Voter Participation with Collusive Parties☆

David K. Levine[1], Andrea Mattozzi[2]

Abstract

We reexamine the theory of rational voter participation where voting is by two collusive parties that can enforce social norms through costly peer punishment. This model nests both the ethical voter model and the pivotal voter model. We initially abstract from aggregate shocks to the population of voters in favor of the original Palfrey and Rosenthal [24] model and analyze the subsequent all-pay auction game. We show that this game has a unique mixed strategy equilibrium in which one party - the advantaged party - gets all the surplus and give a simple formula for determining which party is advantaged. This equilibrium is scale invariant - increasing the size of both parties in proportion has no effect on voter turnout by either party. Our main finding is that when the cost of enforcement of social norms is low and the benefit of winning the election is the same for both parties the larger party is always advantaged. By contrast, when the enforcement of social norms is costly we have a result reminiscent of Olson [27] in which - even if the benefit of winning the election is the same for both parties - the smaller party may be advantaged. We then examine more general contest resolution functions giving conditions under which pure strategy equilibria do and do not exist, examining the robustness of the comparative statics of the all-pay auction model, and give conditions under which participation declines with the size of the electorate.
JEL Classification Numbers: later

Keywords: Voting, Collusion, Organization, Party

☆First Version: June 2, 2013. We are especially grateful to Juan Block, Michele Boldrin, Rohan Dutta, Philipp Kircher, Cesar Martinelli, Salvatore Modica, Massimo Morelli, Salvatore Nunnari, Thomas Palfrey, Kirill Pogorelskiy, Guido Tabellini, participants at the 2015 Priorat Workshop, 2015 Barcelona JOCs, the UCL Theory Workshop, the UECE Meetings in Lisbon, and the Dondena Political Economy Seminar in Bocconi.

*Corresponding author David K. Levine

Email addresses: david@dklevine.com (David K. Levine), andrea.mattozzi@eui.eu (Andrea Mattozzi)

[1]Department of Economics, EUI and WUSTL

[2]Department of Economics, EUI and MOVE

1. Introduction

Woman who ran over husband for not voting pleads guilty. USA Today April 21, 2015

We investigate the idea that voters vote due to peer pressure to adhere to social norms - with the additional feature that these social norms are chosen strategically. We introduce a model based on the well documented fact that social norms play a key role in voter participation.[3] Our model effectively nests the "ethical" voter model of Feddersen and Sandroni [9] and Coate and Conlin [3] in large elections and the pivotal voter model of Palfrey and Rosenthal [24] in small elections while at the same time making a rich set of new predictions about features such as the role of party size in determining participation rates.

Rational theories of voter participation remain controversial. The standard Palfrey and Rosenthal [24] model finds empirical support in the laboratory (see for example Levine and Palfrey [20]) but it has difficulty explaining large scale elections. Coate, Conlin and Moro [4] show that in a sample of Texas referenda, elections are much less close than is predicted by the pivotal voter model, and Coate and Conlin [3] show that a model of "ethical" voters better fits the data than the model of pivotal voters. Indeed, the probability of being pivotal in large elections is very low as documented by Mulligan and Hunger [21]. Moreover, the probability of being pivotal- since it is proportional to standard error - should decline roughly as the square root of the number of voters. The same applies to participation rates if voting costs are non-negative and uniformly distributed.[4]

If we focus on post-war national elections in consolidated democracies with per capita income above the world average and voluntary voting, the relation between voter turnout and the size of the country is hardly consistent with the predictions of the pivotal voter model. In particular, there is a group of small countries with population ranging from 300,000 to 10 million with high voter turnout of 78% to 88% and a group of large countries with population ranging from 35 million to 319 million with lower voter turnout ranging of 55% to 71%. Within these groups of countries there is very little variation or evidence of negative correlation between size and turnout.[5] While it is true that the group of smaller countries generally have higher turnout than the larger countries, within groups turnout is quite homogeneous while population varies by a factor of nearly 10 - this data is in no way consistent with scaling by the square root of the population. In fact, it is not even

[3] See, e.g., Gerber, Green and Larimer [10] and Della Vigna et al [5].

[4] In a two-candidates election with an even number of voters n each casting her vote randomly, the probability of a tie approaches $\sqrt{2/n\pi}$ as n grows large. See Penrose [25] and Chamberlain and Rothschild [2]. Assuming that cost of voting is non-negative and uniformly distributed is standard in the theoretical literature on turnout, see, for example Feddersen and Sandroni [9] and Coate and Conlin [3]. That the cost of voting distribution is relatively flat is also consistent with the well documented fact that turnout is highly sensitive to the importance of the election - for example turnout in U.S. Presidential elections is much greater than in election for local issues only.

[5] Turnout data are averages in the post-war period of OECD countries with voluntary voting and Freedom House Index of political freedom below 3. We included UK and excluded the rest of the EU since in the latter substantial power has passed to the EU itself, so that the significance of "national" elections is different than in fully sovereign nations - in particular for the smaller EU nations. However, including the rest of the EU does not alter the overall picture. Data is taken from http://www.idea.int.

consistent with a monotone relation between turnout and population, which is the main prediction of the pivotal voter model. A similar picture emerges if we turn attention to the dynamics of voter turnout in advanced democracies: turnout declined on average by a mere 10% in the past 50 years in the face of a voting age population which more than doubled.[6] On the other hand, empirical analysis by Gray and Caul [16] relates post-war turnout decrease with the decline of mobilizing actors such as labor parties and trade unions.[7]

The most recent rational voter theories studied by economists have been the social preference model of ethical voters introduced by Feddersen and Sandroni [9] and Coate and Conlin [3]. Roughly speaking these models assume that some or all voters choose to participate based upon whether or not the benefits of their vote to their party justifies the cost of their participation. Here we take the view that these social preferences arise as a social norm - that voters choose whether or not to vote based upon whether social norms call upon them to vote - and we assume that these social norms are endogenous and chosen strategically by political parties. That is, rather than assuming that voters weigh the benefits to the party of their vote against the cost, we assume that each party collusively weighs the benefits of voter turnout against the costs and chooses a social norm that is optimal for the party. In turn this social norm is enforced through costly peer punishment.

Our model of peer punishment originates in Kandori [14]'s work on social norms in repeated game and is a variant of the peer punishment model introduced by Levine and Modica [18]. In this model within each party voters monitor each others voting behavior and punish - through ostracism and social disapproval (and perhaps as the quotation at the top indicates more severely) - those who fail to adhere to the social norm. This is consistent with recent large scale field experiments. For example Della Vigna et al [5] show that an important incentive for voters to vote is to show others that they have voted, and Gerber, Green and Larimer [10] provide evidence that social pressure significantly increase voter turnout.[8] Here we hypothesize that the reason that voters want to show others that they have voted is because either they have internalized a social norm, or they expect to be rewarded for following the social norm or punished for failing to do so. Equally crucial is that we assume that the social norm is endogenous and chosen rationally by a political party that colludes among its members.

The idea of collusive parties is nothing new - a large range of literature in political economy studying parties such as elites and masses and other groups often treats these groups as single players who act in the group interest. Our political parties behave in a similar way although as in Dutta, Levine and Modica [6] they must do so subject to incentive constraints - that is, parties can only collude to make choices that are incentive compatible for its members. If punishment is

[6] In fact in Denmark and Sweden turnout increased by 3% and 6%, respectively in the period 1950-2000.

[7] See also Knack [15] on the decline of American voter turnout and its relation to a weakened enforcement of social norms.

[8] Palfrey and Pogorelskiy [23] provide experimental evidence showing that communication among voters and in particular communication within parties increases turnout.

adequate to induce voter turnout, then that turnout can be chosen by the party - otherwise not.

Initially we abstract from aggregate shocks. We do so not because we believe that aggregate shocks are not likely to be important in practice but to remain as close as possible to the original Palfrey and Rosenthal [24] and also to highlight the role of mixed strategies: Feddersen and Sandroni [9] and Coate and Conlin [3] need to impose restrictions on parameter values in order to avoid the necessity of mixing. Here mixtures play a rather more sensible role because the mixing is done by the parties rather than by individuals. Indeed mixed stragies are essential in contests with opposing interests. In the voting context if one party is expected to win, the second party should not bother to turn out voters, so the first party should make a minimal mobilization effort, in which case the second party should overcome this minimal effort. That is - voting between collusive parties has the flavor of matching pennies and indeed in all pay auctions equilibria as originally shown by Hillman and Riley [13] have only mixed equilibria. This is reflected in the reality of elections. Real political parties engage in the "ground game" or "GOTV" (Get Out The Vote) efforts. This ranges from phone calls reminding people to vote, or the importance of the election to driving people to the polls. We view it as an important part of the peer punishment system establishing the social norm for the particular election, and these GOTV efforts are variable and strategic. Furthermore, political parties have strong incentives not to advertize their GOTV effort, and in fact to keep their GOTV effort secret.[9] Clearly, there is little reason to do that unless indeed GOTV effort is random. Hence, the mere fact that it is secret provides evidence that - consciously or not - political parties engage in randomization when choosing social norms for particular elections.

Our initial setting then is one of collusive parties enforcing costly social norms in an effort to win an all-pay auction. We show that this game has a unique mixed strategy equilibrium in which one party - the advantaged party - gets all the surplus and give a simple formula for determining which party is advantaged. Equilibrium is scale invariant - increasing the size of both parties in proportion has no effect on voter turnout by either party. Our main finding is that when the enforcement of social norms is costless and the benefit of winning the election is the same for both parties the larger party is always advantaged. By contrast, when the enforcement of social norms is costly we have a result reminiscent of Olson [27] in which - even if the benefit of winning the election is the same for both parties - the smaller party may be advantaged. We provide a number of other comparative static results. We show that while being advantaged "ordinarily" results in a higher probability of winning the election it need not do so. In addition we examine when parties will engage in supressing the vote of their rival - finding that disadvantaged parties never will do so, but that advantaged parties generally will.

We also examine general contest resolution functions and incentive constraints that account

[9] Accounts in the popular press document both the surprise over the strength of the GOTV and the secrecy surrounding it. For example "The power of [Obama's GOTV] stunned Mr. Romney's aides on election night, as they saw voters they never even knew existed turn out..." Nagourney et al [22]or "[Romney's] campaign came up with a super-secret, super-duper vote monitoring system [...] to plan voter turnout tactics on Election Day " York [28]. Note that the secrecy at issue is not over whether or not people voted as for example voting pins: we assume that the act of voting is observable. Rather the secrecy is over the social norm that is enforced on election day.

for pivotality. We show that when aggregate shocks are sufficiently large or pivotality sufficiently important pure strategy equilibria exist and that conversely with large electorates and small aggregate shocks only mixed strategy equilibria exist. We examine the robustness of the comparative statics of the all-pay auction to aggregate shocks and pivotality and give conditions under which participation declines with the size of the electorate.

2. Costs of Voting for a Single Party

We follow the approach of Levine and Modica [18] and Dutta, Levine and Modica [6] in modeling a homogeneous collusive party: we treat it as a problem in mechanism design. The party - either by consensus or directed by leaders - moves first and chooses a social norm; the individual party members move second and, given the social norm, make choices about whether or not to vote that are individually optimal. Here we study the cost to the party of inducing a fraction of voters to vote.

2.1. The Model

Each identical party member privately draws a type y from a uniform distribution on $[0, 1]$. This type determines a cost of voting $c(y)$, possibly negative, and based on this the member decides whether or not to vote. The cost of voting c is continuously differentiable, has $c'(y) > 0$ and satisfies $c(\underline{y}) = 0$ for some $0 \leq \underline{y} \leq 1$. Voters for whom $y < \underline{y}$ are called *committed voters.* The (total) *participation cost* of voting is defined as $C(y) = 0$ for $y < \underline{y}$ and $C(y) = \int_{\underline{y}}^{y} c(y)dy$ for $y \geq \underline{y}$. This is a standard formulation: for example Coates and Conlin assume that $c(y)$ is linear so that the participation cost of voting for $y \geq \underline{y}$ is quadratic.

The party can impose punishments $0 \leq P \leq \overline{P}$ on members. The social norm of the party is a threshold φ together with a rule prescribing voting if $y \leq \varphi$. This rule is enforced through peer auditing and punishment. Each member of the party is audited by another party member. The auditor observes whether or not the auditee voted. If the auditee did not vote and the party member did not violate the policy (that is, $y > \varphi$) there is a probability π that the auditor will learn this. The value of π represents the quality of the signal about y: if $\pi = 0$ then the auditor learns nothing about y; if $\pi = 1$ the auditor perfectly observes whether y is above or below the threshold φ. Whatever the quality of the signal, if the auditee voted or is discovered not to have violated the policy, the auditee is not punished. If the auditee did not vote and the auditor cannot determine whether or not the auditee violated the policy, the auditee is punished with a loss of utility P. Initially we are going to assume that the probability of being pivotal is too small to matter. We see immediately that a *social norm is incentive compatible* if and only if $P = c(\varphi)$, in which case any member with $y \leq \varphi$ would be willing to pay the cost $c(y)$ of voting rather than face the certain punishment P, while any member with $y > \varphi$ prefers to pay the expected cost of punishment $(1 - \pi)P$ over the cost of voting $c(y)$.

The overall cost of a punishment P to the party is ψP where $\psi \geq 1$. Naturally the punishment itself as it is paid by a member is a cost to the party. However, there may be other costs: for example, if the punishment is ostracism this may not only be costly to the member punished, but also to other party members who might otherwise have enjoyed the company of the ostracized member. In addition, the audits and the punishments may themselves be costly, and there may be additional rounds of audits and punishments needed so that members are willing to do their share of enforcing the social norm as in Levine and Modica [17]. If so we assume that these costs are proportional to the size of the punishment.[10]

2.2. The Cost of Turning Out Voters

Note that φ is the participation rate of the party, that is, the probability a representative party member votes. Recall that $\underline{y}$ is the (unique) value of y such that $c(\underline{y}) = 0$. The role of committed voters with $y < \underline{y}$, that is, $c(y) < 0$, is quite different from those with $y \geq \underline{y}$, that is $c(y) \geq 0$. Those with $c(y) < 0$ represent voters who out of civic duty or because they enjoy the camaraderie of the polling place will always vote. Since a fraction $\underline{y}$ of the party will vote no matter what, the crucial question for the party is how costly it is for the party to induce additional voters to vote by choosing an incentive compatible social norm $\varphi \geq \underline{y}$. Denote this cost by $D(\varphi)$.

We start by observing that $D(\varphi)$ has two parts. The *participation cost* $C(\varphi) = \int_{\underline{y}}^{\varphi} c(y)dy$ is the total cost of voting to the members who vote. Notice that $C'(\varphi) = c(\varphi)$ and so $C(\varphi)$ is increasing and convex. The *monitoring cost* $M(\varphi) = \int_{\varphi}^{1} \psi(1-\pi)Pdy$ is the (expected) cost of punishing party members who did not vote. As incentive compatibility requires $P = c(\varphi) = C'(\varphi)$, this can be written as $M(\varphi) = \psi(1-\pi)(1-\varphi)C'(\varphi)$. We refer to $\theta \equiv \psi(1-\pi)$ as the monitoring inefficiency. This can be any non-negative number. If the signal quality is high so that π is large monitoring is very efficient. If the costs of issuing punishments ψ is high then monitoring is very inefficient.

Since $c(y)$ is strictly increasing we may define the unique $\overline{y}$ to be such that $c(\overline{y}) = \overline{P}$ where $\overline{P}$ is the maximum feasible punishment, or $\overline{y} = 1$ if $c(1) \leq \overline{P}$. Observe that those for whom $y > \overline{y}$ will not vote regardless of the social norm. The feasible turnout rates φ are therefore those in the range $\underline{y} \leq \varphi \leq \overline{y}$, so our interest is on the behavior of $D(\varphi)$, $C(\varphi)$, $M(\varphi)$ in this range.

The crucial fact is that while $C(\varphi)$ is necessarily convex, $M(\varphi)$ and more importantly $D(\varphi)$ may fail to be so, and indeed may be concave. As we will see the convexity of $D(\varphi)$ is crucial in determining how the cost of turning out a fixed number of voters depends on the size of the party: when $D(\varphi)$ is convex a larger party necessarily has a lower cost of turning out a fixed number of voters, but this need not be the case when $D(\varphi)$ fails to be convex.

Theorem 1. *We have $C(\underline{y}) = M(\underline{y}) = 0$ so $D(\underline{y}) = 0$. The participation cost $C(\varphi)$ is twice continuously differentiable strictly increasing and strictly convex. The monitoring cost $M(\varphi)$ is continuously differentiable. If $\overline{y} = 1$ (that is $c(1) \leq \overline{P}$ so that full participation is possible) the monitoring cost $M(\varphi)$ cannot be concave, must be decreasing over part of its range and $M(1) = 0$ so $D(1) = C(1)$.*

[10] In Levine and Modica [17] it is assumed that audits have a fixed cost component and that all members need to be audited. We do not think in the case of voting that the fixed cost component is terribly significant - for example, it is probably possible to avoid auditing voters - and as it complicates the analysis, we ignore it.

Proof. From the fundamental theorem of calculus $C'(\varphi) = c(\varphi)$; since this is continuously differentiable with $c'(\varphi) > 0$ we see that $C(\varphi)$ is twice continuously differentiable and that $C''(\varphi) > 0$. At $\underline{y}$ we have $M(\underline{y}) = \theta(1 - \underline{y})c(\underline{y}) = 0$ since by definition $c(\underline{y}) = 0$ so $D(\underline{y}) = C(\underline{y}) = 0$. If $c(1) \leq \overline{P}$ we have $M(1) = \theta(1-1)c(1) = 0$ so $D(1) = C(1)$. Since $M(\underline{y}) = 0$ and $M(1) = 0$ and $M(\varphi) > 0$ for $0 < \varphi < 1$ we see that $M(\varphi)$ cannot be concave, must be decreasing over part of its range. □

The key fact is that at $\underline{y}$ there is no punishment cost since punishment is not needed to turn out the committed voters, while at $\overline{y} = 1$ everybody votes, so despite the fact that the punishment is positive, nobody is actually punished. It should be clear that this idea and result is robust to the particular details of the monitoring process. Note also that in addition to the possibility that $D(\varphi)$ may fail to be convex, it is not necessarily increasing.

Example 1. Suppose that for $\varphi \geq \underline{y}$ cost is given by $c(\varphi) = \alpha(\varphi - \underline{y})^{\alpha-1}$ for some $\alpha > 1$, or equivalently that $C(\varphi) = (\varphi - \underline{y})^{\alpha}$. For example, Coate and Conlin [3] consider $\alpha = 2$. From the detailed computation in the Web Appendix we find that $D(\varphi) = (1 - \alpha\theta)(\varphi - \underline{y})^{\alpha} + \alpha\theta(1 - \underline{y})(\varphi - \underline{y})^{\alpha-1}$. If $\alpha\theta > 1$ and $\alpha \leq 2$ then this function is concave. Moreover we have $D'(\varphi) \geq \alpha(1 - \theta)$ so that $D(\varphi)$ is is strictly increasing for $\theta \leq 1$. By way of contrast, if $\theta = 2$ and $\alpha = 2$ at $\phi = 1$ we have $D'(1) = -2(1 - \underline{y}) < 0$ which implies that at $\varphi = 1$, $D(\varphi)$ is decreasing.

3. All Pay Auction

We now suppose that a population of N voters is divided into two parties $k = S, L$ of size $\eta_k N$ where $\eta_S + \eta_L = 1$. These parties compete in an election. We abstract from random variation in voter turnout and assume that the side that produces the greatest expected number of votes wins a prize worth $v_L > 0$ and $v_S > 0$ to each member respectively. We assume that both parties face per capita costs of turning out voters characterized by $\underline{y}_k < \overline{y}_k$ and cost function $D_k(\varphi_k)$. We make the generic assumption that $\eta_S \underline{y}_S \neq \eta_L \underline{y}_L$ and $\eta_S \overline{y}_S \neq \eta_L \overline{y}_L$. We define the *large* party L to be the one with the largest possibility for turning out voters $\eta_L \overline{y}_L > \eta_S \overline{y}_S$, with S the *small* party. We define the *most committed* party to be the one with the largest number of committed voters, i.e., with th largest value of $\eta_k \underline{y}_k$ and the *least committed* party to be the one with the smallest value. We will assume that $D_k'(\varphi) > 0$ since it is the standard assumption in the literature and the non-increasing case is harder to characterize and seems less interesting. For notational convenience we assume that for $\varphi_k < \underline{y}_k$ the cost is $D_k(\varphi_k) = 0$.

A strategy for party k is a probability measure represented as a cumulative distribution function F_k on $[\eta_k \underline{y}_k, \eta_k \overline{y}_k]$ where we refer to $\eta_k \varphi_k$ as the *bid*. A tie-breaking rule is a measurable function B_S from $[\max \eta_k \underline{y}_k, \eta_S \overline{y}_S]^2 \to [0, 1]$ with $B_S(\eta_S \varphi_S, \eta_L \varphi_L) = 0$ for $\eta_S \varphi_S < \eta_L \varphi_L$ and $B_S(\eta_S \varphi_S, \eta_L \varphi_L) = 1$ for $\eta_S \varphi_S > \eta_L \varphi_L$ with $B_L = 1 - B_S$. We say that F_S, F_L are an *equilibrium* if there is a tie-breaking rule B_S such that

$$\int v_k B_k(\eta_k \varphi_k, \eta_{-k} \varphi_{-k}) F_k(d\eta_k \varphi_k) F_{-k}(d\eta_{-k} \varphi_{-k}) - \int D_k(\varphi_k) F_k(d\eta_k \varphi_k) \quad \geq$$
$$\int v_k B_k(\eta_k \varphi_k, \eta_{-k} \varphi_{-k}) \tilde{F}_k(d\eta_k \varphi_k) F_{-k}(d\eta_{-k} \varphi_{-k}) - \int D_k(\varphi_k) \tilde{F}_k(d\eta_k \varphi_k)$$

for all cdfs $\tilde{F}_k$ on $[\eta_k \underline{y}_k, \eta_k \overline{y}_k]$.[11]

Let $\hat{\varphi}_k$ satisfy $D_k(\hat{\varphi}_k) = v_k$ or $\hat{\varphi}_k = \overline{y}_k$ if there is no solution. Hence, $\hat{\varphi}_k$ represents the most fraction of voters the party is willing and able to turn out. We make the generic assumption that the party sizes are such that $\eta_L \hat{\varphi}_L \neq \eta_S \hat{\varphi}_S$. We define the *disadvantaged* party d to be the party for which $\eta_d \hat{\varphi}_d < \eta_{-d} \hat{\varphi}_{-d}$, where $-d$ is the *advantaged* party.

Notice that we have three measures of the "strength" of a party: the overall possibility of turning out voters (large or small), the number of committed voters (most or least committed) and the willingness to turn out voters (advantaged or disadvantaged). Our theorem shows that each of these plays a role in determining the outcome of elections.

Theorem 2. *There is a unique mixed equilibrium. The disadvantaged party earns zero and the advantaged party earns $v_{-d} - D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d) > 0$.*

If $\hat{\varphi}_k \leq (\eta_{-k}/\eta_k)\underline{y}_{-k}$ then the election is uncontested*: the least committed party k is disadvantaged,* concedes *the election by bidding $\eta_k \underline{y}_k$ and the most committed party $-k$* takes *the election by bidding $\eta_{-k}\underline{y}_{-k}$.*

If $\hat{\varphi}_k > (\eta_{-k}/\eta_k)\underline{y}_{-k}$ for $k \in \{S, L\}$ then the election is contested: *in $(\max_k \eta_k \underline{y}_k, \eta_d \hat{\varphi}_d)$ the mixed strategies of the parties have no atoms, and are given by continuous densities*

$$f_k(\eta_k \varphi_k) = D'_{-k}((\eta_k/\eta_{-k})\varphi_k)/(\eta_{-k} v_{-k}).$$

In these contested elections there are three points that may have atoms: each party may turn out only its committed voters and the advantaged party may take the election by turning out $\eta_d \hat{\varphi}_d$ with positive probability. The possible cases are as follows:

1) The only party that concedes the election with positive probability is the disadvantaged party which does so by bidding $\eta_d \underline{y}_d$ with probability $\phi_d(\eta_d \underline{y}_d) = 1 - D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d)/v_{-d} + D_{-d}((\eta_L/\eta_{-d})\underline{y}_L)/v_{-d}$

2) The only time an advantaged party turns out only its committed voters with positive probability is if it is also the most committed party in which case this probability is $\phi_{-d}(\eta_{-d}\overline{y}_{-d}) = D_d((\eta_{-d}/\eta_d)\underline{y}_{-d})/v_d$

3) The advantaged party takes the election by turning out $\eta_d \hat{\varphi}_d$ with positive probability only if $\hat{\varphi}_S = \overline{y}_S$ in which case this probability is $\phi_{-d}(\eta_S \overline{y}_S) = 1 - D_{-d}((\eta_S/\eta_L)\overline{y}_S)/v_S$. This is the only case in which the tie-breaking rules matters: when both parties bid $\eta_S \overline{y}_S$ the large party must win with probability 1.

Proof. See Appendix I. □

Next we examine the comparative statics of the model using a notion of decreased turnout in terms of first-order stochastoc dominance.

Corollary 1. *We have the following*

1. Only the relative sizes of the parties matters.

2. If the value of the prize to the least committed party is small enough then that party is disadvantaged and concedes the election with probability one. If the value of the prize to the large

[11] We note that by the Lesbesgue decomposition theorem the cdf F_k may be decomposed into a density for a continuous random variable f_k and a discrete density ϕ_k along with a singular measure (such as a Cantor measure) that fortunately can be ruled out in equilibrium.

party is large enough then it is advantaged and it takes the election with very high probability, while the small party concedes the election.

In contested elections:

3. Increasing the value of the prize of the advantaged party increases the surplus of the advantaged party (and hence welfare), increases the probability of the advantaged party winning, decreases the turnout of the disadvantaged party and has no effect on the turnout of the advantaged party. The reverse is true for decreasing the valuation of the advantaged party provided it remains advantaged.

4. Decreasing the valuation of the disadvantaged party increases the surplus of the advantaged party (and hence welfare), decreases the turnout of the advantaged party and if $\hat{\varphi}_d < \overline{y}_d$ decreases the turnout of the disadvantaged party. The reverse is true for increasing the valuation of the disadvantaged party provided provided it remains disadvantaged.

Proof. See Appendix I. □

3.1. The Uniform Case

In addition to knowing which party is advantaged and gets all the surplus, it is of interest also to know which party has a better chance of winning the election. We now specialize to the case of identical costs $D_k = D, \underline{y}_k = \underline{y}, \overline{y}_k = \overline{y}$. In general, computing which party has a better chance of winning the election is quite difficult because the continuous part of the distribution given by the density $f_k(\eta_k\varphi_k) = (1/\eta_{-k})D'((\eta_k/\eta_{-k})\varphi_k)/v_{-k}$ depends on the derivative of the non-linear function $D(\varphi)$ and this function is evaluated at different points for the two parties. In the limit case of the polynomial $c(\varphi) = \alpha(\varphi - \underline{y})^{\alpha-1}$ considered above as $\alpha \to 1$ it is possible to give better conclusions. Notice that as $\alpha \to 1$ we converge to the case where all non-committed voters face the same cost. We also assume that $c(1) \leq \bar{P}$, so that $\bar{y} = 1$ and it is possible to turn out all voters. Assume also that monitoring inefficiency $\theta < 1$ so that cost is increasing, and consider only the case in which $\hat{\varphi}_S > (\eta_L/\eta_S)\underline{y}$ so that we can have contested elections in equilibrium. The detailed derivation and proof of the results in this section can be found in the Web Appendix.

For $\varphi > \underline{y}$ the limiting cost is given by $D(\varphi) = (\theta - \underline{y}) + (1-\theta)\varphi$ and since $\theta < 1$ we have that $D'(\varphi) > 0$. Notice that as $\alpha \to 1$, $D(\underline{y}) = 0$ but $D(\underline{y} + \epsilon) \geq \theta(1 - \underline{y}) > 0$ so that the function $D(\varphi)$ is discontinuous at $\underline{y}$ - there is, in effect, a fixed cost of entry - and also that the function $D(\varphi)$ is concave. The discontinuity in $D(\varphi)$ is reflected in the limit of equilibrium in which the small party is advantaged through an additional atom for the small party at $\eta_L\underline{y}$ of size

$$\phi_S(\eta_L\underline{y}) = \frac{\theta(1-\underline{y})}{v_L}$$

with the small party always winning the tie. Moreover the continuous part of the density is now uniform with

$$\lim{}_{\alpha\to 1} f_k(\eta_k\varphi_k) = \frac{1-\theta}{\eta_{-k}v_{-k}}.$$

The next theorem provides a sufficient condition under which a party has a better chance of winning the election.

Theorem 3. *If $\eta_{-d}v_{-d}/\eta_d v_d \geq 1$ then the advantaged party has a higher probability of winning the election.*

On the other hand there are parameters, for example, $\theta = 0$, $\eta_L = 2.5\eta_S$, $v_L = 2/9$, $v_S = 6/9$ and $\underline{y} = 1/9$ for which the large party is advantaged yet the small party has a higher probability of winning the election.

3.2. Common Prize

We continue to assume identical costs and consider now the case of a prize that is of equal value to both parties, that is $\eta_S v_S = \eta_L v_L = V$. We suppose that $D(\varphi)$ is strictly increasing and twice differentiable in $[\underline{y}, \overline{y}]$.

Theorem 4. *If $D(\varphi)$ is convex then the large party is advantaged and has the higher probability of winning the election. If $D(\varphi)$ is concave for $\varphi \geq \underline{y}$, $(\eta_S/\eta_L)\overline{y} > \hat{\varphi}_L$ and for some $\underline{y} < \hat{y} < \overline{y}$ with $D(\hat{y}) < V/\eta_L < D(\overline{y})$ we have*

$$\frac{2\underline{y}D'(\underline{y})}{\hat{y}^2 - \underline{y}^2} < - \max_{\underline{y} \leq y \leq \hat{y}} D''(y)$$

then the small party is advantaged and has the higher probability of winning the election.

Proof. In Appendix I. □

Since the condition for the small party to be advantaged is a complicated one it is useful to summarize the requirements. It must be that $\overline{y}$ is large, $\underline{y}$ is small, that the size disadvantage is not too great, that the prize is of intermediate value and that $D(\varphi)$ is concave enough. Notice that since $C(\varphi)$ is convex, the concavity requirement on $D(\varphi)$ means that monitoring costs must play an important role. There is a straightforward intuition for the requirements. If $\overline{y}$ is small or $\underline{y}$ is large, the election is basically determined by the number of voters. Similarly if the prize is small the election will be determined by the committed voters, while if it is large it will be determined by the party that can turn out the most number of voters - with identical costs this means the larger party. Finally, we note that if $D(\varphi)$ is convex then the average cost of turning out a voter increases with turnout - this favors the larger party which needs to turn out fewer voters to win.

4. Vote Suppression and Mandatory Voting

4.1. Vote Suppression

Suppose that each party can increase the monitoring cost θ of the opposing party to an amount $\overline{\theta} > \theta$ by incurring a cost $G > 0$.

Theorem 5. *[Cesar Martinelli] If $\overline{\theta}$ is sufficiently close to θ then only the advantaged party will suppress votes. If G is sufficiently small it will choose to do so and this will be a strict Pareto improvement.*[12]

Proof. If $\overline{\theta}$ is sufficiently close to θ then there is no change in which party is advantaged regardless of whether votes are suppressed or not. The disadvantaged party therefore gets zero regardless of whether is suppresses votes or not, hence it will not pay a positive cost to do so. On the other hand $D_d(\hat{\varphi}_d) = v_d$. Since suppressing votes by increasing θ raises D and D is increasing in φ we

[12] This theorem was suggested to us by Cesar Martinelli during the 2015 Priorat Workshop.

see that it reduces $\hat{\varphi}_d$. This raises the surplus $v_{-d} - D_{-d}(\hat{\varphi}_d)$ of the advantaged party so if G is small enough it is worth paying. □

Note that if the vote suppression raises the cost to each voter of voting the result about the advantaged party remains the same, but the disadvantaged party will be strictly worse off.

4.2. Mandatory Voting

We now explore the consequences of mandatoring voting, continuing to assume identical costs. Suppose that all non voters are charged a small fee $f > 0$. The fee is collected by an external agency ouside the party (or alternatively we can assume that fees are thrown away). We will make two simplifying assumptions:

1. We assume $\pi = 0$. In this case the auditor cannot determine whether or not the auditee violated the policy and hence any non voter must be punished.

2. We also assume that $\psi = 1$.

Notice that now incentive compatibility requires $P + f = c(\varphi)$, in which case any member with $y \leq \varphi$ would be willing to pay the cost $c(y)$ of voting rather than face the certain punishment P and pay the fee f, while any member with $y > \varphi$ prefers to pay the cost of punishment P plus the fee f over the cost of voting $c(y)$. As before, $D(\varphi)$ has two parts. The participation cost is identical as before. The monitoring cost is $M(\varphi) = (1 - \varphi)(C'(\varphi) - f)$. Hence $D(\varphi) = C(\varphi) + (1 - \varphi)(C'(\varphi) - f)$. Note that since $D(\hat{\varphi}_k) = v_k$, we have that taking derivatives with respect to f

$$\frac{\partial \hat{\varphi}_k}{\partial f} = \frac{1 - \hat{\varphi}_k}{D'(\hat{\varphi}_k)} > 0$$

and from Theorem 2, as it should be, mandatory voting increases turnout.

Consider now the uniform case and let us focus on the case in which the small party is not constrained, $c(1) \leq \bar{P} + f$, so that $\bar{y} = 1$, and $f \in (0,1)$. We say that mandatory voting *enhances parties competition* if

$$\frac{\partial |\hat{\varphi}_k - \hat{\varphi}_{-k}|}{\partial f} < 0,$$

and we have the following result:

Theorem 6. *In the uniform case if the large party is advantaged, mandatory voting enhances parties competition. If the small party is advantaged, mandatory voting enhances parties competition if and only if* $\eta_L/\eta_S < (1 + \underline{y} + v_S)/(1 + \underline{y} + v_L)$.

The proof is in the web appendix. Notice that, Theorem 6 implies that mandatory voting can boost the electoral prospects of a small party with intense preference in particular when the small party is only slightly advantaged.

5. Contests

The all pay auction is the limiting case of a contest which is decided by a conflict resolution function in which the probability of winning the election is a continuous function of the expected number of voters each party turns out. In particular the outcome of the election is decided by the actual number of votes rather than the expected number of votes - in this case the conflict resolution function is derived from the binomial distribution (computed for example in Palfrey and Rosenthal [24] and Levine and Palfrey [20].) However there are other reasons such that the outcome of an election may be random, varying from correlation in the draws of y by voters to random errors in the counting of votes, the way in which ballots are validated or invalidated or intervention by courts. In this section we want to incorporate such randomnes in our model.

A second thing we wish to account for is pivotality in the incentive constraint. This means that the cost function for turning out voters depends on the turnout of the other party, since this affects pivotality. We give the appropriate generalization here, and discuss in more detail how it arises from pivotality below.

Getting to the details, we continue to suppose that a population of N voters is divided into two parties $k = S, L$ of size $\eta_k N$. These parties now compete in a contest. Each party continues to produce an expected number of votes $\eta_k \varphi_k$ and may win a prize worth $v_k > 0$ to each voter respectively. Now, however, the probability of the small party winning the prize is given by a conflict resolution function $p_S(\eta_S\varphi_S, \eta_L\varphi_L) \in [0, 1]$ and we define $p_L(\eta_L\varphi_L, \eta_S\varphi_S) = 1 - p_S(\eta_S\varphi_S, \eta_L\varphi_L)$.

A strategy for party k is a probability measure represented as a cumulative distribution function F_k on $[0, 1]$. Each party faces a per capita costs of turning out voters characterized by a cost function $D_k(\varphi_k, F_{-k})$ which represents the cost of turning out a fraction φ_k of voters when pivotality is determined by φ_k, F_{-k}. For a topology to the space of cumulative distribution functions on $[0, 1]$ we use the weak topology: the corresponding notion of convergence is that $F_k^N \to F_k$ if the expectation of every continous function on $[0, 1]$ converges.[13]

We assume that $p_S(\eta_S\varphi_S, \eta_L\varphi_L)$ and $D_k(\varphi_k, F_{-k})$ are continuous on $[0, 1]$. Note that in general both depend on the absolute size of the population N. The latter assumption amounts to assuming that $\overline{y}_k = 1$, that is, that the punishment is sufficiently large that it is possible (although possibly very expensive) to coerce all voters into voting.[14] Note that we assume nothing regarding the monotonicity of D_k, this is important for the result on convergence to pivotal equilibrium since when we allow pivotality, and monitoring costs are very high, D_k will be continuous but certainly not monotone - it can be expensive to get voters not to vote when they are motivated to vote because of the chance they are pivotal.

We say that F_S, F_L are an *equilibrium* of the conflict resolution model if

$$\begin{aligned}\int v_k p_k(\eta_k\varphi_k, \eta_{-k}\varphi_{-k}) F_k(d\eta_k\varphi_k) F_{-k}(d\eta_{-k}\varphi_{-k}) - \int D_k(\varphi_k, F_{-k}) F_k(d\eta_k\varphi_k) \quad &\geq \\ \int v_k p_k(\eta_k\varphi_k, \eta_{-k}\varphi_{-k}) \tilde{F}_k(d\eta_k\varphi_k) F_{-k}(d\eta_{-k}\varphi_{-k}) - \int D_k(\varphi_k, F_{-k}) \tilde{F}_k(d\eta_k\varphi_k) &\end{aligned}$$

[13] This is called the weak topology by probability theorists and the weak* topology in functional analysis Glicksberg [11].

[14] Otherwise we would have $\overline{y}_k$ depending on the strategy of the other party F_{-k} creating the problem discussed in Dutta, Levine and Modica [6] when there are constraints on party behavior that depend on the choice of the other party.

Theorem 7. *An equilibrium of the conflict resolution model exists.*

Proof. In Appendix II. $\square$

5.1. Continuity

We now consider an infinite sequence of conflict resolution models $p_S^N(\eta_S\varphi_S, \eta_L\varphi_L)$, $D_S^N(\varphi_S, F_L)$, $D_L^N(\varphi_L, F_S)$ and an all-pay auction with costs $D_k(\varphi_k)$ differentiable on $(\underline{y}_k, 1]$ with $(1/\epsilon) > D_k'(\varphi) > \epsilon$ for some $\epsilon > 0$ and $\eta_S\hat{\varphi}_S \neq \eta_L\hat{\varphi}_L$. We say that the sequence of conflict resolution models *converges to the all-pay auction* if for all $\epsilon > 0$ and $\eta_S\varphi_S > \eta_L\varphi_L + \epsilon$ we have $p_S^N(\eta_S\varphi_S, \eta_L\varphi_L) \to 1$ uniformly, and $\eta_S\varphi_S < \eta_L\varphi_L - \epsilon$ implies $p_S^N(\eta_S\varphi_S, \eta_L\varphi_L) \to 0$ uniformly, and $D_k^N(\varphi_k, F_{-k}) \to D_k(\varphi_k)$ uniformly.

Theorem 8. *Suppose that $p_S^N(\eta_S\varphi_S, \eta_L\varphi_L), D_S^N(\varphi_S, F_L), D_L^N(\varphi_L, F_S)$ converge to the all-pay auction $D_S(\varphi_S), D_L(\varphi_L)$, that F_k^N are equilibria of the conflict resolution models and that F_k is the unique equilibrium of the all-pay auction. Then $F_k^N \to F_k$.*

To prove this theorem we use the standard approch to prove that equilibrium correspondences are upper-hemi-continuous: we show that for any strategy by party k, the utility received for large N is uniformly close to the utility received in the limit. Hence, any strictly profitable deviation in the limit would also have to be strictly profitable for large N. If the limit of p_k^N was a continuous function this would be completely straightforward: since the convergence of the objective functions is assumed to be uniform and the integrals defining expected utility would converge for any fixed p_k by a standard argument this would give uniform convergence of the objective functions. The complication is that the limit is not a continuous function: it is discontinuous when there is a tie. Suppose, however, that we can show that the equilibria F_k^N had the property that for any ϵ and large enough N there is a uniform bound such that the probability of an ϵ neighborhood is at most $\underline{\Pi}\epsilon$ - basically that F_k^N is converging to a limit with a bounded continuous density. Then for any choice of φ_{-k} it would be the case that for large enough N the utility of F_k^N would be at most ϵ-different for φ_k outside of an ϵ neighborhood of $(\eta_k/\eta_{-k})\varphi_{-k}$ and that the probability of being in that neighborhood is also of order no more than ϵ. This implies that the utility of F_k^N is of order ϵ different than the utility of F_k^∞, which is what is needed for the standard argument to work. This argument is not completely adequate because there can be two points where there is an atom in the limit, but these cases can be covered by the appropriate choice of tie-breaking rule. The details can be found in Appendix II.

Note that while this is stated as an upper hemi-continuity result, since the equilibrium of the all-pay auction is unique and we have existence it is in fact a continuity result. We can summarize it by saying that if the conflict resolution model is close enough to winner-take-all and pivotality is not important then the equilibria of the conflict resolution model are all close to the unique equilibrium of the all-pay auction.

The theorem allows N to be any abstract index of a sequence. A particularly interesting application is the one in which N represents population size and we consider that the conflict resolution function is binomial arising from independent draws of type by the different voters.

In this case an application of Chebychev's inequality gives the needed uniform convergence of $p_S^N(\eta_S\varphi_S, \eta_L\varphi_L)$. The convergence of costs when pivotality is accounted for is shown below.

5.2. High Value Elections and the Impact of Electorate Size on Turnout

Now we consider an alternative conceptual experiment: we hold fixed the size of the electorate, but allow the size of the prize to be very large.

Theorem 9. *Suppose $v_L \to \infty$. Then $F_L(1-\epsilon) \to 0$.*

Proof. Recall that we have assumed that $\overline{y} = 1$. Suppose that with positive probability party L chooses φ smaller than $1-\epsilon$. Because the small party turns out at least $\underline{y}$ and the conflict resolution is fixed and continuous (N is fixed) this implies that the difference in the probability of losing between φ and 1 is bounded away from 0 by $\kappa > 0$ independent of v_L. Hence party L gains at least $\kappa v_L - \max_{\varphi_L, F_S} D_L^N(\varphi_L, F_S)$ which is positive for v_L large enough. □

This theorem together with Theorem 8 have an important implication about voter turnout and population size. Because the conflict resolution function is continuous, Theorem 9 gives a different result than the all-pay auction. Here as the prize grows large the large party almost certainly turns out all of its voters. In the all-pay auction case we know that it turns out only enough voters to beat the small party, that is $\eta_S/\eta_L < 1$. Consequently, if we first fix N and then make the size of the prize large enough, we obtain that the large party will turn out more than enough voters to beat the small party, that is $\varphi > \eta_S/\eta_L$. If we now fix the size of the prize and increase the number of voters Theorem 8 implies that equilibrium must approach the equilibrium of the all-pay auction, meaning that the turnout of the large party must decline to η_S/η_L.

5.3. Pure Strategy Equilibrium

When the conflict resolution function is "nearly" winner-take-all and pivotality is not too important, we are close to the all-pay auction and hence the equilibrium is in mixed strategies. By contrast if the objective functions $p_k(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})v_k - D_k(\varphi_k, F_{-k})$ are single-peaked in $\eta_k\varphi_k$ (for example: p_k is concave and D_k convex, at least one strictly) then a standard argument shows that there is a pure strategy equilibrium. Indeed, when the objective functions are single-peaked there is a unique optimum for each party and so the party cannot mix - all equilibria must be pure strategy equilibria. This is basically the case considered in Coate and Conlin [3].[15]

It is worth considering what is needed for $p_k(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})$ to be concave in the symmetric case in which $p_L(\eta_L\varphi_L, \eta_S, \varphi_S) = p_S(\eta_L\varphi_L, \eta_S\varphi_S)$. Symmetry implies $p_k(1/2, 1/2) = 1/2$. Concavity implies $p_k(1/4, 1/2) \geq (1/2)p_k(0, 1/2) + (1/2)p_k(1/2, 1/2) \geq 1/4$. In other words, when one party sets a target of a 2-1 majority, it must none-the-less have at least a 25% chance of losing: in this

[15] However, the model inCoate and Conlin [3] differs slightly from the one here: in their model the size of the parties is random. Conceptually this poses a problem for a model of peer punishment within parties - it is not clear how a collusive agreement can be reached among a party whose members are not known. Coate and Conlin [3] also use a different objective function than we do: they assume that the "party" maximizes total utility so that when a random event causes the party to be large the party is "happier" than when it is small. This is not necessary, they could consider (as implicitly we do) a party that maximizes per capita rather than total utility.

sense concavity means “a great deal” of variance in the outcome. Hence we have the broad picture that when pivotality is not important small aggregate uncertainty leads to non-trivial mixing in equilibrium, while large aggregate uncertainty leads to pure strategies in equilibrium.

Example 2. Suppose that the types y_k have both a common and an idiosyncratic component where the common component may be correlated between the two parties. We have indexed types by a uniform distribution on $[0, 1]$. It is convenient in developing an example with a common component to index types by z_k drawn from continous strictly increasing cdfs $G_k(z)$ on $[0, \infty)$. The original index y_k can then be recovered from the formula $y_k = G_k(z_k)$. In our example the objective function will be concave in z_k - this implies that it is single-peaked in y_k. With the index z_k the party chooses a type threshold ζ_k. We assume that N is large - the proof of a formal convergence result here is straightforward since the limiting conflict resolution function is continuous.

The specific example is defined by a parameter $0 < \beta$. We assume that costs are sufficiently high relative to the prize so that $\hat{\varphi}_k < \beta/(1+\beta)$. Each voter i in party k takes an iid draw u_i from a uniform distribution on $[0, 1]$. A single independent common draw ν is taken also from a uniform on $[0, 1]$. We set $\nu_S = \nu^{1/\beta}$ and $\nu_L = (1-\nu)^{1/\beta}$ and a voter's type is $z_{ik} = \beta u_i/(1+\beta)\nu_k$. We let ζ_k denote the threshold for voting in terms of z_k. Because we are assuming that N is large we ignore pivotality so that it follows that the cost of turning out voters is $D_k(G_k(\zeta_k)) \geq 0$.

Conditional on the common shock ν_k the expected fraction of voters that turns out is $\Pr(z_k \leq \zeta_k|\nu_k) = \Pr(u_i \leq ((1+\beta)/\beta)\zeta_k\nu_k|\nu_k)$. For $\zeta_k \leq \beta/(1+\beta)$ this is $\Pr(z_k \leq \zeta_k|\nu_k) = ((1+\beta)/\beta)\zeta_k\nu_k$ (since the RHS is no greater than 1). Observe first that $\Pr(z_k \leq \zeta_k) = \int((1+\beta)/\beta)\zeta_k\nu^{1/\beta}d\nu = \zeta_k$ from which we can conclude that for $\zeta_k \leq \beta/(1+\beta)$ we have $y_k = z_k$. Since it cannot be optimal to choose $\varphi_k > \hat{\varphi}_k$ and $\hat{\varphi}_k \leq \beta/(1+\beta)$ we see that for $\varphi_k \leq \hat{\varphi}_k$ the expected fraction of voters who turn out conditional on the common shock ν_k is $((1+\beta)/\beta)\varphi_k\nu_k$.

Because we are assuming that N is large we suppose that the actual fraction of voters who turn out is exactly $((1+\beta)/\beta)\varphi_k\nu_k$. Hence party k wins the election if $\eta_k\varphi_k\nu_k > \eta_{-k}\varphi_{-k}\nu_{-k}$. Taking logs, this reads $\log(\eta_k\varphi_k/(\eta_{-k}\varphi_{-k})) + (1/\beta)(\log(\nu) - \log(1-\nu) > 0$. Since for a uniform ν on $[0, 1]$ the random variable $\log(\nu) - \log(1-\nu)$ follows a logistic distribution the probability of winning is the Tullock contest success function

$$\frac{1}{1+(\eta_{-k}\varphi_{-k}/(\eta_k\varphi_k))^\beta} = \frac{(\eta_k\varphi_k)^\beta}{(\eta_k\varphi_k)^\beta + (\eta_{-k}\varphi_{-k})^\beta}.$$

A sufficient condition for this to be concave is that $\beta \leq 1$ so that if $D_k(\varphi_k)$ is strictly increasing when it is strictly positive, continous and (at least for $\varphi_k \leq \hat{\varphi}_k$) convex then there are only pure strategy equilibria. By contrast as $\beta \to \infty$ the distribution of $\nu^{1/\beta}$ approaches a point mass at 1 and we approach the case of the all-pay auction and for large β there can be no pure strategy equilibrium. However, under the assumption that the cost function satisfies a diminishing hazard rate condition, Herrera, Morelli and Nunnari [12] show that pure strategy equilibria exist for relatively large β and give a detailed breakdown of the comparative statics of the equilibrium.

5.4. Pivotality

Up to this point we have assumed that voters vote only because either they prefer to do so, the commited voters, or because they face punishment if they do not: these motivations are reflected in the incentive constraint $c_k(y) \leq P_k$ that voters should vote provided the cost of doing so is less than or equal to the cost of being punished for not doing so. This formulation ignores the traditional focus in Palfrey and Rosenthal [24] on the individual incentive to vote based on the chance of being pivotal. We turn now to combining the more traditional incentives of pivotality with the possibility of punishment.

We start by giving a formulation of the contest model that enables us to compute the probability of being pivotal. Rather than a single conflict resolution function we define two partial conflict resolution functions: $P_k^0(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})$, the probability of winning conditional on all voters except one following the social norm φ_k and *the remaining voter not voting*, and $P_k^1(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})$, the probability of winning conditional on all voters except one following the social norm φ_k and *the remaining voter voting*. These should be differentiable and non-decreasing in φ_k. This two functions enable us to compute both the overall conflict resolution function and the probability of being pivotal: the overall conflict resolution function is $p_k(\eta_k\varphi_k, \eta_{-k}\varphi_{-k}) = \varphi_k P_k^1(\eta_k\varphi_k, \eta_{-k}\varphi_{-k}) + (1 - \varphi_k)P_k^0(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})$ and the probability of being *pivotal* is $Q_k(\eta_k\varphi_k, \eta_{-k}\varphi_{-k}) = P_k^1(\eta_k\varphi_k, \eta_{-k}\varphi_{-k}) - P_k^0(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})$. It is convenient in what follows to view the strategies F_k as measures rather than functions.

To analyze incentives with pivotality we start by identifying what and individual voter would like to do in the absence of punishment. This depend on what voters from both parties are doing. For any given social norm φ_k and mixed strategy of the other party F_{-k} we may define the pivotal cutoff $\gamma_k(\varphi_k, F_{-k})$ by the solution to $c_k(\gamma_k) = Q_k(\eta_k\varphi_k, \eta_{-k}F_{-k})v_k$. This represents the type of voter who is indifferent between bearing the cost of voting in order to improve the party's chance of victory and abstaining. Since $c_k(y)$ is differentiable and has a strictly positive derivative the solution is unique and continuous. We can now determine the incentive constraint when there is punishment for not voting. For voters who would not otherwise vote, that is, $y \geq \gamma_k(\varphi_k, F_{-k})$, the incentive constraint is $c_k(y) - Q_k(\eta_k\varphi_k, \eta_{-k}F_{-k})v_k \leq P_k$. This says that the net cost of voting, which is the direct cost c_k minus the benefit because of pivotality $Q_k v_k$, must be less than or equal to the punishment for not voting. Notice that the mixed strategy of the other party F_{-k} appears in the incentive constraint since P_k must be chosen before the realization of φ_{-k} is known.

From the incentive constraint we can derive the monitoring cost for $\varphi_k \geq \gamma_k(\varphi_k, F_{-k})$ as the cost of punishing a non-voter who was not supposed to vote after having received a wrong signal $M_k(\varphi_k, F_{-k}) = \psi(1-\pi)(1-\varphi_k)\left(c_k(\varphi_k) - Q_k(\eta_k\varphi_k, \eta_{-k}F_{-k})v_k\right)$. Notice that if φ_k is pivotal in the sense that $\varphi_k = \gamma_k(\varphi_k, F_{-k})$ then $M_k(\varphi_k, F_{-k}) = 0$ and the function M_k is continuous.

There remains the issue of what happens if the social norm calls for less participation than would be individually optimal in the presence of the pivotality incentive $\varphi_k < \gamma_k(\varphi_k, F_{-k})$. For voters with $\varphi_k < y < \gamma_k(\varphi_k, F_{-k})$ the social norm calls on y to not to vote, but in fact y would like to. This case is not covered by the basic model and there is more than one modeling possibility. One is to assume that there is no cost of getting a voter not to vote, in which case $\varphi_k < \gamma_k(\varphi_k, F_{-k})$ and $M_k(\varphi_k, F_{-k}) = 0$. In this case we may write $M_k(\varphi_k, F_{-k}) = \psi(1-\pi)(1-\varphi_k)\max\{0, (c_k(\varphi_k) - Q_k(\eta_k\varphi_k, \eta_{-k}F_{-k})v_k)\}$ which is obviously continuous, although scarcely linear in F_{-k}. However, all that is required for the results that follow is that $M_k(\varphi_k, F_{-k})$ is non-negative for $\varphi_k < \gamma_k(\varphi_k, F_{-k})$.

To summarize: the goal of the party is to maximize per capita utility $p_k(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})v_k -$

$C_k(\varphi_k) - M_k(\varphi_k, F_{-k})$. We already showed that equilibrium distributions F_S, F_L exist. A useful technical result is that the model chosen of monitoring costs for $\varphi_k < \gamma_k(\varphi_k, F_{-k})$ does not matter because in equilibrium neither party ever chooses such a low participation rate.

Lemma 1. *Basic Lemma: If F_S, F_L are equilibrium distributions then $F_k(\varphi_k < \gamma_k(\varphi_k, F_{-k})) = 0$.*

Proof. Start by assuming that $\varphi_k < \gamma_k(\varphi_k, F_{-k})$ and $M_k(\varphi_k, F_{-k}) = 0$, we will show later that this assunption does not matter. Then the objective function is

$$U_k = \left(\varphi_k P_k^1(\eta_k\varphi_k, \eta_{-k}\varphi_{-k}) + (1-\varphi_k)P_k^0(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})\right) v_k - C_k(\varphi_k)$$

and differentiating with respect to φ_k we get

$$\frac{dU_k}{d\varphi_k} = \left(Q_k(\eta_k\varphi_k, \eta_{-k}\varphi_{-k}) + \varphi_k\eta_k \frac{dP_k^1(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})}{d\eta_k\varphi_k} + (1-\varphi_k)\eta_k \frac{dP_k^0(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})}{d\eta_k\varphi_k}\right) v_k - c(\varphi_k).$$

Since $\varphi_k < \gamma_k(\varphi_k, F_{-k})$ we have $Q_k(\eta_k\varphi_k, \eta_{-k}\varphi_{-k})v_k = c_k(\gamma_k(\varphi_k, F_{-k})) > c_k(\varphi_k)$ so that $dU_k/d\varphi_k > 0$, so that certainly F_k puts no weight on a neighborhood of φ_k. Now we drop the assumption that for $\varphi_k < \gamma_k(\varphi_k, F_{-k})$ we have $M_k(\varphi_k, F_{-k}) = 0$. Notice that we may increase φ_k until the first time that $\tilde{\varphi}_k = \gamma_k(\tilde{\varphi}_k, F_{-k})$ is satisfied, and it follows that $U_k(\tilde{\varphi}_k) - M_k(\tilde{\varphi}_k, F_{-k}) > U_k(\varphi_k) - M_k(\varphi_k, F_{-k})$ since the derivative is strictly positive. But $M_k(\tilde{\varphi}_k, F_{-k}) = 0$ so φ_k is strictly worse than $\tilde{\varphi}_k$. $\square$

One implication of the final step of the proof of the lemma is that the set of equilibria for any definition of $M_k(\varphi_k, F_{-k}) \geq 0$ for $\varphi_k < \gamma_k(\varphi_k, F_{-k})$ contains the equilibria for the corresponding model with $M_k(\varphi_k, F_{-k}) = 0$ for $\varphi_k < \gamma_k(\varphi_k, F_{-k})$ - and in particular since equilibria in the former model exist, they also exist in the latter.

Our main interest is in what equilibrium with pivotality looks like. There are two limits of interest. The first is the large population case which has been our focus. In the case of iid draws of y as $N \to \infty$ the fact that $\underline{y}_k > 0$ forces the probability of being pivotal to converge to zero uniformly. It follows that $M_k(\varphi_k, F_{-k})$ converges uniformly to $M_k(\varphi_k) = \psi(1-\pi)(1-\varphi_k)c_k(\varphi_k)$ and so Theorem 8 tell us that we approach the unique all-pay auction equilibrium: as we expect, pivotality is not important when the electorate is large.

The second limit of interest is the case of large monitoring costs - in this case we expect pivotality to play the decisive role. Specifically we would like to show that as $\psi \to \infty$ we approach a correlated equilibrium of the purely pivotal model. However, this need not be the case - there is no cost of monitoring at $\varphi_k = 1$ and we have assumed $\overline{y} = 1$, so the very high costs of monitoring can potentially be avoided by choosing a very high participation rate. Suppose in particular that $C_k(1) < v_k$ so that it would pay to turn out the entire electorate without monitoring cost. Then we cannot rule out the possibility that even for very large ψ equilibrium might involve participation rates close to 1 and very much higher than the pivotal cutoff $\gamma_k(\varphi_k, F_{-k}^{\psi})$.

By contrast, suppose that $C_k(1) > v_k$. It follows directly from Lemma 1 that

Theorem 10. *If $C_k(1) > v_k$ then as $\psi \to \infty$ we have $F_k^{\psi}(|\varphi_k - \gamma_k(\varphi_k, F_{-k}^{\psi})| \leq \epsilon) \to 1$.*

Notice that this does not necessarily imply that the limit is an equilibrium in the sense of Palfrey and Rosenthal [24] since we allow correlation devices within parties, but rather a correlated equilibrium with pivotality of the type studied by Pogorelskiy [26].

6. Welfare

In our model, effort spent voting and monitoring voters is pure waste. If the prize is of equal value to the two parties and there is no uncertainty about party sizes then a mechanism of flipping a coin (and allowing the committed voters to vote, but ignoring the results of that vote) is a strict welfare improvement over voting. These results are heavily influenced by the fact that increased participation increases costs and thereby tends to reduces welfare. However, there are a number of reasons to believe that higher participation might be welfare improving. In this section we want to explore some of these arguments.[16]

A possible reason for preferring higher participation is that, in a model of incomplete information, higher participation would lead to greater information aggregation. However, once the absolute number of voters is reasonably large the additional improvement in information aggregation from doubling or tripling the number of voters seems quite modest. After all, do we really believe that the United States general elections aggregate information much better than those in the UK because ten times as many voters vote? A more common reason given why high participation is important is because *it increases the sense of participation in government*. What concretely this might mean beyond the relatively meaningless statement that each vote creates a utility benefit for someone is hard to say. One possibility is that it signals the willingness or likelihood that citizens will abide by the democratic rules or that they will fight off attempts at coup d'etats and the like. In our view, one of the most reasonable arguments for preferring higher participation is that when turnout is too low there is a concrete possibility that a minority non-democratic party may come to power. This latter possibility is relatively straightforward to model and not inconsistent with the idea that there might be broader benefits such as signaling.

Suppose that depending on turnout $\eta^k\varphi^k$ there is a probability that a third party who will refer to as *blackshirts* might win the election. Denote this probability by $B(\eta^1\varphi^1, \eta^2\varphi^2)$ and suppose that each party k suffers a per-capita loss of b^k in the event in which the blackshirts win the election. Recall that in our model costs are $D^k(\varphi^k) = C^k(\varphi^k) + M^k(\varphi^k)$. Extending the model to the the possibility of blackshirts winning the election leads to a new cost function $\tilde{D}^k(\eta^1\varphi^1, \eta^2\varphi^2) = D^k(\varphi^k) + b^k B(\eta^1\varphi^1, \eta^2\varphi^2)$. This differs from the original model for two reasons: First, the participation rate of the opposing party $-k$ now enters the cost function of party k. Second, the sizes of the parties have an impact on per capita costs. To see how these differences will affect the welfare implications of the model, consider a simple example in which the number of blackshirts who will cast votes is uncertain and uniformly distributed with density $1/\beta$, and institutions are such that in order to achieve power blackshirts must collect an absolute majority. Then

$$B(\eta^1\varphi^1, \eta^2\varphi^2) = B_0 - \beta(\eta^1(\varphi^1 - \underline{y}^1) + \eta^2(\varphi^2 - \underline{y}^2)),$$

[16]We thank Guido Tabellini for pushing us in this direction and helping us to explore some of these ideas.

where $B_0 \leq 1$ denotes the probability that blackshirts are enough to outnumber the commited voters. Furthermore, we assume that $B(\eta^1\overline{y}^1, \eta^2\overline{y}^2) \geq 0$ so that there is a positive chance that blackshirts are enough to outnumber all the voters in both parties.[17] This additively separable functional form has the advantage that the marginal benefit of reducing the chance of blackshirts winning is independent of the participation rate of the other part so that we may apply our existing analysis of equilibrium. In particular we may analyze equilibrium using the overall per-capita cost function for party k

$$\tilde{D}^k(\eta^1\varphi^1, \eta^2\varphi^2) = D^k(\varphi^k) + b^k \left[B_0 - \beta(\eta^1(\varphi^1 - \underline{y}^1) + \eta^2(\varphi^2 - \underline{y}^2))\right]$$

using our existing analysis. If we are willing to assume b^k, and β small enough and M^k not too big, the $\tilde{D}^k$ function will be increasing, despite the fact that B will be decreasing and M^k might be decreasing as well. In particular, if $\beta\eta^k b^k < dD^k(\varphi^k)/d\varphi^k$ then $\tilde{D}^k$ is increasing in φ^k. In this case our all-pay auction and conflict resolution analysis applies unchanged.

The welfare analysis can change substantially, however, even when $\tilde{D}^k$ is increasing. In fact, if $\beta\left(\eta^1 b^1 + \eta^2 b^2\right) > dD^k(\varphi^k)/d\varphi^k$, the overall per-capita cost of voting

$$\tilde{D}^1(\eta^1\varphi^1, \eta^2\varphi^2) + \tilde{D}^2(\eta^1\varphi^1, \eta^2\varphi^2)$$

is *decreasing* in φ^k. In this case the earlier welfare conclusions will effectively be reversed.

Two further observations may be of interest. First, in addition to parties we may wish to consider an additional social network of "citizens" to which everyone belongs and which also can create incentives for voting.[18] Such a network may have as its objective keeping the blackshirts out of power. Because this social network strictly prefers greater voter participation, we may wish to interpret the "committed" voters $\underline{y}^k$ as those for whom the citizen incentives are binding which is to say that they vote because it is their "civic duty." This would lead to a theory in which the number of committed voters is endogenous and would depend on the strength of the blackshirts and whether the other existing parties are willing to turn out voters. Second, the idea that the strength of blackshirts might vary is an important one. Since the Second World War voter participation has fallen. There are many explanations of this, including the weakening of parties ability to mobilize voters.[19] But note also that during the same period the possibility of a blackshirts surge diminished substantially as democratic institutions became increasingly stronger (ranging from peacefully accepting the election outcome, to the strength of the courts and independence of the press). In the early years of European democracy, institutions were weak, the chance of blackshirts high, and a high voter turnout was perfectly reasonable. As time has gone on and institutions got

[17] It should be noted that turnout in 1933 German federal elections was 89% and the Nazist party fell barely short of an absolute majority with 44% of the votes. Clearly the election was not free from intimidation, but of course this is part of the reason for turning out a great many voters to defeat the blackshirts.

[18] For details on models where individual members belong to several social networks that can induce incentives see Dutta, Levine and Modica [8].

[19] See Gray and Caul [16]and Knack [15] among others.

stronger, better established, and confidence in them has been growing, there has been less reason for high turnout to keep blackshirts out of power.[20]

7. Conclusion

We have examined a simple model of voter turnout where collusive parties choose social norms enforced by peer punishment. This model is broadly consistent with the conclusions of the ethical voter model - for example the dependence of turnout on the size of the electorate is modest and due to the fact that there is a small degree of residual uncertainty about turnout. The key difference from the ethical voter model is that when the cost of punishment is significant, the cost of turning out voters will be concave rather than convex - and as we have seen this advantages the smaller party. There are many elections where special interests do well: for example Indian lotteries, school boards, school salary referenda, prison guard and so forth. In general we would expect that single-issue voting - referenda - which keeps the stakes to the smaller party large while the stakes to the larger party small should favor the smaller party,[21] while general issue voting - for example for Governor or President - will have high stakes disadvantaging the small party. One implication of this is that in the case of referenda the way for a large party to defeat a small party is to make sure that some high stakes issue is on the ballot. A good example of how this works was the passage of Proposition 8 in California in 2008. Here the large party was against gay marriage, and the small party in favor of gay marriage and the black community was especially opposed to gay marriage. Having a black presidential candidate led to very high black voter turnout, and it is generally thought that in a year of more ordinary turnout the proposition would have failed.

The model applies more generally to a situation where two groups compete by turning out members - for example in street demonstrations or strikes. The model potentially also has applications to models of lobbying by bribery as in Hillman and Riley [13], Acemoglu [1], or Levine and Modica [19] - although the welfare analysis is quite different as the "votes" which are wasted in a model of voting (or demonstrations) are income to a politician in a model of lobbying by bribery. There are three additional differences that are potentially important. First, we have examined only a 0-1 decision to participate. In lobbying there is also an intensive margin: participation can be at either a higher or a lower level. Second, rather than committed members of lobbying groups we might expect instead a fixed cost of providing a minimal amount of resources to be useful - this has the opposite effect of committed voters, favoring the smaller party. Third, the prize in lobbying may be fungible (money) that can be used to pay the politician's bribe so that the resource constraint of a small group may not matter so much. These considerations - and the results of Levine and Modica [19] - seem to suggest that lobbying by bribery may be more favorable to a small group

[20] It should be noted that while comparisons are sometimes made between modern populist movements such as UKIP in the UK, 5 Star in Italy and Trump in the US, with the exception of the Golden Dawn in Greece and unlike the Nazi and Fascist movements, these parties do not have anti-democratic paramilitary forces.

[21] Not all referenda are for low stakes. Large issue referenda - such as the recent referendum on Scottish independence or the forthcoming referendum on British membership in the EU - should favor the larger party.

than voting.

References

[1] Acemoglu, S. and J. A. Robinson (2001): “Inefficient Redistribution,” *American Political Science Review* 95: 649-661.

[2] Chamberlain, G., M. Rotschild (1981): “A Note on the Probability of Casting a Decisive Vote,” *Journal of Economic Theory* 25: 152–162.

[3] Coate, S., M. Conlin (2004): “A Group Rule–Utilitarian Approach to Voter Turnout: Theory and Evidence,” *American Economic Review* 94: 1476–1504.

[4] Coate, S., M. Conlin, A. Moro (2008): “The Performance of Pivotal-Voter Model in Small-Scale Elections: Evidence from Texas Liquor Referenda,Journal of Public Economics 92: 582–596.

[5] Della Vigna S., J. A. List, U. Malmendier and G. Rao (2014): “Voting to Tell Others,” NBER Working Paper.

[6] Dutta, R., D. K. Levine and S. Modica (2014): “Collusion, Randomization and Leadership in Groups,” EUI.

[7] Dutta, R., D. K. Levine and S. Modica (2016): “Collusion Constrained Equilibrium,” EUI.

[8] Dutta, R., D. K. Levine and S. Modica (2016): “Damned if you do and damned if you don't: Competing Networks,” EUI.

[9] Feddersen, T., A. Sandroni (2006): “A Theory of Participation in Elections,” *American Economic Review* 96: 1271–1282.

[10] Gerber A. S., D. P. Green, C. W. Larimer (2018): “Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment,” *American Political Science Review* 102: 33-48.

[11] Glicksberg, I. (1952): “A Further Generalization of the Kakutani Fixed Point Theorem, with Application to Nash Equilibrium Points,” *Proceedings of the American Mathematical Society* 3: 170-174.

[12] Herrera, H., M. Morelli, and S. Nunnari (2015): “Turnout Across Democracies,” *American Journal of Political Science.*

[13] Hillman, A. L. and J. Riley (2006): “Politically Contestable Rents and Transfers,” *Economics and Politics* 1: 17-39.

[14] Kandori, M. (1992): “Social Norms and Community Enforcement,” *The Review of Economic Studies* 59: 63-80.

[15] Knack, S. (1992): “Civic Norms, Social Sanctions, and Vote Turnout,” *Rationality and Society* 4: 133-156.

[16] Gray, M. and M. Caul (2000): “Declining Voter Turnout in Advanced Industrial Democracies, 1950 to 1997,” *Comparative Political Studies* 33: 1091-1122.

[17] Levine, D. K. and S. Modica (2013): “Peer Discipline and the Strength of Organizations,” WUSTL.

[18] Levine, D. K. and S. Modica (2014): ”Peer Discipline and Incentives Within Groups,” EUI.

[19] Levine, D. K. and S. Modica (2015): “Size, Fungibility, and the Strength of Organizations,” EUI.

[20] Levine, D., T. Palfrey(2007): “The Paradox of Voter Participation? A Laboratory Study,” *American Political Science Review* 101: 143–158.

[21] Mulligan, C. B. and C. G. Hunter (2003): "The Empirical Frequency of a Pivotal Vote," *Public Choice* 116: 31-54.

[22] Nagourney, A., A. Parker, J. Rutenberg and J. Zelenynov [2012]: "How a Race in the Balance Went to Obama," *New York Times*, November 7.

[23] Palfrey, T. R. and K. Pogorelskiy (2016): "Voting with Communciation: An Experimental Study," Caltech.

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

[25] Penrose, L. S. (1946): "The Elementary Statistics of Majority Voting," *Journal of the Royal Statistical Society* 109: 53-57.

[26] Pogorelskiy, K. (2015): "Correlated Equilibria in Voter Turnout Games," Caltech.

[27] Olson, Mancur (1965): *The Logic of Collective Action: Public Goods and the Theory of Groups.*

[28] York, B. (2012): "In Boston, stunned Romney supporters struggle to explain defeat," W*ashington Examiner*, November 7.

Appendix I: All-Pay Auction

We let ℓ denote the party with the most committed voters.

Theorem. *2 There is a unique mixed equilibrium. The disadvantaged party earns zero and the advantaged party earns* $v_{-d} - D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d) > 0$.

Theorem 11. *If* $\hat{\varphi}_k \leq (\eta_{-k}/\eta_k)\underline{y}_{-k}$ *then the election is* uncontested*: the least committed party* k *is disadvantaged,* concedes *the election by bidding* $\eta_k \underline{y}_k$ *and the most committed party* $-k$ takes *the election by bidding* $\eta_{-k}\underline{y}_{-k}$.

If $\hat{\varphi}_k > (\eta_{-k}/\eta_k)\underline{y}_{-k}$ *for* $k \in \{S, L\}$ *then the election is* contested: *in* $(\max_k \eta_k \underline{y}_k, \eta_d \hat{\varphi}_d)$ *the mixed strategies of the parties have no atoms, and are given by continuous densities*

$$f_k(\eta_k \varphi_k) = D'_{-k}((\eta_k/\eta_{-k})\varphi_k)/(\eta_{-k} v_{-k}).$$

In these contested elections there are three points that may have atoms: each party may turn out only its committed voters and the advantaged party may take the election by turning out $\eta_d \hat{\varphi}_d$ *with positive probability. The possible cases are as follows:*

1) The only party that concedes the election with positive probability is the disadvantaged party which does so by bidding $\eta_d \underline{y}_d$ *with probability* $\phi_d(\eta_d \underline{y}_d) = 1 - D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d)/v_{-d} + D_{-d}((\eta_L/\eta_{-d})\underline{y}_L)/v_{-d}$

2) The only time an advantaged party turns out only its committed voters with positive probability is if it is also the most committed party in which case this probability is $\phi_{-d}(\eta_{-d}\overline{y}_{-d}) = D_d((\eta_{-d}/\eta_d)\underline{y}_{-d})/v_d$

3) The advantaged party takes the election by turning out $\eta_d \hat{\varphi}_d$ *with positive probability only if* $\hat{\varphi}_S = \overline{y}_S$ *in which case this probability is* $\phi_{-d}(\eta_S \overline{y}_S) = 1 - D_{-d}((\eta_S/\eta_L)\overline{y}_S)/v_S$*. This is the only case in which the tie-breaking rules matters: when both parties bid* $\eta_S \overline{y}_S$ *the large party must win with probability* 1.

Proof. No party will never submit a bid $\eta_k \varphi_k$ for which $\eta_k \underline{y}_k < \eta_k \varphi_k < \eta_{-k} \underline{y}_{-k}$ since such a bid will be costly but losing, and neither party will submit a bid for which $\eta_k \varphi_k > \eta_k \hat{\varphi}_k$ since to do so would cost more than the value of the prize. Since $D_k(\underline{y}_k) = 0 < v_k$, then it follows that bids must either be $\max_k \eta_k \underline{y}_k$ or in the range $[\max_k \eta_k \underline{y}_k, \eta_d \hat{\varphi}_d]$. If $v_k \leq D\left(\underline{y}_{-k}\eta_{-k}/\eta_k\right)$, it follows that $\eta_k \hat{\varphi}_k \leq \eta_{-k}\underline{y}_{-k}$. In this case party k will only mobilize committed voters, that is will bid $\eta_k \underline{y}_k$, and the other party can win with probability 1 by bidding $\eta_{-k}\underline{y}_{-k}$. Consider now the case of $v_k > D\left(\underline{y}_{-k}\eta_{-k}/\eta_k\right)$ for both parties. In the range $(\max_k \eta_k \underline{y}_k, \eta_d \hat{\varphi}_d)$ there can be no atoms by the usual argument for all-pay auctions: if there was an atom at $\eta_k \varphi_k$ then party $-k$ would prefer to bid a bit more than $\eta_k \varphi_k$ rather than a bit less, and since consequently there are no bids immediately below $\eta_k \varphi_k$ party k would prefer to choose the atom at a lower bid. This also implies that party k with the least committed voters cannot have an atom at $\eta_{-k}\underline{y}_{-k}$: if $-k$ has an atom there there then k should increase its atom slightly to break the tie. If the $-k$ does not have an atom there then k should shift its atom to $\eta_k \underline{y}_k$ since it does not win either way.

Next we observe than in $(\max_k \eta_k \underline{y}_k, \eta_d \hat{\varphi}_d)$ there can be no open interval with zero probability. If party k has such an interval, then party $-k$ will not submit bids in that interval since the cost of the bid is strictly increasing it would do strictly better to bid at the bottom of the interval. Hence there would have to be an interval in which neither party submits bids. But then, for the same reason, it would be strictly better to lower the bid for bids slightly above the interval.

Let u_k be the equilibrium expected utility of party k. In equilibrium the disadvantaged party must earn zero since it must make bids with positive probability arbitrarily close to $\eta_d \hat{\varphi}_d$, while the advantaged party gets at least $u_{-d} \geq v_{-d} - D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d) > 0$ since by bidding slightly more than $\eta_d\hat{\varphi}_d$ it can win for sure, but gets no more than that since it must make bids with positive probability arbitrarily close to $\eta_d\hat{\varphi}_d$. We conclude that the equilibrium payoff of the advantaged party must be exactly $u_{-d} = v_{-d} - D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d)$.

From the absence of zero probability open intervals in $(\max_k \eta_k \underline{y}_k, \eta_d\hat{\varphi}_d)$ it follows that the indifference condition for the advantaged party

$$v_{-d}F_d(\eta_d\varphi_d) - D_{-d}((\eta_d/\eta_{-d})\varphi_d) = v_{-d} - D_d((\eta_d/\eta_{-d})\hat{\varphi}_d)$$

must hold for at least a dense subset. For the disadvantaged party we have

$$v_d F_{-d}(\eta_{-d}\varphi_{-d}) - D_d((\eta_{-d}/\eta_d)\varphi_{-d}) = 0$$

for at least a dense subset. This uniquely defines the cdf for each party in that range:

$$F_d(\eta_d\varphi_d) = 1 - \frac{D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d) - D_{-d}((\eta_d/\eta_{-d})\varphi_d)}{v_{-d}}$$

for $\eta_d\varphi_d \in (\max_k \eta_k \underline{y}_k, \eta_d\hat{\varphi}_d)$, and

$$F_{-d}(\eta_{-d}\varphi_{-d}) = \frac{D_d((\eta_{-d}/\eta_d)\varphi_{-d})}{v_d}$$

for $\eta_{-d}\varphi_{-d} \in (\max_k \eta_k \underline{y}_k, \eta_d\hat{\varphi}_d)$. As these are differentiable they can be represented by continous density functions which are found by taking the derivative. Evaluating $F_d(\eta_d\varphi_d)$ at $\max_k \eta_k \underline{y}_k$ gives $\phi_d(\eta_d \underline{y}_d) = 1 - D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d)/v_{-d} + D_{-d}(\max_k \eta_k \underline{y}_k/\eta_{-d})/v_{-d}$. Note that $F_d(\max_k \eta_k \underline{y}_k)$ is always strictly positive and may or may not be smaller than 1. To see this, notice that $\eta_{-d}\hat{\varphi}_{-d} > \eta_d\hat{\varphi}_d$ implies $D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d)/v_{-d} < D_{-d}(\hat{\varphi}_{-d})/v_{-d} \leq 1$.

Since the disadvantaged party has an atom at $\max_k \eta_k \underline{y}_k$ if an only if $\max_k \eta_k \underline{y}_k = \eta_d\hat{\varphi}_d$ we see that the disadvantaged party has an atom at $\eta_d \underline{y}_d$ with probability $\phi_d(\eta_d \underline{y}_d) = 1 - D((\eta_d/\eta_{-d})\hat{\varphi}_d)/v_{-d} + D(\max_k \eta_k \underline{y}_k/\eta_{-d})/v_{-d}$ and no other atom.

As for the advantaged party, if -$d = S$ then $\eta_L \overline{y}_L > \eta_S \overline{y}_S \geq \eta_S \hat{\varphi}_S > \eta_L \hat{\varphi}_L$ implies that $F_S(\eta_L\hat{\varphi}_L) = D_L(\hat{\varphi}_L)/v_L = 1$. If instead $-d = L$ then $F_L(\eta_S\hat{\varphi}_S) = D_S(\hat{\varphi}_S)/v_S$. If $\hat{\varphi}_S \leq \overline{y}_S$then this is 1 and there is no atom, otherwise there must be an atom of $1 - D_S(\overline{y}_S)/v_S$.

Turning to $\max_k \eta_k \underline{y}_k$ we see that the atom there is given by $D_d((\eta_{-d}/\eta_d)(\eta_\ell/\eta_{-d})\underline{y}_\ell)/v_d = D_d((\eta_\ell/\eta_d)\underline{y}_\ell)/v_d$. If $\ell = d$ this is $D_d(\underline{y}_d)/v_d = 0$ if $\ell = -d$ this is

$$D_d((\eta_{-d}/\eta_d)\underline{y}_{-d})/v_d > D_d(\underline{y}_d) = 0.$$

□

We make the following observations summarized in Corollary 1 in the text.

1. Examining the derivation and result it is immediate to see that only the relative sizes of the parties matter.

2. If $v_{-\ell} \leq D((\eta_\ell/\eta_{-\ell})\underline{y}_\ell)$, it follows that $\eta_{-\ell}\hat{\varphi}_{-\ell} \leq \eta_\ell \underline{y}_\ell$ and the party with the least committed voters always concedes the election. In other words if the value of the prize to the party with the least committed voters is small enough then that party is disadvantaged and concedes the election. On the other hand as $v_L \to \infty$ then $\hat{\varphi}_k = \overline{y}_L$ so that the large party is advantaged. The probability that the small party concedes is than $P_S(\eta_S \underline{y}_S) = 1 - D_L((\eta_S/\eta_L)\overline{y}_S)/v_L$. Hence, the probability

that the small party concedes goes to one at a rate that is bounded independently of the value of the prize to the small party. In other words, in a very high value election, the small party turns out only its committed voters and the large party acts preemptively turning as many voters as the small party is capable of turning out.

In a contested election:

If v_{-d} increases then $\hat{\varphi}_d$ does not change.The equilibrium payoff of the advantaged party is $v_{-d} - D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d)$ so increases. However, the equilibrium bidding strategy of the advantaged party and its expected payment do not depend on v_{-d}. Hence, it must be the case that the expected probability of winning of the advantaged party increases with v_{-d}. Moreover, since the bidding strategy of the advantaged party does not change, neither does its turnout. Finally, the density of the disadvantaged party $f_d(\eta_d\varphi_d) = D'_{-d}((\eta_d/\eta_{-d})\varphi_d)/(\eta_{-d}v_{-d})$ falls with the extra weight accumulating at the atom where it concedes the election, clearly lowering the turnout (and providing an alternative argument as to why the probability of the advantaged party winning must increase).

Suppose that v_d decreases. Then $\hat{\varphi}_d$ weakly decreases since $D_{-d}(\hat{\varphi}_d) = v_d$ and D_{-d} is assumed to be increasing and strictly decreases if $\hat{\varphi}_d < \overline{y}_d$. The density of the advantaged party $f_{-d}(\eta_{-d}\varphi_{-d}) = D'_d((\eta_{-d}/\eta_d)\varphi_{-d})/(\eta_d v_d)$ increases by a fixed ratio and the probability that it turns out only its committed voters $\phi(\eta_{-d}\overline{y}_{-d}) = D_d((\eta_{-d}/\eta_d)\underline{y}_{-d})/v_d$ increases by exactly the same ratio. This means that the cdf has shifted to the left reducing turnout.

If $if\hat{\varphi}_d < \overline{y}_d$ then $\hat{\varphi}_d$ strictly decreasing the surplus $v_{-d} - D_{-d}((\eta_d/\eta_{-d})\hat{\varphi}_d)$ of the advantaged party. Moreover, $f_d(\eta_d\varphi_d)$ is unchanged, while the range is strictly smaller, with the extra weight accumulating where the disadvantaged party concedes the election so the turnout of the disadvantaged party declines.

Theorem. *[4] IIf $D(\varphi)$ is convex than the large party is advantaged and has the higher probability of winning the election.*

If $D(\varphi)$ is concave for $\varphi \geq \underline{y}$, $(\eta_S/\eta_L)\overline{y} > \hat{\varphi}_L$ and for some $\underline{y} < \hat{y} < \overline{y}$ with $D(\hat{y}) < V/\eta_L < D(\overline{y})$ we have

$$\frac{2\underline{y}D'(\underline{y})}{\hat{y}^2 - \underline{y}^2} < -\max_{\underline{y} \leq y \leq \hat{y}} D"(y)$$

then the small party is advantaged and has the higher probability of winning the election.

Proof. If $\hat{\varphi}_S < (\eta_L/\eta_S)\underline{y}$ the small party is disadvantaged. Otherwise from

$$D(\hat{\varphi}_S)\eta_S = V$$

we find

$$\frac{\partial(\hat{\varphi}_S\eta_S)}{\partial\eta_S} = -\frac{D(\hat{\varphi}_S)}{D'(\hat{\varphi}_S)} + \hat{\varphi}_S$$

so that

$$\frac{\partial(\hat{\varphi}_S\eta_S)}{\partial\eta_S} = \frac{1}{D'(\hat{\varphi}_S)}\left(\hat{\varphi}_S D'(\hat{\varphi}_S) - D(\hat{\varphi}_S)\right) = \frac{1}{D'(\hat{\varphi}_S)}\left(\underline{y}D'(\underline{y}) + \int_{\underline{y}}^{\hat{\varphi}_S} \varphi D"(\varphi)d\varphi\right).$$

If $D(\phi)$ is convex this is positive. Hence, we see that increasing η_S weakly increases $\hat{\varphi}_S\eta_S$ until we reach $\eta_S = \eta_L$ at which point $\hat{\varphi}_S\eta_S = \hat{\varphi}_L\eta_L$ so we conclude that when we started $\hat{\varphi}_S\eta_S \leq \hat{\varphi}_L\eta_L$. Since by assumption the two are not equal at the starting point the small party is disadvantaged there. In the case where $D(\varphi)$ is concave, let $\underline{D}$" denotes the smallest (largest absolute) value of D"(y) for $\hat{y} \leq y \leq \overline{y}$. Since $D(\hat{y}) < V/\eta_L < D(\overline{y})$ we have $\hat{y} < \hat{\varphi}_L < \overline{y}$ and in the limit case where $\eta_S = \eta_L$ we then have $\hat{\varphi}_S\eta_S = \hat{\varphi}_L\eta_L$. Taking derivatives, we have that

$$\frac{\partial(\hat{\varphi}_S\eta_S)}{\partial\eta_S} = \frac{1}{D'(\hat{\varphi}_S)}\left(\hat{\varphi}_S D'(\hat{\varphi}_S) - D(\hat{\varphi}_S)\right) \leq \underline{y}D'(\underline{y}) + \underline{D}\text{"}\frac{\hat{y}^2 - \underline{y}^2}{2} < 0,$$

where the last inequality follows from our assumption on the curvature of D. Hence reducing η_S by a sufficiently small amount must strictly advantage the small party. Moreover,

$$\frac{\partial\hat{\varphi}_S}{\partial\eta_S} = -\frac{D(\hat{\varphi}_S)}{D'(\hat{\varphi}_S)\eta_S} < 0$$

so as η_S decreases $\hat{\varphi}_S$ increases implying that $\partial(\hat{\varphi}_S\eta_S)/\partial\eta_S$ remains negative, so the small party remains advantaged until it hits the boundary at $\hat{\varphi}_S = \overline{y}$ and then remains advantaged still until $\eta_S\overline{y} = \eta_L\hat{\varphi}_L$. Notice that, for η_S close to η_L, a necessary and sufficient condition for the small party being advantaged is

$$D^{-1}(V/\eta_S)D'(D^{-1}(V/\eta_S)) - V/\eta_S < 0.$$

The probability of winning for party k is $\int_{\eta_S\underline{y}}^{\eta_d\hat{\varphi}_d} F_{-k}(b)F_k(db)$ with $F_k(\eta_k\varphi_k) = 1 - \int_{\eta_k\varphi_k}^{\eta_d\hat{\varphi}_d} F_k(db)$. Observe that the density $f_k(b) = D'(b/\eta_{-k})/V$ is higher for the large party if D is convex and higher for the small party if D is concave. In the former case the large party is advantaged, in the latter case we consider only the case in which the small party is advantaged. In either case we have $f_{-d}(b) > f_d(b)$. Since only the advantaqed party can have an atom at $\eta_d\hat{\varphi}_d$ it follows in both cases that for $\eta_k\varphi_k > \eta_L\underline{y}$ we have $\int_{\eta_k\varphi_k}^{\eta_d\hat{\varphi}_d} F_{-d}(db) > \int_{\eta_k\varphi_k}^{\eta_d\hat{\varphi}_d} F_d(db)$. Hence also for $\eta_k\varphi_k > \eta_L\underline{y}$ it follows $\int_{\eta_k\varphi_k}^{\eta_d\hat{\varphi}_d} F_d(b)F_{-d}(db) > \int_{\eta_k\varphi_k}^{\eta_d\hat{\varphi}_d} F_{-d}(b)F_d(db)$ since each term under the integral on the left is larger than on the right.

It remains to evaluate $\int_{\eta_S\underline{y}}^{\eta_L\underline{y}} F_k(b)F_{-k}(db)$, that is, parties turning out only their committed voters. Since $f_{-d}(b) > f_d(b)$ and only the advantaged party can have an atom at $\eta_d\hat{\varphi}_d$ the disadvantaged party must have a higher probability of turning out only its commited voters. Moreover, when it does so it always loses, so $\int_{\eta_S\underline{y}}^{\eta_L\underline{y}} F_d(b)F_{-d}(db) > \int_{\eta_S\underline{y}}^{\eta_L\underline{y}} F_{-d}(b)F_d(db)$. Adding the two inequalities gives the desired result. □

Appendix II: Conflict Resolution Model

Theorem. *7 An equilibrium of the conflict resolution model exists.*

Proof. This is essentially the theorem of Glicksberg (1952), except that we do not require $D_k(\varphi_k, F_{-k})$ to be linear in F_{-k}. However inspection of Glicksberg's proof shows that only continuity in F_{-k} is needed - Glicksberg uses only the fact that the objective function is weakly concave in F_k so that the best-response correpondence is convex-valued and the fact that it is jointly continuous in F_k, F_{-k} so that it is upper hemi-continous. Weak concavity in F_k follows here as it does in Glicksberg because the objective function is linear in F_k - the linearity of the objective in F_{-k} is used by Glicksberg only to establish continuity in F_k, F_{-k} which we have by assumption. □

Theorem. *8 Suppose that $p_S^N(\eta_S\varphi_S, \eta_L\varphi_L), D_S^N(\varphi_S, F_L), D_L^N(\varphi_L, F_S)$ converges to the all-pay auction $D_S(\varphi_S), D_L(\varphi_L)$, that F_k^N are equilibria of the conflict resolution models and that F_k is the unique equilibrium of the all-pay auction. Then $F_k^N \to F_k$.*

Proof. Since the space of distributions is compact in the given topology there is a convergent subsequence. Hence it is sufficient to assume $F_k^N \to F_k^\infty$ and show that $F_k^\infty = F_k$. We do this by showing that F_k^∞ is an equilibrium relative to the tie-breaking rule that advantaged party wins if there is a tie where the disadvantaged party turns out all voters and the party that can turn out the most committed voters wins when it turns out exactly its committed voters. Since the equilibrium of the all-pay auction is unique, it must then be that F_k^∞ is in fact F_k. Note that it then follows that F_k^∞ is also an equilibrium with respect to only the first half of the tie-breaking rule, since that is the case for F_k.

By assumption for any ϵ^2 the convergence is uniform on the set $|\eta_S\varphi_S - \eta_L\varphi_L| \geq \epsilon^2$. Hence we may assume that for any ϵ and for large enough N if $\eta_S\varphi_S > \eta_L\varphi_L + \epsilon^2$ then $p_S^N(\eta_S\varphi_S, \eta_L\varphi_L) > 1-\epsilon^2$ and $\eta_S\varphi_S < \eta_L\varphi_L - \epsilon^2$ then $p_S^N(\eta_S\varphi_S, \eta_L\varphi_L) < \epsilon^2$.

Let d be the disadvantaged party in the all-pay auction. We observe the obvious fact that F_k^∞ places no weight above $(\eta_d/\eta_k)\hat{\varphi}_d$ nor below $\underline{y}_k$ so we certainly have convergence outside these intervals. It is similarly obvious for any $\gamma > 0$ there is an N large enough that F_k^N places no weight above $(\eta_d/\eta_k)\hat{\varphi}_d + \gamma$ nor below $\underline{y}_k - \gamma$, so in examining F_k^N we may restrict attention to those intervals.

The assumption on the slopes of $D_S^N(\varphi_S, F_L), D_L^N(\varphi_L, F_S)$ implies that there are constants $\infty > \overline{D}, \underline{D} > 0$ such that for for any ϵ and κ and all large enough N for $\varphi_k + \kappa\epsilon > \varphi_k' > \varphi_k$

$$D_k^N(\varphi_k', F_{-k}) - D_k^N(\varphi_k, F_{-k}) < \overline{D}\epsilon$$

and for $\varphi_k \geq \underline{y}_k$ and $\varphi_k' > \varphi_k + \epsilon/\kappa$

$$D_k^N(\varphi_k', F_{-k}) - D_k^N(\varphi_k, F_{-k}) > \underline{D}\epsilon$$

Let ℓ denote the party with the largest value of $\eta_k\underline{y}_k$.

Consider first the intervals $((\eta_\ell/\eta_k)\underline{y}_\ell, (\eta_d/\eta_k)\hat{\varphi}_d)$. Fix a point $\tilde{\varphi}_S = (\eta_L/\eta_S)\tilde{\varphi}_L$ in this interval where there is a tie and consider an ϵ open square around of this point, $\Phi_S \times \Phi_L$ (we may assume that these open intervals are entirely contained in the set in question by choosing ϵ small enough). Choose $\epsilon \max_k v_k < \overline{D}$ at least. Consider that one of the parties k has no greater than a 1/2 chance of winning the contest in this interval. If Π_{-k} is the probability F_{-k}^N assigns to Φ_{-k} then if k shifts any weight in Φ_k to the top of the interval he gains at least $(1/2)\Pi_{-k}v_k - \epsilon^2 v_k - \overline{D}\epsilon$, so that if $\Pi_k > 0$ then $\Pi_{-k} \leq (2/v_k)(\overline{D} - \epsilon v_k)\epsilon < \underline{\Pi}\epsilon$. If $\Pi_k = 0$ certainly $\Pi_k \leq \underline{\Pi}\epsilon$. Hencewe see that there is a constant $\overline{\Pi}$ such that in each square of the type described we must have $\Pi_k \leq \underline{\Pi}\epsilon$ for at least one of the two parties k.

Now consider Π_k for the other party for which this bound is not necessarily satisfied, and consider φ_{-k} lying below Φ_{-k}. Shifting to the top of the interval yields a gain by the previous argument of at least $\Pi_k v_{-k} - \epsilon^2 v_{-k} - \overline{D}(\tilde{\varphi}_{-k} - \varphi_{-k} - \epsilon)$. From this we see that there is another constant $\underline{\kappa} > 0$ such that for $\tilde{\varphi}_{-k} - \epsilon \geq \varphi_{-k} \geq \tilde{\varphi}_{-k} - \epsilon - \underline{\kappa}\Pi_k$ party $-k$ places no weight. If $\underline{\kappa}\Pi_k > \epsilon/\kappa$ then shifting all the weight in Φ_k to $\tilde{\varphi}_{-k} - \epsilon - \underline{\kappa}\Pi_k$ causes k to gain $\underline{D}\kappa\Pi_k - (\underline{\Pi}\epsilon + \epsilon^2)v_k \leq 0$. Hence there is a constant $\overline{\Pi}$ such that $\Pi_k \leq \underline{\Pi}\epsilon$ for both parties k.

If $\hat{\varphi}_d < 1 (= \overline{y}_d)$ then exactly the same argument works when we extend the upper limit slightly $((\eta_L/\eta_k)\underline{y}_L, (\eta_d/\eta_k)\hat{\varphi}_d) + \gamma)$, and we already know it is true for $((\eta_d/\eta_k)\hat{\varphi}_d) + \gamma, \eta_d/\eta_k]$ so the bound holds for $((\eta_\ell/\eta_k)\underline{y}_\ell, \eta_d/\eta_k]$.

Suppose instead that one party is turning out all voters at the upper bound: in this case it must be disadvantaged and the other party must be larger. We observe that since for any γ we already know that F_{-d}^{N} places no weight above $(\eta_d/\eta_{-d})+\gamma$ so $F_{-d}^{-\infty}$ places no weight above η_d/η_{-d}. In the intervals $((\eta_d/\eta_k)-\epsilon, \eta_d/\eta_k]$ if W_{-d} is the probability that $-d$ wins conditional on that interval then the gain to $-d$ by shifting to $(\eta_d/\eta_k)+\epsilon$ is $(1-W_{-d})v_{-d}-2\epsilon\overline{D}$. There are two possibilities: either $\Pi_{-d}=0$ in which case we have $\Pi_k \le \underline{\Pi}\epsilon$ for both parties in $((\eta_\ell/\eta_k)\underline{y}_\ell, 1]$, or $W_{-d} \ge 1-\underline{W}\epsilon$.

Now consider the lower bound $(\eta_\ell/\eta_k)\underline{y}_\ell$. The argument above that one party has to satisfy $\Pi_k \le \underline{\Pi}\epsilon$ remain valid since it relies on deviating to the top of the interval and the upper bound on the cost derivative $\overline{D}$ which is globally valid. The party that with the most committed voters does not bid below $(\eta_\ell/\eta_k)\underline{y}_\ell-\gamma$ so if it is the party that satisfies $\Pi_k \le \underline{\Pi}\epsilon$ then the other party has a chance of winning by bidding below $(\eta_\ell/\eta_k)\underline{y}_\ell$ of at most $2\underline{\Pi}\gamma$, while if it were to bid $\underline{y}_S$ it would save nearly $\underline{D}((\eta_\ell/\eta_k)\underline{y}_\ell-\underline{y}_S)$ so for small enough γ it would not choose to bid in this interval. Hence we conclude that the party with the least committed voters must satisfy the bound $\Pi_k \le \underline{\Pi}\epsilon$ in $(\underline{y}_\ell, (\eta_d/\eta_k)\hat{\varphi}_d)$.

Now we are in a position to consider the sequence of equilibrium expected utilities U_k^N which, if necessary by passing to a subsequence may be assumed to converge to some U_k^∞. For any ϵ we observe that in the compact region in which the difference between bids is at least ϵ the objective function in the limit is continuous, so in this region the integral defining expected utility converges to the identical value as computed from the limit distributions. In case $\hat{\varphi}_d < 1(=\overline{y}_d)$ we also see that the region where the difference between bids is smaller than ϵ the probability of that region is at most $C\epsilon^2$ for all large enough N so that this does not matter for computing utility in the limit. Hence in this case $U_k^\infty = U_k$ the utility computed from the limit distributions (and in this case the tie-breaking rule does not matter). In case $\hat{\varphi}_d = 1(=\overline{y}_d)$ if we exclude a neighborhood of the tie at η_d/η_k again utility converges to the right limit, moreover, we have shown that in the region near the tie in equilibrium (for $N<\infty$) $W_{-d} \ge 1-\underline{W}\epsilon$ which gives the same result in the limit as $N\to\infty$ as the tie-breaking rule that $-d$ always wins the tie.

Now consider deviations against the limit distributions. For deviations to a point where the bound $\Pi_k \le \underline{\Pi}\epsilon$ is satisfied by the opponent a strict gain with respect to the limit distribution of the opponent translates immediately in the usual way to a strict gain for large N so this is impossible. The same reasoning applies in case $\hat{\varphi}_d = 1(=\overline{y}_d)$ to deviations by the advantaged party to η_S/η_L since it must win before the limit is reached with probability at least $W_{-d} \ge 1-\underline{W}\epsilon$.

Finally, in the case $\hat{\varphi}_d = 1(=\overline{y}_d)$ if is profitable for the disadvantaged party to deviate to 1 since by the tie-breaking rule it loses for sure it could equally well make a strict profit by bidding slightly less than 1. Nor can it be advantageous for the small party to deviate to $(\eta_L/\eta_S)\underline{y}_L$ since by the tie-breaking rule it loses for sure.

Hence we conclude that F_k^∞ is in fact an equilibrium with respect to the proposed tie-breaking rule. □

Web Appendix: The Uniform Case

Suppose that for $\varphi \ge \underline{y}$ cost is given by $c(\varphi)=\alpha(\varphi-\underline{y})^{\alpha-1}$ for some $\alpha>1$, or equivalently that $C(\varphi)=(\varphi-\underline{y})^\alpha$. Then

$$D(\varphi)=(\varphi-\underline{y})^\alpha+\alpha\theta(1-\varphi)(\varphi-\underline{y})^{\alpha-1}$$

$$D(\varphi)=(\varphi-\underline{y})^\alpha+\alpha\theta(1-\underline{y})(\varphi-\underline{y})^{\alpha-1}-\alpha\theta(\varphi-\underline{y})(\varphi-\underline{y})^{\alpha-1}$$

$$D(\varphi) = (1-\alpha\theta)(\varphi - \underline{y})^{\alpha} + \alpha\theta(1-\underline{y})(\varphi - \underline{y})^{\alpha-1}$$

It follows that

$$D'(\varphi) = \alpha(1-\alpha\theta)(\varphi - \underline{y})^{\alpha-1} + \alpha(\alpha-1)\theta(1-\underline{y})(\varphi - \underline{y})^{\alpha-2}$$

which is smallest when $\phi - \underline{y}$ is biggest so bounded below by

$$D'(\varphi) \geq \alpha\left((1-\alpha\theta) + (\alpha-1)\theta\right)$$

$$D'(\varphi) \geq \alpha\left(1-\theta\right).$$

Next we consider equilibrium in the limiting case of $\alpha \to 1$ where let $c(1) \leq \bar{P}$,so that $\bar{y} = 1$ and $\theta < 1$. In this case for $\varphi > \underline{y}$ we have

$$D(\varphi) = (\varphi - \underline{y}) + \theta(1-\varphi) = \theta(1-\underline{y}) + (1-\theta)(\varphi - \underline{y}) = D(\varphi) = (\theta - \underline{y}) + (1-\theta)\varphi.$$

We consider only the case in which $\hat{\varphi}_S > (\eta_L/\eta_S)\underline{y}$ so that we can have contested elections in equilibrium. In the limit as $\alpha \to 1$ the limiting value of $\hat{\varphi}_k$ is derived from $D(\hat{\varphi}_k) = v_k$ so we may find from $(\theta - \underline{y}) + (1-\theta)\hat{\varphi}_k = v_k$ that

$$\hat{\varphi}_k = \frac{\underline{y} - \theta + v_k}{1-\theta}.$$

This satisfies $\hat{\varphi}_k < \bar{y} = 1$ if and only if $1 - \underline{y} \geq v_k$, so either this is the case or $\hat{\varphi}_k = 1$. If the small party is disadvantaged and cannot mobilize all voters in the party, that is $\hat{\varphi}_S < \min\{1, (\eta_L/\eta_S)\hat{\varphi}_L\}$, we know from Theorem 2 that in equilibrium the small party has an atom at $\eta_S\underline{y}$ of limit size

$$\lim{}_{\alpha\to 1} 1 - \frac{D((\eta_S/\eta_L)\hat{\varphi}_S)}{v_L} = \frac{\eta_L v_L - \eta_S v_S + (\eta_L - \eta_S)(\underline{y} - \theta)}{\eta_L v_L} < 1.$$

Furthermore, in the limit, for $\eta_k\varphi_k \in (\eta_L\underline{y}, \eta_d\hat{\varphi}_d)$ the mixed strategies of the players have no atom and are described by continuous densities which approach the uniform distribution:

$$\lim{}_{\alpha\to 1} f_k(\eta_k\varphi_k) = \frac{1-\theta}{\eta_{-k}v_{-k}}.$$

If the small party is advantaged, the discontinuity of $D(\varphi)$ at $\eta_L\underline{y}$ is reflected through an additional atom for the small party at $\eta_L\underline{y}$ of size

$$\phi_S(\eta_L\underline{y}) = \lim_{\eta_S\varphi_S\to\eta_L\underline{y}} \lim_{\alpha\to 1} F_S(\eta_S\varphi_S) = \lim_{\eta_S\varphi_S\to\eta_L\underline{y}} \lim_{\alpha\to 1} \frac{D((\eta_S/\eta_L)\varphi_S)}{v_L} =$$

$$\lim_{\eta_S\varphi_S\to\eta_L\underline{y}} \lim_{\alpha\to 1} \frac{(1-\alpha\theta)((\eta_S/\eta_L)\varphi_S - \underline{y})^{\alpha} + \alpha\theta(1-\underline{y})((\eta_S/\eta_L)\varphi_S - \underline{y})^{\alpha-1}}{v_L} =$$

$$\lim_{\eta_S\varphi_S\to\eta_L\underline{y}} \frac{(1-\theta)((\eta_S/\eta_L)\varphi_S - \underline{y}) + \theta(1-\underline{y})}{v_L} = \frac{\theta(1-\underline{y})}{v_L}$$

with the small advantaged party always winning the tie.

Observe that in the case of a large advantaged party, having a greater chance of winning a contested election means a greater chance of winning the election. In the case of a small advantaged party we must consider the extra atom of the small advantaged party at $\eta_L \underline{y}$. Call such an election strictly contested if it is contested and the small party bids strictly above $\eta_L \underline{y}$. If the small party has a greater chance of winning a strictly contested election, then the atom at $\eta_L \underline{y}$ for the small party must have lower probability than for the large party (since a small party cannot have an atom at the top these are the only atoms). Hence the overall probability of the small party winning must be greater. Hence the result that the advantaged party wins the election more than half the time when $\eta_{-d} v_{-d} / \eta_d v_d \geq 1$ follows from

$$\lim{}_{\alpha \to 1} f_k(\eta_k \varphi_k) = \frac{1-\theta}{\eta_{-k} v_{-k}},$$

since this implies that the advantaged party wins at least half the time the contested elections if it is the large party, and it wins at least half the time the strictly contested if it is the small party.

Finally we show that when $\theta = 0$, $\eta_L = 2.5 \eta_S$, $v_L = 2/9$, $v_S = 6/9$ and $\underline{y} = 1/9$ the large party is advantaged yet the small party wins the election more than half the time. Note that in this case $\eta_L v_L / \eta_S v_S \equiv (2.5 * 2/9)/(6/9) = 5/6$. That is, the condition that $\eta_{-d} v_{-d} / \eta_d v_d \geq 1$ fails as we know it must. Despite the fact that the large party attaches a lower value to the object than the small party it is never-the-less advantaged. First we recall from above that $\hat{\varphi}_k < 1$ if and only if $1 - \underline{y} \geq v_k$. For the small party $6/9 = v_S \leq 1 - \underline{y} = 8/9$ and for the large party $2/9 = v_L \leq 1 - \underline{y} = 8/9$ so for both partys $\hat{\varphi}_k < 1$. Using

$$\hat{\varphi}_k = \frac{\underline{y} - \theta + v_k}{1 - \theta}$$

we then compute for the small party $\hat{\varphi}_S = (1/(1-\theta))(\underline{y} - \theta + v_S) = 7/9$, while for the large party $\hat{\varphi}_L = (1/(1-\theta))(\underline{y} - \theta + v_L) = 3/9$. Hence, since $\eta_S \hat{\varphi}_S = \eta_S(7/9) = (\eta_L/2.5)(7/9) < \eta_L(3/9) = \eta_L \hat{\varphi}_L$, indeed the small party is disadvantaged. Notice also that $\eta_L \underline{y} = \eta_S(5/18) < \eta_S \hat{\varphi}_S = \eta_S(14/18)$ so that the small party does always not concede. Now we compute the probability the small party wins the election. It is

$$\Pi_S = \int_{\eta_L \underline{y}}^{\eta_S \hat{\varphi}_S} f_S(\eta\varphi) F_L(\eta\varphi) d(\eta\varphi) = \frac{1-\theta}{\eta_L v_L} \int_{\eta_L \underline{y}}^{\eta_S \hat{\varphi}_S} F_L(\eta\varphi) d(\eta\varphi) =$$

$$\frac{1-\theta}{\eta_L v_L} \left(\eta_S \hat{\varphi}_S - \eta_L \underline{y}\right) \left[F_L(\eta_L \underline{y}) + \frac{1-\theta}{2\eta_S v_S} \left(\eta_S \hat{\varphi}_S - \eta_L \underline{y}\right) \right]$$

Since $F_L(\eta_L \underline{y}) + (1-\theta)\left(\eta_S \hat{\varphi}_S - \eta_L \underline{y}\right)/\eta_S v_S = 1$ we have

$$\Pi_S = \frac{1-\theta}{\eta_L v_L} \left(\eta_S \hat{\varphi}_S - \eta_L \underline{y}\right) \left[F_L(\eta_L \underline{y}) + \left(1 - F_L(\eta_L \underline{y})\right) \right] =$$

$$\frac{1-\theta}{2\eta_L v_L}\left(\eta_S\hat{\varphi}_S-\eta_L\underline{y}\right)\left(1+F_L(\eta_L\underline{y})\right)=\frac{\eta_S v_S}{2\eta_L v_L}\left(1-F_L(\eta_L\underline{y})\right)\left(1+F_L(\eta_L\underline{y})\right)=$$
$$\frac{\eta_S v_S}{2\eta_L v_L}\left(1-F_L(\eta_L\underline{y})^2\right).$$

From the proof of Theorem 2 we also have

$$F_L(\eta_L\varphi_L)=\frac{D((\eta_L/\eta_S)\underline{y})}{v_S}=\frac{(1-\theta)\eta_L\underline{y}-\eta_S\left(\underline{y}-\theta\right)}{\eta_S v_S}$$

so that for the given parameters impying $\eta_L\underline{y}=\eta_S(5/18)$

$$F_L(\eta_L\underline{y})=\frac{(5/18)-\underline{y}}{v_S}=\frac{(5/18)-(2/18)}{(12/18)}=1/4$$

Hence $\Pi_S=(3/5)(15/16)=9/16>1/2$.

Theorem. *6 If the large party is advantaged, mandatory voting enhances parties competition. If the small party is advantaged, mandatory voting enhances parties competition if and only if $\eta_L/\eta_S<(1+\underline{y}+v_S)/(1+\underline{y}+v_L)$.*

Proof. In the uniform case we have that $D(\varphi)=\varphi(1+f)-\underline{y}-f$ and

$$\hat{\varphi}_k=\frac{\underline{y}+f+v_k}{1-f}.$$

Hence

$$\frac{\partial\hat{\varphi}_S}{\partial f}>\frac{\partial\hat{\varphi}_L}{\partial f}$$

if and only if

$$\frac{\eta_L}{\eta_S}>\frac{1+\underline{y}+v_S}{1+\underline{y}+v_L},$$

and $\eta_S\hat{\varphi}_S>\eta_L\hat{\varphi}_L$ if and only if

$$\frac{f+\underline{y}+v_S}{f+\underline{y}+v_L}>\frac{\eta_L}{\eta_S}.$$

If the large party is advantaged and $v_L>v_S$ it must be that

$$\frac{\eta_L}{\eta_S}>1>\frac{1+\underline{y}+v_S}{1+\underline{y}+v_L}.$$

If the large party is advantaged and $v_L<v_S$, then

$$\frac{\eta_L}{\eta_S}>\frac{f+\underline{y}+v_S}{f+\underline{y}+v_L}>\frac{1+\underline{y}+v_S}{1+\underline{y}+v_L},$$

since $f<1$and $(f+\underline{y}+v_S)/(f+\underline{y}+v_L)$ is decreasing in f if and only if $v_L<v_S$. On the other hand, tha fact that the small party is advantaged does not necessary imply that

$$\frac{\partial\hat{\varphi}_L}{\partial f}>\frac{\partial\hat{\varphi}_S}{\partial f}$$

which instead follows from $\eta_L/\eta_S<(1+\underline{y}+v_S)/(1+\underline{y}+v_L)$. □