Majority Rule and Utilitarian Welfare

Vijay Krishna* John Morgan†

August 12, 2012

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 is 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 results on the welfare properties of majority rule.

1 Introduction

Pre-election polls leading up to the November 2008 vote on Proposition 8, the California Marriage Protection Act, indicated that it would be easily defeated.[1] If passed, the proposition would make it illegal for same-sex couples to marry. The actual vote count differed sharply from poll predictions—Proposition 8 passed by a 52-48% margin. The results surprised most Californians and were shortly followed by mass protests and lawsuits.[2]

The intent of any referendum, including Proposition 8, is to reflect directly the will of the electorate. But of course, it can only reflect the will of those of the electorate who actually turn out to vote. The election results suggest that the preferences of those who turned out to vote were different from the preferences of the population at large, at least to the extent that the pre-election polls accurately reflected the latter. Precisely, the turnout rates of those in favor of the proposition—that is, against same-sex marriage—were greater than of those opposed. A simple explanation is that those in favor felt more strongly about the matter and turned out in greater numbers.[3] If

*Penn State University, E-mail: `vkrishna@psu.edu`

†University of California, Berkeley, E-mail: `rjmorgan@haas.berkeley.edu`

[1] The three polls closest to the election had Proposition 8 losing by margins of 47-50% (Survey USA), 44-49% (Field Poll) and 44-52% (Public Policy Institute of California).

[2] The proposition was declared unconstitutional by the courts and the matter is, as of now, awaiting consideration by the US Supreme Court.

[3] The vote on Proposition 8 was concurrent with the 2008 presidential election and so one may wonder whether turnout was determined by the latter. But since California voted overwhelmingly for Barack Obama in 2008, this cannot explain the "surprise" positive vote for the proposition.

voters on both sides had come to the polls in proportion to their numbers in the overall populace, there would have been no surprise on election day. When intensity of preference drives turnout, such surprises can, and do, happen.

This paper studies the outcomes produced by majority rule in a setting where the intensity of preference affects turnout. Our starting point is the following well-known conundrum. Suppose that 51% of the populace mildly favors one of two choices. The remainder passionately favors the alternative. If everyone voted, the choice supported by the majority would win; however, a utilitarian social planner would side with the minority since the welfare gains would more than compensate for the modest losses of overruling the majority. In such situations, majority rule would appear to be at odds with utilitarianism.

Or would it? Voting is often a choice rather than a requirement. Moreover, voters incur opportunity (or real) costs in coming to the polls. In other words, voting is inherently costly. Accounting for this casts doubt on our earlier conclusion. Given their intensely held views, the minority may be more motivated to pay the cost of voting than the majority. Thus, the *decision* to vote encodes voters' intensity of preference. But even here the link is, at best, indirect. Both sides are only motivated to turn out to the extent that they are likely to influence the final decision; that is, the benefits from voting are mitigated by the probability that a vote cast is *pivotal*. So even though the minority feels intensely about their favored alternative, were they sufficiently pessimistic about the prospect of casting a decisive vote, this intensity alone would mean little in terms of participation.

We show below that when voting is costly, voluntary voting under majority rule translates societal preferences into outcomes in a consistent way—it always implements the *utilitarian* outcome. Moreover, majority rule is the *only* election rule with this property. Even when voters are strategically sophisticated and can anticipate the effects of the voting rule on outcomes, supermajority rules will not deliver the utilitarian choice. Instead, the outcome disproportionately favors the choice advantaged by the voting rule. The implied welfare weight given to the advantaged choice is equal to the *square* of the required vote ratio. For instance, a 2/3 supermajority rule, which requires a $2:1$ vote ratio to overturn the status quo, would seem to give twice the weight to that choice. We show that such a rule effectively gives *four* times the welfare weight to the status quo than the alternative.

To see why voluntary voting under majority rule is utilitarian, consider the following example: A finite population is to vote on one of two alternatives, A and B. Voters favoring A constitute a fraction $\lambda > 1/2$ of the population and receive payoff $v_A > 0$ when it is selected over the alternative. The remainder favor B and receive payoff $v_B > 0$ when B is chosen. Despite there being a majority of A supporters, a utilitarian planner would prefer B; that is, $\lambda v_A < (1-\lambda) v_B$. Finally, suppose that each voter's cost of coming to the polls is independently drawn from a uniform distribution. For A supporters, the benefits of voting are $v_A \Pr[Piv_A]$, where $\Pr[Piv_A]$ is the probability that an additional A vote is decisive. This, of course, depends on the turnout rates of the two sides. In equilibrium, all A supporters with costs below a threshold c_A will vote, and this is determined by equating it to the benefits of voting:

$$c_A = v_A \Pr[Piv_A]$$

and similarly the cost threshold c_B for B supporters is

$$c_B = v_B \Pr[Piv_B]$$

Next, because costs are uniformly distributed on $[0, 1]$, c_A equals the turnout rate p_A of A supporters; similarly, c_B equals p_B. Using this and then multiplying the first equation by the share of A supporters in the population and the second by the share of B supporters, yields expressions in terms of expected vote shares, λp_A and $(1-\lambda)\,p_B$ for A and B, respectively. Thus, in equilibrium

$$\frac{\lambda p_A}{(1-\lambda)\,p_B} = \frac{\lambda v_A}{(1-\lambda)\,v_B} \times \frac{\Pr[Piv_A]}{\Pr[Piv_B]}$$

The right-hand side is the product of two terms that we call the "welfare ratio" and the "pivot ratio." We claim that if, as assumed, the welfare ratio favors B, then the vote shares must favor B as well. Suppose to the contrary that A enjoys a higher vote share; that is, the left-hand side of the above expression is greater than one. Since, by assumption, the welfare ratio is less than one, it must be that the pivot ratio exceeds one. But the fact that A has a higher vote share implies that the pivot ratio is *less* than one. The reason is that a vote for the candidate that is behind in an election is more likely to be decisive than a vote for the candidate that is ahead—a fact that we term the "underdog" principle. To see why, notice that a vote for the trailing candidate pushes the vote total in the direction of ties or near ties while a vote for the leading candidate pushes the total away. Therefore, the former vote is more likely to be decisive. If the vote share favored A, then both the welfare and pivot ratios would favor B, leading to a contradiction. Thus, the vote share must favor B. An analogous argument would apply if the utilitarian calculus favored A instead. As a consequence, the vote shares always favor the utilitarian choice: $\lambda p_A > (1-\lambda)\,p_B$ if and only if $\lambda v_A > (1-\lambda)\,v_B$. When there are many voters, this implies that A would win with high probability.

Despite its apparent simplicity, the argument above is somewhat subtle. It relies essentially on the fact that costs are randomly distributed (with a lower support at 0). To see why, suppose, along the lines of Palfrey and Rosenthal (1983), that each voter incurred a small, fixed cost of voting, $c > 0$. Assuming that there is an interior solution, participation rates (in this case, the probability of voting) would again be determined by equating the costs and benefits. For A voters, this amounts to

$$c = v_A \Pr[Piv_A]$$

and similarly for B voters

$$c = v_B \Pr[Piv_B]$$

When A receives a greater vote share than B, the "underdog" principle again implies that pivotality considerations favor B. Since the expected benefits of voting must be same for both sides, it then follows that $v_A > v_B$. That is, alternative A prevails when a typical A voter feels more strongly about his favored candidate than a typical B voter. But this pays no attention to the fraction of voters of each type, so the outcome is not utilitarian. Formally, with a fixed cost of voting, $\lambda p_A > (1 - \lambda)\, p_B$ if and only if $v_A > v_B$; in contrast, a utilitarian planner would choose A if and only if $\lambda v_A > (1 - \lambda)\, v_B$.

The difference between the two settings may be better understood by analogy with the difference between "fixed versus flexible prices" in a market environment. When the voting cost c is given exogenously, equilibrium requires that the expected benefits of a vote must equal this cost. As a result, the expected benefits of a vote for each side are also "fixed." When voting costs are random, the costs of the marginal voter on each side, c_A and c_B, are determined endogenously in equilibrium. The expected benefits of a vote for each side are now "flexible" and provide just the right incentives so that the utilitarian outcome results.

Before placing the paper in the context of the extant literature, it is useful to sketch the key features of our model and its main results. As in the example above, there are two alternatives, A and B, and voters know exactly the utility of each outcome to themselves. Unlike the examples, voters may have heterogeneous preferences—some feel passionately about A, others passionately about B, and still others are more or less indifferent between the two choices. Voters also differ in their costs of voting—for some, costs are modest while for others the costs are so large as to dwarf any possible benefit from voting. The distribution of voting costs is arbitrary and costs are orthogonal to preferences. Finally, there is an uncertain populace of potential voters.

In this setting, the main results of the paper are:

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

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

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).

Related Literature Our model is a more general version of that studied by Ledyard (1984). Ledyard is mainly interested in the ideological positioning of candidates when faced with voters with Hotelling-type preferences and privately known costs of voting. If voting were costless, both candidates would co-locate at the preferred point of the median voter. Ledyard's main result is that, with costly voting, both candidates still co-locate, but at the welfare maximizing ideology—and thus there is no incentive to participate. Indeed, in equilibrium, the first-best outcome obtains without any actual voting!

But if candidates have concerns other than merely winning the election, they will not co-locate. This is the starting point for our model. Specifically, we study a situation where candidates' ideological positions are given and different. Here, turnout is positive as the supporters of both sides vie to obtain their preferred choice; nonetheless, the chosen candidate maximizes societal welfare—the utilitarian choice enjoys higher vote share and, in large elections, wins with certainty. The other key difference concerns how outcomes change with the voting rule. Unlike Ledyard, we examine supermajority rules as well and show that the utilitarian property is unique to majority rule. Finally, we allow for correlated voter preferences by introducing aggregate uncertainty into the model. Our main finding here is that the utilitarian property of majority rule is preserved.

Also closely related is Myerson's (2000) reformulation of Ledyard's result when the number of voters is Poisson distributed rather than fixed. Using the asymptotic formulae for the Poisson model, Myerson also links majority rule and utilitarianism as a stepping stone to obtaining Ledyard's "nobody votes" result. Our model allows for general population uncertainty and thus includes Myerson's result as a special case. Like Ledyard, Myerson limits his attention to majority rule in the absence of aggregate uncertainty.

Ledyard and Myerson represent the two standard assumptions about the number of voters. Ledyard assumes that the number of voters is fixed and commonly known whereas Myerson assumes that they are Poisson distributed. Neither assumption is perfectly satisfactory. In large elections, it seems unlikely that voters know the size of the electorate precisely. The Poisson assumption remedies this defect, but comes with the cost the variance in the size of the electorate becomes unbounded as the expected size grows larger. An important contribution of our paper is to do away with the need for either assumption. Our model embeds both cases as well as allowing for more realistic size distributions.

Börgers (2004) compares compulsory and voluntary voting in a completely symmetric special case of our model. His main concern is with the cost of participation. In particular, he shows that voluntary voting, by economizing on voting costs, Pareto dominates compulsory voting. The symmetry in the model, however, allows no scope for examining questions like those we pose. Moreover, Krasa and Polborn (2009) show that Börgers' result may not hold when the symmetry is broken.

Palfrey and Rosenthal (1985) characterize equilibrium properties of large elections with identical preference intensities and random voting costs with lower support at zero. Taylor and Yildirim (2008) study a similar model but where voting costs are bounded above zero. Their main finding is to identify the underdog principle in these settings. In both papers, since preference intensities are identical for both sides, welfare considerations of election outcomes are not investigated.

In a Poisson framework, Feddersen and Pesendorfer (1999) examine majority rule elections with a mix of private and common values and differing preference intensities. Voting is costless; however, owing to the "swing voter's curse", some voters choose to abstain. Their main result is to show that, in large elections, information aggregates in the sense of full information equivalence—the outcome corresponds to what would be obtained were all voters informed about the underlying state. This, however, is not the same as the utilitarian outcome. For instance, the outcome under compulsory voting in our setting also satisfies full information equivalence. Krishna and Morgan (2012) examine a similar model with costly voting and show that, when private value considerations dominate, the utilitarian outcome prevails. Relative to these papers, our contribution is to examine rules other than simple majority and population distributions other than Poisson. Since common value considerations are absent from our model, information aggregation is not a concern.

Mandler (2012) introduces aggregate uncertainty into the pure common values Condorcet model with compulsory voting and shows the existence of equilibria in which information does not aggregate. Myatt (2012) shows that, by incorporating aggregate uncertainty in a private value context, one can explain high turnout in large elections. In Myatt's model, both the costs of voting and the intensity of preferences are identical across voters. As a result, welfare considerations are not examined.

The remainder of the paper proceeds as follows. We sketch the model in section 2. Section 3 establishes that vote shares favor the utilitarian choice when voting costs are uniform. Section 4 generalizes this result to other voting cost distributions and highlights that in large elections majority rule produces the utilitarian outcome with almost certainty. Section 5 studies supermajority rules and shows that they do not satisfy the utilitarian property. Section 6 adds aggregate uncertainty to the model and shows that majority rule robustly implements the utilitarian outcome. Finally, section 7 concludes.

2 The Model

We study a general version of the familiar "private values" voting model. In this setting, two candidates, who differ in their ideology, compete in an election decided by majority rule. Individual voters care only about ideology, thus candidates are not distinguished by "vertical" (valence) characteristics. The generalization comes through allowing for randomness in the intensity of voter preferences, in the preferred ideology of each voter, and possibly in the number of eligible voters. Thus, the model captures electoral settings where ideology is the main driver of voter decisions and where there is possibly considerable uncertainty about the size and preferences of the voting populace at large.

Formally, 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.[4] The size of the electorate is a random variable N which is distributed on $\{0, 1, 2, ...\}$ according to the probability distribution function π^*, with a finite expectation, say n. Thus, the probability that there are exactly m eligible voters (or *citizens*) is $\pi^*(m)$. Of greater interest to an individual voter is the distribution π which determines the

[4] Supermajority rules are considered in Section 5.

probability $\pi(m)$ that there are exactly m *other* eligible voters. Then

$$\pi(m) = \pi^*(m+1) \times \frac{m+1}{n} \tag{1}$$

To see how this is derived, suppose that there is a large pool of M identical potential voters from which the number of eligible voters is drawn according to π^* and each potential voter in the pool has an equal chance of being selected as being eligible. Conditional on the event that a particular voter has been chosen to be eligible, the probability that there are m other eligible voters is

$$\frac{\pi^*(m+1)\frac{m+1}{M}}{\sum_{k=1}^{M}\pi^*(k)\frac{k}{M}} = \frac{\pi^*(m+1)(m+1)}{\sum_{k=1}^{M}\pi^*(k)\,k}$$

and as $M \to \infty$, the denominator converges to n, thus yielding the expression in (1). Note that $\pi = \pi^*$ if and only if π is a Poisson distribution (Myerson, 1998).

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 preference and its intensity will be referred to as her *type*. Types, which are private information, are distributed independently across voters and independently of the number of voters.[5] A citizen knows his own type and that the types of the others are distributed according to λ, G_A and G_B.

Utilitarianism

Before proceeding to study election outcomes, it is helpful to examine a benchmark situation where a social planner selects the winning candidate. Suppose that the planner is utilitarian and gives equal weight to each potential voter. The planner's choice of candidate is made ex ante; that is, the planner only knows the distribution of types, but not their exact realization.

In that case, 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

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

[5] In Section 6, we extend the basic model to allow for aggregate uncertainty about λ, the *ex ante* proportion of A supporters.

If the inequality above holds, we will refer to A as the *utilitarian choice* (and if it is reversed then B will be referred to as such). We will say that a voting rule is *utilitarian* if the candidate elected is the same as the utilitarian choice.

Compulsory Voting

We now turn attention to voting and first examine the case where voting is compulsory—the penalties for not voting are sufficiently stringent that all eligible voters turn out at the polls. On arriving at the polls, each voter can choose between voting for A, voting for B, or abstaining through submitting a blank or spoiled ballot.

Conditional on coming to the polls, a voter's strategy is straightforward—each voter has a strict incentive to vote for her preferred candidate. As a consequence, the election outcome is determined purely by the single parameter, λ, the fraction of A voters. In a large election, candidate A wins if and only if $\lambda > 1/2$. Obviously this is not utilitarian since the average intensity of preferences does not figure into the election outcome at all. For instance, if a majority of voters favor candidate A but are lukewarm in their support (i.e., v_A is only modestly positive) while the minority strongly favor candidate B such that $\lambda v_A < (1 - \lambda) v_B$, then candidate A will still win the election despite the fact that B is the utilitarian outcome. Thus, as a benchmark, it is worthwhile to record the following well-known fact,

Proposition 1 *Under compulsory voting, majority rule is not utilitarian.*

Voluntary and Costly Voting

Suppose that voting were voluntary rather than compulsory. Obviously, if voting costs were zero, all voters would show up and the outcome would be identical to compulsory voting. Here, we examine the situation where voting entails some positive opportunity cost for individuals. Of course, this voting cost will vary from individual to individual depending on their proximity to the polls, job requirements, wages, and so forth.

To model this in the simplest manner, we suppose that a citizen's cost of voting is private information and determined by an independent realization from a continuous probability distribution F satisfying $F(0) = 0$ and with a 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—her type and her cost of voting.

In this circumstance, each voter compares her private costs with the benefits from voting. Since preferences are purely instrumental, the benefits from voting hinge on the chance that a given vote will swing the outcome of the election in favor of the voter's preferred candidate—either from a loss to a tie or from a tie to a win. Thus, the chance that a voter is *pivotal* is critically important in determining an individual's benefit from voting.

Pivotal 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 among 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

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

where

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

denotes the trinomial coefficient.[6] For a realized number of eligible voters, m, the chance of a tie is simply the probability of events of the form (k,k). Formally,

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

Since an individual voter is unaware of the realized number of potential voters, the probability of a tie from that voter's perspective is

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

where the formula reflects a voter's uncertainty about the size of the electorate.

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

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

and, from the perspective of a single voter, the overall probability that A falls one vote short is

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

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

[6] 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,

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

The following proposition is intuitive. It 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 2 $\Pr\left[Piv_A\right] > \Pr\left[Piv_B\right]$ *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 $q_A < q_B$. ■

The difference in the probability of being pivotal when casting an A vote versus a B vote comes down to a comparison of the chance that candidate A is behind by one vote versus ahead by one vote (since the probability of a tied vote is common to both expressions). Obviously, when others are more likely to vote for B than A, then A is more likely to be behind than ahead. With these preliminaries in place, we are now in a position to study equilibria under majority rule with costly voting.

3 Equilibrium

In this section, we will show that there always exists an equilibrium to the voting game. Moreover, in any equilibrium, both A and B supporters participate at strictly positive rates and vote for their preferred candidate regardless of their intensity of preference. That is, participation is not merely confined to those with the most extreme preferences for a candidate (though those with extreme preferences will participate at higher rates).

As with compulsory voting, 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 (1998) has shown that in voting games with population uncertainty, *all* equilibria are type-symmetric.[7] Thus, when we refer to equilibrium, we mean 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:

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

which must hold for all $v \in [0, 1]$.

Letting $p_A(v)$ (resp. $p_B(v)$) be the probability that an A supporter (resp. B supporter) with value v will vote, 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 $q_B = (1 - \lambda) p_B$.

[7] 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.

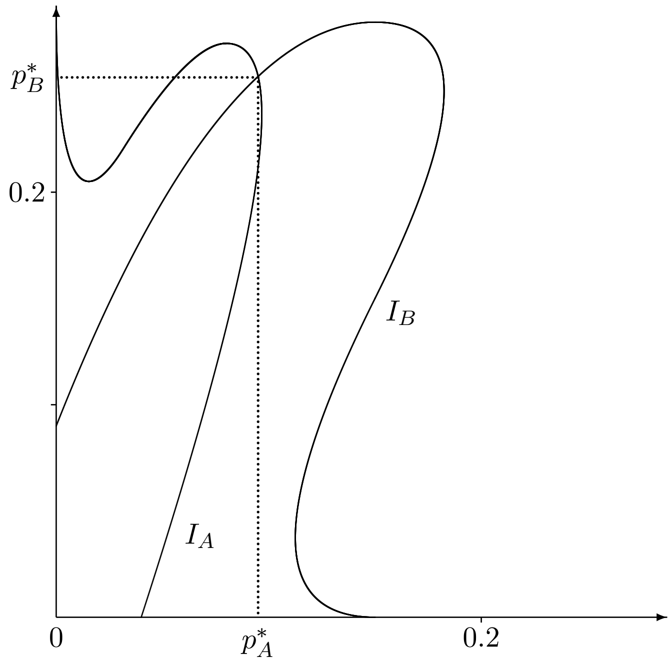

Figure 1: Equilibrium in Example 1

In terms of voting propensities, the equilibrium conditions are

$$q_A = \lambda \int_0^1 F(v \Pr[Piv_A]) \, dG_A(v) \tag{4}$$

$$q_B = (1-\lambda) \int_0^1 F(v \Pr[Piv_B]) \, dG_B(v) \tag{5}$$

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

Proposition 3 *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 (4) and (5). 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, $\Pr[Piv_A] > 0$. Hence for all v, $c_A(v) > 0$ and, in turn, $p_A(v) > 0$ as well. ■

Example 1 *Suppose that the population is distributed according to a Poisson distribution with mean* $n = 100$. *Suppose also that* $\lambda = \frac{2}{3}$, $v_A = \frac{1}{3}$, $v_B = 1$ *and that voting costs are distributed according to* $F(c) = 3c$ *over* $\left[0, \frac{1}{3}\right]$.

Figure 1 indicates the equilibrium participation rates p_A^* and p_B^*. The curve $I_A \equiv \frac{1}{3}p_A - \Pr[Piv_A] = 0$ consists of those participation rates that leave an A voter indifferent between participating and staying home (I_A is obtained from equation (4) after dividing through by λ). I_B is the analogous curve for B voters. Note that given a p_B there may be multiple values of p_A that leave an A voter indifferent. This is because, for fixed p_B, $\Pr[Piv_A]$ is a non-monotonic function of p_A while $F^{-1}(p_A)$ is monotone. Notice also that despite the fact that both curves "bend backwards," there is a unique equilibrium in the example.

Uniform Costs

Proposition 3 establishes that an equilibrium exists and that all types participate at positive rates, but says nothing about the relative participation rates and, in turn, the outcomes of the election. If, however, one temporarily restricts attention to the case where voting costs are uniformly distributed, a more precise characterization is possible. The key is that, with this specification, the equilibrium cost thresholds and the participation rates are one and the same.

Under uniform voting costs, $F(c) = c$, and, in this case, the equilibrium conditions (4) and (5) 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

$$\frac{nq_A}{nq_B} = \frac{\lambda v_A}{(1-\lambda) v_B} \times \frac{\Pr[Piv_A]}{\Pr[Piv_B]} \tag{6}$$

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 A is the utilitarian choice, this expression is greater than 1 whereas it is fractional when B is the utilitarian choice.

Suppose that A is the utilitarian choice. We claim that the expected number of A votes must exceed the expected number of B votes. To see why, suppose to the contrary that $q_A < q_B$. Proposition 2 now implies that a vote for A is more likely to be pivotal than a vote for B and hence $\Pr[Piv_A] / \Pr[Piv_B] > 1$. In that case, both expressions on the right-hand side of (6) exceed 1 while the left-hand side is fractional. Obviously, this is a contradiction. A similar argument establishes that the candidate B enjoys a higher number of expected votes than A when B is the

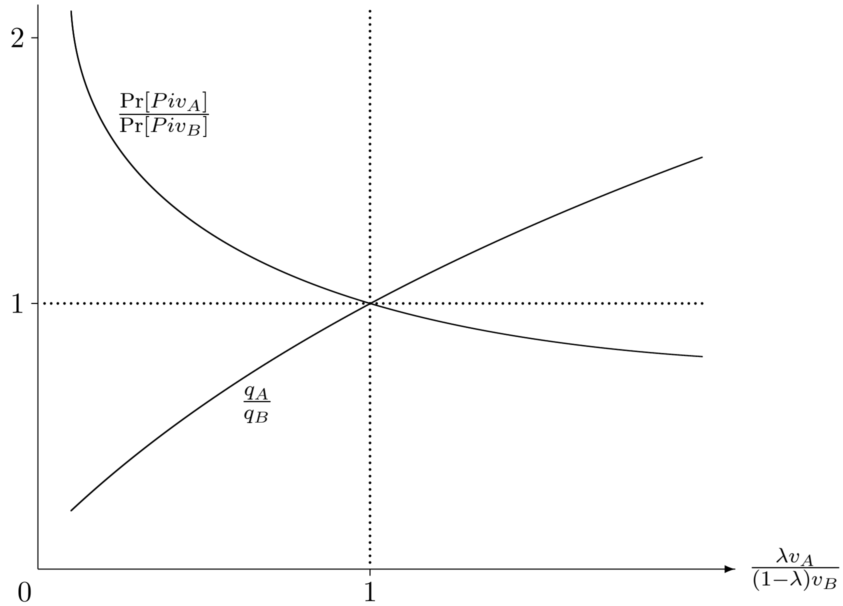

Figure 2: Ratios of Votes and Pivot Probabilities

utilitarian choice. In other words, the vote ratio always mirrors the utilitarian choice under majority rule with uniform costs. Thus, we have shown:

Proposition 4 *Suppose voting costs are uniformly distributed. In any equilibrium, 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* $\lambda v_A > (1 - \lambda) v_B$.

The following example illustrates Proposition 4.

Example 2 *Suppose that the population follows a Poisson distribution with mean* $n = 1000$ *and that voting costs are uniform.*

Figure 2 depicts the equilibrium ratio of the expected number of votes for A versus B, q_A/q_B, as a function of the welfare ratio, $\lambda v_A / (1 - \lambda) v_B$. As the proposition indicates, $q_A > q_B$ if and only if $\lambda v_A > (1 - \lambda) v_B$.

Note that Proposition 4 applies to every equilibrium and regardless of the expected size of the electorate. While it is reassuring that the expected vote shares go in the direction of the utilitarian outcome, this by no means guarantees that the utilitarian candidate will, in fact, be elected. All one can say is that this candidate is more likely to be elected than his rival. Of particular interest are the outcomes of a large election, i.e., where the expected size of the electorate goes to infinity. We explore elections with a large number of potential voters in the next section.

4 Large Elections

The usual rationale for studying large elections is to examine the limiting probabilities that each candidate will be elected. This is of interest in our model as well. We will show that, in large elections, the utilitarian choice is elected with probability approaching one. In our model, the asymptotic case also serves an additional purpose: While Proposition 4 held when costs are uniformly distributed, the analogous asymptotic result, Proposition 6 below, holds regardless of the distribution of costs.

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

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

The second property requires that, for large n, the distribution π^* places almost all the weight on large voter 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.

Many commonly used families of discrete distributions satisfy these properties. Clearly, when the number of potential voters is a fixed size, n, the above properties hold. Likewise, voting populations drawn from Poisson, Negative Binomial, or Geometric distributions also have the required properties. It rules out situations in which, for instance, the population is some fixed amount n' with probability ε and n with the remaining probability.

Consider a sequence of equilibria, one for each n. Let $p_A\left(n\right)$ and $p_B\left(n\right)$ 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 $1/n$. As a result, the expected number of voters of each type is unbounded.

Proposition 5 *In any sequence of equilibria, the participation rates $p_A\left(n\right)$ and $p_B\left(n\right)$ tend to zero, while the expected number of voters $np_A\left(n\right)$ and $np_B\left(n\right)$ 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 participated 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 (2) and (3) is inconclusive.

The limiting pivot probabilities can be derived if we instead rewrite them as follows:

$$\Pr[T \mid m] = \left[\frac{1}{m}\sum_{r=0}^{m-1}\left(\omega^r q_A + \omega^{-r} q_B + q_0\right)^m\right] - (q_A)^m - (q_B)^m \tag{7}$$

where $\omega = \exp(2\pi i/m)$ is a primitive mth root of unity (see Appendix A for a derivation). The advantage of the new formula is that it does not contain any combinatorial terms. To see how this formula is derived, notice that $\Pr[T \mid m]$ is a sum composed of those terms in the trinomial expansion of $(q_A + q_B + q_0)^m$ with the property that the exponent on q_A is the same as the exponent on q_B. An analogous idea may be seen in the case of binomial expansion of $(x+y)^m$. Suppose that one were only interested in the portion of the expansion where the exponent on y is even. One could obtain this by taking the whole expression and subtracting off the odd terms via the formula $\frac{1}{2}(x+y)^m + \frac{1}{2}(x-y)^m$. Equation (7) is analogous in that it starts with the entire trinomial expansion and subtracts off all of the non-tie terms.

Similarly, the chance of a near tie can be written as

$$\Pr[T_{-1} \mid m] = \left[\frac{1}{m}\sum_{r=0}^{m-1}\omega^r\left(\omega^r q_A + \omega^{-r} q_B + q_0\right)^m\right] - m\,(q_A)^{m-1} q_0 \tag{8}$$

(again, see Appendix A).

We use the formulae in (7) and (8) to show that the participation rates indeed converge to zero. To see why, suppose that $q_A(n)$ and $q_B(n)$ were bounded from below by positive numbers in the limit. In equation (7), each term in the sum, save for $r = 0$, has an absolute value strictly less than one. This fact can then be used to show that the term in square brackets goes to zero in the limit. The remaining terms obviously also go to zero. This, however, implies that the benefit from voting goes to zero in the limit and hence the limiting participation rates cannot be strictly positive.

The second part of Proposition 5 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.

Knowing the limiting properties of participation and turnout, we can now extend Proposition 4 to all cost distributions when the election is large. The main idea is that, since the threshold participation rates go to zero in the limit, only the local properties of the cost distribution in the neighborhood of zero matter. Since locally, all distributions are approximately uniform, it follows that voting behavior also mirrors the uniform case—the utilitarian choice will always attract greater vote share than the rival candidate. Formally,

Proposition 6 *Suppose voting costs are distributed according to a continuous distribution F satisfying $F(0) = 0$ and $F'(0) > 0$. In any equilibrium, the expected number of votes for A exceeds the expected number of votes for B if and only if A is the utilitarian choice. Precisely, $q_A > q_B$ if and only if $\lambda v_A > (1-\lambda) v_B$.*

Proof. For a general cost distribution F satisfying $F'(0) > 0$, let $q_A(n) = \lambda p_A(n)$ and $q_B(n) = (1-\lambda) p_B(n)$ be a sequence of equilibrium voting propensities. Since Proposition 5 implies that p_A and p_B go to zero as n increases without bound, it is the case that the pivotal probabilities go to zero as well. This in turn implies that for all v, the cost thresholds $c_A(v)$ and $c_B(v)$ also go to zero. Thus, for large n, the equilibrium conditions (4) and (5) imply[8]

$$\begin{aligned} q_A &\approx \lambda \int_0^1 F'(0)\, v \Pr[Piv_A]\, dG_A(v) = F'(0)\, \lambda v_A \Pr[Piv_A] \\ q_B &\approx (1-\lambda) \int_0^1 F'(0)\, v \Pr[Piv_B]\, dG_B(v) = F'(0)\,(1-\lambda)\, v_B \Pr[Piv_B] \end{aligned}$$

In ratio form, we have

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

which is asymptotically identical to the case of uniform costs and we know from Proposition 4 that $\lambda v_A > (1-\lambda) v_B$ implies $q_A > q_B$. Thus, in large elections we have that $\lambda v_A > (1-\lambda) v_B$ implies $q_A > q_B$. ■

Main Result

We are now in a position to present the main result of the paper: 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 (and of the same order of magnitude), the 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. Formally,

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

Proof. Suppose that $\lambda v_A > (1-\lambda) v_B$ so that A is the utilitarian choice. Proposition 6 implies that in any sequence of equilibria, for all large n, $q_A > q_B$. We now show that the probability that A is elected approaches 1.

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

Denote by T_{-k} the event that the number of votes for B less the number of votes for A is k. Then,

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

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

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

We denote probabilities by $\mathcal{P}$ when they are calculated using Poisson distributions. When m is large, we know that

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

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

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

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

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

As $m \to \infty$,

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

and since in large elections, for any K,

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

we have that

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

This completes the proof. ■

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 utilitarian outcome.

While the main result is robust in many directions, it does require the following key ingredients. The first is that the lower support of the cost distribution be zero. Were costs bounded away from zero, then, provided costs were sufficiently low, nothing would change for finite sized electorates. In the limit, however, the situation would be analogous to the model with fixed voting costs and hence there would be insufficient flexibility in the "prices" of votes to produce the utilitarian outcome. Likewise, the result requires that duty considerations of voting do not overwhelm voting costs. Formally, if there was a positive mass of voters with zero or negative costs of voting then, in large elections, these would be the only voters coming to the polls and, since costs are orthogonal to preferences, the result would be identical to compulsory voting which, as we showed, is not utilitarian. Finally, the costs for A and B voters need to be identical, as least in the neighborhood of zero. If the cost distributions of the two sides were different, say F_A and F_B, and were such that $F_A'(0) \neq F_B'(0)$, then majority rule would maximize a weighted utilitarian welfare function where the weights are determined by $F_A'(0)$ and $F_B'(0)$.

5 Supermajority Rules

Majority rule is probably the most commonly used voting rule in practice; however, there are many situations where the outcome is decided by a supermajority. Many US states, including California and Arizona, require a supermajority vote in the legislature for any tax increase.[9] Moreover, some states, such as Florida and Illinois, require a supermajority among all voters to pass constitutional amendments. We showed that large elections with costly voting produce the utilitarian outcome under majority rule. Since turnout endogenously adjusts to favor the utilitarian choice, one might speculate that the same forces will work in supermajority elections as well. Indeed, Feddersen and Pesendorfer (1998) demonstrate a "voting rule irrelevance" in a Condorcet-type model with pure common values. In their setting, voting behavior endogenously adjusts in response to changes in the voting rule. As a result, information aggregates under all supermajority rules (with the exception of unanimity). In

[9] The required legislative supermajorities differ across states. Arizona and California, among others, require a 2/3 majority. Others—Arkansas and Oklahoma—require a 3/4 majority for certain types of tax increases. Still others, such as Florida and Oregon, require a 3/5 majority.

what follows, we argue that such a rule irrelevance does not hold in our model—in fact, only majority rule is utilitarian

We study supermajority 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 $a \geq b$. In the event of a tie—a situation in which A obtains exactly μ proportion of the votes—the winner is chosen at random, A with probability t and B with probability $1-t$. Note that for majority rule, $a = b = 1$ whereas, say for the two-thirds supermajority rule, $a = 2$ and $b = 1$.

In this section, we suppose that the population of voters follows a Poisson distribution with mean n. We will then show that unless the voting rule is one of simple majority ($a = b = 1$), the outcome of a large election will not coincide with the utilitarian choice. This is sufficient to argue that only majority rule is utilitarian.

The key to our analysis is Proposition 7 which is a generalization, for large Poisson populations, of Proposition 2. This proposition shows again that in large elections, if A is on the losing side, that is, the ratio of voting propensities, q_A/q_B falls short of the required a/b, then the ratio of the pivotal probabilities $\Pr[Piv_A] / \Pr[Piv_B]$ exceeds b/a. Formally,

Proposition 7 *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 $\liminf \frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} \geq \frac{b}{a}$.*

Proof. See Appendix B. ■

To see why Proposition 7 is true, suppose that the vote ratio is approximately the required 2 : 1 under the 2/3 rule with a coin toss determining the winner of a tie. The proposition then implies that a vote for B is *twice* as likely to be pivotal as a vote for A. It would seem that the likelihood of throwing the election into a tie or breaking a tie would be the same for A votes as for B votes and, indeed, this is approximately true. However, unlike A votes, a vote for B can also "flip" the election by swinging the outcome from a sure loss to a sure win. For instance, if the votes of others tally to $(2k-1, k-1)$, favoring A, then one more vote for B will flip the election in B's favor. The chance of flipping the election in this way is approximately equal to the chance of a tie or a near tie. However, the flip term receives twice the weight in a B supporter's calculation of the benefits since it does not lead to or break a tie. As a consequence, $\Pr[Piv_B]$ is approximately twice as large as $\Pr[Piv_A]$.

With Proposition 7 in hand, we are now in a position to offer the main result of this section:

Theorem 2 *Among all supermajority rules only majority rule is utilitarian. Specifically, in a $\frac{a}{a+b}$ supermajority election with a large Poisson population, if*

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

then A wins with probability one. If the reverse inequality holds strictly, then B wins with probability one.

Proof. Suppose that $\lambda v_A > \left(\frac{a}{b}\right)^2 (1-\lambda) v_B$. 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 that there is a sequence of equilibria along which $\frac{q_A}{q_B} \leq \frac{a}{b}$ and so by Proposition 7, along this sequence $\frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} \geq \frac{b}{a}$. If voting costs are uniform, the equilibrium conditions imply

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

But since the left-hand side is less than or equal to a/b while the right-hand side is strictly greater than a/b, this is a contradiction.

The remainder of the proof, showing that when $\frac{q_A}{q_B} > \frac{a}{b}$ holds for all large n, it is the case that $\Pr[A \text{ wins}] \to 1$, is the same as in Theorem 1 and is omitted. ■

When $a > b$, supermajority rules, of course, bias the electoral outcome in favor of B. It is then natural to conjecture that the outcome of a large supermajority election maximizes a *weighted* utilitarian welfare function in which the utilities of B supporters get a weight of a/b relative to the utilities of A supporters. Theorem 2, however, says that supermajority rules exaggerate the bias—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 A supporters. For instance, the 2/3 supermajority rule—which requires A to obtain *twice* as many votes as B—maximizes a weighted utilitarian welfare function in which the weight on the welfare of B supporters is not twice, but *four* times that placed on the welfare of A supporters.

Why do supermajority rules have the "squaring property"? The key is that there are two forces benefiting B. Obviously, making the winning threshold for A higher than that for B directly favors the latter. The indirect effect arises through Proposition 7. Under a supermajority rule, the likelihood that a B vote is pivotal is higher than the likelihood that an A vote is pivotal—when the election is approximately tied, there are more opportunities for B voters to flip the election than A voters. This increases the incentives for B voters to show up and so increases their participation rates as well, thereby making it even harder for A to win. Together, the two effects—the direct one via the voting rule and the indirect one via turnout—lead to the squaring property on the implicit welfare weights. Note that while the first effect is present even when voting is compulsory, the second effect arises only when voting is voluntary and costly.

While Theorem 2 delineates election outcomes in large elections, the following example suggests that the asymptotic results are well-approximated even when the size of the electorate is relatively small.

Example 3 *Consider the 2/3 majority rule. Suppose that the expected size of the population $n = 1000$ and that voting costs are uniform.*

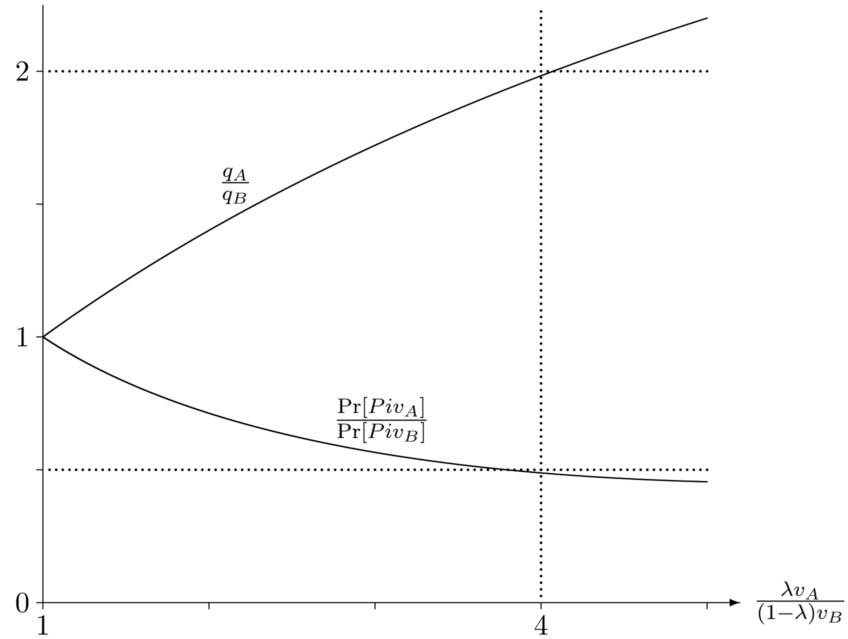

Figure 3: Ratios of Votes and Pivot Probabilities: 2/3 Rule

Figure 3 depicts the equilibrium ratio of the expected number of votes for A versus B, q_A/q_B, as a function of the welfare ratio, $\lambda v_A/ (1-\lambda)\, v_B$. In the example, even with a small number of voters, it is (approximately) the case that $q_A > 2q_B$ if and only if $\lambda v_A > 4 \times (1-\lambda)\, v_B$.

6 Aggregate Uncertainty

The model assumes that the distribution of preferences in the population is commonly known. A consequence of this assumption is that A and B supporters hold identical views about the distribution of preferences in the population. In reality, this seems rather doubtful. More plausibly, A supporters are likely to view the preference distribution as being more favorable toward A and likewise for B supporters. To capture this idea in a parsimonious way, we introduce aggregate uncertainty about the fraction of A supporters, λ, in the population. As a result, voter preferences are now correlated—A supporters place more weight on higher values of λ compared to B supporters. We investigate whether our earlier conclusions about the connection between majority rule and utilitarian outcomes are robust to this type of aggregate uncertainty.

There are good reasons to doubt that positive results concerning majority voting are robust to aggregate uncertainty. In a pure common values model, Mandler (2012) shows that aggregate uncertainty (about the accuracy of voters' signals) leads to a sequence of equilibria where information does not aggregate in the limit.[10] Thus,

[10] Feddersen and Pesendorfer (1997) derive the opposite conclusion in a model with aggregate uncertainty. The difference between the two conclusions stems from orders of limits. Feddersen and

aggregate uncertainty weakens the conclusions about the Condorcet Jury Theorem.

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 a B supporter. Precisely, suppose that λ is distributed according to a continuous density h on $[0,1]$ with mean $\overline{\lambda} \in (0,1)$. An A supporter's posterior density is

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

while a B supporter's posterior is

$$h_B(\lambda) = \frac{h(\lambda)(1-\lambda)}{\int_0^1 h(\theta)(1-\theta)d\theta} = h(\lambda)\frac{1-\lambda}{1-\overline{\lambda}} \tag{10}$$

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

When λ is uncertain, voters participate based on the *expected* pivot probabilities, calculated according to their posterior beliefs h_A or h_B. Our goal here is to explore how majority rule fares under these circumstances. For simplicity, we restrict attention to elections with large Poisson populations.

Two facts are key to our analysis. 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-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 value—for all $\lambda \neq \lambda^*$, the ratio $\mathcal{P}[Piv_A \mid \lambda]/\mathcal{P}[Piv_A \mid \lambda^*]$ goes to zero as n increases. Figure 4 depicts $\sqrt{n}\mathcal{P}[Piv_A \mid \lambda]$ 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 $\mathcal{P}[Piv_A \mid \lambda]$ spikes at λ^* means that for large n, the expected pivot probability $E_\lambda[\mathcal{P}[Piv_A \mid \lambda]]$ is determined solely by the values of the function in a small neighborhood around λ^*. The same is true for $E_\lambda[\mathcal{P}[Piv_B \mid \lambda]]$. 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

Pesendorfer let aggregate uncertainty go to zero and then let the number of voters grow large whereas Mandler does the opposite.

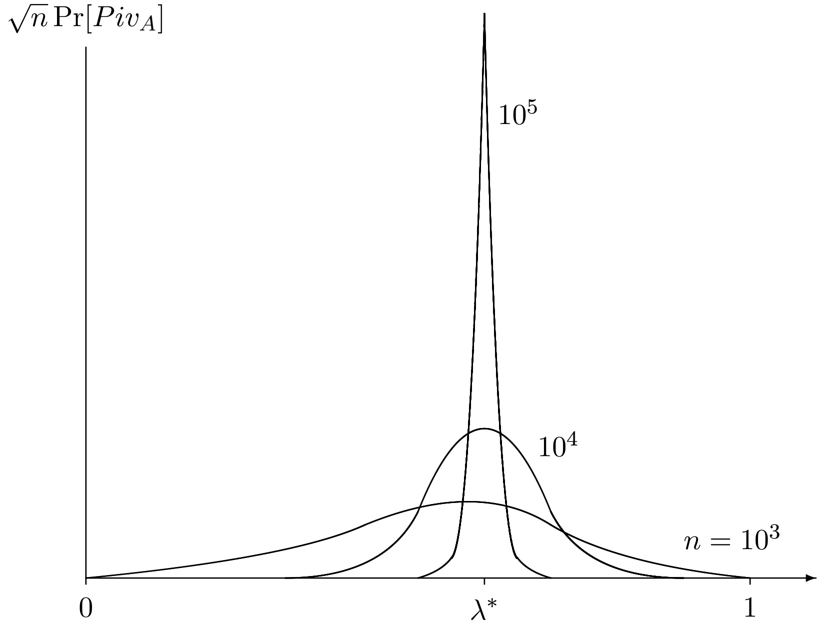

Figure 4: Asymptotic Behavior of Pivot Probabilities

m tails, is

$$\int_0^1 \binom{2m}{m} q^m (1-q)^m \, dq = \frac{1}{2m+1} \tag{11}$$

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

$$\lim_{m\to\infty} 2m \int_0^1 \binom{2m}{m} q^m (1-q)^m h(q) \, dq = h\left(\tfrac{1}{2}\right) \tag{12}$$

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, in this setup, 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 and varies with the expected size of the electorate, n. Consequently, the voting propensities, $q_A = \lambda p_A$ and $q_B = (1-\lambda) p_B$, also vary with n. Hence, we require an analog to the result above that accounts for this dependence. Proposition 8 is the Poisson generalization of equation (12) to situations with endogenous participation.[11]

[11] 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* $(0,1)$,

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

and the same equality holds for Piv_B *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 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 and the critical value of the parameter subject to aggregate uncertainty is the basis of Mandler's negative result. In a Condorcet setting, he shows that there exist equilibria in which voters' behavior is driven almost entirely by the "critical," incorrect, signal precision and, as a consequence, information need not aggregate, even in the limit. Hence, the conclusion of the Condorcet Jury Theorem is thrown into doubt.

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. 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 enough information so that the utilitarian outcome continues to prevail.

The following result is the analog of Proposition 2 when there is aggregate uncertainty. Like the earlier proposition, it shows that a vote on the "losing" side is more likely to be pivotal than one on the "winning" side. Note that because of aggregate uncertainty, "winning" and "losing" are defined relative to the average composition of the population.

Proposition 9 *Suppose that there is a sequence of elections such that* $\frac{p_B}{p_A+p_B} \to \lambda^* \in (0,1)$. *Then for large* n, $\overline{\lambda} p_A < \left(1-\overline{\lambda}\right) p_B$ *if and only if*

$$\int_0^1 \mathcal{P}\left[Piv_A \mid n,\lambda\right] h_A\left(\lambda\right) d\lambda > \int_0^1 \mathcal{P}\left[Piv_B \mid n,\lambda\right] h_B\left(\lambda\right) d\lambda$$

Proof. Proposition 8 implies that for large n,

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

and substituting from (9) and (10) we have

$$\frac{h_A(\lambda^*)}{h_B(\lambda^*)} = \frac{\lambda^*}{1-\lambda^*}\frac{1-\overline{\lambda}}{\overline{\lambda}} \approx \frac{\left(1-\overline{\lambda}\right)p_B}{\overline{\lambda}p_A}$$

This completes the proof. ■

Before proceeding further, it is worthwhile to extend the *ex ante* utilitarian benchmark to settings where aggregate uncertainty is present. Here, a utilitarian planner will choose candidate A if and only if the expected welfare of A supporters is higher than that of B supporters. Since the expected fraction of A supporters is $\overline{\lambda}$, candidate A is the utilitarian choice if and only if

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

When costs are uniform, the equilibrium conditions under aggregate uncertainty are determined in the usual way: Voters participate so long as the benefits from voting outweigh the costs. This yields equilibrium cost thresholds:

$$p_A = v_A \int_0^1 \mathcal{P}\left[Piv_A \mid n, \lambda\right] h_A(\lambda)\, d\lambda \tag{13}$$

$$p_B = v_B \int_0^1 \mathcal{P}\left[Piv_B \mid n, \lambda\right] h_B(\lambda)\, d\lambda \tag{14}$$

Multiplying by $\overline{\lambda}/\left(1-\overline{\lambda}\right)$ expresses this same condition in terms of expected vote shares. That is,

$$\frac{\overline{\lambda}p_A}{\left(1-\overline{\lambda}\right)p_B} \approx \frac{\overline{\lambda}v_A}{\left(1-\overline{\lambda}\right)v_B} \times \frac{\int_0^1 \mathcal{P}\left[Piv_A \mid n, \lambda\right] h_A(\lambda)\, d\lambda}{\int_0^1 \mathcal{P}\left[Piv_B \mid n, \lambda\right] h_B(\lambda)\, d\lambda}$$

The left-hand side of this expression is the expected vote share. The right-hand side is the product of the welfare ratio and the likelihood ratio of supporting A versus B conditional on the critical value, λ^*. Suppose that there was a sequence of elections in which the expected vote share for A exceeded that for B despite the fact that B is the utilitarian outcome. Then the left-hand side is greater than one while the welfare ratio is fractional. But Proposition 9 implies that the ratio of expected pivot probabilities is also fractional. This is a contradiction. An analogous argument establishes that the same holds when A is the utilitarian choice. Thus, we have shown

Proposition 10 *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 A is the utilitarian choice. 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 proposition shows that the introduction of aggregate uncertainty does not fundamentally change our earlier conclusion that vote shares coincide with the utilitarian outcome. Indeed, an analogous line of proof operates in this setting as well. The key is that, while vote propensities are determined by the critical value of λ,the expected benefits from voting depend on posterior beliefs about this value, and these posteriors account for the base rate, $\overline{\lambda}$, as well.

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

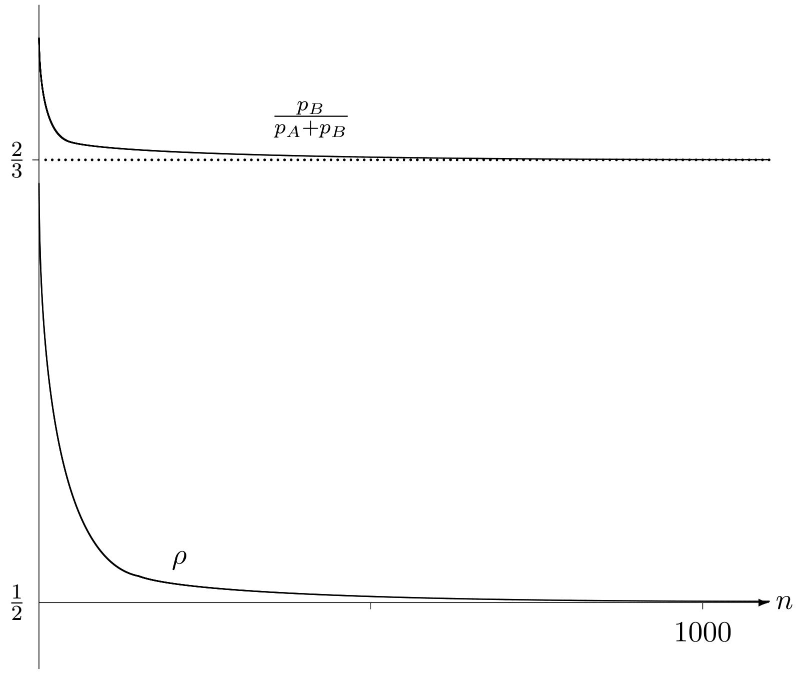

Figure 5: Participation Rates and Expected Pivot Probabilities

Example 4 *Suppose $v_A = 1$, $v_B = 4$ and the fraction of A supporters, λ, is uniformly distributed.*

From (9) and (10), the posterior beliefs of A and B voters in the example are $h_A(\lambda) = 2\lambda$ and $h_B(\lambda) = 2(1-\lambda)$, respectively. Since $\overline{\lambda} = \frac{1}{2}$, the equilibrium conditions (13) and (14) imply 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 5 depicts the equilibrium values of $p_B/(p_A + p_B)$ and ρ as functions of the expected number of voters, n. Notice how rapidly these quantities converge to their limits.

While Proposition 10 shows that the vote shares favor the utilitarian candidate in large elections, they offer no guarantee that this candidate will indeed be elected with certainty. The main result of this section shows that our earlier conclusion that the utilitarian candidate is selected almost certainly extends to the introduction of small amounts of aggregate uncertainty. To establish this, some care is needed with the order of limits. We will examine the limiting properties of majority rule as the distribution of λ becomes degenerate, i.e., as the aggregate uncertainty vanishes.

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*

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

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 (13) and (14) 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) and (14) 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 $r \geq R$,

$$\lim_{n\to\infty} \frac{\overline{\lambda}_r p_A(n,r)}{\left(1-\overline{\lambda}_r\right) p_B(n,r)} > 1 \tag{15}$$

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

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

But now as in the proof of Theorem 1 we have 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. ■

7 Conclusion

Majority rule is, perhaps, the most common means of group decision making. Whether it be mundane problems like a group decision of where to go to lunch, or deeply consequential decisions like the election of a president, the same rule is used. Its ubiquitousness perhaps stems from its simplicity and perceived fairness as compared to other voting rules. However, majority rule is perceived to have a key defect as well. Since it is purely a counting rule, it only reflects the directions and not the intensity of preferences.

When voting is voluntary, we show that majority rule suffers from no such defect. Preference intensity is reflected in participation rates. Of course, this feature is not unique to majority rule—preference intensity is reflected in participation rates for all supermajority rules as well. But majority rule is unique in that it aggregates preference intensity information "correctly" to produce utilitarian outcomes. Other rules distort participation incentives so as to overweight the issue or candidate favored by the rule. Moreover, the utilitarian property of majority rule is quite durable. It survives in contexts where the number of voters, preference intensity, voting costs, and even the fraction of voters on either side of an issue is random.

Of course, for majority rule to work requires that voters be permitted to abstain. It is by this (in)action, as well as coming to the polls, that the intensity of preferences is reflected in outcomes. Regimes that insist on participation, often on grounds that all voices will be heard, merely distort outcomes away from utilitarianism. Ironically, by insisting on participation, such a regime, in effect, mutes the loudest and most passionate of these voices while amplifying the voices of those who feel less strongly. Indeed, the main policy lesson from our results is that majority rule, combined with the freedom to participate (or not), is essential to producing good outcomes.

The coincidence of election outcomes with utilitarianism relies essentially on the "rational voter" assumption. Moreover, voter preferences are instrumental—voters care only about electoral outcomes. If, however, expressive factors also mattered, i.e. a voter's payoff also depended directly on casting a particular vote regardless of the outcome, then the results would change, at least in large elections. In particular, regardless of the weight placed on expressive considerations, in large elections it would be as though these considerations were the only ones relevant to the voting decision. The specification of voting costs also matters. In particular, if a mass of voters has zero or negative voting costs, perhaps because they view voting as a civic duty, then in large elections these would be only voters who show up. In effect, the election would become equivalent to one with compulsory voting with the duty voters as the sole participants.

A Asymptotics

The purpose of this appendix is to provide a proof of Proposition 5. 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 (7) and (8). 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 $r < m$,[12]

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

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,

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

■

Lemma A.2

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

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

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

The following lemma uses the roots of unity formulae to study asymptotic properties of the pivotal probabilities 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,

$$\begin{aligned} \lim_{m\to\infty} \frac{1}{m} \sum_{r=0}^{m-1} \left(\omega^r x + \omega^{-r} y + z\right)^m &\leq \lim_{m\to\infty} \frac{1}{m} \sum_{r=0}^{m-1} \left|\omega^r x + \omega^{-r} y + z\right|^K \\ &= \lim_{m\to\infty} \frac{1}{m} \sum_{r=0}^{m-1} \left|\exp\left(2\pi i \tfrac{r}{m}\right) x + \exp\left(-2\pi i \tfrac{r}{m}\right) y + z\right|^K \\ &= \int_0^1 \left|\exp\left(2\pi i t\right) x + \exp\left(-2\pi i t\right) y + z\right|^K dt \qquad (16) \end{aligned}$$

using the definition of the Riemann integral.

Since

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

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

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

Thus, for all $t \neq 0$, $\left|\exp\left(2\pi i t\right) x + \exp\left(-2\pi i t\right) y + z\right| < 1$. Hence, the integral on the right-hand side of (16) 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

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

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

A.2 Asymptotic Participation Rates

We begin with a lemma that says that both aggregate participation rates cannot remain positive in the limit.

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, $\lim_{m\to\infty} \Pr[Piv_A^* \mid m] = 0$.

Fix any $\varepsilon > 0$. Then there is a K such that for all n large, for all $m > K$, $\Pr[Piv_A^n \mid m] < \varepsilon$. As a result,

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

But since $\lim_{n\to\infty} \sum_{m=0}^{K} \pi_n(m) = 0$, 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 (4) and (5) now imply that along the subsequence, $\lim p_A(n) = 0$ and $\lim p_B(n) = 0$, contradicting the initial supposition. ■

Next we show that in the limit, the participation rates are of the same magnitude.

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 2, $\Pr[Piv_A] > \Pr[Piv_B]$. 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

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

Second, for all m,

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

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 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 the number of voters is a multiple of $a+b$. The set of ties is thus

$$T = \{(la, lb) : l \geq 0\} \tag{17}$$

A2. Candidate A is one vote short of a tie. The set of such events is[13]

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

[13] 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.[14] 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 (17).

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

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

B3. 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

$$\mathcal{P}[T] = \sum_{k=0}^{\infty} e^{-nq_A} \frac{(nq_A)^{ka}}{(ka)!} e^{-nq_B} \frac{(nq_B)^{kb}}{(kb)!} \tag{18}$$

[14] $\lceil z \rceil$ 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 (18), can be approximated as follows:

$$\mathcal{P}[T] \approx \frac{\exp\left((a+b)\left(\frac{nq_A}{a}\right)^{\frac{a}{a+b}}\left(\frac{nq_B}{b}\right)^{\frac{b}{a+b}} - nq_A - nq_B\right)}{\left(2\pi(a+b)\left(\frac{nq_A}{a}\right)^{\frac{a}{a+b}}\left(\frac{nq_B}{b}\right)^{\frac{b}{a+b}}\right)^{\frac{1}{2}}(ab)^{\frac{1}{2}}} \tag{19}$$

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

$$\mathcal{P}[T-(j,k)] \approx \mathcal{P}[T] \times x^{bj-ak} \tag{20}$$

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 (19) and (20):

$$\mathcal{P}[Piv_A] \approx \mathcal{P}[T] \times \left[1-t+tx^b+\sum_{k=1}^{b-1} x^{b\left\lceil \frac{a}{b}k\right\rceil - ak}\right] \tag{21}$$

$$\mathcal{P}[Piv_B] \approx \mathcal{P}[T] \times \left[t+(1-t)x^{-a}+\sum_{j=1}^{a-1} x^{bj-a\left\lceil \frac{b}{a}j\right\rceil}\right] \tag{22}$$

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 (21) and (22) as

$$\mathcal{P}[Piv_A] \approx \mathcal{P}[T] \times \left[1-t+tx^b+\sum_{k=1}^{b-1} x^{k}\right] \tag{23}$$

$$\mathcal{P}[Piv_B] \approx \mathcal{P}[T] \times \left[t+(1-t)x^{-a}+\sum_{j=1}^{a-1} x^{-j}\right] \tag{24}$$

Proof of Proposition 7. Using the formulae in (23) and (24), 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 $x = 1$,

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

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 $\frac{\mathcal{P}[Piv_A]}{\mathcal{P}[Piv_B]} \approx \frac{b}{a}$. 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 other case is analogous. ■

C Aggregate Uncertainty

Our goal in this appendix is to develop asymptotic formulae for the expected pivot probabilities when there is aggregate uncertainty.

In what follows, we make use of the following identity

$$\int_0^1 t^k (1-t)^l \, dt = \frac{1}{k+l+1} \binom{k+l}{k}^{-1} \tag{25}$$

The identity is easily verified by using induction on k (say). Note that (25) is a generalization of (11).

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*

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

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 (25),

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

and so the left-hand side of (26) 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

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

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*

$$
\int_0^1 \sum_{k=0}^{m} \binom{m}{k,k+1} (q_A)^k (q_B)^{k+1} (q_0)^{m-2k-1} \, d\lambda = \frac{1}{m+1} \left(1 - (1-x)^m\right) \tag{27}
$$

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*

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

Proof. Since

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

Lemma C.1 implies that

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

■

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

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

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 Corollary 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*

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

where $\mathcal{P}\left[T_{-1} \mid n,\lambda\right]$ *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*

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

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 probability of Piv_A, observe that for all $\lambda \neq \lambda^*$,

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

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

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

and $n\,(p_A+p_B) \to \infty$, it follows that $n\,(\phi(\lambda)-\phi(\lambda^*)) \to -\infty$. This implies that the ratio $\mathcal{P}\left[Piv_A \mid n,\lambda\right]/\mathcal{P}\left[Piv_A \mid n,\lambda^*\right]$ 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, $\mathcal{P}[Piv_A \mid n, \lambda''] > \mathcal{P}[Piv_A \mid n, \lambda']$. Analogously, for all λ', λ'' satisfying $\lambda^* + \varepsilon < \lambda' < \lambda''$, $\mathcal{P}[Piv_A \mid n, \lambda'] > \mathcal{P}[Piv_A \mid n, \lambda'']$ 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

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

and

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

as well. Thus,

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

using Lemma C.5.

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

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

Thus, we have

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

and so

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

or

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

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

D Poisson Approximations of the Multinomial

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

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

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

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

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

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

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 (1997): "Voting Behavior and Information Aggregation in Elections with Private Information," *Econometrica*, 65 (5), 1029–1058.

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

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

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

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

[13] Myatt, D. (2012): "A Rational Choice Theory of Voter Turnout," Working Paper, London Business School, January.

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

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

[16] Palfrey, T. and H. Rosenthal (1983): "A Strategic Calculus of Voting," *Public Choice*, 41, 7–53.

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

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

[19] 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