

Razor-Thin Elections

David K. Levine¹, Cesar Martinelli²

Abstract

We argue that standard models that solve the paradox of voting do a bad job explaining the frequency of very close elections. We instead model head-to-head elections as a competition between incentive schemes to turn out voters. We show that elections are either heavily contested, and decided by thin margins, or safe, meaning that voters in one of the two sides effectively give in, possibly leading to a landslide in favor of the larger side. In equilibrium, improvements in the quality of polling make contested elections with razor-thin margins more likely.

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

¹Department of Economics, EUI and WUSTL

²Department 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

Head-to-head mass elections decided by a small number of voters are frequent enough as to raise the question of whether they are the deliberate result of political parties' choices. For instance, in the last two presidential runoff elections in Peru, in 2016 and 2021, 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 2016, preelection polls reported close to a tie throughout the campaign.⁴ In 2021, however, the first preelection polls initially 41.5% to Castillo and 21.5% to 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.⁵

Razor-thin margins have occurred elsewhere: for example 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 2,424,946 votes. We analyze Peruvian elections since 2000 and US gubernatorial and senatorial elections since 1995 and show that standard models of ethical voters or peer pressure in which pivotality plays no role do a poor job explaining these exceptionally tight elections. This leads us to hypothesize that the parties intentionally try to arrange that the election be close.

Why would a political party prefer a close election to turning out a few extra voters at minimal cost to win the election? We provide an answer by explicitly modeling elections as competition between incentive schemes. There are two ways to turn out voters: one 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).⁶ 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.

Pivotality as a model of participation in large mass elections has fallen into disfavor. The literature on the paradox of voting starting with Downs (1957) argues that models of pivotality cannot generate the large turnouts seen in practice, or if large turnout is due to committed or expressive voters, then pivotality ceases to matter. Mulligan and Hunter (2003) use a coin toss model to argue that empirically pivotality is too small to matter. Coate, Conlin and Moro (2008) argue that pivotality predicts much closer election results than are ob-

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

⁴See e.g. <https://www.ipsos.com/es-pe/opinion-data-29-de-mayo-de-2016>

⁵IEP Informes de Opinión, <https://iep.org.pe/noticias/informes-de-opinion/>

⁶There is a large literature in economics and political science that we will not attempt to summarize in here. It is common in this literature to contrast peer pressure and other social motivations to strategic voter motives as explanation of turnout; see for example, Blais (2000). In our model, both peer pressure and strategic voter motivations are possible.

served in the data.

Here we push back against these ideas. First, we argue that non-pivotal models produce too few close elections. We then account for both the paradox of voting and the observed closeness of elections by postulating a model in which pivotality sometimes matters. Specifically there is a probability of a signal that enables one party to target the turnout of the other—and as this signal is commonly observed, in these elections pivotality matters, while in others it does not. This highlights a problem in Mulligan and Hunter (2003): their pivotality computations assume *ex ante* knowledge only of the overall distribution of election outcomes. In practice as in our model it is often known in advance that the election will be a tight one, and conditional on this information pivotality calculations change greatly. Indeed, Shachar and Nalebuff (1999) estimate the predictability of elections making the point that indeed pivotality goes up when elections are known *a priori* to be tight. We might also add, as ours is a model of social voting, that pivotality from an incentive point of view is probably not best measured by a single vote swinging an election: if we thought there was a reasonable chance that a handful of votes could swing an election we would certainly put enormous pressure on our immediate friends to make sure they voted.

From a methodological point of view peer pressure is often viewed as an alternative to pivotality as a motivation for voting. We show by contrast how both forces may be present in a coherent model of individual and group behavior. This social mechanism approach can potentially be used as well in setting with relatively few voters where it is likely that both peer pressure and pivotality matter.

We can think of platform convergence, as predicted by spatial models of electoral competition after Hotelling (1929) and Downs (1957), as a force leading to very close elections. However, it is highly implausible that, for instance, in the Peruvian presidential runoff election, parties adjusted their platforms up to the last minute until they converged to the median voter’s ideal policy. Moreover, this line of reasoning would predict that all elections should be close, and polarization should never happen. Recognizing that parties do adjust their platforms strategically in elections still leaves open the question of why voters turn out in the first place.

While close elections based on a spatial model would give equal chance of winning to both parties, close elections in a pivotal model do not necessarily imply that each party has an equal chance of winning. In our model, in particular, parties have an equal chance of winning in close elections with high turnout, while one of the two parties (the second mover) is advantaged in close elections with low turnout. This is important in view of the empirical debate about whether each party has an equal chance of winning close elections (see for example, De la Cuesta and Imai 2016), and provides theoretical support for the possibility of biases.

2. The Model

We base our model on Levine and Mattozzi (2020). Voters are divided into two parties $k = \{S, L\}$, the small and the large. There are a large number N of voters and party k has $\eta_k N$ members with $\eta_L > \eta_S > 0$. The parties compete in an election for a prize worth N to the party that produces the greatest number of votes. Voting is costly and we order voters so that higher numbered voters have higher participation cost. We adopt the Palfrey and Rosenthal (1985) specification of this cost: specifically, we assume that voter n_k in party k faces a participation cost depending on their percentile in the party $n_k/(\eta_k N)$ given by $\gamma C(n_k/(\eta_k N))$. Here $\gamma > 0$ is a parameter and C is non-negative, smooth, and is normalized with $C(1) = 1$. We assume moreover, the regularity conditions $C'(\phi) > 0, C''(\phi) \geq 0$. These are satisfied for $C(\phi) = \phi^\alpha$ for $\alpha \geq 1$. This cost is the *direct cost* of voting net of social pressure: it consists of costs such as those of time, inconvenience, and transportation.

Our core assumption is that each party is able to design a mechanism providing incentives to individual voters in the form of social pressure. We model this as an ability to impose penalties that are costly to individual party members. The mechanism has two parts. The first part is a target fraction ϕ_k together with a rule for party members prescribing voting if $n_k/(\eta_k N) \leq \phi_k$. This means that voters with sufficiently low costs of voting are expected to vote. We refer to this as the *social norm* for the party. The second part of the mechanism is a punishment $P_k \geq 0$ representing social disapproval for failing to comply with the social norm. This punishment may also be costly to party members who carry out the punishment: if Tim is punished by David refusing to have a beer with him this may be costly to David as well as to Tim. Hence the social cost of the punishment is taken to be ψP_k where $\psi > 0$ (possibly greater than one). This punishment takes place after the results of the election are known.

The ability to apply punishments is limited by imperfect information. While there is no difficulty in determining whether or not a party member voted, individual costs of voting are not transparent and the party can only observe a noisy binary signal about whether or not a non-voter followed the social norm. In particular, among voters that failed to vote there are two types: high cost voters who were “excused” according to the social norm and low cost voters who were not. If the non-voting member violated the social norm, that is, $n_k/(\eta_k N) \leq \phi_k$, a negative signal is received with probability $\pi_1 \in [0, 1]$ and a positive signal with corresponding probability $1 - \pi_1$. A non-voting member with $n_k/(\eta_k N) > \phi_k + \mu$ where $\mu > 0$ clearly has a good excuse so always receives a positive signal.⁷ The remainder who did not violate the social norm, that is, $\eta_k N(\phi_k + \mu) \geq n_k > \eta_k N\phi_k$, receive a negative signal with probability

⁷Note that this simplifies the formulation in Levine and Mattozzi (2020) who have an upper bound on the number of voters and assume that all non-voters must be punished. That results in eventually decreasing monitoring costs: here we limit attention to the region of increasing monitoring costs in order to focus on the substitutability between pivotality and peer punishment.

π .

A positive signal should be thought of as “producing a good excuse for not voting” so that a non-voter who is genuinely excused is more likely to be able to produce a good excuse than one who is not. Punishment can then be based on whether or not the party member voted, on the signal received and may also be based on whether or not the election was won or lost. For notational simplicity we assume that only bad signals are punished and that when the probability of losing the election is strictly between zero and one punishment occurs only when there is a loss. We will show in the proof of Theorem 4.1 that this is without loss of generality as it is the optimal punishment scheme.

In addition it is possible that a party might wish to prevent members from voting. Again for notational simplicity we assume that it can do so costlessly, for example because it can perfectly observe and punish voters who were not supposed to vote. As in fact it is never optimal to prevent members from voting, we will show that this assumption also is without loss of generality.

The choice of turnout ϕ_k takes place in two stages. In the first stage each party announces a commitment of ϕ_k^0 voters. One party K by accident or design always attains commitment first, with each party having an equal chance of being that *first mover*. After the first mover is determined and before the *second mover* becomes committed there is a chance Ω that a signal is received (by both parties) revealing information about the actual turnout of the first mover as described below. In this case the second mover may choose an alternative turnout target ϕ_{-K}^1 and we say that *targeting is feasible*.

We want to reflect the fact that depending on the quality of polling and other data it is difficult to know exactly what the turnout of the other party is. We model this by assuming that actual party turnout is a random function of target party turnout: specifically we assume that in addition to the $\phi_k \eta_k N$ intended voters each party costlessly receives a random number of additional votes. Initially these are random variables ζ_k^0 . If targeting is feasible we assume that this uncertainty is reduced so that the random variables are ζ_k . We assume that ζ_S^0, ζ_L^0 are iid as are ζ_S, ζ_L . The actual vote differential is then equal to the intended vote differential $h_k \equiv \phi_k \eta_k N - \phi_{-k} \eta_{-k} N$ plus an error $\xi_k^0 = \zeta_k^0 - \zeta_{-k}^0$ or $\xi_k = \zeta_k - \zeta_{-k}$ depending on whether targeting is feasible.

In the case in which targeting is feasible let $\sigma^2 = \text{var} \xi_k$ and let F denote the cdf for the normalized random variable ξ_k / σ . To cleanly distinguish between *close elections* and *sure elections* we assume that there is a cutoff σL such that $F(L) = 1$ and $F(h_k / \sigma) < 1$ for $h_k < \sigma L$. In this setup by choosing h_k , the second mover determines whether the election is close or sure. We assume, moreover, that the second mover cannot choose $h_k = \sigma L$ exactly but may choose either σL^- slightly below σL or σL^+ slightly above σL . The meaning of this we explain in the next paragraph.

In principle the cdf F should be discrete but we approximate it as continuous except at the cutoff σL . Let f be the density corresponding to F for $\xi_k < \sigma L$. Then the normalized density $(1/\sigma)f(h_k/\sigma)$ is an approximate measure of pivotality: the probability of a tie, the probability of a one vote loss, and the probability of a one vote win. Specifically, in a close election, an individual

who shifts from voting to not voting increases the probability of a loss by half the probability of a tie or a one vote win, which is to say, (approximately) $(1/\sigma)f(h_k/\sigma)$. In case the second mover chooses σL^- we take the probability of winning to be $F^- \equiv \lim_{h_k \uparrow \sigma L} F(h_k/\sigma)$ and the probability of being pivotal to be $(1/\sigma)f^- \equiv (1/\sigma) \lim_{h_k \uparrow \sigma L} f(h_k/\sigma)$. In case $h_k > \sigma L$, including $h_k = \sigma L^+$, the probability of winning is 1 and the probability of being pivotal is 0. Similarly, in case $h_k < -\sigma L$, including $h_k = -\sigma L^+$, the probability of winning is 0 and the probability of being pivotal is 0.

In the case in which targeting is not feasible we also approximate the cdf as continuous but we assume that it is sufficiently noisy that pivotality does not matter.

We assume that individual voting decisions are made after targets are determined and that punishment schemes must satisfy the interim incentive compatibility constraint that conditional on knowing the targets and uncertainty about actual turnout party members who are supposed to vote are willing to do so and those who are not supposed to vote are willing not to do so. Given this, parties attempt to maximize the expected value of winning the prize less the expected costs of turning out voters. For convenience we will normalize this utility dividing by N .

Large Enough Cost

We limit attention to the case in which direct costs are high enough that parties always choose $\phi_k \leq 1 - \mu$ and that the larger party does not overwhelm the smaller party. The former implies that the constraint $\phi_k \leq 1$ never binds and that the fraction of unexcused voters is always exactly μ . Specifically, observe that normalized direct cost of voting is given by $c_k(\phi_k) = (1/N) \int_0^{N\eta_k\phi_k} \gamma C(n_k/(\eta_k N)) dn_k = \eta_k \int_0^{\phi_k} \gamma C(\tilde{\phi}_k) d\tilde{\phi}_k = \eta_k \gamma c(\phi_k)$. No party will choose a target for which $c_k(\phi_k) > 1$ as this is dominated by choosing $\phi_k = 0$. Hence, we assume that costs are sufficiently high that $\eta_S \gamma c(1 - \mu) > 1$: this implies that parties always choose $\phi_k \leq 1 - \mu$. Second, we assume that $\eta_L \gamma c(\eta_S/\eta_L) > 1$: this implies that the large party will never wish to overwhelm the small party by turning out more voters than the small party is able to turn out.

3. Close Elections Without Pivotality

We start by studying the “standard” model in which $\Omega = 0$ so there is no targeting and pivotality does not play a role. Our goal is to establish that if the noise ζ_k^0 is reasonably well behaved then close elections are not likely. The idea is that if the noise is small we are close to an all-pay auction in which substantial randomization by the parties is needed and this prevents close elections, while if the noise is large, the noise itself prevents close elections.

Regular Noise

We will be studying here the distribution of ζ_k^0 and ξ_k^0 . These are random variable over numbers of voters. It will be convenient instead to study the corresponding random variables over the number of voters as a fraction of the population $\bar{\zeta}_k^0 = \zeta_k^0/N$ and $\bar{\xi}_k^0 = \xi_k^0/N$. We denote the densities of these as $\bar{f}^0, \bar{f}^{0\xi}$ respectively.

Key to our results is the assumption of *regular noise*. To understand this, observe that $\bar{\xi}_k^0 = \bar{\zeta}_k^0 - \bar{\zeta}_{-k}^0$ is the random component of the vote fraction differential and observe that $\text{var}\bar{\xi}_k^0 = 2\text{var}\bar{\zeta}_k^0$. If these variances are 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.

To rule clustering at the origin we say that a positive random variable $\bar{\zeta}_k^0$ normalized so that $E\bar{\zeta}_k^0 = 1$ is *regular* if it has a continuous decreasing density function and $R \equiv \int [\bar{f}^0(\bar{\zeta}_k^0)]^2 d\bar{\zeta}_k^0 \leq 1/2$. To understand this, observe that R is the height of the density function of $\bar{\xi}_k^0 = \bar{\zeta}_k^0 - \bar{\zeta}_{-k}^0$ at zero: as indicated we want to make sure that this is not too large. 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).

Large Noise Bound

Fix a regular random variable χ and consider for $s > 0$ the noise $\bar{\zeta}_k^0$ has the distribution of $s\chi$. Let ν denote the equilibrium vote differential as a fraction of the population.

Theorem 3.1. $\Pr(|\nu| \leq w) \leq w/(2s)$.

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

Proof. We do not analyze equilibrium: we simply observe that the highest probability of $|\nu| \leq w$ occurs if both parties choose the same target turnout, in which case $\nu = \bar{\xi}_k^0$. Clearly $\bar{\xi}_k^0$ 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 density $\bar{f}^{0\xi}(\bar{\xi}_k^0) \leq 1/(2s)$. Since $\bar{\xi}_k^0$ is single peaked the greatest probability of $|\nu| \leq w$ occurs when $\bar{\xi}_k^0$ is uniformly of height $1/(2s)$ and multiplying the width of the interval w gives the stated bound on the probability. \square

Small Noise Bound

We approximate the case of small noise with an all-pay auction. 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 v denote the equilibrium vote differential as a fraction of votes cast.

Theorem 3.2. *In an all pay auction with the standard assumption of constant marginal direct cost of participation $\Pr(|v| \leq \omega) \leq \omega$.*

Proof. Greater convexity of the cost function results in more ties, so we are safe to disregard 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 L is the winning party: conditional on ϕ_L the probability that the winning margin is less than or equal to $\omega/2$ is given by $\int_{(1-\omega/2)\phi_L}^{\phi_L} 2\phi d\phi = (1 - (1 - \omega/2)^2)\phi_L^2 = (1 - \omega/4)\omega\phi_L^2$. This is slight overestimate because it does not account for the upper boundary. To get the overall probability we integrate with respect to ϕ_L to find $\int_0^1 (1 - \omega/4)\omega 2\phi_L^3 d\phi_L = (1 - \omega/4)\omega/2 \leq \omega/2$. We double this to account for the equal chance that S is the winning party to find the stated result. \square

When Is Noise Small?

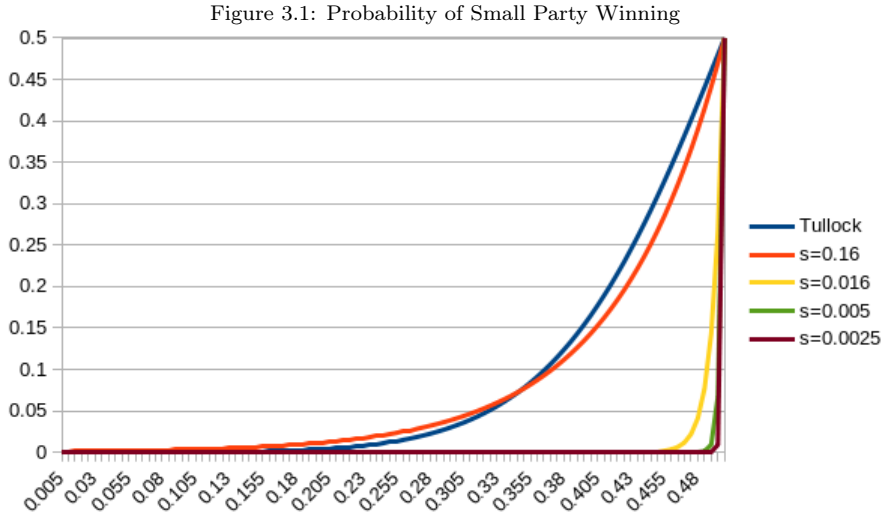
We will want to approximate contests with small values of s by an all-pay auction. But how small we need s to be for this to be a reasonable approximation? Equilibrium strategies depend upon random turnout only through the contest success function: the probability that a party wins as a function of its 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 the 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 $(1/2)e^{-x/s}$. In particular if both parties turnout all their voters then the probability the small party wins is the probability $\xi_k^0 > \eta_L - \eta_S = 1 - 2\eta_S$ and is

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

Contrast this to the Tullock (1967) contest success function $\eta_S^\beta / (\eta_S^\beta + \eta_L^\beta)$ where the probability the small party wins is

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



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

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 $\mathcal{P}_S^t = 1/17$. 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 below 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 3.1.

We should add that the additive case is comparable with 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 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

⁸Levine and Mattozzi (2022).

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

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

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 $s > 0$. Recall that the values of $s = 0.016, 0.005, 0.0025$ correspond to those derived below 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.

Are Razor Thin Elections Common?

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 3.1 $\Pr(|\nu| \leq 0.0020) \leq 0.0010/s$. 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 3.1 the relevant model is one of small noise.

From Theorem 3.2 we see that the probability $\Pr(|\nu| \leq .0025) \leq .0025$. 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.⁹ There were approximately 4 million registered voters for that election.

From Theorem 3.1 $\Pr(|\nu| \leq 0.00003) \leq 0.000016/s$. 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 3.2 we see that the probability $\Pr(|\nu| \leq 0.000044) \leq 0.000044$. 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.¹⁰ We figure turnout at about 75%.

From Theorem 3.1 $\Pr(|\nu| \leq 0.00007) \leq 0.000035/s$. 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 3.2 we see that the probability $\Pr(|\nu| \leq 0.00009) \leq 0.00009$. With a probability of 0.009% there is less than a 4% chance of seeing one such event in 453 trials.

4. Targeted Voting

We see that the model with $\Omega = 0$ and no pivotality does a poor job explaining razor thin elections. We turn now to the opposite case $\Omega = 1$ where electoral targeting is possible. Here we see that there is a viable explanation for razor thin elections.

Parameters and Approximations

We make two assumptions about functional form. First, we assume that the direct cost of voting $C(n_k/(\eta_k N)) = cn_k/(\eta_k N)$ is linear. This implies that the total cost of voting is quadratic and given by $\eta_k N \phi_k^2/2$. Second, we assume that the additional voters ζ_k are drawn from an exponential distribution with mean σ . The latter implies that ξ_k follows a Laplace distribution with mean 0 and scale parameter σ so that for positive x we have $f(x) = (1/2)e^{-x}$ and $F(x) = 1 - (1/2)e^{-x}$.

⁹<https://www.nytimes.com/2004/12/31/us/governorelect-declared-in-washington-recounts.html>

¹⁰There was a third candidate on the ballot; it seems most relevant to the theory not to count those votes as belonging to either party.

Electoral Shock

We assume that size of the random shock to the number of voters as measured by σL is small relative to the the size of the population but reasonably large in absolute terms. Specifically, we assume that

$$\sigma L/N \approx 0,$$

$$e^{-\sigma L} \approx 0,$$

$$\sigma \geq 1.$$

The first assumption is that close elections involve a small fraction of voters, while the second is that at the truncation point the probability of a loss is negligible. The final assumption assures that the probability the second mover loses when a single voter fails to vote $(1/2)e^{-h_k/\sigma} + (1/\sigma)(1/2)e^{-h_k/\sigma} \leq 1$ for all $h_k \geq 0$.

Objective Functions

We now analyze the objective functions and incentive constraints, and introduce a useful approximation for the case of close elections that we will adopt throughout the remainder of the paper. Define $\theta = \mu\psi\pi/\pi_1$, and let the probability that party k wins be Q_k .

Theorem 4.1. *For sure elections with $N\eta_{-K}\phi_{-K} \in [0, N\eta_K\phi_K - \sigma L^+]$ or $N\eta_{-K}\phi_{-K} \in [N\eta_K\phi_K + \sigma L^+, N\eta_{-K}]$, with optimal punishment the normalized objective function of party k is*

$$U_k^s = Q_k - (1/2)c\eta_k(\phi_k)^2 - \eta_k\theta c\phi_k.$$

Define

$$R = \frac{1 - Q_K}{1 - Q_K + (1/\sigma)f(e_{-K}/\sigma)} \max\{0, c\phi_K - (1/\sigma)f(h_{-K}/\sigma)/\eta_K\}.$$

For close elections with $\eta_{-K}N\phi_{-K} = \eta_K N\phi_K + h_{-K}$ and $h_{-K} \in [-\sigma L^-, \sigma L^-]$, with optimal punishment the normalized objective functions are

$$U_K^c = Q_K - (1/2)c\eta_K(\phi_K)^2 - \eta_K\theta R$$

and

$$U_{-K}^c = Q_{-K} - (1/2)c(\eta_K/\eta_{-K})\eta_K(\phi_K)^2 - \eta_K\theta R + O(\sigma L/N).$$

Throughout the remainder of the paper we will adopt the approximation that $U_{-K}^c = Q_{-K} - (1/2)c(\eta_K/\eta_{-K})\eta_K(\phi_K)^2 - \eta_K\theta R$.

Proof. Consider first sure elections. In a sure election either $N\eta_{-K}\phi_{-K} \in [0, N\eta_K\phi_K - \sigma L)$ (including $N\eta_K\phi_K - \sigma L^+$) in which case the probabilities of winning are $Q_K = 1, Q_{-K} = 0$ or $N\eta_{-K}\phi_{-K} \in (N\eta_K\phi_K + \sigma L, \infty)$ (including

$N\eta_K\phi_K + \sigma L^+$) in which case the probabilities of winning are $Q_K = 0, Q_{-K} = 1$. There is no pivotality and punishment P_k is for bad signals. Notice that the cost of punishment ψP_k must be paid for the fraction of the non-voting population for whom bad signals can be received, μ , times the probability π they get bad signals. Hence the total social cost of punishment is $\eta_k N\pi\mu\psi P_k$. The normalized objective function of party k is therefore

$$Q_k - (1/2)c\eta_k(\phi_k)^2 - \eta_k\pi\mu\psi P_k.$$

The marginal voter n_k receives $-cn_k/(N\eta_k) = -c\phi_k$ for voting and $-\pi_1 P$ for not voting. Hence the incentive constraint is

$$P_k \geq c\phi_k/\pi_1.$$

Since punishment is socially costly, optimality requires that it be chosen as small as possible subject to the incentive constraints. Solving the incentive constraints with equality and substituting into (approximate) objective functions gives

$$Q_k - (1/2)c\eta_k(\phi_k)^2 - \eta_k(\pi/\pi_1)\mu\psi c\phi_k$$

for sure elections. Here the final term is the *monitoring cost* due to the need to provide incentives.

Now consider close elections. In a close election, the intended turnout of the first mover is $\eta_K N\phi_K$ and the turnout of the second mover is $\eta_{-K} N\phi_K + h_{-K}$ where $h_{-K} \in [-\sigma L^-, \sigma L^-]$. The probability the second mover wins is $Q_{-K} = F(h_{-K}/\sigma)$, the probability the first mover wins is $1 - Q_{-K}$, and the probability of being pivotal is $(1/\sigma)f(h_{-K}/\sigma)$ for both parties. Suppose that for close elections punishment is conditional on losing the election; we will show later on that this is without loss of generality since it reduces the expected monitoring cost. Conditioning on losing the election, the expected monitoring cost is $(1 - Q_k)\pi\mu\psi P_k$. The normalized objective function of party K is therefore

$$Q_K - (1/2)c\eta_K(\phi_K)^2 - \eta_K(1 - Q_K)\pi\mu\psi P_K.$$

That of the second mover is

$$Q_{-K} - (1/2)c\eta_{-K} \left(\frac{\eta_K N\phi_K + h_{-K}}{\eta_{-K} N} \right)^2 - \eta_{-K}(1 - Q_{-K})\pi\mu\psi P_{-K}.$$

By assumption $|h_{-K}| \leq \sigma L \ll N$ we can approximate h_{-K}/N as being zero, and we have the approximation

$$Q_{-K} - (1/2)c(\eta_K/\eta_{-K})\eta_K(\phi_K)^2 - \eta_{-K}(1 - Q_{-K})\psi\pi\mu P_{-K}.$$

The first mover's marginal voter receives $-cn_K/(N\eta_K) + (1/\sigma)f(h_{-K}/\sigma)/\eta_k$ for voting and $-\pi_1(1 - Q_K + (1/\sigma)f(h_{-K}/\sigma))P_K$ for not voting, so the incentive

constraint is

$$P_K \geq \frac{c\phi_K - (1/\sigma)f(h_{-K}/\sigma)/\eta_K}{\pi_1(1 - Q_K + (1/\sigma)f(h_{-K}/\sigma))}.$$

Notice that if the punishment were not given conditional on losing the election, then the incentive constraint would be $P_K \geq (c\phi_K - (1/\sigma)f(h_{-K}/\sigma)/\eta_K)/\pi_1$, but the expected cost of monitoring would be proportional to P_K rather than to $(1 - Q_K)P_K$, and therefore would be at least as large and strictly larger if the expected cost of punishment were positive. On the other hand, if the punishment occurred only when there was a one-vote loss it would make no difference: to preserve incentive compatibility the size of the punishment would have to be increased to match the reduced chance that it would be received.

The second mover's marginal voter n_{-K} receives

$$\begin{aligned} -cn_{-K}/(N\eta_{-K}) + (1/\sigma)f(h_{-K}/\sigma)/\eta_{-K} = \\ -c\frac{\eta_K N\phi_K + h_{-K}}{N\eta_{-K}} + (1/\sigma)f(h_{-K}/\sigma)/\eta_{-K} \end{aligned}$$

for voting and $-\pi_1(1 - Q_{-K} + (1/\sigma)f(h_{-K}/\sigma))P_{-K}$ for not voting. Again we propose to approximate by ignoring the small term h_{-K}/N , so the incentive constraint is

$$P_{-K} \geq \frac{c(\eta_K/\eta_{-K})\phi_K - (1/\sigma)f(h_{-K}/\sigma)/\eta_{-K}}{\pi_1(1 - Q_{-K} + (1/\sigma)f(h_{-K}/\sigma))}.$$

By a similar argument to the first mover, if the punishment were not conditional on losing the election, the expected cost of monitoring would be at least as large and strictly larger if the expected cost of punishment were positive.

For close elections 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 P_k and substituting $\max\{0, P_k\}$ into (approximate) objective functions. For the first mover we have

$$\begin{aligned} Q_K - (1/2)c\eta_K(\phi_K)^2 \\ - \eta_K(\pi/\pi_1)\mu\psi \frac{1 - Q_K}{1 - Q_K + (1/\sigma)f(h_{-K}/\sigma)} \max\{0, c\phi_K - (1/\sigma)f(h_{-K}/\sigma)/\eta_K\} \end{aligned}$$

and for the second

$$\begin{aligned} Q_{-K} - (1/2)c(\eta_K/\eta_{-K})\eta_K(\phi_K)^2 \\ - \eta_K(\pi/\pi_1)\mu\psi \frac{1 - Q_{-K}}{1 - Q_{-K} + (1/\sigma)f(h_{-K}/\sigma)} \max\{0, c\phi_K - (1/\sigma)f(h_{-K}/\sigma)/\eta_K\}, \end{aligned}$$

as given in the statement of the theorem. \square

Equilibrium

Our main theorem is

Theorem 4.2. *In any equilibrium there are three possible type of commitment: preemptive in which $\phi_L > 0$ and $\phi_S = 0$, concession in which $\phi_K = 0$ and $h_{-K} = \sigma L^-$, or contested in which commitment is $\eta_K \phi_K = (1/2)(1/\sigma)(1/c)$ and the second mover responds with $h_{-K} = 0$. The small group never preempts, and for all $c, \mu, \pi, \pi_1, \eta_L, \eta_S, \sigma$ there is some $\bar{\psi} > 0$ such that for $\psi > \bar{\psi}$ the large group does not preempt and for $\psi < \bar{\psi}$ the large group does preempt. In addition for fixed $c, \mu, \pi, \pi_1, \sigma$ we have $\bar{\psi} \rightarrow 0$ as $\eta_S/\eta_L \rightarrow 1$ and in particular the election is contested whenever $\psi > \pi_1(\sigma + 1)/(\mu\pi)$.*

This follows from Theorems 4.3 and 4.4 below. A key point of the theorem is that decreasing σ , meaning better targeting of voters by the parties, results in closer vote totals and higher turnout as long as groups are not too asymmetric.

Backwards Induction

To solve for subgame perfect equilibrium and prove the theorem, we proceed by backward induction. We first calculate the best responses of the second mover and then we turn to the optimal behavior for the first mover.

Theorem 4.3. *Define \hat{h}_{-k} by*

- (i) *if either $\phi_K < (1/c)(1/\sigma)(1/2)e^{-L}/\eta_K$ or $\theta/(\sigma+1) < 1$, it is $\hat{h}_{-K} = \sigma L^-$.*
- (ii) *if $\phi_K > (1/c)(1/\sigma)(1/2)e^{-L}/\eta_K$ and $\theta/(\sigma + 1) > 1$, it is $\hat{h}_{-K} = -\sigma \log(2\sigma c \eta_K \phi_K)$ if $\sigma c \eta_K \phi_K < 1/2$, and it is $\hat{h}_{-K} = 0$ otherwise.*

There is a cutoff $\bar{\phi}_K \geq (1/2)(1/(\eta_K c[\theta + (1/2)(\eta_K/\eta_{-K}]])$ such that for $\phi_K < \bar{\phi}_K$ the response \hat{h}_{-k} earns positive profits and is the best response of the second mover and such that for $\phi_K > \bar{\phi}_K$ the response \hat{h}_{-k} earns negative profits and the best response of the second mover is $\phi_{-K} = 0$. The cutoff $\bar{\phi}_L < \bar{\phi}_S$ and if $\sigma > (1/c)(1/\eta_K)$ it is decreasing in σ .

Proof. The second mover can guarantee a payoff of zero by choosing $\phi_{-K} = 0$. Therefore the second mover will only choose a close election if it can guarantee itself a positive payoff. The second mover would never choose $h_{-K} < 0$ since given the probability of winning the election would be higher choosing $|h_{-K}|$, the monitoring cost would be the same, given the Laplace distribution assumption, and the cost of voting would be approximately the same. Using the Laplace distribution, in a close election the second mover's objective function for $h_{-K} \geq 0$ is

$$V_K = 1 - (1/2)e^{-h_{-K}/\sigma} - (1/2)c(\eta_K/\eta_{-K})\eta_K(\phi_K)^2 - \eta_K(\pi/\pi_1)\mu\psi \frac{\sigma}{\sigma+1} \max\{0, c\phi_K - (1/\sigma)(1/2)e^{-h_{-K}/\sigma}/\eta_K\}.$$

This is strictly decreasing in ϕ_K and η_K and positive at $\phi_K = 0$. Hence the optimum \hat{V}_K with respect to h_{-K} is also strictly decreasing in these arguments. This shows there is a unique $\bar{\phi}_K$ and that $\bar{\phi}_S < \bar{\phi}_L$.

We can bound \hat{V}_K from below by taking $h_{-K} = 0$. From this we see that $V_K \geq (1/2) - \eta_K c[\theta + (1/2)(\eta_K/\eta_{-K})]\phi_K$ giving the stated lower bound on $\bar{\phi}_K$. Finally from below $\sigma c \eta_K \phi_K < 1/2$ implies that in fact $\hat{h}_{-K} > 0$ so that V_K

is also strictly decreasing in σ . Since $\phi_K \leq 1$ this gives the sufficient condition for $\bar{\phi}_K$ to be decreasing in σ .

Turning to the optimal choice of \hat{h}_{-K} : for $c\phi_K - (1/\sigma)(1/2)e^{-h_{-K}/\sigma}/\eta_K < 0$ we have V_K strictly increasing in h_{-K} , so either $\phi_K < (1/c)(1/\sigma)(1/2)e^{-L}/\eta_K$ in which case the optimum is σL^- or the derivative of the objective function with respect to h_{-K} is

$$\begin{aligned} & (1/\sigma)(1/2)e^{-h_{-K}/\sigma} - \eta_K(\pi/\pi_1)\mu\psi \frac{\sigma}{\sigma+1} (1/\sigma^2)(1/2)e^{-h_{-K}/\sigma}/\eta_K \\ &= \left(1 - (\pi/\pi_1)\mu\psi \frac{1}{\sigma+1}\right) (1/\sigma)(1/2)e^{-h_{-K}/\sigma}. \end{aligned}$$

Hence for $(\pi/\pi_1)\mu\psi/(\sigma+1) < 1$ the solution is $h_{-K} = \sigma L^-$ and for $(\pi/\pi_1)\mu\psi/(\sigma+1) > 1$ the solution entails to reduce h_{-K} until either $e^{-h_{-K}/\sigma} = 2\sigma c\eta_K\phi_K$, so that $c\phi_K - (1/\sigma)(1/2)e^{-h_{-K}/\sigma}/\eta_K = 0$ and the cost of monitoring is zero, or it is given by $h_{-K} = 0$. This finishes the proof of the theorem.

Notice that in order to gain the election with probability near one the second mover can choose $h_{-K} = \sigma L^-$. We are assuming e^{-L} negligible and $\sigma \geq 1$ so that $(1/\sigma)e^{-L}$ is also negligible; this means that we attribute to σL^- no chance of a loss and no chance of being pivotal, but the multiplier

$$\frac{1 - Q_{-K}}{1 - Q_{-K} + (1/\sigma)f(e_{-K}/\sigma)}$$

is still equal to $\sigma/(\sigma+1)$ so that σL^- is better than σL^+ . That is, the second mover will only induce a sure election if it concedes with $\phi_{-K} = 0$. \square

We can revisit now the assumption that voters with costs above the marginal voter can be costlessly deterred from voting. For sure elections, since the marginal voter $n_k = \phi_k N \eta_k$ is indifferent between voting or not, voters with larger costs will not be tempted to vote even if there is no punishment or monitoring for them. For close elections, given the best response behavior of the second mover, both the first mover and the second mover's marginal voters are indifferent between voting or not, so again voters with larger costs than the marginal voter are not tempted to vote in equilibrium. Thus, the assumption is without loss of generality.

Our main Theorem now follows from

Theorem 4.4. *If $\theta < \sigma + 1$, the first mover commit to $\phi_K = 0$ and the second mover responds with $h_{-K} = \sigma L^-$, or the first mover preempts. If $\theta > \sigma + 1$, the first mover commits to $\eta_K\phi_K = (1/2)(1/\sigma)(1/c)$ and the second mover responds with $\eta_{-K}\phi_{-K} = (1/2)(1/\sigma)(1/c)$ and $h_{-K} = 0$. The small group never preempts, and for all $c, \eta_L, \eta_S, \sigma$ there is some $\bar{\theta} \leq \sigma + 1$ such that for $\theta > \bar{\theta}$ the large group does not preempt and for $\theta < \bar{\theta}$ the large group does preempt. Moreover, for all c, σ , we have $\bar{\theta} \rightarrow 0$ as $\eta_S/\eta_L \rightarrow 1$.*

Proof. Consider first the case $\theta/(\sigma+1) < 1$. Given the best response of the second mover in Theorem 4.3, the second mover wins the election with probability

one as long as it can guarantee itself a positive payoff. Thus, the first mover either preempts by choosing ϕ_K large enough so that the second mover is better off conceding and choosing $\phi_{-K} = 0$, or it concedes by choosing $\phi_K = 0$, and the second mover chooses $h_{-K} = \sigma L^-$. Preemption by the first mover requires that the group's turnout choice, say ϕ^* , is large enough to keep the second mover out, that is

$$1 - (1/2)c(\eta_K/\eta_{-K})\eta_K(\phi^*)^2 - \frac{\sigma}{1+\sigma}\eta_K\theta c\phi^* = 0, \quad (4.1)$$

but small enough so that the first mover has a positive gain by not conceding, that is

$$1 - (1/2)c\eta_K(\phi^*)^2 - \eta_K\theta c\phi^* \geq 0.$$

The inequality cannot be satisfied if the first mover is the small group, since for $\eta_K/\eta_{-K} < 1$ the LHS of the inequality is smaller than the LHS of equation 4.1. Hence the small group never preempts for $\theta/(\sigma + 1) < 1$. Consider that the large group is the first mover. As $\theta \rightarrow 0$ as $0 \leq \phi^* \leq 1$ we see that $(1/2)c\eta_L(\phi^*)^2 \rightarrow \eta_S/\eta_L$. Hence the LHS of the inequality approaches $1 - \eta_S/\eta_L > 0$. Thus, the large group will choose to preempt for small enough θ .

We claim that if for some $0 < \bar{\theta} < \sigma + 1$ the large group is indifferent or prefers not to preempt, then the large group prefers not to preempt for all $\bar{\theta} < \theta \leq \sigma + 1$. To see this, notice that the first mover prefers not to preempt if ϕ^* is larger than $\bar{\phi}^*$ given by the unique solution for $\bar{\phi}^* \geq 0$ of

$$1 - (1/2)c\eta_L(\bar{\phi}^*)^2 - \eta_L\theta c\bar{\phi}^* = 0.$$

Applying the implicit function theorem to this equation and equation 4.1 for ϕ^* we have

$$\partial\bar{\phi}^*/\partial\theta = -\bar{\phi}^*/(\bar{\phi}^* + \theta)$$

and

$$\partial\phi^*/\partial\theta = -\phi^*/(((1+\sigma)/\sigma)(\eta_L/\eta_S)\phi^* + \theta).$$

Hence for $\bar{\phi}^* < ((1+\sigma)/\sigma)(\eta_L/\eta_S)\phi^*$ it must be that $\partial\bar{\phi}^*/\partial\theta < \partial\phi^*/\partial\theta$. Suppose that for some $0 < \bar{\theta} < \sigma + 1$ the large group is indifferent or prefers not to preempt: so it must be that $\phi^* \geq \bar{\phi}^*$. So certainly $((1+\sigma)/\sigma)(\eta_L/\eta_S)\phi^* > \bar{\phi}^*$ so $\partial\bar{\phi}^*/\partial\theta < \partial\phi^*/\partial\theta$. Hence for $\theta > \bar{\theta}$ we must have $\phi^* \geq \bar{\phi}^*$ implying that the large group strictly prefers not to preempt.

Now consider the case $\theta/(\sigma + 1) > 1$. The first mover loses for sure if $\phi_K < (1/c)(1/\sigma)(1/2)e^{-L}/\eta_K$ so any such choice is dominated by $\phi_K = 0$. In the range $(1/c)(1/\sigma)(1/2)e^{-L}/\eta_K \leq \phi_K \leq (1/c)(1/\sigma)(1/2)/\eta_K$, given the best response of the second mover, the payoff of the first mover is

$$U_K^c = \sigma c\eta_K\phi_K - (1/2)c\eta_K(\phi_K)^2,$$

which is strictly increasing in ϕ_K since $\phi_K < 1 \leq \sigma$. Hence, if choosing a close

election, the first mover chooses

$$\phi_K = (1/2)(1/\sigma)(1/c)/\eta_K,$$

which is the minimum turnout guaranteeing $h_{-K} = 0$; any higher choice in a close election leads to the same probability of winning, larger turnout costs, and positive monitoring costs. The first mover utility of choosing a close election is therefore

$$\hat{U}_K^c = (1/2) - (1/2)c\eta_K((1/2)(1/\sigma)(1/c)/\eta_K)^2 = 1/2 - (1/8)(1/\sigma^2)(1/(c\eta_K)).$$

Notice that the first mover utility of choosing a close election is always better than conceding since $\sigma \geq 1$ and the lower bound $c > 2\eta_L/\eta_S^2$ implies $c\eta_s > 2$.

Now we need to consider whether it would be better for the first mover to preempt. The payoff of successfully preempting is

$$1 - (1/2)c\eta_K(\phi_K)^2 - \eta_K\theta c\phi_K,$$

where ϕ_K has to be large enough to induce the second mover to concede, thus $\phi_K > (1/2)(1/\sigma)(1/c)/\eta_K$, which implies that the payoff of successful preemption is smaller than

$$1 - (1/8)(1/\sigma^2)(1/(c\eta_K)) - \theta/(2\sigma).$$

This is smaller than the payoff of inducing a contested election given that $\theta/(\sigma+1) > 1$. Thus, the first mover never finds it optimal to preempt for $\theta/(\sigma+1) > 1$.

From previous steps, for all $c, \eta_L, \eta_S, \sigma$ there is some $0 < \bar{\theta} \leq \sigma+1$ such that for $\theta > \bar{\theta}$ the large group does not preempt and for $\theta < \bar{\theta}$ the large group does preempt. Fixing any $1 + \sigma > \bar{\theta} > 0$, for η_S/η_L close enough to one, the LHS of equation 4.1 is larger than the LHS of the positive gain inequality, but the RHS of the inequality is larger than the RHS of equation 4.1 for all $\theta \geq \bar{\theta}$, so that $\bar{\theta} \rightarrow 0$ as $\eta_S/\eta_L \rightarrow 1$. \square

From the statement of Theorem 4.4, for η_S/η_L close enough to one, we have $\bar{\theta} < \mu\pi/\pi_1$. Hence, for enough symmetry in the support of the parties, there are no preemption equilibria for $\psi \geq 1$.

5. The Intermediate Case

We now consider the intermediate case $0 < \Omega < 1$. We continue to assume constant marginal cost: $C(n_k/(\eta_k N)) = cn_k/(\eta_k N)$.

Incentive Compatibility Without Pivotality

If there is no pivotality, that is, if targeting is not feasible, the marginal voter ϕ_k pays $c\phi_k$ to vote and loses $\pi_1 P$ for not voting and must be indifferent so $P = c\phi_k/\pi_1$. A fraction μ of the group who are excused from voting get bad signals π of the time, so as before define $\theta = \mu\psi\pi/\pi_1$. We then see that on

the equilibrium path the social cost of voting per capita, or monitoring cost, is $\theta c\phi_k$ and that the normalized cost is $\theta\eta_k c\phi_k$. Hence the total cost of voting is $v_k(\phi_k) = \eta_k c((1/2)\phi^2 + \theta\phi_k) = g_k(\phi_k)$.

Intermediate Payoffs

Let $\Pi_k(\phi_k)$ be the payoff in the targeted game when k is first mover and the second mover makes the optimal response given in Theorem 4.4. The key point is that this is a continuous function. Hence with probability $\Omega/2$ party k gets an amount that does not depend upon ϕ_k , with probability $\Omega/2$ they get $\Pi_k(\phi_k)$ and with probability $1 - \Omega$ they win 1 per capita with a continuous probability of winning $P(\eta_k\phi_k, \eta_{-k}\phi_{-k})$ and face cost $g_k(\phi_k)$. We assume that ζ_k^0 is noisy enough that in this second game pivotality has no appreciable effect on incentive constraints, but is not so noisy that $P(\eta_k\phi_k, \eta_{-k}\phi_{-k})$ fails to be close to the all-pay auction in which the highest vote count wins for sure.

We may define the pseudo cost $D_k(\phi_k) = -(\Omega/2)\Pi_k(\phi_k) + (1 - \Omega)g_k(\phi_k)$ and we see that this game is just an ordinary contest with contest success $P(\eta_k\phi_k, \eta_{-k}\phi_{-k})$ and costs $D_k(\phi_k)$. This has the unusual feature that $D_k(\phi_k)$ while continuous is not generally increasing because $\Pi_k(\phi_k)$ is not generally decreasing.

In the case in which the targeted game results in large turnout razor thin elections we observe that Ω can be interpreted as the probability of a razor thin election.

Equilibrium

We now wish to argue that the equilibrium exist and that the equilibrium correspondence is upper hemi-continuous. Since the objective functions are continuous, existence follows in the standard way from the Glicksberg fixed point theorem. Upper hemi-continuity almost follows from Theorem 6 in Levine and Mattozzi (2022), but while they allow for $\Omega = 1$ so that the contest success function is completely flat, they require that $D_k(\phi_k)$ be increasing which we do not know to be the case. However, if we consider a sequence $\Omega^n \rightarrow \Omega$ they use this only to show that $D_k^n(\phi_k) \rightarrow D_k(\phi_k)$ uniformly. In our case this follows from the fact that $D_k(\phi_k)$ is a convex combination of two fixed continuous functions, hence the theorem does apply. In particular as $\Omega^n \rightarrow 1$ equilibrium distributions converge weakly to the unique equilibrium we have computed, and as $\Omega^n \rightarrow 0$ they converge to the equilibrium of the contest which is itself by the same theorem close to the usual all-pay auction equilibrium.

When Are Elections Close?

We rejected the non-pivotal theory that $\Omega = 0$ as not generating enough close elections. On the other hand the theory with $\Omega = 1$ generates only close elections or landslides which is also against the evidence. As close elections are rare events, this suggests that the best model is Ω small but not zero. That is, generally elections are non-pivotal, but occasionally an event allows targeting and they become pivotal. What does the model say in this case?

If Ω is small then to a good approximation we know the equilibrium strategies are like the non-pivotal one: for example if the *ex ante* noise is small like those in an all pay auction where each party's equilibrium cdf is proportional to the cost function of the other party, with the small party also have an atom at zero. What happens when in fact the election turns out to be pivotal? The turnout of the first mover ϕ_K is already determined by a draw from the equilibrium distribution for the all pay auction. If this draw is bigger than $\bar{\phi}_K$ then the second mover opts out and the election is a landslide. If it is smaller than $\bar{\phi}_K$ then we get a razor thin margin. Notice that this is more likely if the first mover is the smaller party for two reasons: first, the cutoff for the smaller party as first mover is bigger than for the large, and second, the smaller party makes smaller bids so is more likely to be below the threshold. Regardless, there is a positive probability ρ of a close election when the signal is received in time. As a close election is rather unlikely in the all pay auction, the probability of a close election is given as a good approximation by $\Omega\rho$.

Finally, we observe that for large σ by Theorem 4.3 the cutoffs $\bar{\phi}_K$ are decreasing in σ so that better targeting will indeed increases the chances of a close election.

6. Conclusion

We propose a dynamic model of electoral competition to explain the apparent abundance of knife-edge elections. If the cost of monitoring voters is low, and the small group commits to a turnout level first, it concedes the election, and if the large group commits to a turnout level first, it either concedes, or (if there is enough asymmetry in the electoral support) it preempts and wins by a landslide. If instead the cost of monitoring voters is high, there is a knife-edge election regardless of which group moves first. Moreover, the cost threshold moves down as targeting of voters becomes more precise, which implies that better targeting of voters will make head-to-head elections increasingly knife-edge.

Although there are few elections involved, 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. 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 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.

References

- Blais, André (2000): *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory*. University of Pittsburgh Press.
- 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, 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.
- De la Cuesta, B., K. Imai (2016): “Misunderstandings About the Regression Discontinuity Design in the Study of Close Elections,” *Annual Review of Political Science* 19: 375–396.
- 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.
- 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, David 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.
- Mulligan, C. B. and C. G. Hunter (2003): “The Empirical Frequency of a Pivotal Vote,” *Public Choice* 116: 31–54.
- Palfrey, Thomas R., and Howard Rosenthal (1985): “Voter Participation and Strategic Uncertainty,” *American Political Science Review* 79: 62–78.
- 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.