Razor-Thin Mass Elections with High Turnout

David K. Levine[1], Cesar Martinelli[2]

Abstract

We argue that standard models of voting do a bad job explaining the frequency of very close mass elections with high turnout. We instead model head-to-head elections as a competition between incentive schemes to turn out voters and elucidate conditions under which parties might prefer close elections in which voters are motivated by pivotality rather than providing voters with costly incentives to turn out in an election that is not close. When this is the case, we show that better targeting of voters results in closer votes and higher turnouts and that the smaller of the two parties has a strong incentive to engage in commitment that will drive a close election.

Email addresses: david@dklevine.com (David K. Levine), cmarti33@gmu.edu (Cesar Martinelli)

[1]Department of Economics, EUI and WUSTL

[2]Department of Economics, George Mason University

Acknowledgements: We would like to thank Andrea Mattozzi, Barton Lee, and seminar participants at Carlos III, the RIDGE-LACEA political economy workshop, and the 5th ETH Workshop on Democracy. We gratefully acknowledge support from the MIUR PRIN 2017 n. 2017H5KPLL_01.

1. Introduction

This paper is motivated by two consecutive presidential runoff elections in 2016 and 2021 in Peru, a country with a population of about 35 million. In these elections the margin between the two candidates was 41,057 and 44,263 voters, respectively, out of 17.1 and 17.6 million valid votes.[3] In other words in a large mass election with millions of voters there was high turnout and a razor thin margin deciding the result. There are other examples: the 2004 Washington state governor's race was decided by 129 votes out of 2.9 million votes cast and the 2008 Minnesota senate race was decided by 225 votes out of 2.4 million.

We will argue that standard models of pivotal, ethical, peer pressure, voter mobilization, and expressive voters have a hard time explaining the conjunction of large turnout and razor thin margins in mass elections. We suggest that instead close outcomes might be the deliberate result of political parties' choices. We introduce a model in which pivotality is a substitute for costly peer pressure and show that indeed parties may prefer close elections.

Specifically we introduce a model in which a political party might prefer a close election to turning out a few extra voters at minimal cost to win the election. Our model explicitly models elections as competition between incentive schemes. One way of turning out voters is through costly monitoring and peer pressure, as discussed in Levine and Mattozzi (2020). Alternatively voters may turn out because they are pivotal, as in Palfrey and Rosenthal (1985). Our central contention is that these are substitutes. While turning out a few extra voters may guarantee a win it also assures that voters are not pivotal and increases costs substantially by requiring costly monitoring and peer pressure: hence—depending on circumstances in a way which we elucidate—a close election may be preferable.

An important conclusion from this analysis is that when close elections are preferred better targeting of voters that reduces the noise in election outcomes results in closer votes and higher turnouts. This potentially can explain why close outcomes seem to have become more common in recent decades. A second conclusion from the analysis is that the smaller of the two parties has a strong incentive to make a credible commitment that will drive a close election.

Are existing models adequate?

Models of pivotality, such as that of Palfrey and Rosenthal (1985), have fallen into disfavor as an explanation of participation in large mass elections. Although they do predict close elections, they do not contend well with high turnout: this is the conclusion of the literature on the paradox of voting starting with Downs (1957). This literature argues that models of pivotality cannot generate the large turnouts seen if practice, or if large turnout is due to committed or expressive voters, then pivotality ceases to matter and close margins are merely a coincidence of both parties having equal strength. Specifically,

[3] The official tally is available at https://www.onpe.gob.pe/elecciones/historico-elecciones/

Mulligan and Hunter (2003) use a coin toss model to argue that empirically pivotality is too small to matter. While Coate, Conlin and Moro (2008) argue that pivotality predicts much closer election results than are observed in the data.

As pivotal voter models cannot contend with high turnout, attention has turned to models of voter mobilization such as Shachar and Nalebuff (1999), ethical voters such as Feddersen and Sandroni (2006) or Coate and Conlin (2004) or the related model of peer pressure such as Levine and Mattozzi (2020). However, these models either have pure strategies due to noise that makes close elections unlikely or if noise is small mixed strategy equilibria that also makes close elections unlikely. In the final section of the paper we argue that these models do a poor job empirically of explaining the frequency of close elections.

An alternative type of model are spatial models of platform convergence such as those of Hotelling (1929) and Downs (1957). These predict that parties should be of equal size. They also counterfactually predict platform convergence. Morever, these models are not designed to predict voter turnout, the usual assumption being that all voters turn out, which is not only counterfactual but should result in a tie every time. We also question whether the adjustments taking place in the runup to an election are really due to platform adjustment rather than electoral effort. For example, in Peru 2021 the first preelection polls initially indicated 41.5% for Castillo and 21.5% for Fujimori. During the following two months, up to the election, the voting intention for Castillo hovered around 41% while the voting intention for Fujimori, remarkably, crawled to tie that of Castillo, who ended up winning the election.[4] It seems highly implausible that the parties adjusted their platforms up to the last minute until they converged to the median voter's ideal policy.

Our model also has implications for the likelihood of winning when elections are close: it predicts that each party should have an equal chance of winning. There is an empirical debate, for example, Vogl (2014) and De la Cuesta and Imai 2016, about whether this is the case and the model gives theoretical reason to believe that there should be no bias.[5]

We begin with a stark sequential move model in which one party commits first to a turnout target and the other observing that target responds with its own turnout: our goal is to elucidate when the second mover might prefer a close election. Subsequently we extend this result to the standard model where both parties simultaneously choose turnout targets. Finally, we examine empirically whether non-pivotal models of turnout are consistent with the number of close elections. To do this we need data on a substantial number of broadly comparable elections. We chose US gubernatorial and senatorial elections since 1995 as satisfying these criteria and show that indeed there are more close elections

[4]IEP Informes de Opinión, https://iep.org.pe/noticias/informes-de-opinion/

[5]There is also a literature using close elections for regression discontinuity analysis, discussed, for example, in Caughey and Sekhon (2011) but while relevant to the issue of biases it does not address the issue of how frequent are close elections.

than predicted by non-pivotal turnout models.

2. The Model

Voters are divided into two parties $k = \{1, 2\}$. There are N individuals and party k has $\eta_k N$ members where $\eta_k \geq 0$ and $\eta_1 + \eta_2 = 1$. The parties compete for a common prize worth 1 per capita by turning out their members. This turnout is determined by targets set by the parties, by random noise, and by the individual decisions of members. The party that turns out the most voters wins the prize.

Party turnout is set in two stages. Each party chooses a target fraction of voters to turn out. One party, chosen at random, which without loss of generality we may designate as party 1, or the *first mover*, becomes committed to its target fraction $1 \geq \phi_1 \geq 0$ first. The other party denoted as party 2 and not yet committed is the *second mover*. The second mover observes the commitment of the first mover and chooses a target number of additional members γh_2 to turn out, aiming for a total of $\eta_1 \phi_1 N + \gamma h_2$ voters. Here h_2 is an integer chosen by the second mover and γ is a positive integer called the *granularity*. This reflects the idea that while the second mover can attempt to turn out more or less voters than the first mover, they can only choose between discrete levels of effort. For example, if $\gamma = 12$ then the second mover can choose to match the first party with $h_2 = 0$, or overmatch them with $\gamma h_2 = 12, 24...$ but cannot be so precise as to try to overmatch them by exactly 6 votes. The actual turnout of each party k is equal to targeted turnout plus a random fraction of additional voters ζ_k exponentially distributed with mean σ.

There are two costs of turning out a fraction of voters: *participation costs*, such as time and inconvenience, and the *incentive costs* of getting voters to turn out. Both parties face the same participation costs. We order members so that higher numbered members have a higher participation cost and assume that member i_k in party k faces a participation cost that increases linearly with their percentile in the party $i_k/(\eta_k N)$ so is given by $ci_k/(\eta_k N)$, where c is a positive constant.

Our core assumption is that each party is able to design a mechanism providing incentives to individual members by using social pressure to punish those who fail to do their duty. When the turnout target is ϕ_k the mechanism specifies that members with the lowest costs should vote: that is, member i_k should vote if $i_k/(\eta_k N) \leq \phi_k$. Whether the rule was complied with or not is only imperfectly observed, however. Whether or not a party member voted is perfectly observed, and for those who voted it is perfectly observed whether or not they did what they were supposed to do so. However, there is imperfect information about whether non-voters were supposed to vote. Specifically, for some $\mu > 0$ we assume that it is impossible to determine if a member with $i_k/(\eta_k N) \leq \phi_k + \mu$ was supposed to vote or not. In order for voting to be incentive compatible, all these members are punished with an endogenous utility penalty P_k. Punishment takes place before the results of the election are known. The incentive cost to the party of issuing the punishment P_k is ψP_k where $\psi > 0$.

To reflect the incentives of members, we require that the punishment schemes used by the parties satisfy the interim incentive compatibility constraints that party members who are supposed to vote are willing to do so. Given this, parties attempt to maximize the per capita expected value of winning the prize less the expected per capita costs of turning out voters. Our solution concept is subgame perfect Nash equilibrium.

3. The Main Result

Our main result gives simple sufficient conditions under which these equilibria have high turnout and thin margins: this is in answer to the main question of the paper.

Define $h_1 = -h_2$ and observe that that the actual vote differential is the intended differential γh_k minus the difference in added voters $\xi_k = \zeta_{-k} - \zeta_k$. Because the added voters ζ_k are independent exponentially distributed with mean σ, their difference has the Laplace distribution with cumulative distribution function symmetric around zero given by $F(\xi) = 1 - (1/2)e^{-\xi/\sigma}$ for $\xi \geq 0$ and $F(\xi) = (1/2)e^{\xi/\sigma}$ for $\xi \leq 0$. Hence the probability that k wins the election is $F(\gamma h_k)$ and the *pivotality*, the probability that a voter in party k who is supposed to vote but does not will tip the election from a win to a loss, is $\pi(\gamma h_k) = F(\gamma h_k) - F(\gamma h_k - 1)$.

Theorem: High Turnouts and Razor Thin Margins. *For given η_k, μ, σ and $\gamma \geq 2$ there exists $\underline{c}, \underline{\psi}$ such that for $c \geq \underline{c}, \psi \geq \underline{\psi}$ there exists an $\underline{N}$ such that for $N \geq \underline{N}$ the first mover commits to a target $\hat{\phi}_1$ and the second mover responds with $h_2 = 0$. Moreover as $N \to \infty$*

$$\eta_1 \hat{\phi}_1 = \frac{F(\gamma) - F(0)}{\mu \psi c} + \frac{\pi(\gamma)}{c},$$

the distribution of vote margins in proportion to turned out voters converges in probability to zero, and pivotality $\pi(\gamma)$ and $F(\gamma) - F(0)$ are decreasing and absolute vote margins are increasing in σ.

Proof. We prove the final two statements with the main proof coming subsequently. With $h_2 = 0$, the vote margin is distributed as Laplacean with scale σ. This immediately implies that absolute vote margins are increasing in σ. Expected turnout converges to

$$\left(\frac{F(\gamma) - F(0)}{\mu \psi c} + \frac{\pi(\gamma)}{c} \right) N,$$

so dividing the Laplacean random variable by this we see that the resulting distribution converges in probability to zero.

We have $\pi(\gamma) = F(\gamma) - F(\gamma - 1)$ which like $F(\gamma) - F(0)$ must decrease as the scale parameter σ increases. □

The first key point is that the turnout in percentage terms is at least $\pi(\gamma)/\eta_1 c$ for both groups even for very large N: this is the sense in which turnout is high. The second key point is that in large populations vote margins in proportion to turned out voters are very small. The third key point is that as σ decreases, meaning better targeting, turnout increases and the vote margins grow thinner. This is our first key comparative static result.

Before turning to the proof, we briefly discuss the role of the parameters. There are (apparently) two measures of targeting: γ the granularity and σ the noise and two measures of the social cost of punishment μ the fraction who are punished wrongly and ψ the direct social cost of punishment. Why this apparent redundancy?

Granularity is introduced because the proof of the main theorem fails with granularity $\gamma = 1$, otherwise, despite the plausibility, we would not have introduced it. The reason is this: due to a slight asymmetry between pivotality for small positive and small negative vote differentials the optimal target differential is one not zero. If we allowed this when the second mover reduced incentive costs to zero the first mover would still face positive incentive costs and when ψ was large would prefer to opt out. If granularity is as small as $\gamma = 2$ this cannot happen: zero is better than two. As this seems plausible we adopt this assumption. Note that larger granularity makes no difference: it is the noise σ that is important.

In the incentive cost the parameters μ and ψ always appear as a product, so there seems to be a redundancy. However, the two parameters play quite a different role in the analysis. The parameter μ, the fraction who are wrongly punished, is taken to be reasonably small so as not to overlap with the upper boundary on the population. This greatly simplifies computations. However, we cannot then assert "μ is large" hence we introduce also the direct cost ψ which can be large.

We turn now to proving the main theorem. The problem of finding equilibrium can be divided into two steps: first, punishments should be chosen optimally. Once this is done we are left with reduced form objective functions. The second step is to use these reduced form objective functions to find the equilibrium. First we find the reduced form objective functions. To provide a symmetric definition, define the fraction turned out by the second mover $\phi_2 = (\eta_1 N \phi_1 + \gamma h_2)/(\eta_2 N)$.

Reduced Form Theorem. *For given η_k, μ there exists $\underline{c}$ such that for any $c \geq \underline{c}$ the reduced form objective functions are given by*

$$U_k = F(\gamma h_k) - (1/2)c\eta_k(\phi_k)^2 - \mu\psi \max\{0, c\eta_k\phi_k - \pi(\gamma h_k)\}.$$

Proof. The direct cost of turning out a fraction ϕ_k of voters for party k in per capita terms is $(1/N)\int_0^{\eta_k N \phi_k}(ci/(\eta_k N))di = (1/2)\eta_k c\phi_k^2$. Observe that it is never optimal for party to incur a per capita cost greater than η_k which is the per capita value of the prize to each party member. Denote by S, L the small and the large party, respectively. Hence, if $c > 2/\left(\eta_S(1-\mu)^2\right)$ then no party chooses a turnout target $\phi_k > 1-\mu$. If $c > 2\eta_L/\eta_S^2$ then the large party will not choose a turnout target with more voters than the small party has. It follows that the incentive cost per party member is $\mu\psi P_k$. Consequently, the per capita objective function of party k is the probability of winning minus the direct cost, minus the incentive cost of punishment:

$$F(\gamma h_k) - (1/2)c\eta_k(\phi_k)^2 - \eta_k\mu\psi P_k.$$

Now we examine optimal punishments. Party k's marginal voter receives $-c\phi_k$ for voting and $-P_k - \pi(h_k)/\eta_k$ for not voting, so the incentive constraint is

$$P_k \geq c\phi_k - \pi(h_k)/\eta_k.$$

Optimal punishment must be as small as possible. However, we must take account of the fact that the lower bound on punishment might be negative due to pivotality. Hence we solve the incentive constraints with equality to get $\underline{P}_k$ and with the actual punishment being $P_k = \max\{0, \underline{P}_k\}$. Substituting this back into the objective function gives the reduced form objective function shown. □

Hereafter it is assumed that c is sufficiently large that the Reduced Form Theorem holds. The steps below complete the proof of Theorem: High Turnouts and Razor Thin Margins.

Proof. Since N is picked last, we assume it is sufficiently large in all steps.

First: The reduced form objective function of the second mover can be written as

$$F(\gamma h_2) - (1/2)c\eta_2\left(\frac{\eta_1 N\phi_1 + \gamma h_2}{\eta_2 N}\right)^2 - \mu\psi\max\left\{0, c(\eta_1\phi_1 + \frac{\gamma h_2}{N}) - \pi(\gamma h_2)\right\}. \tag{3.1}$$

Differentiate this with respect to h_2 to find

$$(1/\gamma)dU_2/dh_2 =$$

$$F'(\gamma h_2) - \frac{c}{N}\frac{\eta_1 N\phi_1 + \gamma h_2}{\eta_2 N} - \mu\psi\mathbf{1}\left(c(\eta_1\phi_1 + \frac{\gamma h_2}{N}) \geq \pi(\gamma h_2)\right)\left(\frac{c}{N} - \pi'(\gamma h_2)\right). \tag{3.2}$$

Second: For $\gamma h_2 \geq 1$ we calculate the needed derivative in the final expression above

$$\pi'(\gamma h_2) = -(1/(2\sigma))e^{-\gamma h_2/\sigma}\left(e^{1/\sigma} - 1\right).$$

It follows that if

$$\phi_1 \leq \tilde{\phi}_1 = \frac{\pi(0)}{c\eta_1},$$

there is a unique solution denoted by $\hat{h}_2$ in $h_2 \geq 0$ to

$$c(\eta_1\phi_1 + \frac{\gamma h_2}{N}) - \pi(\gamma h_2) = 0.$$

From

$$c\left(\eta_1\phi_1 + \frac{\gamma \hat{h}_2}{N}\right) - \pi(\gamma \hat{h}_2) = 0, \tag{3.3}$$

we see that the LHS is increasing in both $\hat{h}_2$ and ϕ_1 so it follows that as ϕ_1 rises $\hat{h}_2$ must fall. That is, $\hat{h}_2$ is decreasing in ϕ_1.

Third: If $\phi_1 \leq \tilde{\phi}_1$, then the optimal choice h_2^* in the region $h_2 \geq 0$ satisfies $h_2^* \in \left\{\lfloor \hat{h}_2 \rfloor, \lceil \hat{h}_2 \rceil\right\}$. For $0 \leq h_2 < \hat{h}_2$ we have

$$(1/\gamma)dU_2/dh_2 = (1/(2\sigma))e^{-\gamma h_2/\sigma} - \frac{c}{N}(\frac{\phi_1\eta_1}{\eta_2} + \frac{\gamma h_2}{\eta_2 N})$$

$$\geq (1/(2\sigma))e^{-\gamma \hat{h}_2/\sigma} - \frac{c}{N}(\frac{\phi_1\eta_1}{\eta_2} + \frac{\gamma \hat{h}_2}{\eta_2 N}).$$

From equation 3.3 and using $\pi(\gamma h_k) = (1/2)e^{-\gamma h_k/\sigma}\left(e^{1/\sigma} - 1\right)$, the right-hand side of the inequality is equal to

$$(1/(2\sigma))e^{-\gamma \hat{h}_2/\sigma} - \frac{1}{\eta_2 N}(1/2)e^{-\gamma \hat{h}_2/\sigma}\left(e^{1/\sigma} - 1\right),$$

which is positive for sufficiently large N: it follows that the optimum is at least $\lfloor \hat{h}_2 \rfloor$.

For $h_2 \geq \hat{h}_2$ we have

$$(1/\gamma)dU_2/dh_2 = (1/(2\sigma))e^{-\gamma h_2/\sigma} - \frac{c}{N}(\frac{\phi_1\eta_1}{\eta_2} + \frac{\gamma h_2}{\eta_2 N}) - \mu\psi\left(\frac{c}{N} + (1/(2\sigma))e^{-\gamma h_2/\sigma}\left(e^{1/\sigma} - 1\right)\right)$$

$$\leq (1/(2\sigma))e^{-\gamma h_2/\sigma} - \mu\psi\left((1/(2\sigma))e^{-\gamma h_2/\sigma}\left(e^{1/\sigma} - 1\right)\right)$$

$$= (1/(2\sigma))e^{-\gamma h_2/\sigma}\left(1 - \mu\psi\left(e^{1/\sigma} - 1\right)\right).$$

This is negative for $\mu\psi\left(e^{1/\sigma} - 1\right) > 1$; hence, the optimum is at most $\lceil \hat{h}_2 \rceil$ for $\psi > 1/(\mu(e^{1/\sigma} - 1))$.

Finally, utility from h_2^* is

$$F(\gamma) - (1/2)c\eta_2\left(\frac{\eta_1\phi_1}{\eta_2} + \frac{\gamma h_2^*}{\eta_2 N}\right)^2$$

while utility from $h_2 = 0$ is

$$F(0) - (1/2)c\eta_2 \left(\frac{\eta_1 \phi_1}{\eta_2}\right)^2,$$

so that for sufficiently large N the choice h_2^* is better except when $h_2^* = 0$.

Fourth: The optimum in $0 \geq \gamma h_2 \geq \log(1/\log N)$ is $h_2 = 0$. For $h_2 \leq 0$ we have $\pi(\gamma h_2) = (1/2)e^{\gamma h_2/\sigma}(1 - e^{-1/\sigma})$. Hence

$$\frac{c}{N} - \pi'(\gamma h_2) = \frac{c}{N} - (1/(2\sigma))e^{\gamma h_2/\sigma}(1 - e^{-1/\sigma})$$

so for $\gamma h_2 \geq \log(1/\log N)$ and sufficiently large N this is negative so from 3.2 U_2 is increasing in h_2 for sufficiently large N.

Fifth: Define $\hat{\phi}_1$ by

$$\eta_1 \hat{\phi}_1 = \frac{F(\gamma) - F(0) - (1/2)\frac{c\gamma^2}{\eta_2 N^2} - \mu\psi c\frac{\gamma}{N} + \mu\psi\pi(\gamma)}{\frac{c\gamma}{\eta_2 N} + \mu\psi c}.$$

Then at $\hat{\phi}_1$ and $h_2 = 0$ both parties get at least $F(\log(1/\log N))$ in utility.

Utility at $\hat{\phi}_1$ and $h_2 = 0$ is given by

$$U_1 = 1/2 - (1/2)c\eta_1\hat{\phi}_1^2$$

$$U_2 = 1/2 - (1/2)c\left(\frac{\eta_1}{\eta_2}\right)\eta_1\hat{\phi}_1^2,$$

for large enough ψ, since incentive costs are zero at $\hat{\phi}_1$ and $h_2 = 0$ for large enough ψ. Let $\nu = \max\{\eta_k/\eta_{-k}\}$: we must show

$$1/2 - (1/2)c\nu\eta_1\hat{\phi}_1^2 > F(\log(1/\log N))$$

For N large this means showing that

$$1/2 - (1/2)\frac{1}{c}\nu\eta_1\left(\frac{F(\gamma) - F(0)}{\eta_1\mu\psi} + \frac{\pi(\gamma)}{\eta_1}\right)^2 > 0$$

which is true for large enough c.

Sixth: When the first mover sets $\phi_1 \geq \hat{\phi}_1$ and the second mover chooses $h_2 \leq F(\log(1/\log N))$ the first mover gets less utility than at $\hat{\phi}_1$ and $h_2 = 0$. If the second mover chooses $h_2 \leq F(\log(1/\log N))$ and $\phi_1 \geq \hat{\phi}_1$ the greatest possible utility for the first mover is

$$1 - (1/2)c\eta_1\hat{\phi}_1^2 - \mu\psi\left(c\eta_1\hat{\phi}_1 - \pi((\gamma/2)(1/\log N)^{1/\sigma})\right) \tag{3.4}$$

as against

$$1/2 - (1/2)c\eta_1\hat{\phi}_1^2. \tag{3.5}$$

Observe that for large enough N and for $\psi \geq 1$,

$$\hat{\phi}_1 \geq \frac{\pi(\gamma)}{c\left(\frac{\eta_1}{\eta_2}\right)\left(\frac{\gamma}{\mu}\right) + \eta_1 c}.$$

Hence for sufficiently large ψ and N the former utility 3.4 is less than the latter 3.5

$$1 - (1/2)c\eta_1\hat{\phi}_1^2 - \mu\psi\left(c\eta_K\hat{\phi}_1 - \pi((\gamma/2)(1/\log N)^{1/\sigma})\right) < 1/2 - (1/2)c\eta_1\hat{\phi}_1^2.$$

Seventh: For $\phi_1 \leq \hat{\phi}_1$ the best response satisfies $h_2 \geq 0$. From FOUR if $h_2 < 0$ then it is also no greater than $(1/\gamma)\log(1/\log N))$ yielding the second mover no more utility than $F(\log(1/\log N))$. Since for fixed h_2 the second mover utility strictly decreases in ϕ_1, the same is true for the optimal h_2. Hence the second mover utility is bigger at ϕ_1 than at $\hat{\phi}_1$, but from FIVE the second mover utility at $\hat{\phi}_1$ is greater than $F(\log(1/\log N))$, so the optimum cannot be $h_2 \leq (1/\gamma)\log(1/\log N))$.

Summary: We know from SIX that the optimal commitment of the first mover $\phi_1 \leq \hat{\phi}_1$ and from SEVEN that the best response of the second mover in this range satisfies $h_2 \geq 0$. We now examine the best response h_2 of the second mover in this range showing that it weakly declines with ϕ_1 until $h_2 = 1$. It then remains constant until $\hat{\phi}_1$ is reached at which point it is $h_2 = 0$.

Eighth: As ϕ_1 increases from zero $\hat{h}_2$ decreases until $\hat{h}_2 = 0$ and h_2^* is optimal in this range. We know from TWO that $\hat{h}_2$ is decreasing in ϕ_1. At $\phi_1 = 0$ for large N,

$$c\frac{\gamma}{N} - \pi(\gamma) < 0$$

implies that $\hat{h}_2 > 1$. As ϕ_1 rises from 0, then, $\hat{h}_2$ decreases until it eventually reaches 0. We know from THREE that as long as $\hat{h}_2 \geq 0$, $h_2^* \geq 0$ is optimal.

Ninth: Denote the point $\underline{\phi}_1$ where $\hat{h}_2 = 1$. In the range $\underline{\phi}_1 \leq \phi_1 < \hat{\phi}_1$ the best response is $h_2 = 1$. At $\hat{\phi}_1$, $h_2 = 0$ is also a best response: since it is the unique best response for slighly higher ϕ_1 we assume that it must be chosen as well at $\hat{\phi}_1$.

At $h_2 = 1$ the second mover utility is given by

$$F(\gamma) - (1/2)c\eta_2\left(\frac{\eta_1 N\phi_1 + \gamma}{\eta_2 N}\right)^2 - \mu\psi c\frac{\eta_1 N\phi_1 + \gamma}{N} + \mu\psi\pi(\gamma).$$

As long as $c\eta_1\phi_1 - \pi(0) \leq 0$, the utility benefit over $h_2 = 0$ is given by

$$F(\gamma)-F(0)-(1/2)c\eta_2\left(\frac{\eta_1 N\phi_1+\gamma}{\eta_2 N}\right)^2+(1/2)c\eta_2\left(\frac{\eta_1 N\phi_1}{\eta_2 N}\right)^2-\mu\psi c\frac{\eta_1 N\phi_1+\gamma}{N}+\mu\psi\pi(\gamma)$$

$$=F(\gamma)-F(0)-(1/2)\frac{c}{\eta_2 N^2}\left(2\eta_1 N\phi_1\gamma+\gamma^2\right)-\mu\psi c\frac{\eta_1 N\phi_1+\gamma}{N}+\mu\psi\pi(\gamma),$$

which is decreasing in ϕ_1 and equal to zero at $\hat{\phi}_1$.

As $N\to\infty$

$$\hat{\phi}_1\to\frac{F(\gamma)-F(0)}{\eta_1\mu\psi c}+\frac{\pi(\gamma)}{c\eta_1},$$

so $\hat{\phi}_1<\tilde{\phi}_1$ for big enough ψ and the utility benefit calculation is exact.

Tenth: From EIGHT, the best response of the second mover is $h_2^*\geq 1$ for $\phi_1\leq\underline{\phi}_1$. Hence the utility of the first mover is given by

$$U_1=(1/2)e^{-\gamma h_2^*/\sigma}-(1/2)c\eta_1(\phi_1)^2-\mu\psi\max(0,c\eta_1\phi_1-\pi(-\gamma h_2^*)).$$

Let

$$\hat{U}_1=(1/2)e^{-\gamma\hat{h}_2/\sigma}-(1/2)c\eta_1(\phi_1)^2$$

and

$$\bar{U}_1=(1/2)e^{-\gamma\lfloor\hat{h}_2\rfloor/\sigma}-(1/2)c\eta_1(\phi_1)^2.$$

Note that $U_1\leq\bar{U}_1$ for $\phi_1\leq\underline{\phi}_1$ with equality if $h_2^*=\hat{h}_2$ and $c\eta_1\phi_1-\pi(-\gamma h_2^*)\leq 0$, and $\bar{U}_1=\hat{U}_1$ for $\phi_1\leq\underline{\phi}_1$ if $h_2^*=\hat{h}_2$. Moreover, choosing ϕ_1 such that $h_2^*\neq\hat{h}_2$ cannot be a best response, since choosing a slightly smaller ϕ_1 reduces participation and incentive costs for the first mover without changing the probabililty of winning and pivotality. We claim that $\hat{U}_1$ is strictly increasing in ϕ_1 for $0\leq\phi_1<\hat{\phi}_1$. To see this, differentiating $\hat{U}_1$ with respect to ϕ_1 we find

$$d\hat{U}_1/d\phi_1=-(1/2)(\gamma/\sigma)e^{-\gamma\hat{h}_2/\sigma}d\hat{h}_2/d\phi_1-c\eta_1\phi_1.$$

From the implicit function theorem and $c(\eta_1\phi_1+\frac{\gamma\hat{h}_2}{N})-\pi(\gamma\hat{h}_2)=0$,

$$d\hat{h}_2/d\phi_1=-\frac{c\eta_1}{\gamma c/N+(\gamma/(2\sigma))(e^{1/\sigma}-1)e^{-\gamma\hat{h}_2/\sigma}}.$$

Hence,

$$d\hat{U}_1/d\phi_1=(\gamma/(2\sigma))e^{-\gamma\hat{h}_2/\sigma}\frac{c\eta_1}{\gamma c/N+(\gamma/\sigma)(e^{1/\sigma}-1)e^{-\gamma\hat{h}_2/\sigma}}-c\eta_1\phi_1.$$

This is strictly positive for large N if

$$\phi_1 < \frac{1}{(\gamma/\sigma)(e^{1/\sigma} - 1)e^{-\gamma/\sigma}}.$$

Since for large N we have

$$\hat{\phi}_1 \to \frac{F(\gamma) - F(0)}{\eta_1 \mu \psi c} + \frac{\pi(\gamma)}{\eta_1 c},$$

we see that this is true for sufficiently large c.

Eleventh: We know that the optimal commitment satisfies $\phi_1 \leq \hat{\phi}_1$, and for every $\phi_1 \leq \underline{\phi}_1$ that may be a best response the utility is bounded above by $\hat{U}_1$ which is strictly increasing in ϕ_1. The last step is to note that the utility at $\hat{\phi}_1$ is equal to $\hat{U}_1$. □

4. Simultaneous Moves

We turn now to the situation where rather than one party committing both parties move simultaneously. Our goal is to establish that the smaller of the two parties has strong reason to avoid the simultaneous move game by making a credible commitment to $\hat{\phi}_S$.

Our simultaneous move results are driven by the proof of Theorem: High Turnouts and Razor Thin Margins. There we computed the best responses of the two parties. By intersecting these best responses we can use those results to establish the important fact that if the game is played simultaneous move, that is, both parties simultaneously choose a commitment ϕ_k, then there are pure strategy equilibria with high turnout and razor thin margins. After presenting this result we will discuss the connection to the sequential move case, implications for the efficiency of equilibrium and the incentives to commit. Recall that we use the subscripts S, L to refer to the small and large party respectively.

Theorem: Simultaneous Pure Strategy Equilibria. *For given η_k, μ, σ and $\gamma \geq 2$ there exists $\underline{c}, \underline{\psi}$ such that for $c \geq \underline{c}, \psi \geq \underline{\psi}$ there exists an $\underline{N}$ such that for $N \geq \underline{N}$ there exists $\hat{\phi}_L$, the same as in Theorem: High Turnouts and Razor Thin Margins, and $\overline{\phi} > \hat{\phi}_L$ such that there is a pure strategy equilibrium of the simultaneous move game if and only if $\eta_1\phi_1 = \eta_2\phi_2 \in [\eta_L\hat{\phi}_L, \eta_L\overline{\phi}]$.*

In the proof we observe that $\eta_S\hat{\phi}_S < \eta_L\hat{\phi}_L$ so that we have the following picture. If the small party moves first the equilibrium is $\eta_S\hat{\phi}_S$ while if the large party moves first the equilibrium is $\eta_L\hat{\phi}_L$ and this is also a simultaneous move game equilibrium. However, there are also simultaneous move equilibrium for larger values of $\eta_1\phi_1 = \eta_2\phi_2$. These equilibria (with both parties aiming to turn out the same number of voters) are Pareto ranked, as lower $\eta_k\phi_k$ reduces costs for both parties and does not change the chances of winning. Hence, as some coordination is needed in the simultaneous move case to hit the exact same target, it seems that the two parties might agree on $\eta_L\hat{\phi}_L$. However, they would both do even better if the small party moved first, so indeed if they could coordinate on this they would.

Notice that the theorem characterizes pure strategy equilibria and does not address the existence of other mixed equilibria. We can get a handle on mixed equilibria by considering the limiting case $N = \infty$ where there are a continuum of voters. With a continuum of voters there is a sense in which no voter is pivotal, so the analysis is that of an all-pay auction: these auctions have been extensively studied—see Levine, Mattozzi and Modica (2024) for a review of the literature and a detailed set of results that include the cost functions studied here. The upshot is that there is a unique equilibrium in which the larger party mixes continuously while the smaller party has an atom at zero turnout and mixes continuously above that point.

A curious point is that this limit model fails upper hemi-continuity. That is, the finite N models for the parameters in Theorem: Simultaneous Pure Strategy Equilibria converge to pure strategies in the limit game, but these are not equilibria. The problem, however, is that the limit game is not properly defined. In the finite population games if $\phi_S \neq \phi_L$ it is true that pivotality converges to zero, but if $\phi_S = \phi_L$ is remains fixed at $\pi(0)$. Hence the correct limit game should assume pivotality is $\pi(0)$ when there is an exact tie. With this modification upper hemi-continuity is restored: a small upward deviation from the tie increases the chances of winning by $1/2$ but increases the incentive cost by $\mu\psi\pi(0)$ so that for $\psi \geq 1/(2\mu\pi(0))$ we have an equilibrium.

Notice, however, that the all-pay auction strategies remain a mixed equilibrium: since the probability of a tie is zero pivotality is in fact zero, so the all-pay auction calculations under the assumption of zero pivotality remain correct. We can also say a bit about the possibility of other mixed equilibria. The unique all-pay auction equilibrium is the only possibility if pivotality is zero. For pivotality to be non-zero with mixing there must be a positive probability of a tie at more than one point. This means that there is a possibility of miscoordination in which one party chooses one mass point and the other party chooses the other: and indeed for one of the two parties conditional on the choice of mass point the probability of miscoordination is at least $1/2$. This means that this choice results in pivotality of at most $\pi(0)/2$ so that if $\psi < 1/(\mu\pi(0))$ it is profitable to deviate. In other words, in the range $1/(\mu\pi(0)) > \psi \geq 1/(2\mu\pi(0))$ the pure strategy and all-pay auction equilibria constitute all equilibria.

We can now say something about the incentive of the small party to commit: in the all-pay auction equilibrium they receive utility of zero, while if they make the optimal commitment they get positive utility so they would clearly like to commit. Moreover, if the size of the two parties is equal, the other party strictly agrees this is a good idea, and so would also agree provided the two parties are of sufficiently similar size.

To what extent do these results about mixed strategy equilibria extend to the finite case? Because the equilibrium correspondence is upper hemi-continuous there cannot be equilibria substantially different than those in the limit case - but is there an equilibrium that resembles the all-pay auction equilibrium? This question of lower hemi-continuity is mathematically far more difficult than that of upper hemi-continuity and we do not know the answer. From a practical point of view, however, approximate equilibria, that is equilibria in which deviations yield at most trivial gains, are no less descriptive of reality than exact equilibria. For approximate equilibria lower hemi-continuity is easy. In particular, fixing the equilibrium strategies of the limit game when population N is large, the probability of being pivotal with those strategies is trivially small, so the loss to an individual in ignoring it is also trivially small.

Hence our second main conclusion is that the smaller party has strong reason to try to make a credible commitment turning the game into a sequential rather than simultaneous move game: and indeed for a range of ψ the larger party has no reason to try to prevent this.

Proof. The exact value of $\hat{\phi}_k$ was established in the proof of Theorem: Simultaneous Pure Strategy Equilibria as

$$\eta_1\hat{\phi}_1 = \frac{F(\gamma) - F(0) - (1/2)\frac{c\gamma^2}{\eta_2 N^2} - \mu\psi c\frac{\gamma}{N} + \mu\psi\pi(\gamma)}{\frac{c\gamma}{\eta_2 N} + \mu\psi c}.$$

This is decreasing in η_2 so largest at η_S, that is when the large party is first mover.

From the summary we know that for $\phi_1 \leq \hat{\phi}_1$ and the best response h_2 of the second mover declines with ϕ_1 until $h_2 = 1$. It then remains constant until $\hat{\phi}_1$ is reached at which point it is $h_2 = 0$. From THIRD and FOURTH and FIFTH we know that for larger ϕ_1 the best response is $h_2 = 0$ provided that

$$1/2 - (1/2)c\nu\eta_1\phi_1^2 > F(\log(1/\log N)).$$

We showed that by choosing c large enough we can satisfy

$$1/2 - (1/2)c\nu\eta_1\hat{\phi}_1^2 > 2F(\log(1/\log N))$$

so that there is a $\phi' > \hat{\phi}_1$ independent of N (for N large enough) for which the best response to $\phi_1 \in [\hat{\phi}_1, \phi']$ is $h_2 = 0$.

Consider that for N large enough $\eta_S\hat{\phi}_S$ and $\eta_L\hat{\phi}_L$ are arbitrarily close together, so in particular $\eta_L\hat{\phi}_L < \min\{\eta_S\phi'_S, \eta_L\phi'_L\}$. In particular if

$$\eta_1\phi_1 = \eta_2\phi_2 \in [\eta_L\hat{\phi}_L, \min\{\eta_S\phi'_S, \eta_L\phi'_L\}$$

both are playing a best response to each other so this is a pure strategy Nash equilibrium.

As we try to increase $\eta_k\phi_k$ above this range utility for both parties decreases until one becomes indifferent to choosing some h_2 less than $\log(1/\log N))$ at which point (and above) we no longer have an equilibrium. This defines $\overline{\phi}$.

If $h_k \neq 0$ for some k this cannot be a pure strategy equilibrium by the argument in ELEVENTH that for large ψ one party must be getting negative utility so would be better off at 0.

Finally, if $\eta_1\phi_1 = \eta_2\phi_2 < \eta_L\hat{\phi}_L$ then the best response of the small party is $h_2 \geq 1$ by EIGHTH, so this is not be an equilibrium either. $\square$

5. Are Elections Close?

We have proposed a novel model of razor thin margins with high turnout. We argued in the introduction that such a model is needed because existing models cannot explain such narrow margins along with high turnout. Here we elaborate on that by examining existing models of voter turnout.

The most obvious source of close elections is the standard pivotal voter model of Palfrey and Rosenthal (1985). The well known problem with this model is that while it is consistent with narrow margins and is consistent with high turnout (if there are many committed voters) it is not consistent with both: indeed the difficult with explaining strategic high turnout by the model is the main reason for the subsequent development of ethical voter and peer pressure models. The empirical point was made strongly in Coate, Conlin and Moro (2008)

As there are no existing models that combine random turnout with pivotality, we examine random turnout ethical voter/peer pressure models. Recall that if parties set targets ϕ_1, ϕ_2 then ζ_k is the random vote added to $\eta_k \phi_k N$ to determine the outcome of the election. The standard model (see for example Coate and Conlin (2004)) is a simultaneous move version of the targeted model in which pivotality plays no role: we simply assume that parties provide effort with some cost $C_k(\phi_k)$ which is continuous and increasing when positive. It is presumed that ζ_k has a continuous density and the work of Levine and Mattozzi (2022) shows that there is an equilibrium. However, as shown there and in the work of Ewerhart (2017) if the noise is relatively small then the equilibrium involves mixed strategies providing a second source of noise.

To analyze the closeness of elections in a non-parametric way, we compute two bounds for this model, one for larger noise and one for smaller noise. The first is computed from the fact that when noise is large close elections cannot be that common. The second is computed by approximating the equilibrium by a mixed strategy Tullock equilibrium and showing that the endogenous noise is sufficiently great that close elections cannot be that common. Subsequently we apply both bounds to election data to show that razor thin elections with high turnout are indeed difficult to reconcile with existing models.

Larger Noise Bound

Let f denote the common continuous density of ζ_k/N.Key to our results is the assumption of *regular noise*. We say that a positive random variable χ with $E\chi = 1$ is *regular* if it has a continuous decreasing density function and $R \equiv \int [f(\chi)]^2 d\chi \leq 1/2$. A wide variety of distributions satisfy this assumption: the uniform, the triangular, the gamma, the exponential and the squared normal (chi-squared with two degrees of freedom). Noise is regular if ζ_k has the distribution of $s\chi$ for some regular χ and $s > 0$.

To understand why we require noise to be regular, observe that $\xi_1 = \zeta_1 - \zeta_2$ is the random component of the vote fraction differential and that $\text{var}\xi_1 = 2\text{var}\zeta_1$. If this variance is small then there can be close elections. However, large variance does not rule out close elections: For example we could have a substantial point mass at the origin and tails that are Pareto like so that the variance is quite large, indeed infinite. This would result in close elections, but also makes no sense as a model of noise. The assumption of regularity rules out this kind of clustering at the origin. Observe that R/s is the height of the density function of ξ_1 at zero: the assumption of regularity imposes the requirement that this not be too large.

As we know that an equilibrium exists, let ν denote the equilibrium vote differential as a fraction of the population - conditional on the realization of randomization if the parties are using mixed strategies. Note that the expected number of added random voters is sN.

Theorem 5.1. *For any $w > 0$ we have* $\Pr(|\nu| \leq w) \leq w/s$.

Notice that this bound is not very useful for small s and in particular completely useless if $2s \leq w$.

Proof. Observe that the highest probability of $|\nu| \leq w$ occurs if both parties choose the same target turnout, in which case $\nu = \xi_1$. Clearly $\overline{\xi}_k^0 \xi_1$ is symmetric around zero and from the convolution we see that it is single peaked with a maximum at zero. Since χ is regular the maximum of the density $f(\xi_1) \leq 1/(2s)$. Since ξ_1 is single peaked the greatest probability of $|\nu| \leq w$ occurs when ξ_1 is uniformly of height $1/(2s)$ and multiplying the width of the interval $2w$ gives the stated bound on the probability. □

Small Noise Bound

We approximate the case of small noise with an all-pay auction. In the Appendix we argue this is a good approximation for $s < 0.16$. We know from Levine and Mattozzi (2022) that for small s the equilibrium probability of ties will be small regardless of the presence of the noise. We study the benchmark case of constant direct marginal cost used in empirical studies such as Shachar and Nalebuff (1999) and Coate and Conlin (2004).

Let υ denote the equilibrium vote differential as a fraction of votes cast. (Note that the denominator of υ is different than the large noise ν.)

Theorem 5.2. *In an all pay auction with the standard assumption of constant marginal direct cost of participation* $\Pr(|\upsilon| \leq \omega) \leq \omega$.

Proof. Greater convexity of the cost function results in more ties, so we are safe to disgard the monitoring cost as this decreases convexity. With constant marginal cost the greatest probability of a tie is when the two parties are equal, so we compute the symmetric equilibrium.

We may normalize so that total direct cost is equal to 1 at $\phi = 1$. Hence marginal cost $c\phi$ integrated from 0 to 1 which is $c/2$ should be equal to the per capital value of the prize which is one, so $c = 2$. Then the equilibrium cdf is ϕ^2 and the density is 2ϕ.

Since we are dealing with near ties, we may assume that the vote differential relative to votes cast by a single party is $\omega/2$. Suppose 1 is the winning party:

conditional on ϕ_1 the probability that the winning margin is less than or equal to $\omega/2$ is given by

$$\int_{(1-\omega/2)\phi_1}^{\phi_1} 2\phi d\phi = (1-(1-\omega/2)^2)\phi_1^2 = (1-\omega/4)\omega\phi_1^2.$$

This is slight overestimate because it does not account for the upper boundary. To get the overall probability we integrate with respect to ϕ_1 to find

$$\int_0^1 (1-\omega/4)\omega 2\phi_1^3 d\phi_1 = (1-\omega/4)\omega/2 \leq \omega/2.$$

We double this to account for the equal chance that 2 is the winning party to find the stated result. $\square$

6. Are Razor Thin Elections Common?

In this section, we argue that if elections are simultaneous contests, as in existing models, electoral margins as tight as those in the motivating examples are too unlikely. This motivates our interest in the possibility of targeting ocurring during the election cycle.

Peruvian Elections

We take as our starting point the year 2000 when Fujimori resigned: elections cannot be considered free and fair during his term in office. Since that time there have been five elections in Peru. Two of them have razor thin vote margins relative to votes cast of 0.0024 and 0.0025. Turnout in 2016 was 82% so the corresponding margin as a fraction of the (voting) population is 0.0020.

From Theorem 5.1 $\Pr(|\nu| \leq 0.0020) \leq 0.0010/s$. We observe that 40% of elections were this close and this probability is attained when $s = 0.25\%$. As this is more than an order of magnitude less than the critical cutoff of 16 as can be seen in Figure 7.1 in the Appendix the relevant model is one of small noise.

From Theorem 5.2 we see that the probability $\Pr(|\upsilon| \leq .0025) \leq .0025$. With a probability of 0.25% it is extremely unlikely we would see one of these events in five elections, let alone two.

US Gubernatorial Elections

In the US we take our starting point as 1995 when the election of Newt Gingrich as speaker of the House of Representatives indicates the culmination of the historical realignment of the two parties: the Republicans, historically the party of the north and Democrats, historically the party of the south, shifting regions and platforms.

During this period there were 329 gubernatorial elections. One of these elections, the 2004 Washington state election, was extremely close, being decided by 129 votes out of 2.9 million votes cast.[6] There were approximately 4 million registered voters for that election.

From Theorem 5.1 $\Pr(|\nu| \leq 0.00003) \leq 0.000016/s$. We observe that 0.3% of the elections were this close and this probability is attained when $s = 0.5\%$, like Peru's, more than an order of magnitude less than the critical cutoff of 16%.

From Theorem 5.2 we see that the probability $\Pr(|\upsilon| \leq 0.000044) \leq 0.000044$. With a probability of 0.0044% there is less than a 1.5% chance of seeing one such event in 329 trials.

US Senatorial Elections

During the 27 years between 1995 and 2022 there were 453 senatorial elections. One of these elections, the 2008 Minnesota election, was extremely close being decided by 225 votes out of 2,424,946 votes cast for one of the two leading candidates.[7] We figure turnout at about 75%.

From Theorem 5.1 $\Pr(|\nu| \leq 0.00007) \leq 0.000035/s$. We observe that 0.2% of the elections were this close and this probability is attained when $s = 1.6\%$ larger than Peru or for gubernatorial elections, but still an order of magnitude less than the critical cutoff of 16%.

From Theorem 5.2 we see that the probability $\Pr(|\upsilon| \leq 0.00009) \leq 0.00009$. With a probability of 0.009% there is less than a 4% chance of seeing one such event in 453 trials.

7. Conclusion

We propose a dynamic model of electoral competition to explain the apparent abundance of knife-edge elections. If the cost of monitoring (ψ) and turning out (c) voters are high, there is a knife-edge election in the sequential move case, and in the simultaneous move case there are also equilibria of this sort. In addition, the smaller group prefers to commit, and for a range of parameter values the larger group will not object. When moves are sequential we find that as the targeting of voters becomes more precise (σ declines) turnout increases and elections become tighter.

It is worth looking at the four close elections in Peru, Washington and Minnesota to see if they fit the narrative of targeting by one party, while the other remains committed. Of course targeting can happen at any time, but if the signal is received early in the electoral cycle we ought to see one party maintaining stable support while if the small party targets the large it has support that creeps up to the large party level, while if the large party targets the small it has support that creeps down to the small party level. Two of our elections follow this narrative. As indicated in the introduction: the small party creeping

[6] https://www.nytimes.com/2004/12/31/us/governorelect-declared-in-washington-recounts.html

[7] There was a third candidate on the ballot; it seems most relevant to the theory not to count those votes as belonging to either party.

up to a stable large political group is exactly what happened in Peru in 2021. The opposite occurred in Minnesota in 2008. There support for Franken remained relatively stable while support for the front-runner Coleman gradually declined. We should mention as well that prior to the election it was expected to be close: all of the major polling organizations rated the election a toss-up.

Our model of targeting provides a stylized, tractable view of changes in political support during the election campaign. As illustrated by the examples, the consideration of those changes seems extremely useful for understanding close elections.

Appendix: When Is The Tullock Approximation Good?

We have approximated elections with small values of s by an all-pay auction. But how small we need s to be for this to be a reasonable approximation? As in other contests, equilibrium strategies depend upon random turnout only through the contest success function: the probability that a party wins as a function of its intended turnout. We know from Levine and Mattozzi (2022) if contest success functions are close in the sense of pointwise convergence then the equilibrium strategies will be close in the sense of weak convergence. To get an idea of what constitutes small s we turn our attention here to contest success functions.

We start by considering our benchmark case in which χ is a standard exponential. In this case the vote differential follows a Laplace distribution, and the probability of a draw greater than x is given by $(1/2)e^{-x/s}$. Let S, L denote the smaller and larger parties respectively. Then if both parties turn out all their voters then the probability the smaller party wins is the probability

$$\xi_k > \eta_L - \eta_S = (1 - \eta_S) - \eta_S = 1 - 2\eta_S$$

and is

$$\mathcal{P}_S^e = (1/2)e^{-(1-2\eta_S)/s}.$$

Contrast this to the widely used Tullock (1967) contest, for which good quantitative results are known. Given the Tullock contest success function $\eta_S^\beta/(\eta_S^\beta + \eta_L^\beta)$, the probability the small party wins is

$$\mathcal{P}_S^t = \frac{1}{1 + ((1/\eta_S) - 1)^\beta}.$$

Continuing to take the benchmark case of linear marginal cost Ewerhart (2017) and Levine and Mattozzi (2022) show that in the Tullock model for $\beta < 4$ there is a unique pure strategy equilibrium, while for $\beta > 4$ all equilibria require mixing and the payoffs to both parties are exactly the same as for the all pay auction. If the small party is half the size of the large party so $\eta_S = 1/3$ at the critical cutoff $\beta = 4$ we find using the expression above $\mathcal{P}_S^t = 1/17$. Solving

$$\mathcal{P}_S^e = (1/2)e^{-(1-2\eta_S)/s} = 1/17$$

for s with $\eta_S = 1/3$ we find

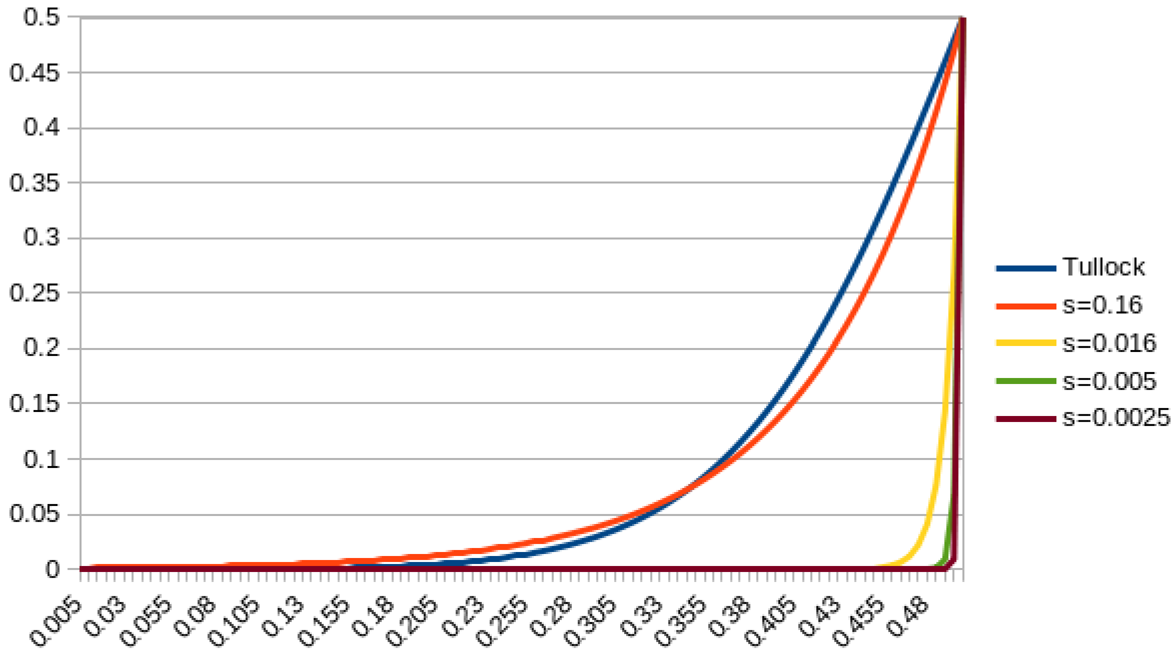
$$(1/2)e^{-(1/3)/s} = 1/17$$

$$s = \frac{1/3}{\log(17/2)} = 16\%.$$

In other words, with noise even as large as $s = 0.16$ the equilibrium already resembles that of the all-pay auction.

To better visualize the contest success functions—the functions from which equilibria are derived—we plot both the Tullock and exponential for different values of s. The values of $s = 0.016, 0.005, 0.0025$ correspond to those derived

Figure 7.1: Probability of Small Party Winning



The horizontal axis represents η_S and the vertical axis the probability of the small party winning.

above for specific elections: as can be seen they are quite close to an all-pay auction.

We should comment briefly on the fact that the Tullock (1967) model is derived from a multiplicative random turnout while we have assumed additive random turnout. We have done so in order to enable a tractable model of targeting. There are not great differences, however, between the multiplicative and additive models as can be seen from the blue Tullock line and orange $s = 0.16$ line in Figure 7.1.

We should add that the additive case fits the Hirshleifer (1989) model, which in turn is compatible with our exponential assumption. Unfortunately, unlike the Tullock model, good quantitative results about mixed equilibrium are not known for the Hirshleifer model. A key point is this: the Hirshleifer model leads to more randomization by parties than the Tullock model—in particular, pure strategy equilibria are possible only when one party makes no effort at all.[8] In this sense we think that computing the critical cutoff for the Tullock model is relatively conservative.

Approximate Equilibrium

Nash equilibrium in a parametric model is at best an approximation to reality. It says that equilibrium strategies are such that no player can gain by deviating. Theorists have weakened this to the notion of ϵ-equilibrium which asserts that strategies are such that no player can gain more than ϵ by deviating. From an applied point of view an ϵ-equilibrium with small ϵ is as valid a theory as exact equilibrium. This leads us to ask: suppose that the parties employ their equilibrium strategies for the all-pay auction when in fact there is small noise measured by s? In particular, how great is the loss ϵ? Before computing a bound, observe that the probability of a tie if parties use their all-pay auction equilibrium strategies and there is noise is lower than in the all-pay auction itself, so those bounds remain valid.

To get a bound on the greatest possible loss given the other player is employing their all-pay auction equilibrium strategy we can simply bound the difference between the approximate (all-pay) and true ($s > 0$) objective functions: the greatest gain to deviating can be no greater than this. Since costs of intended turnout are the same in the approximate and true model, we need only consider the differences between the winning probabilities.

Continuing to normalize players equilibrium choices to run between 0 and 1 the probability the opponent bids less than or equal to ϕ_k is ϕ_k^2. That is the probability of winning in the all pay model with a bid of ϕ_k is simply ϕ_k^2. For small ϕ_k this understates the chances of winning since a good draw could bring a win, but the probability the opponent is bidding there is only $2\phi_k$ which is not large. For intermediate ϕ_k there are extra chances of winning due to good draws, but extra chances of losing due to bad draws and these tend to cancel out. We see, then, that the worst case is high ϕ_k since there are extra chances

[8] Levine and Mattozzi (2022).

of losing and the opponent is likely to bid there. Hence we may bound the loss by computing the loss to the bid $\phi_k = 1$. For the exponential/Laplace model the added chance of losing when the opponent bids ϕ_{-k} is $(1/2)e^{-(1-\phi_{-k})/s}$ and to get a bound we integrate over all opponent bids from 0 to 1.

$$\epsilon = \int_0^1 2\phi_{-k}(1/2)e^{-(1-\phi_{-k})/s} d\phi_{-k} = e^{-1/s}\int_0^1 \phi_{-k}e^{\phi_{-k}/s} d\phi_{-k}.$$

With change of variable this becomes

$$\epsilon = e^{-1/s}\int_0^{1/s}(sy)e^y s dy,$$

which we may integrate by parts to find

$$\epsilon = s(1-s) \leq s.$$

In conclusion s measures the maximum loss from employing the all-pay auction strategy against an opponent who does the same when the noise is given by $s > 0$. The values of $s = 0.016, 0.005, 0.0025$ correspond to those derived above for specific elections. In particular, we may wonder if, given all the practical difficulties of organizing a party, it would engage in elaborate strategizing in order to increase the chances of winning by at most a quarter to one and half percent.

References

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

Blais, A. (2000): *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory.* University of Pittsburgh Press.

Caughey, D. and J. S. Sekhon (2011): "Elections and the regression discontinuity design: Lessons from close US house races, 1942–2008," *Political Analysis* 19: 385-408.

Coate, S. and M. Conlin (2004): "A Group Rule–Utilitarian Aapproach 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.

De la Cuesta, B. and K. Imai (2016): "Misunderstandings About the Regression Discontinuity Design in the Study of Close Elections," *Annual Review of Political Science* 19: 375–396.

Della Vigna, S., J. A. List, U. Malmendier and G. Rao (2017): "Voting to Tell Others," *Review of Economic Studies* 84: 143–181.

Downs, A. (1957): *An Economic Theory of Democracy*, Harper and Row, N.Y.

Ewerhart, C. (2017): "Contests with Small Noise and the Robustness of the All-Pay Auction," *Games and Economic Behavior* 105: 195–211.

Feddersen, T., and A. Sandroni (2006): "A theory of participation in elections," *American Economic Review* 96: 1271-1282.

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

Harsanyi, J. C. (1973): "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points," *International Journal of Game Theory* 2, 1 23.

Hirshleifer, J. (1989): "Conflict and Rent-Seeking Success Functions: Ratio vs. Difference Models of Relative Success," *Public Choice* 63: 101–112.

Hotelling, H. (1929): "Stability in Competition," *Economic Journal* 153: 41–57.

Levine, D. K. and Andrea Mattozzi (2020): "Voter Turnout with Peer Punishment," *American Economic Review* 110: 3298–3314.

Levine, D. K. and A. Mattozzi (2022): "Success in Contests," *Economic Theory* 73: 595–624.

Levine, D. K., A. Mattozzi and S. Modica (2022) *Social Mechanisms and Political Economy: When Lobbyists Succeed, Pollsters Fail and Populists Win*, draft.

Levine, D. K., A. Mattozzi and S. Modica (2024): "The Tripartite Auction Folk Theorem," *European Economic Review*, forthcoming.

Mauss, M. (1990): *The Gift. The Form and Reason for Exchange in Archaic Societies*, W.W. Norton, London.

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

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

Nagourney, A., A. Parker, J. Rutenberg and J. Zelenynov (2012): "How a Race in the Balance Went to Obama," *New York Times*, November 8.

Needham, R. (1962): *Structure and Sentiment: A Case Study in Social Anthropology*, University of Chicago, Chicago.

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

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

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

Tullock, G. (1967): "The Welfare Costs of Tariffs, Monopolies, and Theft," *Western Economic Journal* 5: 224–232.

Vogl, T. S. (2014): "Race and the politics of close elections," *Journal of Public Economics* 109: 101-113.