_{r=0}^{m-1} \left(\omega^r x + \omega^{-r} y + z \right)^m &\leq \lim_{m \to \infty} \frac{1}{m} \sum_{r=0}^{m-1} \left| \omega^r x + \omega^{-r} y + z \right|^K \\ &= \lim_{m \to \infty} \frac{1}{m} \sum_{r=0}^{m-1} \left| \exp\left(2\pi i \tfrac{r}{m} \right) x + \exp\left(-2\pi i \tfrac{r}{m} \right) y + z \right|^K \\ &= \int_0^1 \left| \exp(2\pi i t) \, x + \exp(-2\pi i t) \, y + z \right|^K dt \qquad (15) \end{aligned}$$

using the definition of the Riemann integral.

Since

$$\begin{aligned} \left| \exp(2\pi i t) \, x + \exp(-2\pi i t) \, y + z \right| &\leq \left| \exp(2\pi i t) \right| x + \left| \exp(-2\pi i t) \right| y + z \\ &= x + y + z \\ &= 1 \end{aligned}$$

with a strict inequality unless $t = 0$. To see this, first note that the inequality above is strict for $t = \frac{1}{2}$. For all $t \neq 0, \frac{1}{2}$, observe that

$$\begin{aligned} \left|\exp(2\pi it)\,x + \exp(-2\pi it)\,y\right| &= \sqrt{x^2 + y^2 + 2xy\cos(4\pi t)} \\ &< |x + y| \end{aligned}$$

Thus, for all $t \neq 0$, $|\exp(2\pi it)\,x + \exp(-2\pi it)\,y + z| < 1$. Hence, the integral on the right-hand side of (15) is decreasing in K and converges to zero as $K \to \infty$. ■

Lemma A.4 *For x, y, z positive, satisfying $x + y + z = 1$,*

$$\lim_m \frac{1}{m} \sum_{r=0}^{m-1} \omega^r \left(\omega^r x + \omega^{-r} y + z\right)^m = 0$$

Proof. Note that

$$\begin{aligned} \frac{1}{m} \sum_{r=0}^{m-1} \left|\omega^r \left(\omega^r x + \omega^{-r} y + z\right)^m\right| &= \frac{1}{m} \sum_{r=0}^{m-1} |\omega^r| \left|\left(\omega^r x + \omega^{-r} y + z\right)^m\right| \\ &= \frac{1}{m} \sum_{r=0}^{m-1} \left|\left(\omega^r x + \omega^{-r} y + z\right)^m\right| \end{aligned}$$

since $|\omega^r| = 1$. The result now follows by applying the previous lemma. ■

A.2 Asymptotic Participation Rates

Lemma A.5 *Along any sequence of equilibria, either* $\lim p_A = 0$ *or* $\lim p_B = 0$ *(or both).*

Proof. Suppose to the contrary that neither is zero. Then there exists a subsequence such that $\lim p_A(n) = p_A^* > 0$ and $\lim p_B(n) = p_B^* > 0$. Setting $x = \lambda p_A^*$ and $y = (1 - \lambda)\, p_B^*$ in Lemmas A.3 and A.4 implies that, along this subsequence, $\lim_{m\to\infty} \mathbb{P}[Piv_A \mid m] = 0$.

Fix any $\varepsilon > 0$. Then there is a K such that for all $m > K$, $\mathbb{P}[Piv_A \mid m] < \varepsilon$. As a result,

$$\begin{aligned} \Pr[Piv_A] &= \sum_{m=0}^{K} \pi_n(m)\,\mathbb{P}[Piv_A \mid m] + \sum_{m=K+1}^{\infty} \pi_n(m)\,\mathbb{P}[Piv_A \mid m] \\ &< \sum_{m=0}^{K} \pi_n(m) + \varepsilon \sum_{m=K+1}^{\infty} \pi_n(m) \end{aligned}$$

But since $\lim_{m\to\infty} \sum_{m=0}^{K} \pi_n(m) = 0$, for all ε, $\limsup \Pr[Piv_A] < \varepsilon$ and so $\lim \Pr[Piv_A] = 0$.

A similar argument shows that $\lim \Pr[Piv_B] = 0$ as well. But the equilibrium conditions (5) and (6) now imply that along the subsequence, $\lim p_A(n) = 0$ and $\lim p_B(n) = 0$, contradicting the initial supposition. ■

Lemma A.6 *Along any sequence of equilibria,* $0 < \liminf \frac{p_A}{p_B} \leq \limsup \frac{p_A}{p_B} < \infty$.

Proof. Suppose that for some subsequence, $\lim \frac{p_A}{p_B} = 0$. This implies that for all n large enough, along the subsequence, $q_A = \lambda p_A(n) < (1-\lambda) p_B(n) = q_B$ and so from Lemma 1, $\Pr[Piv_A] > \Pr[Piv_B]$. It now follows from the equilibrium conditions: for all v :

$$\begin{aligned} c_A(v) &= v \Pr[Piv_A] \\ c_B(v) &= v \Pr[Piv_B] \end{aligned}$$

that when n is large enough, for all v,

$$c_A(v) > c_B(v)$$

and hence, for all v,

$$p_A(v) = F^{-1}(c_A(v)) > F^{-1}(c_B(v)) = p_B(v)$$

The fact that $\lim \frac{p_A}{p_B} = 0$ implies that $\lim p_A = 0$ and since $p_A = \int_0^1 p_A(v)\, dG_A(v)$, for almost all values of v, $\lim p_A(v) = 0$. Since $p_A(v)$ is continuous in v, we have that for all v, $\lim p_A(v) = 0$. Now because $p_A(v) > p_B(v)$, it is the case that $\lim p_B(v) = 0$ as well. This in turn implies that $\lim c_A(v) = 0 = \lim c_B(v)$.

Thus, along the subsequence, when n is large enough,

$$p_A = \int_0^1 F(c_A(v))\, dG_A(v) \approx \int_0^1 F'(0)\, v \Pr[Piv_A]\, dG_A(v) = F'(0) \Pr[Piv_A]\, v_A$$

Similarly, $p_B \approx F'(0) \Pr[Piv_B]\, v_B$. Thus, for all large n,

$$\frac{p_A(n)}{p_B(n)} \approx \frac{\Pr[Piv_A]}{\Pr[Piv_B]} \frac{v_A}{v_B} > \frac{v_A}{v_B}$$

since $\Pr[Piv_A] > \Pr[Piv_B]$. Since the right-hand side of the inequality above is independent of n, this contradicts the assumption that $\lim \frac{p_A}{p_B} = 0$. ■

Lemma A.7 *In any sequence of equilibria, the participation rates* $p_A(n)$ *and* $p_B(n)$ *tend to zero.*

Proof. Lemmas A.5 and A.6 together complete the proof of Lemma A.7. ■

Lemma A.8 *In any sequence of equilibria, the expected number of voters* $np_A(n)$ *and* $np_B(n)$ *tend to infinity.*

Proof. Suppose to the contrary that there is a sequence of equilibria in which, say, $\lim np_A < \infty$. Lemma A.6 then implies that $\lim np_B < \infty$ as well. First, recall that

$$\Pr[T] = \sum_{m=0}^{\infty} \pi_n(m)\, \mathbb{P}[T \mid m]$$

Second, for all m,

$$|\mathbb{P}[T \mid m] - \mathcal{P}[T \mid m]| \leq q_A + q_B$$

where $\mathcal{P}[Piv_A \mid m]$ is the probability of Piv_A calculated according to a Poisson multinomial distribution with an expected population size of m (see Appendix D). Combining these, we can write

$$\left| \Pr[T] - \sum_{m=0}^{\infty} \pi_n(m) \mathcal{P}[T \mid m] \right| \leq q_A + q_B$$

But if $\lim_{m\to\infty} mq_A = M_A$ and $\lim_{m\to\infty} mq_A = M_B$, then using the formula for tie events using Poisson probabilities,

$$\lim_{m\to\infty} \mathcal{P}[T \mid m] = e^{-M_A - M_B} \sum_{k=0}^{\infty} \frac{(M_A)^k}{k!} \frac{(M_B)^k}{k!} > 0$$

Since for all K, $\lim_{n\to\infty} \sum_{m=K}^{\infty} \pi_n(m) = 1$ and $\lim_{n\to\infty} q_A = 0 = \lim_{n\to\infty} q_B$

$$\lim_{n\to\infty} \Pr[T] > 0$$

and thus $\lim_{n\to\infty} \Pr[Piv_A] > 0$ as well.

But now from the equilibrium conditions it follows that $\lim_{n\to\infty} p_A > 0$, contradicting Lemma A.7. ■

B Supermajority Rules

This appendix provides a proof of Proposition 7 and hence of Theorem 2. Throughout, we assume that the population is *Poisson* distributed with mean n. We will then show that as long as $a > b$, no $\frac{a}{a+b}$ supermajority rule is utilitarian in large elections. Thus we will have shown that Theorem 1 does not extend to general supermajority rules: only majority rule is utilitarian.

Pivot Probabilities As before, an event (j, k) is *pivotal for* A if a single additional vote for A will affect the outcome of the election and denote the set of such events by Piv_A. Given a supermajority rule, the events in Piv_A can be classified into three separate categories:

A1. There is a tie and so a single vote for A will result in A winning. A tie can occur only if number of voters is a multiple of $a + b$. The set of ties is thus

$$T = \{(la, lb) : l \geq 0\} \tag{16}$$

A2. Candidate A is one vote short of a tie. The set of such events is[8]

$$T - (1, 0) = \{(la - 1, lb) : l \geq 1\}$$

[8] Of course, we assume that the number of votes cast is nonnegative, so that the point $(-1, 0)$ is excluded from this set.

A3. A is losing but a single additional vote will result in his winning. For any integer k such that $1 \leq k < b$, events in sets of the form

$$T - (\left\lceil \frac{a}{b}k \right\rceil, k) = \left\{ (la - \left\lceil \frac{a}{b}k \right\rceil, lb - k) : l \geq 1 \right\}$$

have the required property.[9] This is because for any $k < b$ the condition that

$$\frac{la - \left\lceil \frac{a}{b}k \right\rceil}{lb - k} < \frac{a}{b} < \frac{la - \left\lceil \frac{a}{b}k \right\rceil + 1}{lb - k}$$

is equivalent to

$$\left\lceil \frac{a}{b}k \right\rceil > \frac{a}{b}k > \left\lceil \frac{a}{b}k \right\rceil - 1$$

Similarly, events that are *pivotal for B* can also be classified into three categories:

B1. There is a tie and so a single vote for B will result in B winning. This occurs for vote totals in the set T as defined above in (16).

B2. Candidate B is one vote short of a tie. The set of such events is

$$T - (0, 1) = \{(la, lb - 1) : l \geq 1\}$$

B is losing but a single additional vote will result in her winning. For any integer j such that $1 \leq j < a$, events in sets of the form

$$T - (j, \left\lceil \tfrac{b}{a}j \right\rceil) = \left\{ (la - j, lb - \left\lceil \tfrac{b}{a}j \right\rceil) : l \geq 1 \right\}$$

have the required property. This is because for any $j < a$, the condition that

$$\frac{la - j}{lb - \left\lceil \frac{b}{a}j \right\rceil} > \frac{a}{b} > \frac{la - j}{lb - \left\lceil \frac{b}{a}j \right\rceil + 1}$$

is equivalent to

$$\left\lceil \frac{b}{a}j \right\rceil - 1 < \frac{b}{a}j < \left\lceil \frac{b}{a}j \right\rceil$$

(Under majority rule, of course, there are no events of the kind listed in A3. and B3.)

As usual, let q_A be the probability of a vote for A and q_B the probability of a vote for B. Under the $\frac{a}{a+b}$-supermajority rule, the probability of a tie is

$$\mathcal{P}[T] = \sum_{k=0}^{\infty} e^{-nq_A} \frac{(nq_A)^{ka}}{(ka)!} e^{-nq_B} \frac{(nq_B)^{kb}}{(kb)!} \tag{17}$$

[9] $\lceil z \rceil$ denotes the smallest integer greater than z.

Approximations Now suppose that we have a sequence $(q_A(n), q_B(n))$ such that both $nq_A(n) \to \infty$ and $nq_B(n) \to \infty$. Myerson (2000) has shown first that in that case, for large n, the probability of a tie in state a, given in (17), can be approximated as follows:

$$\mathcal{P}[T] \approx \frac{\exp\left((a+b)\left(\frac{nq_A}{a}\right)^{\frac{a}{a+b}}\left(\frac{nq_B}{b}\right)^{\frac{b}{a+b}} - nq_A - nq_B\right)}{\left(2\pi(a+b)\left(\frac{nq_A}{a}\right)^{\frac{a}{a+b}}\left(\frac{nq_B}{b}\right)^{\frac{b}{a+b}}\right)^{\frac{1}{2}}(ab)^{\frac{1}{2}}} \tag{18}$$

Second, Myerson (2000) has also shown that the probability of "offset" events of the form $T-(j,k)$ can be approximated as follows

$$\mathcal{P}[T-(j,k)] \approx \mathcal{P}[T] \times x^{bj-ak} \tag{19}$$

where

$$x = \left(\frac{q_B}{q_A}\frac{a}{b}\right)^{\frac{1}{a+b}}$$

The probabilities of the pivotal events can then be approximated by using (18) and (19):

$$\mathcal{P}[Piv_A] \approx \mathcal{P}[T] \times \left[1-t+tx^b+\sum_{k=1}^{b-1} x^{b\left\lceil \frac{a}{b}k\right\rceil - ak}\right] \tag{20}$$

$$\mathcal{P}[Piv_B] \approx \mathcal{P}[T] \times \left[t+(1-t)x^{-a}+\sum_{j=1}^{a-1} x^{bj-a\left\lceil \frac{b}{a}j\right\rceil}\right] \tag{21}$$

where t is the probability that a tie is resolved in favor of A. Next, using the fact that $\left\{b\left\lceil\frac{a}{b}k\right\rceil - ak : k=1,2,...,b-1\right\} = \{1,2,...,b-1\}$ and similarly, that $\left\{a\left\lceil\frac{b}{a}j\right\rceil - bj : j=1,2,...,a-1\right\} = \{1,2,...,a-1\}$, we can rewrite (20) and (21) as

$$\mathcal{P}[Piv_A] \approx \mathcal{P}[T] \times \left[1-t+tx^b+\sum_{k=1}^{b-1} x^k\right] \tag{22}$$

$$\mathcal{P}[Piv_B] \approx \mathcal{P}[T] \times \left[t+(1-t)x^{-a}+\sum_{j=1}^{a-1} x^{-j}\right] \tag{23}$$

Lemma B.1 *If for all n large, $\frac{q_A(n)}{q_B(n)} \geq \frac{a}{b}$, then $\limsup \frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} \leq \frac{b}{a}$. Similarly, if for all n large, $\frac{q_A(n)}{q_B(n)} \leq \frac{a}{b}$, then $\limsup \frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} \geq \frac{b}{a}$.*

Proof. Using the formulae in (22) and (23) we have that the ratio of the pivotal probabilities

$$\frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} \approx \frac{1-t+tx^b+\sum_{k=1}^{b-1} x^k}{t+(1-t)x^{-a}+\sum_{j=1}^{a-1} x^{-j}}$$

Now note that the numerator is increasing in x, while the denominator is decreasing. Thus, the ratio of the pivotal probabilities is increasing in x. Also, when $x = 1$,

$$\frac{\mathcal{P}\left[Piv_A\right]}{\mathcal{P}\left[Piv_B\right]} \approx \frac{b}{a}$$

If, for all n large, $\frac{q_A(n)}{q_B(n)} > \frac{a}{b}$, then $x < 1$ and so for all n large, $\frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} < \frac{b}{a}$. If there is a subsequence along which $\frac{q_A}{q_B} = \frac{a}{b}$ and along this subsequence $\lim \frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} > \frac{b}{a}$, then this contradicts the fact that $x = 1$ implies that $\frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} \approx \frac{b}{a}$. Thus, if for all n large, $\frac{q_A(n)}{q_B(n)} \geq \frac{a}{b}$, then $\limsup \frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} \leq \frac{b}{a}$. ■

The equilibrium conditions imply

$$\frac{q_A}{q_B} = \frac{\lambda v_A}{(1-\lambda)\, v_B} \frac{\mathcal{P}\left[Piv_A\right]}{\mathcal{P}\left[Piv_B\right]}$$

Note that under supermajority rules, voluntary voting has more of a status quo bias than does compulsory voting (assuming that $\frac{\lambda}{1-\lambda} > \frac{a}{b}$).

Lemma B.2 *If*

$$\frac{\lambda v_A}{(1-\lambda)\, v_B} > \left(\frac{a}{b}\right)^2$$

then, in large elections, A is elected with probability 1.

Proof. We first claim that for all large n, $\frac{q_A}{q_B} > \frac{a}{b}$; that is, the vote shares favor A.

Suppose to the contrary there is a sequence along which $\frac{q_A}{q_B} \leq \frac{a}{b}$ and so by Lemma B.1, along this sequence $\frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} \geq \frac{b}{a}$. The equilibrium conditions now imply that

$$\frac{q_A}{q_B} = \frac{\lambda v_A}{(1-\lambda)\, v_B} \frac{\mathcal{P}\left[Piv_A\right]}{\mathcal{P}\left[Piv_B\right]} > \frac{a}{b}$$

which is a contradiction. Thus we have argued that in all large elections it must be that

$$x = \left(\frac{q_B}{q_A} \frac{a}{b}\right)^{\frac{1}{a+b}} < 1$$

We also claim that $\liminf \frac{q_A}{q_B} > \frac{a}{b}$ or equivalently that, $x > 1$. Otherwise, there is a sequence x_n such that $\lim x_n \geq 1$. So along the sequence we have

$$\begin{aligned} \frac{b q_A(n)}{a q_B(n)} &= \frac{\lambda v_A}{(1-\lambda)\, v_B} \frac{b}{a} \frac{\mathcal{P}\left[Piv_A\right](n)}{\mathcal{P}\left[Piv_B\right](n)} \\ x_n^{-(a+b)} &> \left(\frac{a}{b}\right)^2 \frac{b}{a} \frac{\mathcal{P}\left[Piv_A\right](n)}{\mathcal{P}\left[Piv_B\right](n)} \\ x_n^{-(a+b)} &> \frac{a}{b} \frac{\mathcal{P}\left[Piv_A\right](n)}{\mathcal{P}\left[Piv_B\right](n)} \end{aligned}$$

which is a contradiction since $\liminf \frac{\mathcal{P}[Piv_A](n)}{\mathcal{P}[Piv_B](n)} \geq \frac{b}{a}$ while $\lim x_n^{-(a+b)} \leq 1$. ■

Lemma B.3 *If*

$$\frac{\lambda v_A}{(1-\lambda)\, v_B} < \left(\frac{a}{b}\right)^2$$

then, in large elections, B is elected with probability 1.

Proof. The proof is identical to that of Lemma B.2 and is omitted. ■

C Aggregate Uncertainty

In what follows, we make use of the following identity

$$\int_0^1 (\alpha - t)^k\, t^l dt = \frac{1}{k+l+1}\binom{k+l}{k}^{-1} \tag{24}$$

The identity is easily verified by using induction on k (say).

Lemma C.1 *For all* $x \in (0,1)$, *if* $q_A = (1-x)\,\lambda$, $q_B = x\,(1-\lambda)$ *and* $q_0 = 1 - q_A - q_B$, *then*

$$\int_0^1 \sum_{k=0}^{m} \binom{m}{k,k} (q_A)^k\, (q_B)^k\, (q_0)^{m-2k}\, d\lambda = \frac{1}{m+1} \tag{25}$$

Proof. Note that since $q_0 = 1 - q_A - q_B = x\lambda + (1-x)\,(1-\lambda)$, the binomial theorem implies that

$$(q_A)^k\, (q_B)^k\, (q_0)^{m-2k} = \sum_{j=0}^{m-2k} \binom{m-2k}{j} (x\lambda)^{j+k}\, ((1-x)\,(1-\lambda))^{m-k-j}$$

Thus, using (24),

$$\begin{aligned}
&\int_0^1 (q_A)^k\, (q_B)^k\, (q_0)^{m-2k}\, d\lambda \\
&= \sum_{j=0}^{m-2k} \binom{m-2k}{j} x^{j+k}\, (1-x)^{m-k-j} \int_0^1 \lambda^{j+k}\, (1-\lambda)^{m-k-j}\, d\lambda \\
&= \frac{1}{m+1} \sum_{j=0}^{m-2k} \binom{m-2k}{j} \binom{m}{j+k}^{-1} x^{j+k}\, (1-x)^{m-k-j}
\end{aligned}$$

and so the left-hand side of (25) equals

$$\begin{aligned}
&\frac{1}{m+1} \sum_{k=0}^{\lfloor m/2 \rfloor} \binom{m}{k,k} \sum_{j=0}^{m-2k} \binom{m-2k}{j} \binom{m}{j+k}^{-1} x^{j+k}\, (1-x)^{m-k-j} \\
&= \frac{1}{m+1} \sum_{k=0}^{\lfloor m/2 \rfloor} \sum_{j=0}^{m-2k} \binom{j+k}{k} \binom{m-k-j}{k} x^{j+k}\, (1-x)^{m-k-j} \\
&= \frac{1}{m+1} \sum_{k=0}^{\lfloor m/2 \rfloor} \sum_{j=k}^{m-k} \binom{l}{k} \binom{m-l}{k} x^{l}\, (1-x)^{m-l}
\end{aligned}$$

Interchanging the order of summation, the last expression can be rewritten as

$$\begin{aligned} & \frac{1}{m+1}\sum_{l=0}^{m}\sum_{k=0}^{\min\{m-l,l\}}\binom{l}{k}\binom{m-l}{k}x^{l}\left(1-x\right)^{m-l} \\ = & \frac{1}{m+1}\sum_{l=0}^{m}\binom{m}{l}x^{l}\left(1-x\right)^{m-l} \end{aligned}$$

which follows from the fact that for all $l \leq m$,

$$\sum_{k=0}^{l}\binom{l}{k}\binom{m-l}{k}=\binom{m}{l}=\sum_{k=0}^{m-l}\binom{l}{k}\binom{m-l}{k}$$

a consequence of the Vandermonde combinatorial identity (see, for instance, Feller, 1968, p. 64). ■

Lemma C.2 *For $x \in (0,1)$, if $q_A = x(1-\lambda)$, $q_B = (1-x)\lambda$ and $q_0 = 1 - q_A - q_B$, then*

$$\int_0^1 \sum_{k=0}^{m}\binom{m}{k,k+1}\left(q_A\right)^k\left(q_B\right)^{k+1}\left(q_0\right)^{m-2k-1}d\lambda = \frac{1}{m+1}\left(1-\left(1-x\right)^m\right) \tag{26}$$

Proof. The proof is almost identical to that of Lemma C.1 and is omitted. ■

Corollary C.1 *For $x \in (0,1)$, if $q_A = (1-x)\lambda$, $q_B = x(1-\lambda)$, then*

$$\int_0^1 e^{-n(q_A+q_B)}\sum_{k=0}^{\infty}\frac{\left(nq_A\right)^k}{k!}\frac{\left(nq_B\right)^k}{k!}d\lambda = \frac{1}{n}\left(1-e^{-n}\right)$$

Proof. Since

$$e^{-n(q_A+q_B)}\sum_{k=0}^{\infty}\frac{\left(nq_A\right)^k}{k!}\frac{\left(nq_B\right)^k}{k!} = \sum_{m=0}^{\infty}e^{-n}\frac{n^m}{m!}\sum_{k=0}^{m}\binom{m}{k,k}\left(q_A\right)^k\left(q_B\right)^k\left(q_0\right)^{m-2k}$$

Lemma C.1 implies that

$$\begin{aligned} \int_0^1 e^{-n(q_A+q_B)}\sum_{k=0}^{\infty}\frac{\left(nq_A\right)^k}{k!}\frac{\left(nq_B\right)^k}{k!}d\lambda &= \sum_{m=0}^{\infty}e^{-n}\frac{n^m}{m!}\frac{1}{m+1} \\ &= \frac{1}{n}\left(1-e^{-n}\right) \end{aligned}$$

■

Corollary C.2 *For $x \in (0,1)$, if $q_A = (1-x)\lambda$ and $q_B = x(1-\lambda)$, then*

$$\int_0^1 e^{-n(q_A+q_B)}\sum_{k=0}^{\infty}\frac{\left(nq_A\right)^k}{k!}\frac{\left(nq_B\right)^{k+1}}{(k+1)!}d\lambda = \frac{1}{n}\left(1-\frac{e^{-nx}-xe^{-n}}{1-x}\right)$$

Proof. Follows easily by using the fact that

$$e^{-n(q_A+q_B)}\sum_{k=0}^{\infty}\frac{(nq_A)^k}{k!}\frac{(nq_B)^{k+1}}{(k+1)!}=\sum_{m=0}^{\infty}e^{-n}\frac{n^m}{m!}\left(\sum_{k=0}^{m}\binom{m}{k,k+1}(q_A)^k(q_B)^{k+1}(q_0)^{m-2k-1}\right)$$

and applying Lemma C.2. ■

Lemma C.3 *Suppose that there is a sequence of elections for which* $\lim\frac{p_B}{p_A+p_B}\in(0,1)$. *Then*

$$\lim_{n\to\infty}n(p_A+p_B)\int_0^1\mathcal{P}[T\mid n,\lambda]\,d\lambda=1$$

where $\mathcal{P}[T\mid n,\lambda]$ *denotes the Poisson probability of a tie when the voting propensities are* $\lambda p_A(n)$ *and* $(1-\lambda)p_B(n)$, *respectively.*

Proof. If we set $x=\frac{p_B}{p_A+p_B}$, then

$$\begin{aligned}
&\int_0^1\mathcal{P}[T\mid n,\lambda]\,d\lambda\\
=&\int_0^1 e^{-n(\lambda p_A+(1-\lambda)p_B)}\sum_{k=0}^{\infty}\frac{(n\lambda p_A)^k}{k!}\frac{(n(1-\lambda)p_B)^k}{k!}d\lambda\\
=&\int_0^1 e^{-n(p_A+p_B)((1-x)\lambda+x(1-\lambda))}\sum_{k=0}^{\infty}\frac{(n(p_A+p_B)(1-x)\lambda)^k}{k!}\frac{(n(p_A+p_B)x(1-\lambda))^k}{k!}d\lambda\\
=&\frac{1-e^{-n(p_A+p_B)}}{n(p_A+p_B)}
\end{aligned}$$

where the last equality follows from Lemma C.1, using $n(p_A+p_B)$ in place of n.

Since $n(p_A+p_B)\to\infty$ as $n\to\infty$, taking limits yields the result ■

Lemma C.4 *Suppose that there is a sequence of elections for which* $\lim\frac{p_B}{p_A+p_B}\in(0,1)$. *Then*

$$\lim_{n\to\infty}n(p_A+p_B)\int_0^1\mathcal{P}[T_{-1}\mid n,\lambda]\,d\lambda=1$$

where $\mathcal{P}[T_{-1}\mid n,\lambda]$ *denotes the Poisson probability that* A *is one vote behind when the voting propensities are* $\lambda p_A(n)$ *and* $(1-\lambda)p_B(n)$, *respectively.*

Proof. The proof is almost the same as that of Lemma C.3 and is omitted. ■

Lemma C.5 *Suppose that there is a sequence of elections for which* $\lim\frac{p_B}{p_A+p_B}\in(0,1)$. *Then*

$$\lim_{n\to\infty}n(p_A+p_B)\int_0^1\mathcal{P}[Piv_A\mid n,\lambda]\,d\lambda=1$$

Proof. Follows immediately from Lemmas C.3 and C.4. ■

Proof of Proposition 8. We prove the result for Piv_A. The proof for Piv_B is analogous.

First, using the asymptotic formulae for the Poisson probablity of Piv_A, observe that for all $\lambda \neq \lambda^*$,

$$\begin{aligned} & \frac{\mathcal{P}\left[Piv_A \mid n, \lambda\right]}{\mathcal{P}\left[Piv_A \mid n, \lambda^*\right]} \\ \approx\ & \frac{e^{-\left(\sqrt{n\lambda p_A}-\sqrt{n(1-\lambda)p_B}\right)^2}}{\sqrt{4\pi n\sqrt{\lambda p_A(1-\lambda)p_B}}}\left(1+\sqrt{\frac{(1-\lambda)p_B}{\lambda p_A}}\right) \div \frac{e^{-\left(\sqrt{n\lambda^* p_A}-\sqrt{n(1-\lambda^*)p_B}\right)^2}}{\sqrt{4\pi n\sqrt{\lambda^* p_A(1-\lambda^*)p_B}}}\left(1+\sqrt{\frac{(1-\lambda^*)p_B}{\lambda^* p_A}}\right) \\ \approx\ & e^{-\left(\sqrt{n\lambda p_A}-\sqrt{n(1-\lambda)p_B}\right)^2+\left(\sqrt{n\lambda^* p_A}-\sqrt{n(1-\lambda^*)p_B}\right)^2} \times K\left(\lambda, \lambda^*\right)^{\frac{1}{4}} \\ =\ & e^{n(\phi(\lambda)-\phi(\lambda^*))} \times K\left(\lambda, \lambda^*\right)^{\frac{1}{4}} \end{aligned}$$

where $\phi(\lambda) = 2\sqrt{p_A p_B}\sqrt{\lambda(1-\lambda)} - \lambda p_A - (1-\lambda)p_B$ and $K(\lambda, \lambda^*)$ is a rational function that does not depend on n, p_A or p_B. It is routine to verify that the strictly concave function $\phi(\lambda)$ is uniquely maximized at $\lambda = \frac{p_B}{p_A+p_B}$. Since $\frac{p_B}{p_A+p_B} \to \lambda^*$, for all large n, $\phi(\lambda) < \phi(\lambda^*)$. Moreover, since

$$\begin{aligned} n\phi(\lambda) &= n\left(p_A+p_B\right)\left(2\sqrt{\frac{p_A}{p_A+p_B}\frac{p_B}{p_A+p_B}}\sqrt{\lambda(1-\lambda)} - \lambda\frac{p_A}{p_A+p_B} - (1-\lambda)\frac{p_B}{p_A+p_B}\right) \\ &\approx n\left(p_A+p_B\right)\left(2\sqrt{\lambda^*(1-\lambda^*)}\sqrt{\lambda(1-\lambda)} - \lambda(1-\lambda^*) - (1-\lambda)\lambda^*\right) \end{aligned}$$

and $n(p_A+p_B) \to \infty$, it follows that $n(\phi(\lambda) - \phi(\lambda^*)) \to -\infty$. This implies that the ratio $\mathcal{P}[Piv_A \mid n, \lambda] / \mathcal{P}[Piv_A \mid n, \lambda^*]$ converges to zero as $n \to \infty$.

Fix an $\varepsilon > 0$ and let n be large enough so that $\frac{p_B}{p_A+p_B} > \lambda^* - \varepsilon$. As in the expression above, we can write for all $\lambda' < \lambda'' < \lambda^* - \varepsilon$

$$\frac{\mathcal{P}\left[Piv_A \mid n, \lambda''\right]}{\mathcal{P}\left[Piv_A \mid n, \lambda'\right]} = e^{n(\phi(\lambda'')-\phi(\lambda'))} \times K\left(\lambda'', \lambda'\right)^{\frac{1}{4}}$$

and since $\phi(\lambda)$ is strictly concave and reaches a maximum at $\frac{p_B}{p_A+p_B} > \lambda^* - \varepsilon$, $\phi(\lambda'') > \phi(\lambda')$. Thus, for n large enough, $\mathcal{P}\left[Piv_A \mid n, \lambda''\right] > \mathcal{P}\left[Piv_A \mid n, \lambda'\right]$. Analogously, for all λ', λ'' satisfying $\lambda^* + \varepsilon < \lambda' < \lambda''$, $\mathcal{P}\left[Piv_A \mid n, \lambda'\right] > \mathcal{P}\left[Piv_A \mid n, \lambda''\right]$ once n is large enough.

For any $\varepsilon > 0$, as $n \to \infty$, $\mathcal{P}[Piv_A \mid n, \lambda]$ converges to zero uniformly for all $\lambda \in [0, \lambda^* - \varepsilon]$. Similarly, for any $\varepsilon > 0$, $\mathcal{P}[Piv_A \mid n, \lambda]$ converges to zero uniformly for all $\lambda \in [\lambda^* + \varepsilon, 1]$. As a result, if we denote by $I(\varepsilon)$ the interval $[\lambda^* - \varepsilon, \lambda^* + \varepsilon]$, then

$$\lim n\left(p_A+p_B\right)\int_{I(\varepsilon)^c} \mathcal{P}\left[Piv_A \mid n, \lambda\right] d\lambda = 0$$

and

$$\lim n\left(p_A+p_B\right)\int_{I(\varepsilon)^c} \mathcal{P}\left[Piv_A \mid n, \lambda\right] h(\lambda)\, d\lambda = 0$$

as well. Thus,

$$\begin{aligned}\lim n\left(p_A+p_B\right)\int_{I(\varepsilon)}\mathcal{P}\left[Piv_A \mid n,\lambda\right]d\lambda &= \lim n\left(p_A+p_B\right)\int_0^1\mathcal{P}\left[Piv_A \mid n,\lambda\right]d\lambda \\ &= 1\end{aligned}$$

using Lemma C.5.

Since h is continuous, for any $\delta>0$, we can pick an ε small enough so that for all $\lambda\in\left[\lambda^*-\varepsilon,\lambda^*+\varepsilon\right]$,

$$h\left(\lambda^*\right)-\delta\leq h\left(\lambda\right)\leq h\left(\lambda^*\right)+\delta$$

Thus, we have

$$\begin{aligned}&\left(h\left(\lambda^*\right)-\delta\right)n\left(p_A+p_B\right)\int_{I(\varepsilon)}\mathcal{P}\left[Piv_A \mid n,\lambda\right]d\lambda \\ \leq\ & n\left(p_A+p_B\right)\int_{I(\varepsilon)}\mathcal{P}\left[Piv_A \mid n,\lambda\right]h\left(\lambda\right)d\lambda \\ \leq\ & \left(h\left(\lambda^*\right)+\delta\right)n\left(p_A+p_B\right)\int_{I(\varepsilon)}\mathcal{P}\left[Piv_A \mid n,\lambda\right]d\lambda\end{aligned}$$

and so

$$h\left(\lambda^*\right)-\delta\leq\lim n\left(p_A+p_B\right)\int_{I(\varepsilon)}\mathcal{P}\left[Piv_A \mid n,\lambda\right]h\left(\lambda\right)d\lambda\leq h\left(\lambda^*\right)+\delta$$

or

$$h\left(\lambda^*\right)-\delta\leq\lim n\left(p_A+p_B\right)\int_0^1\mathcal{P}\left[Piv_A \mid n,\lambda\right]h\left(\lambda\right)d\lambda\leq h\left(\lambda^*\right)+\delta$$

and since δ was arbitrary, the proof is complete. ■

D Poisson Approximations of the Multinomial

We are interested the distribution of the sum of independent Bernoulli vector variables (X_A, X_B) where

$$\begin{aligned}\Pr\left[(X_A,X_B)=(1,0)\right] &= q_A \\ \Pr\left[(X_A,X_B)=(0,1)\right] &= q_B \\ \Pr\left[(X_A,X_B)=(0,0)\right] &= 1-q_A-q_B\end{aligned}$$

where $q_A+q_B\leq 1$. If $q_0=1-q_A-q_B$, then the probability that after m draws, the sum of the variables (X_A, X_B) is (k,l) is

$$\mathbb{P}\left[(k,l)\mid m\right]=\binom{m}{k,l}\left(q_A\right)^k\left(q_B\right)^l\left(q_0\right)^{m-k-l}$$

Now consider a multivariate Poisson distribution with means mq_A and mq_B, respectively. The probability $\mathcal{P}[(k,l)]$ that the total number of occurrences of A and B will be k and l, respectively, is

$$\mathcal{P}[(k,l) \mid m] = e^{-mq_A - mq_B} \frac{(mq_A)^k}{k!} \frac{(mq_B)^l}{l!}$$

Roos (1999, p. 122) has shown that

$$\sup_{S \subset \mathbb{Z}_+^2} |\mathbb{P}[S \mid m] - \mathcal{P}[S \mid m]| \leq q_A + q_B$$

References

[1] Bayes, T. (1763): "An Essay towards Solving a Problem in the Doctrine of Chances," (Communicated by R. Price) *Philosophical Transactions of the Royal Society*, 53, 370–418 .

[2] Börgers, T. (2004): "Costly Voting," *American Economic Review*, 94, 57–66.

[3] Chamberlain, G. and M. Rothschild (1981): "A Note on the Probability of Casting a Decisive Vote," *Journal of Economic Theory*, 25, 152–162.

[4] Feddersen T. and W. Pesendorfer (1998): "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts Under Strategic Voting," *American Political Science Review*, 92 (1), 23–35.

[5] Feddersen T. and W. Pesendorfer (1999): "Abstention in Elections with Asymmetric Information and Diverse Preferences," *American Political Science Review*, 93, 381–398.

[6] Feller, W. (1968): *An Introduction to Probability Theory and Its Applications* (3rd. ed.), New York: Wiley.

[7] Ghosal, S. and B. Lockwood (2009): "Costly Voting when both Information and Preferences Differ: Is Turnout Too High or Too Low?" *Social Choice and Welfare*, 33, 25–50.

[8] Good, I. and L. Mayer (1975): "Estimating the Efficacy of a Vote," *Behavioral Science*, 20 (1), 25–33.

[9] Herrera, H. and M. Morelli (2009): "Turnout and Power Sharing," Working Paper, Columbia University, August.

[10] Krasa, S. and M. Polborn (2009): "Is Mandatory Voting better than Voluntary Voting?" *Games and Economic Behavior*, 66, 275–291.

[11] Krishna, V. and J. Morgan (2009): "Voluntary Voting: Costs and Benefits," Working paper, Penn State University and UC Berkeley, November.

[12] Krishna, V. and J. Morgan (2011): "Overcoming Ideological Bias in Elections," *Journal of Political Economy*, 119, No. 2, 183–221.

[13] Ledyard, J. (1984): "The Pure Theory of Large Two-candidate Elections," *Public Choice*, 44, 7–41.

[14] Mandler, M. (2012): "The Fragility of Information Aggregation in Large Elections," *Games and Economic Behavior*, 74, 257–268.

[15] Myerson, R. (1998a): "Extended Poisson Games and the Condorcet Jury Theorem," *Games and Economic Behavior*, 25, 111–131.

[16] Myerson, R. (1998b): "Population Uncertainty and Poisson Games," *International Journal of Game Theory*, 27, 375–392.

[17] Myerson, R. (2000): "Large Poisson Games," *Journal of Economic Theory*, 94, 7–45.

[18] Palfrey, T. and H. Rosenthal (1985): "Voter Participation and Strategic Uncertainty," *American Political Science Review*, 79, 62–78.

[19] Riker, W. and P. Ordeshook (1968): "A Theory of the Calculus of Voting," *American Political Science Review*, 62, 25–42.

[20] Roos, B. (1999): "On the Rate of Multivariate Poisson Convergence," *Journal of Multivariate Analysis*, 60, 120–134.

[21] Taylor, C. and H. Yildirim (2010): "A Unified Analysis of Rational Voting with Private Values and Group Specific Costs," *Games and Economic Behavior*, 70, 457–471.