Is Mandatory Voting Better than Voluntary Voting?

Stefan Krasa   Mattias Polborn*

September 8, 2005

Abstract

We use a costly voting model to investigate whether policies that increase voter turnout are socially beneficial. The model generalizes Börgers [2] by assuming that the probability of being an A-supporter, α, is initially drawn from a distribution. We show that (generically) when the number of citizens is large, voter participation is inefficiently low (unlike in Börgers). Specifically, there exists a subsidy for voters that implements the efficient choice with probability 1, for a per capita voting cost of (approximately) 0. We also derive conditions when mandatory voting dominates voluntary participation voting from an ex-ante welfare perspective.

JEL Classification Numbers: C70, D72.
Keywords: Costly voting, mandatory voting, compulsory voting, externalities

1 Introduction

Empirically, many societies appear to encourage voter participation in elections and meetings. For example, in the recent 2004 U.S. elections, many states expanded the opportunities for early and absentee voting, and in the U.K. election day is a public holiday. Several other countries have tried to increase turnout by making participation in elections "mandatory".[1] For example, the Australian parliament enacted mandatory voting in 1924, because voter turnout had dropped below 60 percent. According to the law all Australian citizens over the age of 18 must be registered to vote and show up at the polling place on election day. A citizen who does not go to the polling place is subject to a $15 fine.[2] The Australian mandatory voting law appears to be successful, as voter turnout has been consistently above 90%.

The introduction of mandatory voting is also occasionally discussed in US editorials, especially around election time; see, e.g. Olbermann [10], Dean [4] and Weiner [11].[3] Supporters of mandatory voting see voting as a civic duty similar to paying taxes and argue that a higher level of participation increases the legitimacy of government. "The most important [argument] is that compulsory voting ensures that government does indeed represent the will of the whole population, not merely the section of the population that decides to express their opinions" (Wikipedia [12]).

We use a costly voting model to investigate whether policies that increase voter turnout are socially beneficial. Specifically, we address the following questions: Does subsidizing voters lead to different election outcomes than voluntary voting? If election outcomes are different, and if subsidized voting improves the social decisions, do these benefits outweigh the increased voting costs that are a consequence of higher voter turnout?

In our model N citizens have to make a decision among two candidates A and B. Each citizen's preference is private information and is independently drawn from a common distribution that assigns probability α to being an A-supporter (and $1 - \alpha$ to being a B-supporter). In addition, each individual has private information concerning the cost c that he must pay if he votes.

The parameter α is drawn at an interim stage, so that individuals know α when they decide whether or not to vote. For example, pre-election opinion polls of potential voters can provide this information. However, institutional choices (e.g., if voting should be subsidized or made compulsory) cannot be conditioned on the realization of α. This appears to be a reasonable assumption, as institutional choices usually apply for a longer time period than just a single issue election, and a rule that explicitly conditions on α (say, "choose

---

[1]These include most South American countries, as well as Australia and several European and Asian countries. See, for example, Wikipedia [12] for a list.

[2]All non-voters receive a letter asking them to pay the fine. Instead of paying the fine, non-voters can also provide a written excuse. Thus, the actual cost of non-voting is the minimum over the disutility of paying $15 or writing the letter (see Weiner [11] for details).

[3]There were a few experiments with mandatory voting laws before 1900 in the US. In 1896, the Supreme Court of Missouri struck down a Kansas City charter provision as unconstitutional that assessed a $2.50 poll tax on every man twenty-one years of age or older who failed to vote in the general city election. See Dean [4].

A without an election if $\alpha > \frac{1}{2}$") would likely lead to large controversies among A and B-supporters as to what α is.

We show that voting should always be subsidized, if there are sufficiently many citizens. In this case, a subsidy equal to the minimal voting cost can implement the first best decision at a negligible per capita cost. Subsidized voting dominates voluntary voting for *any* (non-degenerate) distribution of α. In fact, we show that, at the interim stage, citizens are better off with subsidized voting, unless $\alpha = 0.5$ is realized. We also characterize conditions under which compulsory voting (where citizens are forced to vote) is better from a welfare perspective than voluntary participation in elections. Furthermore, we show numerically that our main insights remain valid also for a smaller number of citizens.

The question of optimal policy towards voter participation has recently been analyzed in Börgers [2]. His analysis suggests that the externality from voting is negative, and participation in voting should therefore be discouraged at the margin rather than encouraged. In particular, he proves that (in his model), voting with voluntary participation dominates voting with mandatory participation from an ex-ante welfare perspective. Our model generalizes his setup by including the ex-ante stage where α is drawn, while $\alpha = \frac{1}{2}$ with certainty in Börgers [2]. Our theoretical results focus on the case that N, the number of potential voters, is large, while the results derived by Börgers [2] hold for any N, if $\alpha = \frac{1}{2}$. Numerically, we show that our main result holds for relatively small N as long as α is not too close to $\frac{1}{2}$. Our papers can therefore be interpreted as complementary: Börgers' results show that, when the expected absolute number of supporters is very similar for both candidates, then voting should be discouraged, while our results show that voting should be encouraged, if the expected difference in the absolute number of A- and B-supporters is large (which is generically true when the number of voters is large).

To understand why our results differ from Börgers [2] it is useful to recall the intuition for his result. The equilibrium in his model is characterized by a cost cutoff such that potential voters go to vote if and only if their cost of voting is lower than the cost cutoff. When deciding whether or not to vote, a given (say) A-supporter calculates the probability that he is pivotal if he votes for A, multiplies this probability with his benefit from a changed decision, and goes to the election if and only if this expected benefit is larger than the voting cost. The A-supporter is pivotal in only two cases: (i) If the number of A-voters and B-voters among the other individuals is exactly the same (in which case the tie will be broken in candidate A's favor), and (ii) if there is one more B-voter than A-voter (so that a tie is reached if the A-supporter we consider votes).

How do the private and social incentives to vote differ in this scenario? Consider the externality that an A-supporter imposes on other individuals if he decides to vote. In case (i), the positive externality on other A-voters and the negative externality on B-voters cancel each other, because there is the same number of A and B-voters (excluding our A-supporter). In case (ii) there is one more B-voter, and as a consequence the negative externality dominates. Börgers assumes that each individual is equally likely to favor either candidate. Our A-supporter therefore does not impose an externality on non-voters in expected terms. In summary, any person who votes imposes a negative externality on the other individuals, which leads to an inefficiently high level of voting participation. As a consequence, taxing voters can improve welfare.

Now consider our model, and suppose for example that the probability α that an individual supports candidate A is 0.6 at the interim stage. If our A supporter is pivotal, then the number of A and B-voters is approximately the same. As a consequence, even more than 60 percent of non-voters are A-supporters. Therefore, an A-voter imposes a positive externality and a B-voter imposes a negative externality on non-voters. Thus, from society's perspective there are too few A-voters and too many B-voters, and any policy that increases the expected number of A-voters more than the expected number of B-voters increases welfare. We show that both subsidized and compulsory voting have this desired effect.

The costly voting literature dates back to Ledyard [8]. Other papers that study related costly voting environments include Goeree and Grosser [7], Ghosal and Lockwood [6] and Campbell [3]. Goeree and Grosser [7] analyze a model similar to Börgers [2] and ours, except that the voting cost in their model is deterministic and equal for all voters. Their main focus is on the effects of opinion polls on turnout and welfare. They find that public opinion polls increase turnout, but decrease welfare, relative to the case that each voter only knows his own preference type. While they also show that the externality from voting is positive when $\alpha \neq \frac{1}{2}$, they do not study the optimal institutional setup to deal with this problem. Furthermore, since all players have the same voting costs, the subsidy solution that implements the first best in our model is less beneficial or even inferior to the voluntary participation equilibrium in their model.

Ghosal and Lockwood [6] also analyze a model with costly voting in which voluntary participation can be inefficiently low, but the cause of this inefficiency differs from ours. In their model, each individual has a "private value" preference for one of the politicians, but also a "common value" preference to select the politician who matches the unknown state of the world. Individuals have a utility function that is a convex combination of a private and common value components. Individuals first decide whether to participate in the costly voting, in which case they receive a private signal that partially reveals the state of the world. If the common value component dominates in the utility function so that individuals vote according to their private signals, there is always too little participation, because more voters lead in expectation to better decisions, and individuals neglect this positive externality. If, instead, the private value component dominates so that individuals vote according to their preference type, there is too much participation. The latter result is due to the assumption that each individual prefers A with a probability of exactly $\frac{1}{2}$, as in Börgers [2].

Campbell [3] analyzes a costly voting model and shows that, if supporters of different candidates have voting costs drawn from different distributions, then the candidate whose supporters constitute a majority among those agents with the least voting costs wins almost certainly in large electorates, independently of the percentage of people who prefer A to B. As this outcome may be inefficient, Campbell's effect may provide another reason for subsidizing participation (Campbell [3] does not analyze the effect of subsidies in his paper). We assume that A and B supporters' cost of voting are drawn from the same distribution and therefore Campbell's effect is absent in our model.

We present the model in the next section. Sections 3 and 4 contain the results. Concluding remarks are in Section 5. Proofs are in the Appendix.

2 Model

There are N individuals who have the right to vote for one of the two candidates, A or B. The probability α that an individual prefers candidate A to candidate B is chosen by nature according to a probability density function $g(\alpha)$, and becomes public information before the election. One can interpret the public information about α as the result of pre-election opinion polls (cf., Goeree and Grosser [7]). Preferences for candidates A and B, respectively, are then drawn independently across individuals according to probability α.

Participating in the election is costly. In particular, each individual's costs c is drawn independently according to a probability density function $f(c)$. We assume that $f(c)$ is strictly positive on its support $[\underline{c}, \bar{c}]$, where $\underline{c} \geq 0$. We write $F(c)$ for the cumulative distribution function that corresponds to $f(c)$.

The outcome of the election is determined by majority rule. In case of a tie, each candidate has an equal probability of winning of $\frac{1}{2}$. Individual i receives a benefit normalized to 1 if his preferred candidate is elected, and has a utility cost c_i if he participates in the election. Formally, let individual i's type be (P_i, c_i), where $P_i \in \{A, B\}$ is i's preferred candidate and c_i is i's voting cost. Let $v_i = 1$ and $v_i = 0$ be the decision whether or not to vote. If E is the candidate who is elected, then individual i's utility is given by

$$u(E, v_i; P_i, c_i) = \begin{cases} 1 - v_i c_i & \text{if } E = P_i; \\ -v_i c_i & \text{if } E \neq P_i. \end{cases}$$

We compare the benefits of four different types of social decision making:

1. *Voluntary Voting*: As in the standard costly voting model, each individual chooses whether or not to vote.

2. *Subsidized Voting*: All individuals who choose to vote receive a subsidy s, or equivalently, non-voters must pay a fine s. Of course, mandatory voting laws cannot physically force individuals to participate, but rather encourage participation through fines for non-votes. Thus, our notion of subsidized voting corresponds to how mandatory voting laws work in practice.

3. *Compulsory Voting*: All individuals are forced to participate. Compulsory voting is equivalent to subsidized voting with a subsidy that is larger than $\bar{c}$.

4. *Ex-ante Decision Making*: Individuals choose one of the two candidates before information about the candidates (i.e., about α), is revealed.

3 Results

We first establish the existence of an equilibrium for the four different social decision rules. We then characterize the winning probabilities of the two candidates under voluntary voting, and compare the expected utility under each of the four social decision rules. Since our theoretical welfare results obtain for the case that the number of voters is large, we provide some numerical results for a small electorate in Section 4.

Whenever an individual votes (and hence, in particular, with compulsory voting), his weakly dominant strategy is to vote for his preferred candidate. Under ex-ante decision making, all agents will choose candidate A if $E[\alpha] > \frac{1}{2}$, they will select candidate B if $E[\alpha] < \frac{1}{2}$ and either of the two candidates if $E[\alpha] = \frac{1}{2}$. For voluntary and subsidized voting there exists a symmetric equilibrium in pure strategies, which is characterized by a simple cutoff value rule for the voting costs: A-supporters choose to vote if and only if their voting costs are no higher than c_A, and B-supporters have an analogous cost threshold c_B that determines their voting behavior.

Proposition 1 *Under both voluntary voting and subsidized voting there exists a symmetric equilibrium in pure strategies. An equilibrium is characterized by cutoff values c_A and c_B such that individual i votes for his preferred candidate P_i if $c_i \leq c_{P_i}$, and abstains otherwise.*

We now turn to a more detailed characterization of the symmetric cutoff equilibrium. The following proposition shows that independent of the proportion α of A-supporters in the population, the equilibrium probability that candidate A wins under voluntary voting converges to $\frac{1}{2}$ as the number of potential voters N goes to infinity, provided that the minimum voting cost $\underline{c} > 0$.

Suppose, for example, that $\alpha > \frac{1}{2}$ so that the number of A-supporters is (almost certainly) greater than the number of B-supporters for N sufficiently large. This implies that the efficient social decision would be to elect candidate A. However, in equilibrium B-supporters are more likely to vote — in the limit exactly balancing the higher number of A-supporters — so that each candidate's winning probability converges to $\frac{1}{2}$. Hence, voting with voluntary participation very often leads to the wrong social decision.

Proposition 2 *Suppose that $\underline{c} > 0$. Then the probability that candidate A wins the election under voluntary voting converges to $\frac{1}{2}$ as $N \to \infty$.*

The intuition for this result is as follows. Consider a particular A-supporter and let V_A and V_B be number of other A and B supporters who participate in the election. Our A-supporter is pivotal for A if and only if $V_A - V_B$ is either 0 or -1. If our A-supporter is pivotal, then he increases the probability that candidate A wins by $\frac{1}{2}$ (either from 0 to $\frac{1}{2}$ if he brings A into a tie, or from $\frac{1}{2}$ to 1 if he breaks a tie). Our A-supporter will vote if and only if this expected gross benefit of voting is larger than his cost of voting c_i. Thus, the cost cutoffs c_A must fulfill $\frac{1}{2} P(\{V_A - V_B \in \{-1, 0\}) = c_A$ In equilibrium, as N grows, only people with very low voting costs will vote (otherwise, there would be infinitely many voters and the probability of being pivotal would go to zero). Thus, c_A must converge to the lowest possible cost, $\underline{c}$, which implies that the pivot probabilities $P(\{V_A - V_B \in \{-1, 0\})$ converges to $2\underline{c}$. The same is true for a B-supporter's pivot probability $P(\{V_A - V_B \in \{0, 1\})$.

In order to make conclusions about winning probabilities for the two candidates, we must determine the limit distribution of V_A and V_B. We show that V_A and V_B converge to Poisson distributions, where the parameters are the expected number of voters $E[V_A]$ and $E[V_B]$, respectively. This implies that $E[V_A] = E[V_B]$ in the limit, because the pivot probabilities for A and B would differ otherwise. Therefore, V_A and

V_B are independent and identically distributed random variables. Hence, $P(\{V_A > V_B\}) = P(\{V_B > V_A\})$, which implies that each candidate has the same probability of winning the election.

The following corollary shows that expected per capita welfare with voluntary voting goes to $\frac{1}{2}$, as N goes to infinity. Independent of α, the gross per capita surplus is equal to the expected proportion of people who agree with the social decision, $\frac{1}{2}\alpha + \frac{1}{2}(1 - \alpha) = \frac{1}{2}$. Note that, since the participation rate goes to zero, the per capita voting cost that has to be deducted to get the net surplus, is zero.

Corollary 1 *If $\underline{c} > 0$ and voting is voluntary, then the ex ante expected per capita surplus converges to $\frac{1}{2}$, as $N \to \infty$.*

We now compare the surplus obtained under voluntary voting to the surplus achieved under compulsory voting.[4] If N is sufficiently large, the percentage of citizens who prefer A is almost certainly close to α, and hence A is implemented if $\alpha > \frac{1}{2}$, and B is implemented if $\alpha < \frac{1}{2}$. Consequently, for a given α, expected per-capita utility is

$$\max(\alpha, 1 - \alpha) - E(c), \tag{1}$$

where $E(c) = \int_{\underline{c}}^{\overline{c}} cf(c)\, dc$ is the average voting cost. Taking expectations over α implies that the ex-ante expected per-capita utility under compulsory voting is given by

$$\frac{1}{2} + E\left(\left|\alpha - \frac{1}{2}\right|\right) - E(c), \tag{2}$$

where $E\left(\left|\alpha - \frac{1}{2}\right|\right) = \int_0^{\frac{1}{2}}(\frac{1}{2} - \alpha)g(\alpha)d\alpha + \int_{\frac{1}{2}}^1(\alpha - \frac{1}{2})g(\alpha)d\alpha$ is the expected deviation of α from $\frac{1}{2}$. Note that the expected margin for the victorious candidate under compulsory voting is just $2E\left(\left|\alpha - \frac{1}{2}\right|\right)$. We have therefore proved the following Proposition 3:

Proposition 3 *When the number of voters N goes to infinity, compulsory voting leads to a higher ex-ante expected utility than voting with voluntary participation if and only if the expected margin of victory for the winning candidate under compulsory voting is larger than twice the average voting cost.*

While compulsory participation may be better than voluntary participation, it is intuitively clear that forcing everyone to vote is socially wasteful: As long as the voters participating in the election are representative for the population as a whole, the correct decision can still be made (with a probability close to 1), and for a considerably smaller per-capita voting cost. To analyze this case in more detail, consider what happens when a subsidy s is paid to all individuals who choose to vote (or, equivalently, a fine s is imposed on all those individuals who do not vote).[5]

---

[4]Remember that, as a practical matter, "mandatory" voting laws do not literally make voting compulsory; rather, they impose some penalty as the price of not voting, and some people may choose to pay this price rather than their voting cost. Still, it is interesting to analyze the limit case of a very large penalty that would effectively enforce participation by everybody.

[5]Note that the amount of the actual subsidy or penalty paid is irrelevant for welfare considerations, as it is a mere redistribution between different citizens and therefore cancels out on average. The relevant question for welfare purposes is how s changes the voters' behavior (i.e., for which cost realizations they go to vote).

Consider what happens when we set $s = \underline{c} + \varepsilon$, where $\varepsilon > 0$ but small. Clearly, all citizens with cost realizations between $\underline{c}$ and $\underline{c} + \varepsilon$ find it optimal to vote. For finite N, there may be in addition some types with costs slightly higher than s who vote, because the probability of being pivotal outweighs their net cost $c - s$. However, for any N, the expected number of participants is at least $F(s)N$, and hence goes to infinity as N increases. Consequently, the probability to be pivotal goes to zero and hence the equilibrium cost cutoffs for participation must converge to s. Therefore, the percentage of votes cast for A is almost certainly close to α, and A will be chosen if $\alpha > \frac{1}{2}$, and B will be chosen if $\alpha < \frac{1}{2}$, implementing the socially correct decision. Moreover, the per-capita voting cost under this scheme is about $F(s)s$, which is close to zero if s is sufficiently close to $\underline{c}$.

Proposition 4 *Suppose that a subsidy of $s = \underline{c} + \varepsilon$ is paid to every citizen who chooses to vote. As N goes to infinity, expected per-capita utility goes to*

$$\frac{1}{2} + E\left(\left|\alpha - \frac{1}{2}\right|\right) - \int_{\underline{c}}^{\underline{c}+\varepsilon} cf(c)dc \tag{3}$$

In particular, as ε can be chosen arbitrarily small, per-capita expected utility can reach (almost) $\frac{1}{2} + E\left(\left|\alpha - \frac{1}{2}\right|\right)$.

Hence, through an appropriately chosen subsidy, it is possible to implement the correct social decision with probability 1, while per capita expected voting costs can be made arbitrarily small.

Note that Proposition 4 requires $F(c)$ to be continuous (near $\underline{c}$). Only then is it possible to fine-tune the number of voters through subsidy s, such that enough voters participate to secure the correct result while simultaneously avoiding large voting costs. For example, in the model of Goeree and Grosser [7], all individuals are assumed to have the same voting cost c, so that subsidies could not implement the first best in their model.

Finally, we analyze if ex-ante decision making can be optimal. As argued above, ex-ante decision making may be unconstitutional in some applications, but it may be an available option in other cases. One way to think about ex-ante decision making is that voters do not yet know whether they prefer A or B, but they know is the distribution, $g(\alpha)$. Society can choose between either making the decision in the present meeting, or waiting for the realization of α and the voters' types and re-convene in another meeting with costly participation to make the decision.

If candidate A is selected ex-ante, the expected per-capita utility is $E[\alpha]$. Note that

$$E[\alpha] = \frac{1}{2} + E\left[\alpha - \frac{1}{2}\right] \leq \frac{1}{2} + E\left(\left|\alpha - \frac{1}{2}\right|\right),$$

where the equality is strict if $\alpha < \frac{1}{2}$ with positive probability. Similarly, suppose that candidate B is selected the ex-ante. Then the per-capita utility is

$$E[1 - \alpha] = \frac{1}{2} + E\left[\frac{1}{2} - \alpha\right] \leq \frac{1}{2} + E\left(\left|\alpha - \frac{1}{2}\right|\right),$$

where the equality is strict if $\alpha > \frac{1}{2}$ with positive probability. In both cases Proposition 4 implies that subsidized voting dominates choosing a candidate ex-ante if N is sufficiently large.

Now compare ex-ante decision making to voluntary voting. Note that $\max(E(\alpha), E(1-\alpha)) \geq \frac{1}{2}$, where the inequality is generically strict. If $N \to \infty$ then the law of large numbers implies that it is socially desirable to choose candidate A if $\alpha > \frac{1}{2}$, and candidate B if $\alpha < \frac{1}{2}$. In contrast, Proposition 2 implies that the wrong candidate will be chosen with probability $\frac{1}{2}$ under voluntary voting. Thus, ex-ante decision making strictly dominates voluntary voting for large N when $\alpha \neq \frac{1}{2}$.

We have therefore proved the following result.

Proposition 5

1. *Suppose that $0 < \int_0^{\frac{1}{2}} g(\alpha)\, d\alpha < 1$. Then subsidized voting with a subsidy $s = \underline{c} + \varepsilon$ strictly dominates ex-ante decision making if N is sufficiently large.*

2. *If $E(\alpha) \neq \frac{1}{2}$ then ex-ante decision making dominates voluntary voting if N is sufficiently large.*

4   Subsidies in small electorates

In the last section, we have established that, for a sufficiently large society, social welfare is always higher with an appropriately chosen subsidy than with voluntary participation, if the minimum voting cost is positive. In contrast, for $\alpha = \frac{1}{2}$ Börgers [2] shows that the negative externality dominates and voter turnout is inefficiently large. It is therefore interesting to see which of these effects dominates for intermediate values of N, given a distribution of voting costs and a distribution over α.

We computed an example with $N = 100$ voters. Voting costs are uniformly distributed between 0.03 and 0.06 (i.e., between 3 percent and 6 percent of an individual's benefit from getting the preferred candidate elected). In the left panels of Figure 1 we vary the probability α that an individual supports candidate A between 0.5 and 0.8. Figure 1 shows the values of the displayed variables per capita surplus, winning probabilities and participation rates, as a function of the level of α (i.e., at the interim stage, when α is known).

First consider the three panels on the left. The top panel displays per-capita surplus at the interim. For α close to 0.5 voluntary voting dominates compulsory voting, reflecting the result from Börgers [2]. However, once α reaches 0.57, compulsory voting starts to dominate. This has the following implication about the ex-ante optimality of the voting system. If α takes only values between (about) 0.43 and 0.57, then voluntary voting is certain to dominate compulsory voting at the ex-ante stage in a society of 100 voters. On the other hand, if the distribution of α puts sufficiently much weight on more extreme values of α, then compulsory voting will dominate voluntary voting.

The mid panel displays the probability that A wins under compulsory voting (i.e., the probability that A is the efficient choice because the number of A supporters is higher than the number of B supporters) and voluntary voting. Note that the probability that A wins under voluntary voting is considerably below
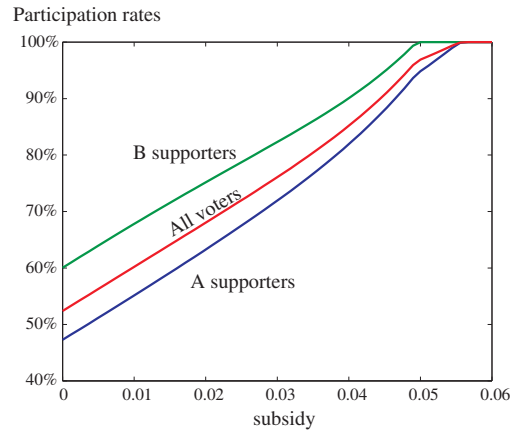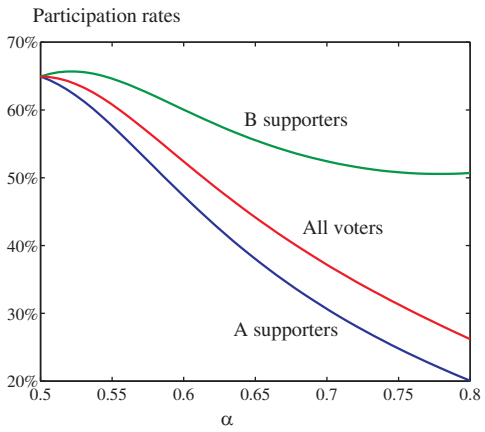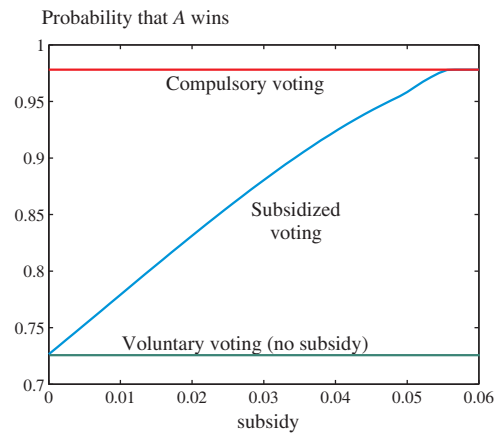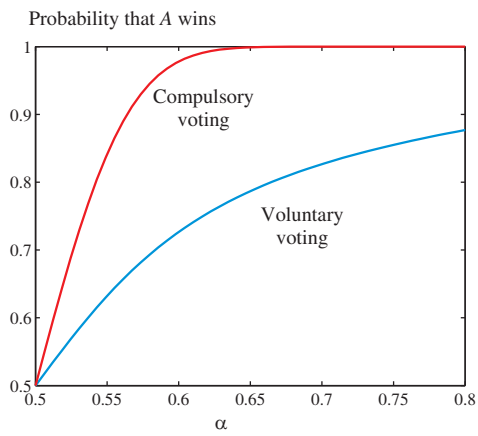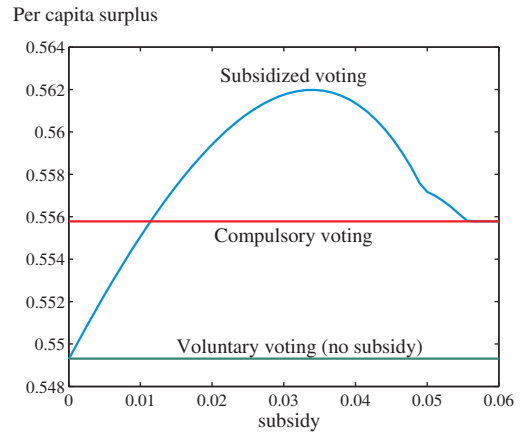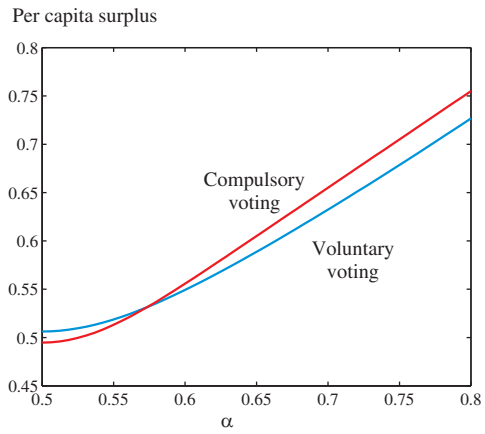
Figure 1: Comparing voluntary, compulsory, and subsidized voting at the interim stage

9

the probability that A is the efficient choice, and the difference is probability of an inefficient choice under voluntary voting. The reason why A might lose even when α is greater than $\frac{1}{2}$ is that B supporters are more likely to vote than A supporters. This effect can be seen in the bottom panel.

Now consider the effect of subsidized voting. In the right panels of Figure 1, we fix $\alpha = 0.6$ and vary the subsidy between 0 and 0.06. A subsidy of about 0.034 maximizes per-capita surplus for this parameter constellation. Subsidies significantly increase voter participation and dominate voluntary voting. The bottom panel on the right indicates that voter participation increases by about the same rate for A and B supporters. Because there are more A supporters than B supporters this implies that the subsidy increases the number of A supporters more strongly than the number of B supporters who go to the poll. As explained in the introduction, A-voters impose a positive externality and B-voters a negative externality on non-voters. As a consequence, it is the stronger increase of A supporters among all voters that is responsible for the increase in surplus.

5   Conclusion

Many societies provide incentives for voters to participate in elections. In this paper, we have provided a model in which costly voting induces suboptimal equilibrium participation and frequently leads to wrong choices. In such a world, providing incentives for citizens to vote will increase the quality of electoral decisions and social welfare. Specifically, we show that a subsidy for voters or a penalty for non voters that is slightly higher than the minimum voting cost implements the first best.

Our setup is probably the easiest model in which questions of costly voting can be studied: Citizens know which candidate they prefer, they only have to decide whether or not to vote, and the voting costs are drawn from the same distribution for both A and B supporters. Extending our model to allow for incomplete information about candidates and differential voting costs for A and B supporters should reinforce our qualitative finding that voting is similar to providing a public good. For example, when individuals fundamentally agree on which candidate is better in a given state of the world, so that voting aggregates information as in Ghosal and Lockwood [6], an individual citizen who becomes informed and votes increases the quality of public decisions for all his compatriots and hence too few citizens will provide this public good voluntarily. This suggests that the conclusion that it is efficient to subsidize rather than tax voters is quite robust.

Following Börgers [2], we have assumed that each voter's benefit from the election of his preferred candidate is normalized to one, while costs are random draws for individual voters. More generally, one could assume that each voter's benefit is also random. If both costs and benefits are drawn from the same distribution for A and B-supporters, the analysis for voluntary and compulsory voting is largely unchanged (up to a re-normalization): In particular, let θ be the citizen's benefit of getting his preferred candidate elected, and let c be the voting cost. Then a citizen votes in the alternative model if and only if a citizen with cost c/θ votes in our model. Thus, compulsory voting dominates voluntary voting under the same (slightly adjusted) condition as in Proposition 3. The analysis of subsidized voting is somewhat more complicated

10

than in our model, because a subsidy affects the direct voting costs c and not the normalized voting costs c/θ; in other words, two voters who have the same value of c/θ will not necessarily have the same value of "normalized net cost" $(c-s)/\theta$. Specifically, a subsidy encourages voters with a low voting cost and a low benefit stronger than those with a high cost and a high benefit. While this is an undesirable effect from the point of view of social welfare, the qualitative results of our model regarding subsidized voting obtain in this setting, too. In particular, subsidized voting with a subsidy of $\underline{c}/\overline{\theta}$ (where $\overline{\theta}$ is the highest possible benefit realization) dominates voluntary voting for sufficiently large N.

6 Appendix

Proof of Proposition 1. Let c_A be given. We construct c_B such that an individual who prefers B is indifferent between voting and not voting if $c_i = c_B$.

Consider a supporter of B. The probability that a of the remaining $N - 1$ individuals support candidate A and that k of these A supporters participate at the election, is given by

$$\text{Prob}\{\#A\text{-supporters} = a, \#A\text{-voters} = k\} = \binom{N-1}{a}\alpha^a(1-\alpha)^{N-1-a}\binom{a}{k}F(c_A)^k(1-F(c_A))^{a-k}. \quad (4)$$

If there are a supporters of A, of which k participate in the election then our B supporter's expected benefit of voting including subsidy s but excluding voting costs is

$$\text{Benefit}(a,k) = 0.5\left[\binom{N-a-1}{k-1}F(c_B)^{k-1}(1-F(c_B))^{N-a-k}\right.$$
$$\left. + \binom{N-a-1}{k}F(c_B)^k(1-F(c_B))^{N-a-k-1}\right] + s. \quad (5)$$

It follows immediately that $\text{Prob}\{\#A = a, \#A\text{-voters} = k\}$ and $\text{Benefit}(a,k)$ are continuous in c_A. The expected benefit from voting for a B-supporter with voting costs c_B is

$$\text{EB}_B(c_A, c_B) = \sum_{a=0}^{N-1}\sum_{k=0}^{a}\text{Prob}\{\#A = a, \#A\text{-voters} = k\}\text{Benefit}(a,k),$$

which is continuous in c_A and c_B. Similarly, an A-supporters gross benefit, $\text{EB}_A(c_A, c_B)$, is continuous. We now define the function $T : [\underline{c}, \bar{c}]^2 \to [\underline{c}, \bar{c}]^2$ by

$$T(c_A, c_B) = \left(\max\{\min\{\text{EB}_A(c_A, c_B), \bar{c}\}, \underline{c}\}, \max\{\min\{\text{EB}_B(c_A, c_B), \bar{c}\}, \underline{c}\}\right).$$

Clearly, T is continuous. As a consequence, Brouwer's fixed point theorem implies that there exist $\bar{c}_A, \bar{c}_B$ with $T(\bar{c}_A, \bar{c}_B) = (\bar{c}_A, \bar{c}_B)$. Consider $\bar{c}_A$. If $\underline{c} < \bar{c}_A < \bar{c}$ then the gross benefit of an A supporter with costs $\bar{c}_A$ who participates in the election is exactly $\bar{c}_A$. As a consequence an A supporter with cost $\bar{c}_A$ is indifferent between voting and not voting, and every A supporter with a lower cost will strictly prefer to vote. Now let $\bar{c}_A = \underline{c}$. Then $\text{EB}_A(c_A, c_B) \leq \underline{c}$. Thus, no A supporter with cost $c > \underline{c}$ will participate. Because the distribution of costs is continuous, this implies that with probability one no A supporter will participate. Finally, $\bar{c}_A = \bar{c}$ implies that all A supporters participate in the voting. Therefore, $\bar{c}_A$ is the cost cutoff for A supporters. Similarly, it follows that $\bar{c}_B$ is equilibrium cutoff for B supporters. ∎

Lemma 1 *Suppose that $\underline{c} > 0$. Then the expected number of A and B voters, $\bar{v}_A(N)$ and $\bar{v}_B(N)$, are bounded away from ∞, i.e., there exists an M such that $\bar{v}_A(N), \bar{v}_B(N) \leq M$ for all $N \in \mathbb{N}$.*

Proof of Lemma 1. The strategy of the proof is to show that if the expected number of voters goes to infinity as $N \to \infty$ then the pivot probabilities go to zero. This provides a contradiction because the voting costs c are always strictly positive, i.e., $c \geq \underline{c} > 0$.

Suppose by way of contradiction that the expected number of A and B voters are both infinite (the case where only the expected number of one type of voter is finite is similar and omitted). Then because the expected number of A and B voters are given by $\bar{v}_A(N) = F(c_A(N))\alpha N$ and $\bar{v}_B(N) = F(c_B(N))(1-\alpha)N$, respectively, we have

$$\lim_{N \to \infty} F(c_A(N))N = \infty, \text{ and } \lim_{N \to \infty} F(c_B(N))N = \infty. \tag{6}$$

Let $N_A(N)$ be the realized number of A-supporters out of N citizens. Let $N_B(N) = N - N_A(N)$.

Claim 1. The expected number of A and B voters goes to infinity for almost all realizations N_A and N_B, i.e., $\lim_{N \to \infty} N_A(N)F(c_A(N)) = \lim_{N \to \infty} N_B(N)F(c_B(N)) = \infty$.

Let X_i be a random variable such that $X_i = 1$ if person i is an A supporter and $X_i = 0$ if agent i is a B supporter. Then since the X_i are i.i.d., the central limit theorem implies that

$$\lim_{N \to \infty} P\left(\left\{\frac{\sum_{i=1}^{N}(X_i - \alpha)}{\sqrt{N\alpha(1-\alpha)}} \le \lambda\right\}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-\frac{x^2}{2}} \, dx. \tag{7}$$

Thus, for every $\varepsilon > 0$ there exists a $\lambda > 0$ such that

$$\lim_{N \to \infty} P\left(\left\{N\alpha - \lambda\sqrt{N\alpha(1-\alpha)} \le \sum_{i=1}^{N} X_i \le N\alpha + \lambda\sqrt{N\alpha(1-\alpha)}\right\}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} e^{-\frac{x^2}{2}} \, dx > 1 - \varepsilon. \tag{8}$$

Hence, with probability arbitrarily close to 1,

$$N_A(N) \in [N\alpha - \lambda\sqrt{N\alpha(1-\alpha)}, N\alpha + \lambda\sqrt{N\alpha(1-\alpha)}]. \tag{9}$$

so that $N_A(N)F(c_A(N))$ and $N_B(N)F(c_B(N)) \to \infty$, proving claim 1.

Let $Y_i^{N_A}$ be the random variable which assumes the value 1 if the ith A supporter votes and 0, otherwise. Similarly, define $Z_i^{N_B}$ for B supporters. The probability that a particular agent is pivotal is less or equal to $P\left(\left\{\sum_{i=1}^{N_A} Y_i^{N_A} - \sum_{i=1}^{N_B} Z_i^{N_B} \in \{-1, 0, 1\}\right\}\right)$. To determine this upper bound for the pivot probability, we next show that the limit distribution is normal. Care must be taken in applying the central limit theorem, because $Y_i^{N_A}$ and $Z_i^{N_B}$ converge to zero a.e., as N_A and $N_B \to \infty$.[6]

Claim 2. Suppose that $N_A, N_B \to \infty$ such that $N_A F(c_A(N)), N_B F(c_B(N)) \to \infty$. Then

$$\lim_{N_A, N_B \to \infty} P\left(\left\{\frac{\sum_{i=1}^{N_A} Y_i^{N_A} - \sum_{i=1}^{N_B} Z_i^{N_B} - N_A F(c_A(N)) + N_B F(c_B(N))}{\sqrt{N_A \operatorname{var}[Y_i^{N_A}] + N_B \operatorname{var}[Z_i^{N_B}]}} \le \lambda\right\}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\lambda} e^{-\frac{x^2}{2}} \, dx,$$

where the convergence is uniform in λ.

---

[6]In fact, the distinction between this result and that in Proposition 2 is as follows. Here we show that if the expected number of voters were to go to infinity, then the limit distribution would be normal (which, as shown below, leads to a contradiction). In contrast the Poisson limit distribution in Proposition 2 is compatible with strictly positive voting costs.

Let $C_N = \sum_{i=1}^{N_A} E[|Y_i^{N_A} - F(c_A(N))|^{2+\delta}] + \sum_{i=1}^{N_B} E[|Z_i^{N_B} - F(c_B(N))|^{2+\delta}]$, for some $\delta > 0$. According to Theorem 4.4 in Doob [5] it is sufficient to check that

$$\lim_{N_A, N_B \to \infty} \frac{C_N}{\left(N_A \operatorname{var}[Y_i^{N_A}] + N_B \operatorname{var}[Z_i^{N_B}]\right)^{1+\delta/2}} = 0. \tag{10}$$

Recall that $Y_i^{N_A}$ and $Z_i^{N_B}$ assume the value 1 with probabilities $F(c_A(N))$ and $F(c_B(N))$, respectively; and the value 0 otherwise. Thus, we get

$$\lim_{N_A, N_B \to \infty} \frac{C_N}{\left(N_A \operatorname{var}[Y_i^{N_A}] + N_B \operatorname{var}[Z_i^{N_B}]\right)^{1+\delta/2}}$$

$$\leq \lim_{N_A \to \infty} \frac{\sum_{i=1}^{N_A} E[|Y_i^{N_A} - F(c_A(N))|^{2+\delta}]}{\left(N_A \operatorname{var}[Y_i^{N_A}]\right)^{1+\delta/2}} + \lim_{N_B \to \infty} \frac{\sum_{i=1}^{N_B} E[|Z_i^{N_B} - F(c_B(N))|^{2+\delta}]}{\left(N_B \operatorname{var}[Y_i^{N_B}]\right)^{1+\delta/2}}$$

$$= \lim_{N_A \to \infty} \frac{N_A F(c_A(N))(1 - F(c_A(N))) \left[(1 - F(c_A(N)))^{1+\delta} + F(c_A(N))^{1+\delta}\right]}{\left[N_A F(c_A(N))(1 - F(c_A))\right]^{1+\delta/2}}$$

$$+ \lim_{N_B \to \infty} \frac{N_B F(c_B(N))(1 - F(c_B(N))) \left[(1 - F(c_B(N)))^{1+\delta} + F(c_B(N))^{1+\delta}\right]}{\left[N_B F(c_B(N))(1 - F(c_B(N)))\right]^{1+\delta/2}}$$

$$= \lim_{N_A \to \infty} \frac{(1 - F(c_A(N)))^{1+\delta} + F(c_A(N))^{1+\delta}}{\left[N_A F(c_A(N))(1 - F(c_A))\right]^{\delta/2}} + \lim_{N_B \to \infty} \frac{(1 - F(c_B(N)))^{1+\delta} + F(c_B(N))^{1+\delta}}{\left[N_B F(c_B(N))(1 - F(c_B(N)))\right]^{\delta/2}} = 0,$$

because $N_A F(c_A(N))$ and $N_B F(c_B(N)) \to \infty$ by claim 1. Thus, condition 10 is satisfied, proving claim 2.

In the remainder of the proof we use claim 2 to derive a contradiction. In particular, if an agent is pivotal then $\sum_{i=1}^{N_A} Y_i^{N_A} - \sum_{i=1}^{N_B} Z_i^{N_B} \in \{-1, 0, 1\}$. However, the normalized sum (i.e., the expression in claim 2) then converges to zero because the standard deviation goes to infinity. Because the limit distribution is continuous, this implies that the probability of being pivotal converges to zero, which is incompatible with strictly positive voting costs.

Formally, define

$$b = \lim_{N \to \infty} \frac{N_A F(c_A(N)) - N_B F(c_B(N))}{\sqrt{N_A \operatorname{var}[Y_i^{N_A}] + N_B \operatorname{var}[Z_i^{N_B}]}}. \tag{11}$$

Note that we allow for the possibility that b is negative or positive infinity. Furthermore, we can assume without loss of generality that the sequence converges. Otherwise, we can take a converging subsequence. Let $\epsilon > 0$ be arbitrary. Then there exists $\lambda > 0$ such that

$$\frac{1}{\sqrt{2\pi}} \int_{b-\lambda}^{b+\lambda} e^{-\frac{x^2}{2}} dx < \epsilon. \tag{12}$$

Furthermore $\sqrt{N_A \operatorname{var}[Y_i^{N_A}] + N_B \operatorname{var}[Z_i^{N_B}]} \leq \sqrt{N_A \operatorname{var}[Y_i^{N_A}]} + \sqrt{N_B \operatorname{var}[Z_i^{N_B}]}$, where each of the summands converges to ∞ because of Claim 1, i.e., because $N_A F(c_A(N))$ and $N_B F(c_B(N))$ converge to ∞. This and (11) imply that for sufficiently large N a necessary condition for being pivotal is that

$$b - \lambda \leq \frac{\sum_{i=1}^{N_A} Y_i^{N_A} - \sum_{i=1}^{N_B} Z_i^{N_B} - N_A F(c_A(N)) + N_B F(c_B(N))}{\sqrt{N_A \operatorname{var}[Y_i^{N_A}] + N_B \operatorname{var}[Z_i^{N_B}]}} \leq b + \lambda. \tag{13}$$

14

Thus, (12), (13), and claim 2 imply that the probability of being pivotal is less than ϵ. Because ϵ was chosen arbitrarily, the pivot probability must converge to zero, for almost all realizations N_A and N_B. Taking the expectation over all possible realizations of N_A and N_B we can therefore conclude that the pivot probability is zero.

Now recall that voting costs $c \geq \underline{c} > 0$. Because the pivot probability converges to zero, the payoff to a voter is therefore strictly negative as N gets large. This is a contradiction, since the citizen would better off not voting. ■

Proof of Proposition 2. Let $N_A(N)$ be a sequence that satisfies (9). Define the random variables $X^A_{N_A(N),i}$ and $X^B_{N_A(N),i}$ as in lemma 1. We first prove that $\sum_{i=1}^{N_A(N)} X^A_{N_A(N),i}$ converges to a Poisson distribution.

The expected number of A voters given that there are N_A supporters of A is given by $\bar{v}_A(N|N_A) = N_A F(c_A(N))$. Thus,

$$
\binom{N_A(N)}{k} F(c_A(N))^k (1 - F(c_A(N)))^{N_A(N)-k}
$$
$$
= \frac{(N_A(N) - 1) \dots (N_A(N) - k + 1)}{N_A(N)^{k-1}} \frac{\bar{v}_A(N|N_A(N))^k}{k!} \left(1 - \frac{\bar{v}_A(N|N_A(N))}{N_A(N)}\right)^{N_A(N)-k}. \tag{14}
$$

Note that

$$
\lim_{N \to \infty} \frac{\bar{v}_A(N|N_A)}{\bar{v}_A(N)} = \lim_{N \to \infty} \frac{N_A(N) F(c_A(N))}{\alpha N F(c_A(N))} = 1, \tag{15}
$$

because $N_A(N)$ satisfies (9).

Lemma 1 implies that there exist a subsequence $\bar{v}_A(N_n)$ of $\bar{v}_A(N)$ such that $\lim_{n \to \infty} \bar{v}_A(N_n) = \bar{v}_A$. We a slight abuse of notation we denote the subsequence again by $\bar{v}_A(N)$. Thus, (14) and (15) imply

$$
\lim_{N \to \infty} P\big(\{\#A\text{-voters} = k\} | \{\#A\text{-supporters} = N_A(N)\}\big) = \frac{\bar{v}_A^k}{k!} e^{-\bar{v}_A}.
$$

where the convergence is uniform for all sequences $N_A(N)$ that satisfy (9). Thus, (8) implies

$$
\lim_{N \to \infty} \left| P\big(\{\#A\text{-voters} = k\} | N\big) - \frac{\bar{v}_A^k}{k!} e^{-\bar{v}_A} \right| < \epsilon.
$$

Because $\epsilon > 0$ was chosen arbitrarily, it follows that

$$
\lim_{N \to \infty} P\big(\{\#A\text{-voters} = k\} | N\big) = \frac{\bar{v}_A^k}{k!} e^{-\bar{v}_A}, \tag{16}
$$

Similarly, it follows that $\lim_{N \to \infty} P\big(\{\#B\text{-voters} = k\} | N\big) = \frac{\bar{v}_B^k}{k!} e^{-\bar{v}_B}$,

Suppose that there are N individuals. An A supporter with voting costs c_i will vote if and only if

$$
\frac{1}{2} P\big(\#A\text{-voters} - \#B\text{-voters} \in \{0, -1\} | N\big) \geq c_i, \tag{17}
$$

where $P(\#A\text{-voters} - \#B\text{-voters} \in \{0, -1\} | N)$ is the probability that candidate A gets either one less vote than B or the same number of votes as candidate B (so that our citizen is pivotal and can increase the

probability of victory for A by 0.5), given that there are N individuals. Lemma 1 implies that for arbitrary $\epsilon > 0$, there exists K such that $P(\{\#A\text{-voters} \geq K\}|N) < \epsilon$ and $P(\{\#B\text{-voters} \geq K\}|N) < \epsilon$ for all sufficiently large N. Thus, (16) implies

$$\lim_{N \to \infty} P(\#A\text{-voters} - \#B\text{-voters} \in \{0, -1\}|N) = \sum_{i=0}^{\infty} \frac{\bar{v}_A^i}{i!} e^{-\bar{v}_A} \left(\frac{\bar{v}_B^i}{i!} e^{-\bar{v}_B} + \frac{\bar{v}_B^{i+1}}{(i+1)!} e^{-\bar{v}_A} \right), \quad (18)$$

By formula 9.6.10 in Abramowitz and Stegun [1] (see also Myerson [9]), we get

$$\lim_{N \to \infty} P(\#A\text{-voters} - \#B\text{-voters} \in \{0, -1\}|N) = \sqrt{\frac{\bar{v}_A}{\bar{v}_B}} \frac{I_1(2\sqrt{\bar{v}_A \bar{v}_B})}{e^{\bar{v}_A + \bar{v}_B}} + \frac{I_0(2\sqrt{\bar{v}_A \bar{v}_B})}{e^{\bar{v}_A + \bar{v}_B}}, \quad (19)$$

where I_k is a modified Bessel function. Similarly, the pivot probability for a B supporter is

$$\lim_{N \to \infty} P(\#A\text{-voters} - \#B\text{-voters} \in \{0, 1\}|N) = \sqrt{\frac{\bar{v}_B}{\bar{v}_A}} \frac{I_1(2\sqrt{\bar{v}_A \bar{v}_B})}{e^{\bar{v}_A + \bar{v}_B}} + \frac{I_0(2\sqrt{\bar{v}_A \bar{v}_B})}{e^{\bar{v}_A + \bar{v}_B}}. \quad (20)$$

As $N \to \infty$ both pivot probabilities in (19) and (20) must converge to $2\underline{c}$, by (17) and Lemma 1. Thus,

$$\sqrt{\frac{\bar{v}_A}{\bar{v}_B}} \frac{I_1(2\sqrt{\bar{v}_A \bar{v}_B})}{e^{\bar{v}_A + \bar{v}_B}} + \frac{I_0(2\sqrt{\bar{v}_A \bar{v}_B})}{e^{\bar{v}_A + \bar{v}_B}} = \sqrt{\frac{\bar{v}_B}{\bar{v}_A}} \frac{I_1(2\sqrt{\bar{v}_A \bar{v}_B})}{e^{\bar{v}_A + \bar{v}_B}} + \frac{I_0(2\sqrt{\bar{v}_A \bar{v}_B})}{e^{\bar{v}_A + \bar{v}_B}}. \quad (21)$$

However, since the Bessel function I_1 is never zero, (21) implies that $\bar{v}_A = \bar{v}_B$, i.e., in the limit the number of A and B voters are drawn from the same Poisson distribution. As a consequence, each candidate wins with probability $\frac{1}{2}$, independent of α. ∎

References

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Tables*. New York, Dover Publications, 1965.

[2] T. Börgers. Costly voting. *American Economic Review*, 94(1):57–66, 2004.

[3] C. M. Campbell. Large electorates and decisive minorities. *Journal of Political Economy*, 107:1199–1217, 1999.

[4] J. Dean. Is it time to consider mandatory voting laws? *Findlaw.com*, 2003. Posted Friday, Feb. 28, 2003.

[5] J. Doob. *Stochastic Processes*. New York, John Wiley & Sons, 1990.

[6] S. Ghosal and B. Lockwood. Costly voting and inefficient participation. University of Warwick, 2004.

[7] J. Goeree and J. Grosser. False consensus voting and welfare reducing polls. Caltech, 2004.

[8] J. Ledyard. Pure theory of two-candidate competition. *Public Choice*, 44:7–41, 1984.

[9] R. B. Myerson. Large poisson games. *Journal of Economic Theory*, 94:7–45, 2000.

[10] K. Olbermann. Make voting mandatory! *Salon.com*, 2002. Posted Nov. 5, 2002.

[11] E. Weiner. You must vote. It's the law. Australia requires citizens to vote. should the U.S.? *Slate*, 2004. Posted Friday, Oct. 29, 2004.

[12] Wikipedia. Compulsory voting. http://en.wikipedia.org/wiki/Compulsory_voting, 2004.