Voter Turnout with Peer Punishment☆

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

Abstract

We introduce a model of turnout where social norms, strategically chosen by competing political parties, determine voters' turnout. Social norms must be enforced through costly peer monitoring and punishment. When the cost of enforcement of social norms is low, the larger party is always advantaged. Otherwise, in the spirit of Olson (1965), the smaller party may be advantaged. Our model shares features of the "ethical" voter model and it delivers novel and empirically relevant comparative statics results.

JEL Classification Numbers: later

Keywords: Voting, Turnout, Party, Peer Punishment

☆First Version: June 2, 2013. Previously titled "Voter Participation with Collusive Parties." We are especially grateful to Juan Block, Michele Boldrin, Rohan Dutta, Tim Feddersen, Andrea Galeotti, Piero Gottardi, Johannes Hörner, Philipp Kircher, Cesar Martinelli, Salvatore Modica, Massimo Morelli, Salvatore Nunnari, Pietro Ortoleva, Thomas Palfrey, Kirill Pogorelskiy, John Roemer, Guido Tabellini, Luigi Tenco, Francesco Trebbi, participants at the 2015 Priorat Workshop, 2015 Barcelona JOCs, 2016 Cornell Political Economy Conference, 2016 ASSET, the UCL Theory Workshop, the UECE Meetings in Lisbon, the Dondena Political Economy Seminar in Bocconi, the Yale theory seminar, the MEDS theory seminar, the editor and four referees. Giovanni Andreottola provided excellent research assistance.

[1]Department of Economics, EUI and WUSTL, david@dklevine.com

[2]Department of Economics, EUI and MOVE, andrea.mattozzi@eui.eu

1. Introduction

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

Social norms and peer pressure play a key role in voter participation. To take a few of many pieces of evidence, Della Vigna et al (2014) demonstrate that an important incentive for citizens to vote is to show others that they have voted. Gerber, Green and Larimer (2008) show that social pressure significantly increase turnout. Amat et al (2018) using historical elections data of Spain's Second Republic show that turnout was driven by political parties and trade unions' social pressure.[3]

In this paper we introduce a model where social norms determine voters' turnout and investigate the theoretical relation between party size and electoral advantage. We build on the ethical voter model described in Feddersen and Sandroni (2006) and Coate and Conlin (2004). In that model, individuals adopt a social norm in the form of a participation rate for its members that *ex ante* their party would most prefer. We instead assume that the social norm is chosen collectively and must be enforced through costly peer monitoring and punishment. We show that monitoring costs play a key role in determining the outcome of elections: low monitoring costs - zero in the existing literature - favor the large party. By contrast, in intermediate stakes elections with few committed voters (those with a negative cost of voting), high monitoring costs favor the small party.

Typically social norms are maintained by various forms of social disapproval and ostracism (Ostrom (1990)). While the news article mentioned in the incipit is clearly an extreme case of punishment, a less traumatic example is represented by Ted Cruz's campaign strategy in the 2016 Iowa Presidential primaries. Voters who were most likely to support Cruz received mailings with information about their own past voting behavior and that of their neighbors and were encouraged

[3]The relation between voter participation and peer pressure is also widely discussed in the sociology literature, see for example Coleman (1988).

to caucus.[4]

That social norms are chosen collectively to maximize a group objective is also founded in existing research: Coleman (1988) and Ostrom (1990) as well as Olson (1965) all provide evidence that - within the limits of available monitoring and punishment - peer pressure mechanisms do a good job in solving public goods problems. Palfrey and Pogorelskiy (2016) provide experimental evidence that communication among voters and in particular communication within parties increases turnout - that is, enables them to attain more advantageous social norms.

To capture these stylized facts our model supposes that individual party members, given the social norm, optimally choose whether or not to vote. Voters imperfectly monitor each others voting behavior within each party and punish deviators. The total cost to parties of choosing an incentive compatible social norm is the sum of the participation cost of voting born by voters, and the monitoring costs, which is the expected cost of punishing party members who did not vote. Each party decides as a collective to implement the social norm that is to their greatest advantage. Because both parties incur the cost of turning out voters but only the one with larger turnout wins, the resulting game is an all-pay auction.

In the existing turnout models - both ethical and pivotal voter models - a large party has a natural advantage. Since the marginal cost of turning out an additional voter is increasing in the fraction of party members who turned out, a large party can turn out the same number of voters as a small party at lower cost. In our model this is indeed the case when monitoring is perfect and there are no equilibrium punishment costs. When monitoring costs are important, however, for any given number of voters turned out, a small party must punish fewer non-voters. Since monitoring costs favor the small party, the small party may have an advantage over the large party: a possibility not present in existing

[4]The mailing contained the following statement: "You are receiving this election notice because of low expected voter turnout in your area. Your individual voting history as well as your neighbors' are public record. Their scores are published below, and many of them will see your score as well. Caucus on Monday to improve your score and please encourage your neighbors to caucus as well."

turnout models. We investigate this possibility and show that when monitoring costs are high, committed voters are few, and the value of winning election is intermediate, the small party has indeed an advantage.

Our model delivers a rich set of intuitive comparative statics on equilibrium turnout with respect to the value of the election, its closeness, and to changes in the relative size of the groups. We clarify which results carry over from standard models and which do not. Furthermore, with our model we are able to address questions that could not be asked with existing turnout models. In particular, we examine the consequences of changes in monitoring costs. We observe, for example, that rule changes, which lower turnout costs, may also raise monitoring costs and so have the perverse effect of lowering turnout. This prediction matches with empirical evidence on the effect of postal voting on turnout in Swiss elections (Funk (2010)) and on the effect of the monitoring capacity of parties' local mobilizers on turnout buying in Mexico (Larreguy, Marshal and Querubin (2016)). Our predictions are also consistent with the failure of the 2000 UK policy experiment of setting in-store poll booths to reduce voting costs and hence increase turnout.[5]

The rest of the paper is organized as follows. In Section 2 we present our model. In Section 3, we characterize the equilibrium and provide sufficient conditions for a party to be advantaged. We discuss the comparative statics of the model in Section 4. We review our assumptions, the robustness of our results and how they relate to the literature in Section 5. Section 6 concludes.

2. The Model

A continuum of voters is divided into two parties $k = \{S, L\}$ denoting *Small* and *Large*, respectively. The fraction of the voting population belonging to party k is η_k where $0 < \eta_S < \eta_L$ and $\eta_S + \eta_L = 1$. The two parties compete in an election for a common prize worth V to the party that produces the greatest number of votes and V/η_k to each member of party k.[6] We assume, in other

[5] See https://www.theguardian.com/uk/2000/aug/17/juliahartleybrewer.

[6] A party is generally made up of individuals with different interests: we do not attempt to model the internal decision making of the party, but simply note that in practice V represents

words, that the prize is fungible - such as taxes, subsidies, or government jobs - so that the collective value of the prize to a party does not depend on the size of the party.[7] Parties - either by consensus or directed by leaders - move first, and simultaneously and non-cooperatively choose a social norm in the form of a participation rate for their members. The individual party members move second and, given the social norm, optimally choose whether or not to vote in an election.

Participation costs are borne by individual voters who are assumed to face the same distribution of voting costs independent of party. Each identical voter privately and independently draws a type y from a uniform distribution on $[0, 1]$. This type determines a *net participation cost* of voting $c(y)$ and, based on this, the voter decides whether or not to vote. This cost of voting consists of the direct cost and inconvenience (costs of time and transportion) minus the direct personal benefits such as fulfilling civic duty, the camaraderie of the polling place or expressive voting. The participation cost of voting c is continuously differentiable, strictly increasing and satisfies $c(\underline{y}) = 0$ for some $\underline{y} \in [0, 1]$. Voters for whom $y < \underline{y}$, those with a negative net cost of voting, are called *committed voters* and will always vote.

The social norm of the party is a threshold φ_k together with a rule for party members prescribing voting if $y \leq \varphi_k$. Hence φ_k is the probability that a representative party member votes, and since there are a continuum of voters is also the turnout rate of the party. This social norm is enforced through peer auditing and the possibility of imposing punishments on party members.

Each member of the party is audited by another party member. The auditor recieves a binary signal of whether or not the auditee followed the social norm. If the auditee voted then the auditor receives a positive signal. If the auditee did not vote and the auditee violated the social norm, i.e. $y < \varphi_k$, the auditor receives a negative signal with probability $1 > \pi_0 > 0$. If the auditee did not vote and the auditee did not violate the social norm, i.e. $y > \varphi_k$, the auditor receives a negative signal with probability $0 \leq \pi_1 \leq \pi_0$. Upon receiving a

a composite of those interests.

[7]We discuss this assumption in Section 5.

negative signal, the auditor punishes the auditee with a loss of utility $P_k \geq 0$.

The punishment level is chosen by the party and must be chosen to be *incentive compatible,* that is, given the punishment level it must be individually optimal to follow the social norm. Hence, incentive compatibility requires $\pi_0 P_k = c(\varphi_k)$. Any member with $y \leq \varphi_k$ would be willing to pay the participation cost $c(y)$ of voting rather than face the expected punishment $\pi_0 P_k$, while any member with $y > \varphi_k$ prefers to pay the expected cost of punishment $\pi_1 P_k$ over the participation cost of voting $c(y)$. The punishment itself, as it is paid by a member, is a cost to the party and we assume that the overall cost of a punishment P_k to the party is exactly P_k.

We are now ready to determine the *expected cost* $C(\varphi_k)$ to the party of imposing a social norm φ_k on a voter. By convention $C(\varphi_k) = 0$ for $\varphi_k \leq \underline{y}$ since the committed voters always vote. We can decompose the expected cost $C(\varphi_k)$ into two additive components. The first component is the *turnout cost* $T(\varphi_k) = \int_{\underline{y}}^{\varphi_k} c(y)dy$, which is the expected direct cost of participation when the member votes.[8] The second component is the *monitoring cost* $M(\varphi_k) = \int_{\varphi_k}^{1} \pi_1 P_k dy$, which is the expected cost of punishment when the member does not vote. Defining the *monitoring difficulty* as $\theta = \pi_1/\pi_0 \in [0, 1]$, we may rewrite the monitoring cost using the incentive constraint as $M(\varphi_k) = \int_{\varphi_k}^{1} \theta c(\varphi_k)dy = \theta(1 - \varphi_k)c(\varphi_k)$. Note that for $\varphi_k > \underline{y}$, total expected cost is increasing since $C'(\varphi_k) = (1 - \theta)c(\varphi_k) + \theta c'(\varphi_k) > 0$.

The outcome of the election is determined by the fraction of the electorate $b_k = \eta_k \varphi_k$ that each party turns out, and sometimes we will refer to this as the *bid* of party k. The party that turns out more of its members wins. In case of a tie we assume that the large party wins.[9] This is similar to an all-pay auction:

[8] Coate and Conlin (2004) assume that $c(y)$ is linear so that the turnout cost of voting for $y \geq \underline{y}$ is quadratic.

[9] If we assume that in case of a tie each party has an equal chance of winning an equilibrium may fail to exist for an uninteresting reason. As we shall see there can be equilibria in which the large party with positive probability bids preemptively - and this is the only case in which the tie-breaking rule matters. The large party bids preemptively when it mobilizes more voters than there are in the small party. However, it would always benefit from mobilizing slightly fewer voters. Hence we must allow the large party to bid preemptively by mobilizing the exact number of voters in the small party - meaning if there is a tie it must win. An

the party that "submits the highest bid" wins, but each party pays the cost for their bid. In terms of bids, if a party bids only its committed voters it bids $\underline{b}_k \equiv \eta_k \underline{y}$ and if it bids all its voters it bids η_k.

A strategy for party k is a probability measure represented by a cumulative distribution function F_k over bids, that is, on $[\underline{b}_k, \eta_k]$. We take the objective function of each party to be the total utility of members $\Pi_k(b_k, F_{-k})V - \eta_k C(b_k/\eta_k)$ where $\Pi_k(b_k, F_{-k})$ is the probability that a bid b_k wins. An *equilibrium* consists of strategies for both parties such that the strategy of each party is optimal given the other.

3. Main Results

What is gained by adding monitoring to an otherwise standard group-turnout model? A large group has a natural advantage since, for a given social norm, it can turn more voters out. In particular, to mobilize a fraction b of the electorate, the large party can choose a smaller social norm $\varphi_L = b/\eta_L$ than the small party $\varphi_S = b/\eta_S$. Hence, focusing only on turnout costs, the large party has a cost advantage in turning out voters since the marginal cost of turnout $T'(\varphi) = c(\varphi)$ is increasing. By contrast, the large party faces a disadvantage in monitoring cost since it will need to monitor and punish a greater proportion $(1 - \varphi_L)$ of non-voters.

The total expected cost of mobilizing a fraction b of the electorate is the sum of turnout and monitoring costs and equals

$$\eta_k C(b/\eta_k) = b\frac{C(b/\eta_k)}{(b/\eta_k)} = bAC(b/\eta_k)$$

where $AC(\varphi_k)$ denotes the average cost. We see immediately that the party with the lower average cost will have a cost advantage. If $AC(\varphi_k)$ is declining in φ this will be the small party while if $AC(\varphi_k)$ is increasing it will be the large party. Alternatively, we can think in terms of the concavity of the expected cost

alternative approach is to follow Simon and Zame (1990) an allow the tie-breaking rule to be endogenous: this leads to the same equilibrium described here.

function $C(\varphi_k)$. If $C(\varphi_k)$ is convex, then $AC(\varphi_k)$ is increasing and the large party has an advantage. If it is concave, then $AC(\varphi_k)$ is declining and the small party has an advantage.

Since the expected turnout cost $T(\varphi_k)$ is necessarily convex, when monitoring is perfect, i.e. $\theta = 0$, we are in the existing world of the ethical voter model. In this case, $AC(\varphi_k)$ is increasing and the large party always has an advantage. By contrast, if $\theta > 0$, the monitoring cost $M(\varphi_k) = \theta(1 - \varphi_k)c(\varphi_k)$ is non-negative, takes on strictly positive values, yet at the endpoints is equal to zero. In particular, when only committed voters vote, no monitoring is needed, while on the other hand if everyone votes there is nobody to punish.[10] Because the monitoring cost cannot be convex, it might be the case that average costs are decreasing giving the small party a cost advantage.

3.1. Equilibrium

In order to characterize the equilibrium in the simplest possible way, we should first determine the highest fraction of the electorate a party is willing to turn out. Intuitively each party is willing to reach an upper bound $\overline{b}_k$ where either it reached full turnout or the utility from winning the election is equal to 0. That is, if $\eta_k C(1) < V$ then the party is willing to turn out all its voters and $\overline{b}_k = \eta_k$. If instead $\eta_k C(1) \geq V$, then $\overline{b}_k$ is the unique solution to $\eta_k C(\overline{b}_k/\eta_k) = V$. We refer to $\overline{b}_k$ as the *willingness to bid.* We say that the party with the smaller willingness to bid is *disadvantaged*, denoted by d, and the party with the higher value is *advantaged*, denoted by $-d$. Except where explictly stated we also assume that the small party is willing to turn out at least the number of committed voters of the large party, that is, $\overline{b}_S > \eta_L \underline{y}$.[11] In the next theorem we characterize payoffs in the unique equilibrium.

Theorem 1. *There is a unique equilibrium. In this equilibrium neither party uses a pure strategy, the utility of the disadvantaged party is* 0 *and the utility of the advantaged party is* $V - \eta_{-d} C(\overline{b}_d/\eta_{-d})$.

[10] Notice that this property of monitoring cost is robust to the details of the particular monitoring process.

[11] When $\overline{b}_S < \eta_L \underline{y}$ there is a unique equilibrium in which each party turns out only committed voters. This is the only case in which there is an equilibrium in pure strategies.

While a complete proof of this theorem with a characterization of the equilibrium strategies can be found in the Online Appendix, the intuition for the result is fairly straightforward and follows from basic properties of auction theory. In a second price auction the disadvantaged party loses the auction and gets 0 while the advantaged party gets the difference between the value of the prize and the cost of matching the bid of the disadvantaged party. We know that this result continues to hold for the all-pay auction - although the equilibrium strategies are mixed rather than pure. Note that since there cannot be an equilibrium in pure strategies, the small party has always a positive chance of winning.[12] We can now move to one of the main results of our paper: what determines party advantage.

3.2. Which party is advantaged?

Intuitively, the large party having a large number of committed voters is naturally advantaged. It is certainly the case if the prize is sufficiently large, since it will more than compensate the cost of turning out all votes. On the other hand, we also know that increasing average costs of turning out supporters favors the small party. The expected cost function $C(\varphi)$, however, is not a fundamental of our model: it depends on the distribution of costs in the population and on the monitoring difficulty θ. We aim here to establish how these economic fundamentals interact to determine party advantage. Specifically, we will establish that, regardless of the distribution of costs, for the small party to be advantaged three conditions must be satisfied: *i)* monitoring costs must be high; *ii)* the small party must not be "too small"; *iii)* the value of winning the election must be of intermediate size. Not only do these three conditions lead to small party advantage, but the failure of any one of them leads to large party advantage.

Theorem 2. *For* any *individual cost function* $c(y)$ *with corresponding committed voters* $\underline{y}$ *there exist* $\underline{\theta}^S < 1, \underline{\eta}_S^S < 1/2$ *and* $\overline{V} > \underline{V}^S > 0$ *such that if all the*

[12]An implication of this is that regardless of the polls, events such as "Brexit wins" or "Trump wins" always have a positive probability of occurring in equilibrium.

conditions $\theta > \underline{\theta}^S, \eta_S \geq \underline{\eta}_S^S$ and $\underline{V}^S < V < \overline{V}$ hold the small party is advantaged. Conversely if $\underline{y} > 0$ for any values of the other parameters there exist $\underline{\theta}^L > 0, \underline{\eta}_S^L > 0$ and $\underline{V}^L > 0$ such that if any of the conditions $\theta < \underline{\theta}^L$, $\eta_S < \underline{\eta}_S^L$, $V < \underline{V}^L$ or $V > \overline{V}$ are satisfied then the large group is advantaged.

Notice that the reason for the intermediate value of the prize is very intuitive: if the prize is small parties turnout will be low and committed voters will play a disproportionate role favoring the large party. If the value of the prize is high parties will be willing to turn out many voters and as the large party has more voters this also favors the large party. It is only in the intermediate case that the small party may be advantaged. As the proof is simple and instructive we give it here.[13]

Proof. For the first half, observe that marginal cost is $C'(\varphi) = (1-\theta)c(\varphi) + \theta(1-\varphi)c'(\varphi)$ so if $\theta = 1$ then $C'(1) = 0$. Since average cost at $\varphi = 1$ is $C(1) > 0$, average cost is strictly larger than marginal cost at $\theta = 1, \varphi = 1$. Therefore from continuity it must be so for θ, φ both sufficiently close to 1. That is for $1 \geq \theta > \underline{\theta}^S$ and $1 > \overline{\varphi} > \varphi > \underline{\varphi}$ average cost is declining. If we choose η_S large enough, that is, $\eta_S \geq \underline{\eta}_S^S$ then in this region the small party is able to outbid $\overline{\varphi}$. Hence if we choose the prize V so that the large party maximal willingness to mobilize lies in this range, that is, $\overline{\varphi} > \overline{b}_L/\eta_L > \underline{\varphi}$ then the small party must be advantaged as is able to outbid the large party and has a lower average cost of matching the large party bid. For the second half of the Theorem, the large party is advantaged for $\theta = 0$ hence by continuity for small θ. For $\eta_S < \eta_L\underline{y}$ the small party is unable to overcome the committed voters of the large party. If $V < C(\eta_L\underline{y}/\eta_S)$ then the small party is unwilling to bid. if $V > \eta_L C(\eta_S/\eta_L) = \overline{V}$ then $\overline{b}_L > \eta_S$ so the large party is surely advantaged. □

Observe that the construction in the first half of the proof has the following implication: if monitoring costs are high, the two parties are relatively close in size and willing to turn out most but not all of their voters then it is the

[13]For completeness we allow in this theorem the possibility that $\overline{b}_S < \eta_L\underline{y}$, that is the small party may or may not be willing to turn out at least the number of committed voters of the large party.

small and not the large party that is advantaged. One may reflect whether it might be the case that the smaller of the two similar size parties in the U.S., the Republican party, is advantaged for this reason.

3.3. Who Wins?

Advantage is defined in terms of willingness to mobilize supporters. From Theorem 1 we know that this is the same as a utility advantage: the advantaged party receives a positive utility and the disadvantaged party receives no utility. Does it translate also into an advantage in terms of winning the election? To what extent does the advantaged party turn out more voters and have a better chance of winning? As turnout is stochastic for both parties, a natural measure is first order stochastic dominance (FOSD). If the equilibrium bidding function of one party FOSD that of the other then it has a higher chance of winning the election and in a strong sense it turns out more voters.

Party advantage, as we shall see, is not enough to guarantee FOSD. Hence we introduce notion of *strong advantage.* Convexity of $C(\varphi)$ is a simple sufficient condition for large party advantage and it is natural to view this as a strong advantage. On the other hand, in the presence of committed voters $C(\varphi)$ cannot be concave, so for the small party we introduce the weaker notion of incremental concavity - that $C(\varphi)$ be concave for $\varphi \geq \underline{y}$. We define strong advantage for the small party as the combination of small party advantage (that is a larger willingness to mobilize supporters) and incremental concavity.[14] Equipped with this definition we have the following result:

Theorem 3. *The equilibrium bidding function of a strongly advantaged party FOSD that of the disadvantaged party.*

To relate strong party advantage with the distribution of costs in the population and the monitoring difficulty θ, we denote by $G(c)$ the cdf of costs for an individual so that $c(\varphi) = G^{-1}(\varphi)$, $\varphi = G(c)$ and the support is $[c(0), c(1)]$.

[14]We discuss the case of incrementally concave costs and an advantaged large party in the Online Appendix. Indeed, with incrementally concave costs and an advantaged large party, it might be the case that the small disadvantaged party turns out more members in expectation and has a higher probability of winning than the large advantaged party.

We denote the density of $G(c)$ by $g(c)$, and we assume it is continuously differentiable, strictly positive, and has a single "top" in the sense that it is either single peaked or a it is a limiting case such as the uniform where the density is flat at the top.

We will show that the key determinant of strong advantage is how many relatively low cost and relatively high cost voters there are. In order to do so we introduce two measures based on the density of relatively low cost voters $\underline{\gamma} \in [0, 1/2]$ and of relatively high cost voters $\overline{\gamma} \in [0, 1/2]$. While the formal definition of these measures can be found in the Appendix, they are particularly useful for our purposes since they satisfy a number of key properties: If the density is weakly decreasing so there are many relatively low cost voters then $\underline{\gamma} = 1/2$. Similarly for weakly increasing densities $\overline{\gamma} = 1/2$. There are the most voters of both types with uniform distribution and $\underline{\gamma} = \overline{\gamma} = 1/2$.[15]

Intuitively we expect that having many relatively low cost voters, that is high $\underline{\gamma}$, is similar to having many committed voters and so it should favor the large party. The next theorem makes this precise and also shows that, conversely, having many relatively high cost voters, that is high $\overline{\gamma}$ favors the small party.

Theorem 4. *The large party is strongly advantaged if and only if $\theta < \underline{\gamma}$. Cost is incrementally concave (a necessary condition for small party strong advantage) if and only if $\theta > 1 - \overline{\gamma}$.*

In particular a necessary condition for the large party to be strongly advantaged is $\theta < 1/2$ and similarly $\theta > 1/2$ is necessary for the small party to be strongly advantaged. These conditions are sufficient in the uniform case. More broadly for a downward sloping density $\theta < 1/2$ is sufficient for the large party to be strongly advantaged and for a upward sloping density $\theta > 1/2$ is neces-

[15]Furthermore, in the Online Appendix we show when there is a single peak in the interior and we increase the dispersion of the density around that peak, that is, make the density flatter and more like a uniform, then there are more relatively low and high cost voters and both $\underline{\gamma}$ and $\overline{\gamma}$ increase. If the peak is shifted right by shifting the density so there are now more voters who have non-negative costs to the left of the peak then the number of relatively low cost voters $\underline{\gamma}$ increases while the number of relatively high cost voters $\overline{\gamma}$ does not change. If we hold fixed the upper bound $c(1)$ then as the peak and concentration of voters approaches the upper cost bound the number of relatively high cost voters $\overline{\gamma}$ increases.

sary for the small party to be strongly advantaged. When there is a single peak in the interior, increasing dispersion by raising both $\underline{\gamma}$ and $\overline{\gamma}$ favors whichever party has the monitoring cost advantage. That is, if $\theta < 1/2$ it favors the large party and if $\theta > 1/2$ it favors the small party.

The connection between monitoring difficulty, the distribution of voting cost, and strong advantage becomes particularly transparent in the next example.

Example. Suppose participation costs c are normalized to lie in $[0, 1]$ and have a density function $g(c) = \alpha c^{\alpha-1}$ where $\alpha > 0$. When $\alpha = 1$ we have the uniform distribution. When $\alpha < 1$ the density is downward sloping so $\underline{\gamma} = 1/2$ hence the large party is strongly advantaged if and only if $\theta < 1/2$. The density also drops very rapidly near zero (the slope of the density there is $-\infty$) and this concentration of voters near 0 cost means few relatively high cost voteres $\overline{\gamma} = 0$ - so that incremental concavity fails regardless of θ and the small party is never strongly advantaged. When $\alpha > 1$ we get the reverse case. The density is upward sloping $\overline{\gamma} = 1/2$ and as there are no committed voters the small party is strongly advantaged if and only if $\theta > 1/2$ and $V < \overline{V}$. Moreover, the density is very flat near zero and this means $\underline{\gamma} = 1/2$ so that the large party is never strongly advantaged.

Discussion

A unique feature of our theory is that, when the enforcement of social norms is costless, it delivers predictions consistent with the ethical voter and follow-the-leader type of theories. In particular, in equilibrium the large party is strongly advantaged, turns out a higher expected number of voters and has a better chance of winning the election. However, things change drastically when the enforcement of social norms is costly. In this case, contrary to existing theories of political participation and much in the spirit of Olson (1965), our model predicts that when there are many relatively high cost voters and the prize is of intermediate value, the small group turns out a higher expected number of voters, it has a better chance of winning the election and, as Theorem 1 shows, it has a higher equilibrium payoff.

While our results are based on the neutral assumption that costs are the same for both parties this is not essential. Our earlier working paper analyzed the case of differential costs: the all-pay auction is still equivalent to the second price auction while anything that lowers a party's costs are to their advantage. In particular: small party advantage rests not on high monitoring costs, but on high monitoring costs for the large party - if the small party has lower monitoring costs this is also to its advantage. So, for example, a rural minority may have an advantage because urban voters have high monitoring costs although the rural voters have low monitoring costs.

Examples of a smaller group prevailing over a larger one, are not uncommon, but, since our theory predicts a positive probability of the disadvantaged party winning, we cannot drawn conclusions about advantage by examining the results of a single election. One case where we have data on many similar elections is that of teacher unions capturing school boards. These have been studied by Moe (2003) and Moe (2006) who indicates that these elections are often the only ballot issue and that the unions - the small party - are consistently successful at defeating the parents - the large party. Since the stakes are control over budgetary resources the common prize model is not unreasonable and turnout is low indicating that civic duty is probably not an important reason for voting. Although other explanations are possible it seems likely that the fact that the interested voters in the large party (the parents) are a scattered fraction of the overall population makes monitoring difficult. That is, these elections seem likely to satisfy our conditions for small party advantage.

There is also a strategic lesson here for small parties. Consider a fixed cost per voter of turning out: for example, the cost of busing voters to the polls or a voter ID law. An increase in the fixed cost shifts the distribution of voting costs to the right, raising $\underline{\gamma}$ and leaving $\overline{\gamma}$ unchanged. It also decreases the number of committed voters $\underline{y}$. The former decreases the chances of large party strong advantage, and the latter increases the chances of small party advantage: that is, higher fixed costs for both parties favor the small party.

It is natural to try to raise costs for the other party. What this analysis shows is that it is enough for the small party to raise the fixed costs of voting - make it more difficult and unpleasant - for everyone. However, for this to work two other things must be true: the stakes must be sufficiently low and monitoring costs sufficiently high. We argue below in section 4.4 that over time monitoring costs have probably increased. If so it would pay small parties to try to increase participation costs, and this may explain why the small party in the U.S., the Republican party, has increasingly engaged in efforts to raise costs through voter ID laws and the like. Moreover, under these circumstances there is less reason for a party to be large and this may explain why the Republican party has been willing to become more extreme and shrink in size. If, however, the stakes become large enough this policy can fail catastrophically: a small change from $V < \overline{V}$ to $V > \overline{V}$ will abruptly shift party advantage from the small to the large. It may be that with the election of Donald Trump the Republican party has managed to do this - time will tell.

4. Comparative Statics

We will now investigate the effects on turnout and closeness of elections of three important variables: the value of election, the relative size of parties and the efficiency of the monitoring technology. In our model turnout is stochastic, so the meaning of greater turnout and closeness must be qualified. As measures of turnout we consider FOSD (first order stochastic dominance) of the bid distributions, expected turnout, expected turnout cost and peak turn out. By *peak turnout* we mean the highest equilibrium turnout of a party: we know that this is the same for both parties and equal to the disadvantaged party's willingness to bid $\overline{b}_d$. As measures of the closeness of elections we consider the expected vote differential and the *bid differential* $\overline{b}_{-d} - \overline{b}_d$.

4.1. High and Low Value Elections

If $V > \eta_L C(\eta_S/\eta_L)$ the large party is willing to outnumber the entire small party: in this case we say the election is a *high value election.* This is a natural model when the stakes are high such as elections for national leader or important referenda such as Brexit. Notice, however, that while in a high value election both parties are willing to mobilize the number of voters in the small party, in equilibrium neither party does so.

To further analyze turnout in high value elections, let us say that the small party *concedes the election* if it mobilize only its committed voters. Furthermore, we say that the large party *preempts the election* if it mobilize the most voters feasible for the small party, that is η_S. Then while both parties are willing to mobilize all their voters the next result shows that what they do mobilize in equilibrium is far less.

Theorem 5. *In a high value election the probabilities that the small party concedes and the large party preempts the election increase in V, and approach 1 in the limit. As V increases the bid distribution of the small party declines in FOSD and the bid distribution of the large party increases in FOSD. The expected vote differential increases in V while the expected turnout cost remains constant.*

We refer to the fact that mobilization of the small party is decreasing and its probability of concession increasing in V as the *discouragement effect*, which is standard in all pay auctions. Since the large party is willing and able to outnumber the small party, the small party becomes discouraged and, as the stakes increase, turns out fewer and fewer voters. Nevertheless, because the stakes are very high and despite the fact that the small party is turning out very few voters, with very high probability the large party turns out enough voters to guarantee victory against the small party. Notice that the expected turnout cost of the small party declines and the expected turnout cost of the large party increases, but the two effects exactly offset each other.

It is interesting to contrast a *low value election* in which $V < \eta_S C(\eta_L \underline{y}/\eta_S)$ and only committed voters turn out with a high value election. Turnout of both parties in a high value election is substantially higher than in a low value election. This is consistent with suggestive evidence of higher participation in national than in local elections, and with empirical evidence showing that electoral participation will be higher in elections where stakes are high.[16]

In the high value election we can see clearly that there is a discontinuity in the surplus when the parties are of near equal size. As $\eta_L \to 1/2$ the surplus

[16]See, Andersen, Fiva and Natvik (2014).

of the large party approaches $V - (1/2)C(1)$, that is, in the limit it does not approach zero. Hence a small change in party size shifting a small party into a large party causes the surplus of that party to jump from zero to a strictly positive value and conversely. Moreover neither the probability of concession nor the probability of the large party taking the election approach zero and for large V both are close to one. In other words, a small change in the party size causes a party that was conceding with positive probability to stop conceding and instead preempt the election with positive probability. The discontinuity is important if we step back from the model and consider a broader setting in which parties choose platforms in an effort to compete for members prior to the election: we see that a small shift in the relative sizes of the parties can have disproportionate consequences, suggesting that the competition over platforms may be a fierce one.

4.2. Close Elections with Small Party Advantage

A measure of the closeness of the election is the bid differential $\bar{b}_{-d} - \bar{b}_d$. In a standard all-pay auction as the bid differential approaches zero the surplus of the advantaged party vanishes. In our setting when the small party is strongly advantaged this need not be the case.

A stongly advantaged small party implies that the cost to the large party of matching the maximum turnout of the small party is greater than the cost to the small party of turning out all voters. If the value of winning the elections is intermediate, the small party is advantaged but constrained: it would like to turn out more voters but cannot do so. As V increases the willingess to bid of the small party does not change, while the willingness to bid of the large party increases, reducing the bid differential - increasing the closeness of the election. Nevertheless the surplus accruing to the small party does not approach zero. As V further increases, the advantage switches to the large party and the surplus of the small party drops abruptly to zero.

These results are interesting from the point of view of agenda setting, for example a referendum proposed by the small party - they want on the one hand to ask for a large prize, but if they make it just a bit too big they can lose everything. One example of this may have been the heavy defeat of Proposition 16 in California in 2010: This was a ballot initiative sponsored by Pacific Gas and Electric Company that was designed to reduce competition from local governments.

4.3. Disagreement

Following Feddersen and Sandroni (2006), we next consider the effect of the level of "disagreement" meaning how closely divided the electorate is. Specifically since $\eta_S = 1 - \eta_L$, as η_S grows we approach a situation where party supporters are evenly divided, that is, the level of disagreement in society increases with η_S. Intuitively we expect that greater disagreement should mean more fiercely contested elections that are close and have higher turnout. In the ethical voters model of Feddersen and Sandroni (2006) or in the group-turnout model with aggregate shocks studied in Herrera, Morelli and Nunnari (2015) greater disagreement does indeed lead to higher turnout and closer elections, and the same is true in pivotal voter models such as Castanheira (2003). A similar result holds here:

Theorem 6. *When either party is strongly advantaged disagreement increases the peak turnout, the expected turnout cost and decreases the bid differential.*

4.4. Monitoring Difficulty in High Value Elections

Finally, we turn to monitoring difficulty θ. Here we focus on the important interesting case of a high value election. Our intuition is that increasing monitoring difficulty should decrease turnout. The following theorem shows that if the small party is neither too large nor too small this is true and that in addition elections are closer.

Theorem 7. *In a high value election, an increase in monitoring difficulty θ decreases the turnout of the advantaged (large) party in terms of FOSD. Furthermore, there exists $0 < \underline{\eta} < \overline{\eta} \leq 1/2$ such that for $\underline{\eta} < \eta_S < \overline{\eta}$ the expected turnout of the disadvantaged (small) party decreases in monitoring difficulty in terms of FOSD while the expected vote differential also decreases.*

There is some direct data on the effect of monitoring inefficiency on turnout cost: Larreguy, Marshal and Querubin (2016) found that increased monitoring inefficiency of local mobilizers decreases turnout buying for two parties of similar size as Theorem 7 suggests. Another application concerns the idea that in Western Europe, over the period since World War II, the social ties underlying the party system have broken down. One possible interpretation of this is that monitoring has become more inefficient. For example, in the old days labor union members in the UK socialized in pubs and old money socialized in clubs, with the resulting strong social ties keeping monitoring costs low for the Labor and Conservative party, respectively. This is consistent with the concept of "mass parties" in the political science literature - see, for example, the discussion of the literature in Katz and Mair (1995). For a considerable period after World War II, Western Europe was dominated by large mildly left-wing parties of various flavors of labor or Christian Democrats. Theorem 7 supports the idea that there is a connection between the breakdown in social ties - meaning less efficient monitoring - and the decline in these parties as measured by declining turnout, more competitive elections.[17]

That increased monitoring difficulty decreases turnout may also help to explain why measures designed to increase turnout by lowering participation costs may actually have the perverse effect of decreasing turnout because they also raise monitoring costs. Voting at a polling place is a relatively visible and easy to monitor activity. Voting by post, internet, or indeed in the supermarket is not so much so. Hence lowering the inconvenience of voting by allowing it to take place away from the polling place is an example of a reform that may have the perverse effect of reducing turnout. There is evidence that this is indeed the case: Funk (2010) shows this was the case when postal voting was introduced in Switzerland, and the 2000 UK policy experiment of setting in-store poll booths also failed to increase turnout.

[17]Gray and Caul (2000) relates post-war turnout decrease with the decline of mobilizing actors such as labor parties and trade unions, and Knack (1992) connects the decline of American voter turnout with a weakened enforcement of social norms.

5. Discussion of the Model and the Literature

In this section we examine the four key elements of the model: symmetry of the fundamentals, the use of peer punishment to enforce an endogenous social norm, the absence of exogenous uncertainty, and the fact that costs are bounded.

5.1. Symmetry of the Fundamentals

As a benchmark we have assumed symmetry in voting costs between the two parties and in that they compete for a common prize worth the same to each party. This is a useful benchmark model: voters in both parties face identical *ex ante* participation costs and a common prize is efficiency neutral. Naturally if one side has a cost advantage or values the prize more highly than the other party it will be more able and willing to turn out voters and this will give it an electoral advantage.

A common prize makes sense when the outcome of the election are taxes, subsidies and other transfer payments. By contrast, if civil rights and law changes are at stake it may make sense to assume that the benefit of winning is the same for all members of a party. It is less certain in this case that the benefit should be the same for both parties - is the benefit of depriving another the right to sit in the front of the bus equal to cost of being deprived of that right?

In the literature, assumptions about a common prize vary. Shachar and Nalebuff (1999) and Herrera, Morelli and Nunnari (2015) who are interested in the same turnout issues we are assume a common prize. Coate and Conlin (2004) assume that the per capita value of the prize is independent of the size of the party, but may differ between the two parties: this is appropriate in their setting of liquor referenda. Feddersen and Sandroni (2006) assume the per capita value of the prize is the same for everyone - but their interest is primarily in information not turnout. Palfrey and Rosenthal (1985) and Levine and Palfrey (2007) also assume the per capita value of the prize is the same for everyone - but study a pivotal voter model in which the aggregate size of the prize to the party is of less importance.

The case of a common per capita prize make a useful contrast to that of a common prize: this provides an additional advantage to a large party which has more members to enjoy the per capita value. Indeed it provides enough advantage to the large party that it is advantaged in our model regardless of monitoring costs.[18] In all likelihood reality lies in between a common prize and a per capita prize: typically elections involve a mix of issues, some involving taxes and transfers, other involving rights. Esteban and Ray (2011) consider a model with both types of prizes to describe the surge of ethnic conflicts and Esteban, Mayoral and Ray (2012) bring their theoretical predictions to the data showing that both types of prizes are important.[19]

5.2. Endogenous Social Norms and Peer Punishment

Our model of peer punishment is meant to capture in a very stylized way an elaborate informal process. In reality, the political hierarchy (candidates, party officials, donors, and activists) chooses strategically to devote resources to mobilization, and voters use this to determine an appropriate social norm through pub and dinner table conversations (not to speak of social media). Nevertheless we think that the model captures the important feature that the social norm adapts to the circumstances of the election.

Our formal model of peer punishment originates in Kandori (1992)'s work on social norms in repeated games. It is a variant of the mechanism design approach to collective action of Levine and Modica (2014) and Dutta, Levine and Modica (2014). In these models agents monitor each others behavior and punish deviators through ostracism and social disapproval and this gives rise to monitoring costs - the new feature of our model.

We assume that peer pressure and social norms take place within a party rather than globally. That is, if the social norm is that voting is a civic duty then this is not a party specific social norm. Our view is that both global social

[18]When the prize $V_k = \eta_k v$ the objective function is equivalent to $\Pi_k(b_k, F_{-k})v - C(b_k/\eta_k)$. Since for a given bid b the smaller party must turn out more voters $b/\eta_S > b/\eta_L$ and expected costs are increasing, it follows that the cost of a bid is always lower for the large party, hence it is advantaged.

[19]In the Online Appendix we show that our results about small group advantage are robust in the sense that they hold for any non-trivial mix of a common and per capita prize. It is only in the extreme case of a pure per capita prize that the large group is advantaged regardless of monitoring cost.

norms and party social norms matter. Here we incorporate the general social norm of civic duty into the committed voters, treating it as exogenous - as it is standard in previous literature - in order to focus on the party social norm. There are a number of reasons why the party social norm should be relevant. First, many voters are more inclined to enforce a social norm within their own party - that is with voters who are likely to vote as they do. Certainly in a high stakes election many people will put a lot of pressure on like-minded family and friend to vote, and will be less likely to bother with those who take the "wrong position." This has been observed both in the laboratory (Grosser and Schram (2006)), in the field (Bond et al (2012)), and Shachar and Nalebuff (1999) show that parties' effort (measured by the number of calls and visits to individuals to encourage their turnout) is positively correlated with group membership and parents' involvement in politics. More to the point: there is a high correlation of political beliefs within the social networks important for enforcing social norms. We see this in the positive correlation of political beliefs within families (Jennings, Stoker and Bowers (2009)), in the geographic concentration of political preferences (Chen and Rodden (2013)), and in the strength of party identity (Dunham, Arechar and Rand (2016)).[20]

With our assumption of collective decision making in a large group there is no room for pivotality of the individual voter to play a role. Although they do well in the laboratory for small elections (see for example Levine and Palfrey (2007)) pivotal voters models such Palfrey and Rosenthal (1985) have difficulty in explaining turnout in mass elections.[21] Consequently attention has turned to follow-the-leader models such as Shachar and Nalebuff (1999) or models of ethical voters such as Feddersen and Sandroni (2006) and Coate and Conlin (2004).[22] Herrera, Morelli and Nunnari (2015) examine all three models.

[20]One theory of the strength of these social networks is the skill selection model of Penn (2015).

[21]Coate, Conlin and Moro (2008) show that in a sample of Texas liquor referenda, elections are much less close than what would be predicted by the pivotal voter model, and Coate and Conlin (2004) show that a model of "ethical" voters better fits that data than the model of pivotal voters. Not surprisingly, the probability of being pivotal in large elections is very low as documented by Mulligan and Hunger (2003) and Shachar and Nalebuff (1999).

[22]Ali and Lin (2013) extend the Feddersen and Sandroni (2006) model by introducing

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. Our model of collective decision making by the party is in a similar vein. While Shachar and Nalebuff (1999) focus on the costs to the leaders, we follow the ethical voter literature in focusing on the costs to the followers.

5.3. Endogenous versus Exogenous Uncertainty

There are different assumption used in the literature about the way in which voting determines the outcome of an election. Palfrey and Rosenthal (1985) assume as do we that the parties are of fixed size and the party with the most votes wins. The remaining models introduce aggregate shocks and assume that these are sufficiently large to guarantee the existence of a pure strategy equilibrium. In our model voter turnout is also random but this is endogenous due to the use of mixed strategies by the parties. This is reflected in the reality of elections as in the case of "GOTV" (Get Out The Vote) efforts. Our view is that these efforts are an important part of establishing the social norm for the particular election, and indeed, GOTV efforts are variable and strategic. Furthermore, political parties have strong incentives not to advertize their GOTV effort, and in fact to keep it secret.[23] 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.

"pragmatic" voters alongside ethical voters. A pragmatic voter votes only because she wishes others to think of her as being ethical. Morris and Shadmer (2017) study a model where "voters" with heterogeneous beliefs are encouraged to participate by a "party" that provides incentives. In their model incentives have an opposite effect on optimistic and pessimistic voters, while our incentives encourage all voters.

[23] 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 (2012) or "[Romney's] campaign came up with a super-secret, super-duper vote monitoring system [...] to plan voter turnout tactics on Election Day " York (2012). 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.

In general and in Shachar and Nalebuff (1999) and Coate and Conlin (2004) exogenous random turnout leads to a model in which the probability of winning depends not only on bids but also on the size of the two parties. Herrera, Morelli and Nunnari (2015) use a more standard contest resolution function in which the probability of winning depends only on the bids and this is also the case in the specific application of Feddersen and Sandroni (2006).

Despite this wide variety of assumptions on conflict resolution the existing literature assumes that monitoring costs are absent and conclude that the large party is advantaged. They all study pure strategy equilibria. As each paper makes special assumptions we give a general result in the Online Appendix for a common prize, convex common costs, and a standard contest resolution function where the probability of winning depends only on the bids: pure strategy equilibrium advantages the large party which turns out more voters and a gets greater utility than the small party. This is the same result we find for our mixed equilibrium. There is an analogous result for concave costs, but it is of lesser interest since pure strategy equilibria are not so likely to exist in that case.

5.4. Bounded Costs

We have assumed that costs are bounded by $c(1)$. A consequence of this is that if the prize V is sufficiently large a party that could insure victory by doing so would turn out all of its voters. Moreover while we show in the Online Appendix that the equilibrium probability that either party turns out all of its voters is zero in a high stakes election, there is a small probability that the turnout of the small party is close to 100%. Empirically this has little meaning since measured turnout is expressed as a fraction of the voting age population and not as a fraction of people actively contemplating voting, as it is the case in the model. Never-the-less, it might be judged unreasonable that a party would turn out nearly all of its voters: in reality we would expect that some voters would have such high costs that it would not be worth turning them out regardless of how high the stakes might be. It turns out that this does not matter very much for the theory. In the Online Appendix we show that while the equilibrium is more difficult to compute with unbounded costs the equilibrium strategies and utilities are close to those with bounded costs even when V is very large.

6. Conclusion

We have examined a model that captures the importance of social norms and peer pressure in voter turnout. The resulting theory does not discard the major existing theories - the ethical voter model corresponds to the special case in which monitoring costs are zero and when the electorate is small and pivotality can be incorporated into the incentive constraints for individual voters. The theory also makes a rich new set of predictions - relating, for example, monitoring cost to turnout. One key prediction concerns the case in which monitoring cost is large and committed voters few: in this case, unlike in ethical voting and follow-the-leader theories, the small group may be advantaged.[24] This may explain why there are many referenda where special interests do well: for example, the type of commercial gambling permitted on Indian reservations, school budgets, the working environment for prison guards and so forth.

Our 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 (2006), Acemoglu (2001), or Levine and Modica (2015).

References

Acemoglu, S. and J. A. Robinson (2001): "Inefficient Redistribution," *American Political Science Review* 95: 649-661.

Akerlof, G. A. (1990): "A Theory of Social Custom, of Which Unemployement May Be One Consequence," *Quarterly Journal of Economics* 94: 749–775.

Ali, S. N. and C. Lin (2013): "Why People Vote: Ethical Motives and Social Incentives," *American Economic Journal: Microeconomics* 5: 73–98.

[24]We should mention that minority advantage is also present in Casella and Turban (2014) albeit for an entirely different reason - they study a model in which votes can be bought and sold in market.

Amat, F., Boix C., Munoz J. and T. Rodon (2018): "From Political Mobilization to Electoral Participation: Turnout in Barcelona in the 1930s," Mimeo.

Andersen, J. J., J. H. Fiva and G. J. Natvik (2014): "Voting When the Stakes are High," *Journal of Public Economics* 110: 157-166.

Baye, M. R., D. Kovenock and C. De Vries (1996): "Rigging the Lobbying Process: An Application of the All-Pay Auction," *American Economic Review* 83: 289-294.

Bond, R. M., C. J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle and J. H. Fowler (2012): "A 61-million-person experiment in social influence and political mobilization," *Nature* 489: 295-298.

Casella, Alessandra, and Sébastien Turban (2014): "Democracy undone. Systematic minority advantage in competitive vote markets," *Games and Economic Behavior* 88: 47-70.

Castanheira, Micael (2003): "Victory margins and the paradox of voting," *European Journal of Political Economy* 19: 817-841.

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

Chen, J. and J. Rodden (2013): "Unintentional gerrymandering: Political geography and electoral bias in legislatures," *Quarterly Journal of Political Science* 8: 239-269.

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

Coate, S., M. Conlin and 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.

Coleman, James S. (1988): "Social Capital in the Creation of Human Capital," *American Journal of Sociology* 94: S95-S120.

Della Vigna, S., J. A. List, U. Malmendier and G. Rao (2016): "Voting to Tell Others," *Review of Economic Studies* forthcoming.

Dunham, Y., A. A. Arechar and D. G. Rand (2016): "Unity for Democrats But Not Republicans: The Temporal Dynamics of Intra-Party Bias in US Electoral Politics," SSRN 2846915.

Dutta, R., D. K. Levine and S. Modica (2014): "Collusion, Randomization and Leadership in Groups," EUI.

Dutta, R., D. K. Levine and S. Modica (2016): "Collusion Constrained Equilibrium," EUI.

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

Esteban, J., L. Mayoral and D. Ray (2012): "Ethnicity and Conflict: An Empirical Study," *American Economic Review* 102(4): 1310–1342.

Esteban, J. and D. Ray (2011): "Linking Conflict to Inequality and Polarization," *American Economic Review* 101 (4): 1345–74.

Feddersen, T., A. Sandroni (2006): "A Theory of Participation in Elections," *American Economic Review* 96: 1271–1282.

Funk, P. (2010): "Social Incentives and Voter Turnout: Evidence from the Swiss Mail Ballot System," *Journal of the European Economic Association* 8: 1077–1103.

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.

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

Grosser, J. and A. Schram (2006): "Neighborhood information exchange and voter participation: An experimental study," *American Political Science Review* 100: 235-248.

Herrera, H., M. Morelli, and S. Nunnari (2016): "Turnout Across Democracies," *American Journal of Political Science* 60: 607-624.

Hillman, A. L. and J. Riley (2006): "Politically Contestable Rents and Transfers," *Economics and Politics* 1: 17-39.

Jennings, M. K., L. Stoker, and J. Bowers (2009): "Politics across Generations: Family Transmission Reexamined," *The Journal of Politics* 71:782-799.

Katz, Richard S., and Peter Mair (1995): "Changing models of party organization and party democracy the emergence of the cartel party," *Party Politics* 1: 5-28.

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

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

Kobach, K. (1993): *The Referendum: Direct Democracy in Switzerland,* Aldershot.

Larreguy, H., J. Marshall and P. Querubin (2016): “Parties, Brokers and Voter Mobilization: How Turnout Buying Depends Upon the Party’s Capacity to Monitor Brokers,” *American Political Science Review* 110: 160-179.

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

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

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

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

Moe, T. M. (2003): “Teacher Unions and School Board Elections,” Hoover Institute.

Moe, T. M. (2006): “Political Control and the Power of the Agent,” *Journal of Law, Economics, and Organization* 22: 1-29.

Morris, S., M. Shadmehr (2017): “Reward and Punishment in a Regime Change Game,” mimeo.

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

Nagourney, A., A. Parker, J. Rutenberg and J. Zelenynov (2012): “How a Race in the Balance Went to Obama,” *New York Times*, November 7.

Olson, M. (1965): *The Logic of Collective Action: Public Goods and the Theory of Groups,* Harvard University Press.

Ostrom, E. (1990): *Governing the Commons: The Evolution of Institutions for Collective Action,* Cambridge University Press.

Palfrey, T. R. and K. Pogorelskiy (2016): “Voting with Communciation: An Experimental Study,” Caltech.

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

Penn, Maggie (2015): "Inequality, Social Context, and Value Divergence," *Journal of Politics,* forthcoming.

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

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

Rawnsley, G. (2005): *Political Communication and Democracy,* Palgrave Macmillan.

Shachar, R. and B. Nalebuff (1999): "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review* 89: 525-547.

Siegel, R. (2014): "Asymmetric Contests with Head Starts and Nonmonotonic Costs," *American Economic Journal: Microeconomics* 6: 59-105.

Simon, L. K. and W. R. Zame (1990): "Discontinuous Games and Endogenous Sharing Rules," *Econometrica* 58: 861-872.

Szech, N. (2015): "Tie-Breaks and Bid-Caps in All-Pay Auctions," *Games and Economic Behavior* 92: 138-149.

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

Appendix

To define $\underline{\gamma}, \overline{\gamma}$ we first define

$$\lambda(c) = -\frac{(1 - G(c))\, g'(c)}{(g(c))^2}$$

with $\underline{\lambda} = \min_{c \geq 0} \lambda(c) \leq 0$ the smallest possible value of $\lambda(c)$ and $\overline{\lambda} = \max_{c \geq 0} \lambda(c) \geq 0$ the largest. In the uniform case $g'(c) = 0$ so $\lambda(c) = 0$. If the density is increasing then $\lambda(c) \leq 0$ so $\overline{\lambda} = 0$ and if it is decreasing $\underline{\lambda} = 0$.

Define

$$\underline{\gamma} = \frac{1}{2 - \underline{\lambda}}$$

and define $\overline{\gamma} = 0$ if $\overline{\lambda} > 1$ and

$$\overline{\gamma} = 1 - \frac{1}{2 - \overline{\lambda}}$$

otherwise. Hence $\underline{\gamma}$ is an increasing function of $\underline{\lambda}$ and $\overline{\gamma}$ is a decreasing function of $\overline{\lambda}$. Since $\underline{\lambda} \leq 0$ and $\overline{\lambda} \geq 0$ we have $0 \leq \underline{\gamma}, \overline{\gamma} \leq 1/2$. The properties of $\underline{\gamma}, \overline{\gamma}$ for the uniform, increasing and decreasing cases can be read directly from the results for $\underline{\lambda}, \overline{\lambda}$: both $1/2$ for the uniform case, $\overline{\gamma} = 1/2$ in the increasing case, and $\underline{\gamma} = 1/2$ in the decreasing case. For the singel-peaked case in the Online Appendix we prove the proposition below.

Proposition. *If the density shifts to the right then $\overline{\gamma}$ is constant and $\underline{\gamma}$ decreases; if the density shifts to the right holding fixed $c(1)$ then $\overline{\gamma}$ increases. Furthermore, increasing dispersion by a change of scale around the mode increases both $\underline{\gamma}$ and $\overline{\gamma}$.*

Online Appendix (Not for publication)

Equilibrium

We here characterize equilibria in the all-pay auction model. We do not assume common prize, we allow arbitrary $V_k \geq 0$ and common cost $C(\varphi)$ which is continuous and strictly increasing for $\varphi \geq \underline{y}$. As in the text we assume $\overline{b}_L \neq \overline{b}_S$.

Proposition 1. [1 in text] *There is a unique equilibrium. In this equilibrium neither party uses a pure strategy, the utility of the disadvantaged party is* 0 *and the utility of the advantaged party is* $V_{-k} - \eta_{-d} C(\overline{b}_d/\eta_{-d})$.

We record the additional facts not reported in the text but are used subsequently in the Appendix which are the equilibrium strategies. Let $F_k^0(b)$ record the size of the atom at b (if any). In $(\eta_L \underline{y}, \overline{b}_d)$:

$$F_d(b_d) = 1 - \frac{\eta_{-d} C(\overline{b}_d/\eta_{-d}) - \eta_{-d} C(b_d/\eta_{-d})}{V_{-k}}$$

$$F_{-d}(\eta_{-d}\varphi_{-d}) = \frac{\eta_d C(b_{-d}/\eta_d)}{V_k}$$

The disadvantaged party has a single atom at $F^0_d(\eta_d \underline{y}) = 1 - \eta_{-d} C(\overline{b}_d/\eta_{-d})/V_{-k} + \eta_{-d} C(\eta_L \underline{y}/\eta_{-d})/V_{-d}$. The advantaged party if it is large has an atom at $F^0_L(\overline{b}_S) = 1 - \eta_S C(\overline{b}_S/\eta_S)/V_L$, and whichever party is advantaged has an atom at $F^0_{-d}(\underline{y}\eta_L/\eta_d) = \eta_d C(\underline{y}\eta_L/\eta_d)/V_d$.

Proof. S will never submit a bid b_k for which $\eta_S \underline{y} < b_k < \eta_L \underline{y}$ since such a bid will be costly but losing, and neither party will submit a bid for which $b_k > \overline{b}_k$ since to do so would cost more than the value of the prize. It follows that k must either bid $\eta_k \underline{y}$ or in the range $[\eta_L \underline{y}, \overline{b}_d]$. If $V_S \leq \eta_S C\left(\underline{y}\eta_L/\eta_S\right)$, it follows that $\overline{b}_S \leq \eta_L \underline{y}$. In this case S will only mobilize committed voters, that is will bid $\eta_S \underline{y}$, and L wins with probability 1 by bidding $\eta_L \underline{y}$. This case is ruled out in the text.

Consider now the case $V_S > \eta_S C\left(\underline{y}\eta_L/\eta_S\right)$. In the range $(\eta_L \underline{y}, \overline{b}_d)$ there can be no atoms by the usual argument for all-pay auctions: if there was an atom at b_k then party $-k$ would prefer to bid a bit more than b_k rather than a bit less, and since consequently there are no bids immediately below b_k party k would prefer to choose the atom at a lower bid. This also implies that S cannot have an atom at $\eta_L \underline{y}$: if L has an atom there, then S should increase its atom slightly to break the tie. If L does not have an atom there, then S should shift its atom to $\eta_S \underline{y}$ since it does not win either way.

Next we observe that in $(\eta_L \underline{y}, \overline{b}_d)$ 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 $\overline{b}_d$, while the advantaged party gets at least $U_{-d} \geq V_{-d} - \eta_{-d} C(\overline{b}_d/\eta_{-d}) > 0$ since by bidding slightly more than $\overline{b}_d$ it can win for sure, but gets no more than that since it must make bids with positive probability arbitrarily close to $\overline{b}_d$. We conclude that the equilibrium payoff of the advantaged party must be exactly $U_{-d} \geq V_d - \eta_{-d} C(\overline{b}_d/\eta_{-d})$.

From the absence of zero probability open intervals in $(\eta_L \underline{y}, \overline{b}_d)$ it follows that the indifference condition for the advantaged party

$$F_d(b_d)V_d - \eta_{-d}C(b_d/\eta_{-d}) = V_d - \eta_{-d}C(\overline{b}_d/\eta_{-d})$$

must hold for at least a dense subset. Similarly for the disadvantaged party

$$F_{-d}(b_{-d})V_{-d} - \eta_d C(b_{-d}/\eta_d) = 0$$

for at least a dense subset. This uniquely defines the cdf for each party in $(\eta_L \underline{y}, \overline{b}_d)$:

$$F_d(b_d) = 1 - \frac{\eta_{-d}C(\overline{b}_d/\eta_{-d}) - \eta_{-d}C(b_d/\eta_{-d})}{V_d}$$

$$F_{-d}(b_{-d}) = \frac{\eta_d C(b_{-d}/\eta_d)}{V_{-d}}.$$

As these are differentiable they can be represented by continuous density functions which are found by taking the derivative.

Evaluating $F_d(b_d)$ at $\eta_d \underline{y}$ gives $F_d^0(\eta_d \underline{y}) = 1 - \eta_{-d}C(\overline{b}_d/\eta_{-d}/V_{-d}) + \eta_{-d}C(\eta_L \underline{y}/\eta_{-d})/V_{-d}$. Since $F_d(\overline{b}_d) = 1$ and we already proved that S has no atom at $\eta_L \underline{y}$ this is in fact the only atom for the disadvantaged party.

As for the advantaged party, if -d = S then $\eta_L > \eta_S \geq \overline{b}_S > \overline{b}_L$ implies that $F_S(\overline{b}_L) = \eta_L C(\overline{b}_L/\eta_L)/V_L = 1$. If instead $-d = L$ then $F_L(\overline{b}_S) = \eta_S C(\overline{b}_S/\eta_S)/V_L$. If $\overline{b}_S < \eta_S$ then this is 1 and there is no atom, otherwise there must be an atom of size $F_L^0(\overline{b}_S) = 1 - \eta_S C(\overline{b}_S/\eta_S)/V_L$. Turning to $\eta_L \underline{y}$ we see that the atom there is given by

$$F_{-d}^0 = \frac{\eta_d C(\underline{y}\eta_L/\eta_d)}{V}$$

since the advantaged group never bids less. $\square$

Who Wins?

Theorem. [3 in text] *The equilibrium bidding function of a strongly advantaged party FOSD that of the disadvantaged party.*

Proof. At $\bar{b}_d$ we have $F_d(\bar{b}_d) = \hat{F}_{-d}(\bar{b}_d) = 1$ so this is irrelevant for FOSD. For $\eta_S \underline{y} \le b < \eta_L \underline{y}$ we have $F_L(b) = 0$ while $F_S(b) > 0$ if and only if S is disadvantaged. Hence when S is disadvantaged its bidding schedule cannot FOSD that of L, while if it is advantaged this range is irrelevant for FOSD.

It remains to examine the range $\eta_L \underline{y} \le b < \bar{b}_d$. In this range the equilibrium bid distributions are given by

$$F_d(b) = 1 - \frac{\eta_{-d} C(\bar{b}/\eta_{-d})}{V} + \frac{\eta_{-d} C(b/\eta_{-d})}{V}$$

$$F_{-d}(b) = \frac{\eta_d C(b/\eta_d)}{V}.$$

Hence for FOSD of the advantaged party, we must have

$$1 - \frac{\eta_{-d} C(\hat{b}_d/\eta_{-d})}{V} + \frac{\eta_{-d} C(b/\eta_{-d})}{V} - \frac{\eta_d C(b/\eta_d)}{V} > 0.$$

Moreover since $\eta_d v_d \ge \eta_d C(\bar{b}_d/\eta_d)$ this is true if

$$1 - \frac{\eta_{-d} C(\bar{b}_d/\eta_{-d})}{\eta_d C(\bar{b}_d/\eta_d)} + \frac{\eta_{-d} C(b/\eta_{-d})}{\eta_d C(\bar{b}_d/\eta_d)} - \frac{\eta_d C(b/\eta_d)}{\eta_d C(\bar{b}_d/\eta_d)} > 0$$

and if and only if the disadvantaged party is not constrained in bidding. This is equivalent to

$$(\eta_{-d} C(b/\eta_{-d}) - \eta_d C(b/\eta_d)) - \left(\eta_{-d} C(\bar{b}_d/\eta_{-d}) - \eta_d C(\bar{b}_d/\eta_d)\right) > 0.$$

Let $t(\eta, b) \equiv \eta C(b/\eta)$. The derivative with respect to η is $t_\eta(\eta, b) = C(b/\eta) - (b/\eta)C'(b/\eta)$ so the cross partial is $t_{\eta b}(\eta, b) = -(b/\eta^2)C''(b/\eta)$. Observe that the sufficient condition may be written as

$$\begin{aligned} 0 &< (t(\eta_{-d}, b) - t(\eta_d, b)) - \left(t(\eta_{-d}, \bar{b}_d) - t(\eta_d, \bar{b}_d)\right) \\ &= \int_{\eta_d}^{\eta_{-d}} \left(t_\eta(\eta, b) - t_\eta(\eta, \bar{b}_d)\right) d\eta \end{aligned}$$

$$= -\int_{\eta_d}^{\eta_{-d}} \int_{b}^{\bar{b}_d} t_{\eta b}(\eta, b') d\eta db' = \int_{\eta_d}^{\eta_{-d}} \int_{b}^{\bar{b}_d} (b'/\eta^2) C''(b'/\eta) d\eta db'$$

This is positive if $\eta_{-d} > \eta_d$ and C is convex or if $\eta_d > \eta_{-d}$ and C is concave, which gives the primary result. On the other hand, in the case of a common prize, L advantaged, and S unconstrained, it is negative and gives the exact sign of $F_d(b) - F_{-d}(b)$ (it is necessary and sufficient). Hence, since the difference between F_d and F_{-d} is positive for $\eta_S \underline{y} \le b < \eta_L \underline{y}$ and negative for $\eta_L \underline{y} \le b < \hat{b}_d$ neither bidding schedule FOSD the other. □

The next proposition studies the case where costs are incrementally concave and yet the large party is advantaged. It shows how the *FOSD* result can fail in the strong sense that the disadvantaged small party turns out more members in expectation and has a higher probability of winning than the large advantaged party.

Proposition 2. *Suppose that cost is quadratic so that for $\theta > 1/2$ it is incrementally concave. For any η_S there exists a $\underline{\varphi} > 0$ such that for any $\underline{y} < \underline{\varphi}$ there is an open set of V's and for any such V there are bounds $1/2 < \underline{\theta} < \theta^* < \overline{\theta} \le 1$ such that*

a. for $\overline{\theta} > \theta > \theta^*$ the small party is advantaged

b. for $\theta^* > \theta > \underline{\theta}$ the large party is advantaged yet the small party turns out more expected voters and has a higher probability of winning the election.

Proof. Recall the quadratic case $C(\varphi_k) = (1-2\theta)(\varphi_k - \underline{y})^2 + 2\theta(1-\underline{y})(\varphi_k - \underline{y})$. Hence $C'(\underline{y}) = 2\theta(1-\underline{y})$ and $C''(\varphi_k) = 2(1-2\theta)$. We fix $\theta > 1/2$ so that cost is incrementally concave.

We first establish that for sufficiently small $\underline{y}$ there is a range of $V's$ such $\eta_L \underline{y} < \overline{b}_S < \eta_S$ for $1/2 \le \theta \le 1$ and such that S is advantaged at $\theta = 1$.

Since the derivative of C with respect to θ is $2(\varphi_k - \underline{y})(1-\varphi_k) > 0$ the greatest willingess to bid is at $\theta = 1/2$ and the least is at $\theta = 1$. At $\theta = 1/2$ the utility of S is $V - (1-\underline{y})(\varphi_k - \underline{y})$ and so $\overline{b}_k < 1$ for $V < (1-\underline{y})^2 = \overline{V}_S$. At $\theta = 1$ the utility of S is $V + (\varphi_k - \underline{y})^2 - 2(1-\underline{y})(\varphi_k - \underline{y})$ so $\overline{b}_k > \eta_L \underline{y}$ for $V > 2(1-\underline{y})(\eta_L \underline{y}/\eta_S - \underline{y}) - (\eta_L \underline{y}/\eta_S - \underline{y})^2 = \underline{V}_S$.

Set $\theta = 1$ and let φ^* be defined by $A(\varphi^*) = A(\eta_L \varphi^*/\eta_S)$. Some algebra yields $\varphi^* = \sqrt{(\eta_S/\eta_L)\underline{y}(2-\underline{y})}$. This will be less than η_S/η_L provided $\underline{y}(2-\underline{y}) < \eta_S/\eta_L$. At $\theta = 1$ the utility for L is $V - \eta_L\left(-(\varphi_L - \underline{y})^2 + 2(1-\underline{y})(\varphi_L - \underline{y})\right)$. Hence L would like to bid greater than $\eta_L\varphi^*$when

$$V > \eta_L\left(-(\sqrt{(\eta_S/\eta_L)\underline{y}(2-\underline{y})} - \underline{y})^2 + 2(1-\underline{y})(\sqrt{(\eta_S/\eta_L)\underline{y}(2-\underline{y})} - \underline{y})\right) = \underline{V}_L.$$

It is smaller than η_S when $V < \eta_L\left(-(\eta_S/\eta_L - \underline{y})^2 + 2(1-\underline{y})(\eta_S/\eta_L - \underline{y})\right) = \overline{V}_L$. Hence for V in this range and $\theta = 1$ S is advantaged.

We observe that when $\underline{y} = 0$ we have $\overline{V}_S = 1, \underline{V}_S = 0, \overline{V}_L = \eta_S\left(2 - \eta_S/\eta_L\right) > \eta_S$ and $\underline{V}_L = 0$. This establishes that for sufficiently small $\underline{y}$ there is a range of $V's$ such that $\eta_L\underline{y} < \overline{b}_S < \eta_S$ for $1/2 \leq \theta \leq 1$ and such that tS is advantaged at $\theta = 1$. Fix such a V.

Define the desire to bid as the solution of

$$(1-2\theta)(b_k/\eta_k - \underline{y})^2 + 2\theta(1-\underline{y})(b_k/\eta_k - \underline{y}) = V/\eta_k$$

and for S at least this is also the willingness to bid, and it will be the willingness to bid of L provided the constraint $b_L < \eta_L$ is satisfied. Since the last equation is quadratic in b_k it can be solved by the quadratic formula from which it is apparent that $\overline{b}_k(\theta)$ is a continuous function. This implies as well that the strategies are continuous in θ, since the support of the continuous part of the density is continuous as is the upper bound. We can also conclude that $\overline{b}_S = \overline{b}_L = b$ if and only if

$$(1-2\theta)(b - \eta_S\underline{y})^2 + 2\theta(1-\underline{y})\eta_S(b - \eta_S\underline{y}) = \eta_S V$$

and

$$(1-2\theta)(b - \eta_S\underline{y})^2 + 2\theta(1-\underline{y})\eta_S(b - \eta_S\underline{y}) - \eta_S V =$$
$$(1-2\theta)(b - \eta_L\underline{y})^2 + 2\theta(1-\underline{y})\eta_L(b - \eta_L\underline{y}) - \eta_L V.$$

The latter equation is linear in b since the b^2 terms are the same on both sides. Hence the equation has a unique solution $b(\theta)$ which is a rational function of

θ. Substituting that into the first equation we find that those values of θ for which $\overline{b}_S = \overline{b}_L$ are zeroes of a rational function. Hence, either there must be a finite number of zeroes or the function must be identically equal to zero. But it cannot be identically zero since $\overline{b}_S - \overline{b}_L$ is negative at $\theta = 1/2$ and positive at $\theta = 1$. We conclude that there is some point θ^* at which $\overline{b}_S = \overline{b}_L$ and S is advantaged for $\theta^* < \theta < \overline{\theta}$ for some $\overline{\theta}$, while L is advantaged for $\theta_0 < \theta < \theta^*$ for some θ_0.

Since C is incrementally concave in $\theta^* < \theta < \overline{\theta}$ and S is advantaged there, it follows that S follows a strategy that FOSD that of Ly. Hence in the limit at θ^* the strategy of the small party either FOSD that of the large party or is the same as that of the large party. However, for $\theta > \theta^*$, S plays $\eta_L \underline{y}$ with probability zero while L plays it with probability

$$1 - \frac{C((b_L/\eta_S))}{\eta_S V} + \frac{C((\eta_L/\eta_S)\underline{y})}{\eta_S V} \to \frac{C((\eta_L/\eta_S)\underline{y})}{\eta_S V} > 0$$

so in the limit the two strategies are not identical. Since at θ^* the strategy of S FOSD that of L, it has a strictly higher probability of winning and strictly higher expected turnout. Since the probability of winning and expected turnout are continuous functions of the strategies which are continuous in θ it follows that this remains true in an open neighborhood of θ^*. □

When is Advantage Strong?

Recall the definition of $\underline{\lambda}, \overline{\lambda}$

$$\lambda(c) = -\frac{(1 - G(c))\, g'(c)}{(g(c))^2}$$

with $\underline{\lambda} = \min_{c \geq 0} \lambda(c) \leq 0$ the smallest possible value of $\lambda(c)$ and $\overline{\lambda} = \max_{c \geq 0} \lambda(c) \geq 0$ the largest. Recall also the definition of $\underline{\gamma}, \overline{\gamma}$

$$\underline{\gamma} = \frac{1}{2 - \underline{\lambda}}$$

and define $\overline{\gamma} = 0$ if $\overline{\lambda} > 1$ and

$$\overline{\gamma} = 1 - \frac{1}{2 - \overline{\lambda}}$$

otherwise. Hence $\underline{\gamma}$ is an increasing function of $\underline{\lambda}$ and $\overline{\gamma}$ is a decreasing function of $\overline{\lambda}$. The properties of $\underline{\gamma}$ and $\overline{\gamma}$ follow directly from the properties of $\underline{\lambda}$ and $\overline{\lambda}$. Here we restrict attention to the single-peaked case.

Proposition 3. *a. If the density shifts to the right then $\overline{\lambda}$ is constant and $\underline{\lambda}$ decreases ($\underline{\gamma}$ decreases); if the density shifts to the right holding fixed $c(1)$ then $\overline{\lambda}$ decreases ($\overline{\gamma}$ increases);*

b. Increasing dispersion by a change of scale around the mode increases $\underline{\lambda}$ ($\underline{\gamma}$ increases) and decreases $\overline{\lambda}$ ($\overline{\gamma}$ increases).

Proof. (a) We consider first the case of shifting the density to the right holding fixed $c(1)$. The only interesting case is when the peak c_g is interior, that is, satisfies $c(1) > c_g > 0$. Consider a $h(c)$ also with upper support $c(1)$ with mode c_h. Suppose that for some positive constants Δ, ζ we have $c_h > c_g + \Delta$ and for $c > c_h$ we have $h(c) = \zeta g(c - \Delta)$ (density shifts right). Notice the scaling factor ζ is needed since holding fixed the upper bound $c(1)$ mass is lost as we shift g to the right. We prove that $\overline{\lambda}_h < \overline{\lambda}_g$.

Notice that since by assumption of a single peak $g'(c_g) = h'(c_h) = 0$, we can define $\overline{\lambda}$ without loss of generality only for values of c to the right of the mode. Hence we have that

$$\overline{\lambda}_h = \max_{c(1) \geq c \geq c_h} -\frac{\int_c^{c(1)} h(\xi) d\xi h'(c)}{(h(c))^2} = \max_{c(1) \geq c \geq c_g + \Delta} -\frac{\int_c^{c(1)} \zeta g(\xi - \Delta) d\xi \zeta g'(c - \Delta)}{(\zeta g(c - \Delta))^2}$$

and after a change of variable $\tilde{c} = c - \Delta$ we have

$$= \max_{c(1) - \Delta \geq \tilde{c} \geq c_g} -\frac{\int_c^{c(1) - \Delta} g(\tilde{\xi}) d\tilde{\xi} g'(\tilde{c})}{(g(\tilde{c}))^2} < \max_{c(1) \geq c \geq c_g} -\frac{\int_c^{c(1)} g(\tilde{\xi}) d\tilde{\xi} g'(\tilde{c})}{(g(\tilde{c}))^2} = \overline{\lambda}_g.$$

This gives the result for fixed $c(1)$. Focus on the key result

$$\overline{\lambda}_h = \max_{c(1)-\Delta \geq \tilde{c} \geq c_g} -\frac{\int_c^{c(1)-\Delta} g(\tilde{\xi}) d\tilde{\xi} g'(\tilde{c})}{\left(g(\tilde{c})\right)^2}; \overline{\lambda}_g = \max_{c(1) \geq c \geq c_g} -\frac{\int_c^{c(1)} g(\tilde{\xi}) d\tilde{\xi} g'(\tilde{c})}{\left(g(\tilde{c})\right)^2}$$

For $\overline{\lambda}$ there are two effects of a right shift: the range over which the integral of $g(\tilde{\xi})d\tilde{\xi}g'(\tilde{c})$ in the numerator is taken is shorter for h and the maximum is taken over a narrower range. There is no analagous result for $\underline{\lambda}$. For $\underline{\lambda}$ the range of the integral remains the same, but rather than a maximum over $c(1) - \Delta \geq \tilde{c} \geq c_g$ we take minimum over $c_g \geq \tilde{c} \geq c(0) - \Delta$. Hence the minimum is taken over a larger range, offsetting the effect of the shorter range of the integral and the combination of the two is ambiguous.

For an ordinary right shift (that is, not holding fixed $c(1)$) the range of the integral does not change. For $\overline{\lambda}$ the range over which the maximum is is taken does not change, so the right shift is neutral. For $\underline{\lambda}$ the range over which the minimum is taken increases so the minimum becomes more negative.

(b) We first prove the result for $\overline{\lambda}$. Consider a $h(c)$ also with upper support $c(1)$ with mode $c_h = c_g$. Suppose that for some positive constants $\sigma > 1, \zeta$ for $c > c_h$ we have $h(c) = \zeta g(c_g + (c - c_g)/\sigma)$ (greater dispersion to the right of the mode).

We have

$$\overline{\lambda}_h = \max_{c(1) \geq c \geq c_g} -\frac{\int_c^{c(1)} h(\xi) d\xi h'(c)}{\left(h(c)\right)^2}$$

$$= \max_{c(1) \geq c \geq c_g} -\frac{\int_c^{c(1)} \zeta g(c_g + (\xi - c_g)/\sigma) d\xi (1/\sigma) \zeta g'(c_g + (c - c_g)/\sigma)}{\left(\zeta g(c_g + (c - c_g)/\sigma)\right)^2}$$

and after a change of variable $\tilde{c} = c_g + (c - c_g)/\sigma$ we have

$$= \max_{c_g + (c(1) - c_g)/\sigma \geq \tilde{c} \geq c_g} -\frac{\int_c^{c_g + (c(1) - c_g)/\sigma} g(\tilde{\xi}) d\tilde{\xi} g'(\tilde{c})}{\left(g(\tilde{c})\right)^2}.$$

Furthermore, since $\sigma > 1$ and $c_g < c(1)$ we have $c_g + (c(1) - c_g)/\sigma = [(\sigma -$

$1)/\sigma]c_g + [1/\sigma]c(1) < c(1)$ so

$$\overline{\lambda}_h < \max_{c(1)\geq c\geq c_g} -\frac{\int_c^{c(1)} g(\tilde{\xi})d\xi g'(\tilde{c})}{(g(\tilde{c}))^2} = \overline{\lambda}_g.$$

Here again there are two effects, a shorter range of integral and a shorter range over which the maximum is taken, both lowering $\overline{\lambda}$. In the case of $\underline{\lambda}$ it is also the case that both the range of the integral and range over which the minimum is taken shrink: hence the minimum must increase. $\square$

Theorem 4 is equivalent to

Proposition 4. *a. cost is incrementally convex if and only if $\theta < 1/(2-\underline{\lambda})$*
b. cost is incrementally concave if and only if $\overline{\lambda} < 1$ and $\theta > 1/(2-\overline{\lambda})$.

Proof. We report expected cost $C(\varphi) = \int_{\underline{y}}^{\varphi} c(y)dy + \theta(1-\varphi)c(\varphi)$ and its first two derivatives $C'(\varphi) = (1-\theta)c(\varphi) + \theta(1-\varphi)c'(\varphi)$, and $C''(\varphi) = (1-2\theta)c'(\varphi) + \theta(1-\varphi)c''(\varphi)$. Observe that $c(\varphi) = G^{-1}(\varphi)$ so

$$c'(\varphi) = \frac{1}{g(G^{-1}(\varphi))}$$

$$c''(\varphi) = -\frac{g'(G^{-1}(\varphi))}{(g(G^{-1}(\varphi)))^3}$$

and hence we can rewite $C''(\varphi)$ as

$$C''(\varphi) = \frac{1-\theta(2-\lambda(c))}{g(c)}.$$

Hence $C''(\varphi) > 0$ if and only if $\theta < 1/(2-\lambda(c))$ from which the result follows. $\square$

Proposition 5. [Example in text] *Suppose participations costs c are normalized to lie in $[0,1]$ and have a density function $g(c) = \alpha c^{\alpha-1}$ where $\alpha > 0$. If $\alpha < 1$ then $\underline{\gamma} = 1/2$ and $\overline{\gamma} = 0$. If $\alpha > 1$, then $\underline{\gamma} = 0$ and $\overline{\gamma} = 1/2$.*

Proof. For $\alpha < 1$ the density is decreasing so $\underline{\lambda} = 0$ and for $\alpha > 1$ it is increasing so $\overline{\lambda} = 0$. We have

$$\lambda(c) = -(\alpha - 1)\frac{(1 - c^{\alpha})\,\alpha c^{\alpha-2}}{(\alpha c^{\alpha-1})^2} = -\frac{\alpha - 1}{\alpha}\,(1 - c^{\alpha})\,c^{-\alpha}$$

which goes to infinity in absolute value as $c \to 0$. Hence $\overline{\lambda} = \infty$ and $\underline{\lambda} = -\infty$ giving the required result. □

High and Low Value Elections

Theorem. [Theorem 5 in text] *In a high value election the probabilities that the small party concedes and the large party preempts the election increase in* V*, and approach* 1 *in the limit. As* V *increases the bid distribution of the small party declines in FOSD and the bid distribution of the large party increases in FOSD. The expected vote differential increases in* V *while the expected turnout cost remains constant.*

Proof. In a high value election S is constrained and L is advantaged. The probability L preempts is $F_L^0(\eta_S) = 1 - (\eta_S/V)C(1)$, increasing in V. The probability of concession by S is $F_S^0(\eta_S\underline{y}) = 1 - \eta_L C(\eta_S/\eta_L)/V$ increasing in V.

Since changing V with $\overline{b}_d = \eta_S$ the support and shape of the cost function in the mixing range do not change, so raising V simply lowers the densities by a common factor, meaning that these shifts reflect stochastic dominance as well. The FOSD result implies the increased vote differential.

Total surplus is $V - \eta_{-d}C(\overline{b}_d/\eta_{-d})$. Since some party certainly gets the prize this implies the expected turnout cost is $\eta_{-d}C(\overline{b}_d/\eta_{-d})$ and in a high value election $\overline{b}_d$ remains constant at η_S, so expected turnout cost is $\eta_L C(\eta_S/\eta_L)$ independent of V . □

Disagreement

Theorem. [Theorem 6 in text] *If the advantaged party is strongly advantaged disagreement increases the peak turnout, the expected turnout cost and decreases the bid differential.*

Proof. The case in which $C(1)\eta_L \leq V$ is immediate since $\overline{b}_S = \eta_S < \eta_L = \overline{b}_L$ and the result follows. If instead, $C(1)\eta_S > V$, neither party is constrained.

Given the definition of willingness to bid $\eta_k C(\bar{b}_k/\eta_k) - V = 0$, we can apply the implicit function theorem and find that

$$\frac{d\bar{b}_k}{d\eta_k} = -\frac{C(\bar{b}_k/\eta_k) - \hat{b}_k C'(\bar{b}_k/\eta_k)/\eta_k}{C'(\bar{b}_k/\eta_k)} =$$
$$(\bar{b}_k/\eta_k)\frac{C'(\bar{b}_k/\eta_k) - C(\bar{b}_k/\eta_k)/(\bar{b}_k/\eta_k)}{C'(\bar{b}_k/\eta_k)}.$$

If C is convex, $\bar{b}_L > \bar{b}_S$, and marginal cost is larger than average cost for both parties. As a result, we have that

$$\frac{d\bar{b}}{d\eta_k} > 0.$$

Hence increasing η_S and decreasing η_L implies that the bid differential decreases and peak turnout increases. If instead C is concave and S is advantaged (that is, $\bar{b}_S > \bar{b}_L$), marginal cost is smaller than average cost for both parties. As a result

$$\frac{d\bar{b}_k}{d\eta_k} < 0.$$

Hence increasing η_S and decreasing η_L implies that the bid differential increases and peak turnout increases. Total surplus is $V - \eta_{-d}C(\bar{b}_d/\eta_{-d})$. Since one party gets the prize for certain, expected turnout cost is $\eta_{-d}C(\bar{b}_d/\eta_{-d})$. Differentiate this with respect to η_{-d} to find

$$C(\bar{b}_d/\eta_{-d}) - (\bar{b}_d/\eta_{-d})C'(\bar{b}_d/\eta_{-d}) + C'(\bar{b}_d/\eta_{-d})d\bar{b}_d/d\eta_{-d}$$
$$(\bar{b}_d/\eta_{-d})\left(\frac{C(\bar{b}_d/\eta_{-d})}{\bar{b}_d/\eta_{-d}} - C'(\bar{b}_d/\eta_{-d})\right) + C'(\bar{b}_d/\eta_{-d})d\bar{b}_d/d\eta_{-d}$$

If C is convex the bracketed expression is negative and $d\bar{b}_d/d\eta_{-d} \leq 0$, hence the entire expression is negative when disagreement decreases. If C is concave and S is advantaged the bracketed expression is positive and $d\bar{b}_d/d\eta_{-d} \geq 0$, hence the entire expression is positive when disagreement increases. □

Monitoring Difficulty in High Value Elections

Theorem. [Theorem 7 in text] *In a high value election, an increase in monitoring difficulty θ decreases the turnout of the advantaged (large) party in terms of FOSD. Furthermore, there exists $0 < \underline{\eta} < \overline{\eta} \leq 1/2$ such that for $\underline{\eta} < \eta_S < \overline{\eta}$ the expected turnout of the disadvantaged (small) party decreases in monitoring difficulty in terms of FOSD while the expected vote differential also decreases.*

Proof. If the election is not high value the disadvantaged party is unconstrained. Hence, given the definition of willingness to bid $\eta_k C(\overline{b}_k/\eta_k) - V = 0$, we can apply the implicit function theorem and find that

$$\frac{d\overline{b}_d}{d\theta} = -\frac{\eta_d dC(\overline{b}_d/\eta_d)/d\theta}{C'(\overline{b}_d/\eta_d)} = -\frac{\eta_d \theta(1-\overline{b}_d/\eta_d)c(\overline{b}_d/\eta_d)}{C'(\overline{b}_d/\eta_d)} < 0$$

Hence as θ decreases, that is as monitoring efficiency increases, so it does peak turnout. In a high value election the peak turnout $\overline{b}_d$ is fixed at η_S and S is disadvantaged. Examining the equilbrium bid distributions we have

$$F_S(b) = 1 - \frac{\eta_L C(\eta_S/\eta_L)}{V} + \frac{\eta_L C(b/\eta_L)}{V}$$

$$F_L(b) = \frac{\eta_S C(b/\eta_S)}{V}$$

while $C(\varphi_k) = T(\varphi_k) + \theta(1-\varphi_k)T^{'}(\varphi_k)$. Examining $F_L(b)$ first, we see that $dF_L/d\theta > 0$ which is the condition for a decrease in FOSD. For $F_S(b)$ we have

$$\frac{dF_S}{d\theta} = -\frac{\eta_L}{V}\Big((1-\eta_S/\eta_L)T^{'}(\eta_S/\eta_L) - (1-b/\eta_L)T^{'}(b/\eta_L)\Big)$$

Notice that for φ_k sufficiently close to $\underline{y}$ we must have $(1-\varphi_k)T^{'}(\varphi_k)$ increasing, say for $\underline{y} < \varphi_k < \varphi_0$. Hence for $\eta_S/\eta_L < \varphi_0$ we have $dF_S/d\theta < 0$ for $b \leq \eta_S$. This is the condition for an increase in FOSD. Since F_L stochastically dominates F_S and F_L decreases while F_S increases it follows that the expected vote differential must decrease.

Consider next that as $\eta_S \to 1/2$, it follows that $(1-\eta_S/\eta_L)T^{'}(\eta_S/\eta_L) \to 0$. Hence for any fixed b it is eventually true that $dF_S(b)/d\theta > 0$. It follows that, for sufficiently large η_S, the expected turnout of$S$ must decline with θ. Since the derivative of expected turnout is a continous function of θ, it follows that there is a value of $\underline{\eta}$ such that expected turnout of S is constant with θ while for larger η_S it declines. At $\underline{\eta}$ the expected vote differential must decline with θ since S expected turnout is constant and L expected turnout declines. Since the derivative of the expected vote differential is also continuous in η_S it follows that for η_S larger than but close enough to $\underline{\eta}$, S expected turnout declines and the expected vote differential does as well. $\square$

Symmetry of the Fundamentals

Let $\rho \in [0, 1]$ be a measure of the mix of issues between transfers and laws where $\rho = 0$ mans the election is purely about transfers and $\rho = 1$ means it is purely about laws. Examples of transfers include control over natural resources, the division of government jobs, the division of a fixed budget, taxes and subsidies and limitations on competition such as trade restrictions or occupational licensing. Examples of laws include civil rights, laws concerning abortion, criminal law, defense expenditures, non-trade foreign policy and policies concerning monuments. We suppose that $V_k = v(\eta_k, \rho)$ where $v(1/2, \rho) = V$. We take pure transfers to mean a common prize so that $v(\eta_k, 0) = V$ and pure laws to mean a common per capita prize so that $v(\eta_k, 1) = 2V\eta_k$. We assume that $v(\eta_k, \rho) \geq 0$ twice continuously differentiable with $v_\eta(\eta_k, \rho) \geq 0$. Define the prize elasticity with respect to party size $\gamma(\eta_k, \rho) = d(\log v(\eta_k, \rho)/d \log \eta_k) = v_\eta(\eta_k, \rho)\eta_k/v(\eta_k, \rho)$. Then for pure transfers we have $\gamma(\eta_k, 0) = 0$ for for pure laws we have $\gamma(\eta_k, 1) = 1$. It is natural to assume then that $\gamma_\rho(\eta_k, \rho) > 0$: that as the importance of laws as an issue increases the prize elasticity with respect to party size goes up. This implies in addition that $v_\rho(\eta_k, \rho) > 0$ for $\eta_k > 1/2$ and $v_\rho(\eta_k, \rho) < 0$ for $\eta_k < 1/2$. That is, as the importance of laws as an issue increases the value of prize to the large party goes up and to the small party goes down. It follows directly that increasing the importance of laws improves the advantage (positive or negative) of the large party by raising its willingness to bid and lowering that of the small party.

Example. Suppose that the election has a mix of transfer and legal issues so tht $v(\eta_k, \rho) = (1 - \rho) + 2\rho\eta_k$ where $0 \leq \rho \leq 1$ is the relative importance of legal issues. Then $\gamma(\eta_k, \rho) = 2\rho\eta_k / \left((1-\rho) + 2\rho\eta_k\right)$ and $\gamma(\eta_k, 0) = 0$, $\gamma(\eta_k, 1) = 1$ and the derivative is

$$\gamma_\rho(\eta_k, \rho) = \frac{2\eta_k\left((1-\rho) + 2\rho\eta_k\right) + 2\rho\eta_k(1-2\eta_k)}{\left((1-\rho) + 2\rho\eta_k\right)^2} > 0.$$

Proposition 6. *If $\rho > 1$ then there are cost functions, prize values, party sizes, and monitoring difficulty for which the small party is advantaged.*

Proof. Willingness to bid is $v(\eta_k, \rho) - \eta_k C(b_k/\eta_k) = 0$ or $1 - (\eta_k/v(\eta_k,\rho))C(b_k/\eta_k) = 0$. Using the implicit function theorem we find

$$db_k/d\eta_k = -\frac{(v(\eta_k, \rho) - v'(\eta_k, \rho)\eta)C(b_k/\eta_k)/v(\eta_k, \rho)^2 - (1/v(\eta_k, \rho))C'(b_k/\eta_k)(b_k/\eta_k)}{C'(b/\eta)/v(\eta_k, \rho)}$$

so that the sign determined by $C'(\varphi_k)\varphi_k - (1-\gamma(\eta_k, \rho))C(\varphi_k)$. If the parties are of near equal size and this is positive or $V > (1/2)C(1)$ then L is advantaged, if the parties are of near equal size, $V < (1/2)C$ and this is negative, S is advantaged. If $\rho = 0$ so the election is purely about transfers then this is $C'(\varphi_k)\varphi_k - C(\varphi_k)$ so which party is advantaged depends on whether average cost is increasing or decreasing as we know. If $\rho = 1$ so the election is purely about laws this is $C'(\varphi_k)\varphi_k$ which is always positive so L is always advantaged. In the intermediate cases there are always parameter values for which S is advantaged. Take the quadratic case with no committed voters where $C(\varphi_k) = (1-2\theta)\varphi_k^2 + 2\theta\varphi_k$ At $\theta = 1$ this is $C(\varphi_k) = -\varphi_k^2 + 2\varphi_k$ and $C'(\varphi_k) = -2\varphi_k + 2$. Hence

$$C'(\varphi_k)\varphi_k - (1-\gamma(\eta_k, \rho))C(\varphi_k) = (-2\varphi_k + 2)\varphi_k - (1-\gamma(\eta_k, \rho))\left(-\varphi_k^2 + 2\varphi_k\right).$$

$$= -\varphi_k^2 + \gamma(\eta_k, \rho))\left(-\varphi_k^2 + 2\varphi_k\right)$$

$$= -(1 + \gamma(\eta_k, \rho))\varphi_k^2 + \gamma(\eta_k, \rho))2\varphi_k.$$

Notice that for positive $\gamma(\eta_k, \rho)$ and small φ_k this is necessarily positive. However, as $\varphi_k \to 1$ this approaches $-(1-\gamma(\eta_k, \rho))$ which is strictly negative for $\rho < 1$, so also for $\varphi_k < 1$ but close to 1. $\square$

The proof shows that with quadratic cost given $\rho < 1$ if there are sufficiently few committed voters, if $V < (1/2)C(1)$ but close enough (intermediate size prize), parties of similar enough size (small party not too small) and θ near enough 1 (high monitoring costs) the small party is advantaged. This is the same qualitatively as in the $\rho = 0$ case: however, quantitatively the criteria are much more stringent.

Endogenous versus Exogenous Uncertainty

Suppose that the probability of winning the election for party k is given by $P(b_k, b_{-k})$ non-decreasing in b_k. This must satisfy the identity $P(b_k, b_{-k}) = 1 - P(b_{-k}, b_k)$. Suppose there is a common prize the value of which we may normalize to 1 and common cost $C(\varphi)$. The objective function of party k is therefore $P(b_k, b_{-k}) - \eta_k C(b_k/\eta_k)$.

Proposition 7. *In any pure strategy equilibrium b_k, b_{-k} (if one exists) if $C''(\varphi) > 0$ then $b_L > b_S$ and the large party receives strictly greater utility than the small party; if $b_L \leq \eta_S$ and $C''(\varphi < 0)$ then $b_S > b_L$ and the small party receives strictly greater utility than the large party.*

Proof. In the convex case if $b_L > \eta_S$ then certainly L turns out more than Sy, so in both cases we may assume $b_L \leq \eta_S$. Consider that the utility to party k from playing b_{-k} rather than b_k must not yield an improvement in utility. That is

$$P(b_k, b_{-k}) - \eta_k C(b_k/\eta_k) \geq (1/2) - \eta_k C(b_{-k}/\eta_k)$$

or

$$P(b_k, b_{-k}) - (1/2) \geq \eta_k C(b_k/\eta_k) - \eta_k C(b_{-k}/\eta_k).$$

For party $-k$ this reads

$$P(b_{-k}, b_k) - (1/2) \geq \eta_{-k} C(b_{-k}/\eta_{-k}) - \eta_{-k} C(b_k/\eta_{-k})$$

and using $P(b_{-k}, b_k) = 1 - P(b_k, b_{-k})$

$$(1/2) - P(b_k, b_{-k}) \geq \eta_{-k} C(b_{-k}/\eta_{-k}) - \eta_{-k} C(b_k/\eta_{-k})$$

or

$$P(b_k, b_{-k}) - 1/2 \le \eta_{-k}C(b_k/\eta_{-k}) - \eta_{-k}C(b_{-k}/\eta_{-k})$$

so the inequalities for the two parties are

$$\eta_k C(b_k/\eta_k) - \eta_k C(b_{-k}/\eta_k) \le \eta_{-k}C(b_k/\eta_{-k}) - \eta_{-k}C(b_{-k}/\eta_{-k}).$$

Suppose without loss of generality that $b_k \ge b_{-k}$ so both sides are non-negative.

We work through the convex case. If $k = S$ we see that we must have

$$\eta_S C(b_k/\eta_S) - \eta_S C(b_{-k}/\eta_S) \le \eta_L C(b_k/\eta_L) - \eta_L C(b_{-k}/\eta_L).$$

Consider the function $\eta_k C(b_k/\eta_k) - \eta_k C(b_{-k}/\eta_k)$ and differentiate it with respect to η_k to find

$$C(b_k/\eta_k) - C(b_{-k}/\eta_k) - ((b_k/\eta_k)C'(b_k/\eta_k) - (b_{-k}/\eta_k)C'(b_{-k}/\eta_k))$$

which may also be written as

$$C(b_k/\eta_k) - (b_k/\eta_k)C'(b_k/\eta_k) - (C(b_{-k}/\eta_k) - (b_{-k}/\eta_k)C'(b_{-k}/\eta_k)).$$

Consider the function $C(\varphi) - \varphi C'(\varphi)$ and differentiate with respect to φ to find

$$-\varphi C''(\varphi) < 0.$$

This implies

$$C(b_k/\eta_k) - (b_k/\eta_k)C'(b_k/\eta_k) - (C(b_{-k}/\eta_k) - (b_{-k}/\eta_k)C'(b_{-k}/\eta_k)) < 0$$

which in turn implies

$$\eta_L C(b_k/\eta_L) - \eta_L C(b_{-k}/\eta_L) < \eta_S C(b_k/\eta_S) - \eta_S C(b_{-k}/\eta_S)$$

a contradiction, so we conclude that $k = L$, that is, $b_L > b_S$.

If $b_L > b_S$ suppose that L were to lower its bid to b_S. It would then have a

1/2 chance of winning - at least the equilibrium utility of S - and a cost lower than the equilibrium cost of S, so bidding b_S yields L more than the equilibrium utility of S. Hence the equilibrium utility of L must be larger than that of S.

Finally if $C(\varphi)$ is concave then the role of the two parties in determining the equilibrium bids is reversed, so we conclude that $b_S > b_L$. □

Bounded Costs

We compare two participation cost functions: $c(y)$, $\xi(y)$ where for some $\eta_S/\eta_L < \overline{y} < 1$ and $y \leq \overline{y}$ we have $c(y) = \xi(y)$ while for $\overline{y} < y \leq 1$ we have $c(y) < \xi(y)$. The cost function $c(y)$ is bounded, but we allow $\xi(1) = \infty$. It follows that the corresponding expected cost functions $C(y), \Xi(y)$ share the same property that $y \leq \overline{y}$ we have $C(y) = \Xi(y)$ while for $\overline{y} < y \leq 1$ we have $C(y) < \Xi(y)$ and $C(y)$ is bounded while $\Xi(y)$ need not be

Proposition 8. *If c has high stakes so $V > \max\{\eta_L C(\eta_S/\eta_L), \eta_S C(1)\}$ and ξ has high costs $\Xi(1) > V/\eta_S$ then the large party is advantaged. The equilibrium strategies and payoffs of the small party are the same for c, ξ. For the large party for low bids $b \leq \eta_S\overline{y}$ the strategies are the same for c, ξ. The probability of a high bid under ξ is approximately the same as the atom at η_S under c*

$$\left[1 - F_L^{\xi}(\eta_S\overline{y})\right] - F_L^{0c}(1) = \eta_S\left[C(1) - C(\overline{y})\right]/V$$

as are the equilibrium payoffs

$$\eta_L\left(C(1) - C(\overline{y})\right) > U_L^{\xi} - U_L^{c} > 0.$$

Proof. As L never bids more than η_S and $\overline{y} > \eta_S/\eta_L$ only c is relevant for computing the payoffs of L; this implies in particular that the strategy of S is the same for c or ξ. Moreover, L is advantaged for both c, ξ. This follows from $V > \eta_L C(\eta_S/\eta_L)$ meaning L is willing to bid more than η_S which is the most S can bid. Since L is advantaged for c, ξ, S gets 0 in either case. For L bids below $\eta_S\overline{y}$ we have $F_L^c(b) = \eta_S C(b/\eta_S)/V = \eta_S\Xi(b/\eta_S)/V = F_L^{\xi}(b)$.

Under c, S is willing to bid η_S (by high stakes) while under χ, S is willing to bid $\eta_S\overline{y} < \overline{b}_S < \eta_S$. The first part $\eta_S\overline{y} < \overline{b}_S$ follows from $V - \eta_S\Xi(\overline{y}) =$

$V - \eta_S C(\overline{y}) > V - \eta_S C(1) > 0$ and the second part $\overline{b}_S < \eta_S$ follows from the high cost assumption $V - \eta_S \Xi(1) < 0$.

We now compute the probability L makes a high bid $1 - F_L^{\xi}(\eta_S \overline{y})$. Since $F_L^{\xi}(\eta_S \overline{y})V - \eta_S D(\overline{y}) = 0$ we have $1 - F_L^{\xi}(\eta_S \overline{y}) = 1 - \eta_S C(\overline{y})/V$. By contrast $F_L^{0c}(1)$ satisfies $(1 - F_L^{0c}(1))V - \eta_S C(1) = 0$ so $F_L^{0c}(1) = 1 - \eta_S C(1)/V$. These two give the desired result

$$\left[1 - F_L^{\xi}(\eta_S \overline{y})\right] - F_L^{0c}(1) = \eta_S \left[C(1) - C(\overline{y})\right]/V$$

Finally we compute $U_L^{\xi} - U_L^c = \eta_L \left(C(1) - C(\overline{b}_S/\eta_S)\right)$. Hence indeed $\eta_L \left(C(1) - C(\overline{y})\right) > U_L^{\xi} - U_L^c > 0$ $\square$