Do Polls create Momentum in Political Competition?

Philipp Denter[1]
Dana Sisak[2]

[1] University of St. Gallen, Switzerland;
[2] Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute, The Netherlands.

Do Polls Create Momentum in Political Competition?[*]

Philipp Denter

University of St.Gallen

School of Economics and Political Science

philipp.denter@unisg.ch

Dana Sisak

Erasmus University Rotterdam &

Tinbergen Institute

sisak@ese.eur.nl

This version: October 2013. First draft: June 2009.

Abstract

We explore how public opinion polls affect candidates' campaign spending in political competition. Generally, polls lead to (more) asymmetric behavior. Under a majority rule there always exists an equilibrium in which the initially more popular candidate invests more in the campaign and thereby increases her lead in expectation: polls create momentum. When campaigning is very effective and the race is very close, a second type of equilibrium may exist: the trailing candidate outspends and overtakes his opponent. Regardless of the type of equilibrium, polls have a tendency to decrease expected total campaigning expenditures by amplifying ex-ante asymmetries between candidates and thus defusing competition. When candidates care also for their vote share in addition to having the majority, candidates' incentives crucially depend on the distribution of voters' candidate preferences.

JEL Codes: D02, D72, D74, D83

Keywords: polls, political campaigns, feedback, momentum

1 Introduction

While an informed electorate is generally considered essential for a well-functioning democracy, one exception concerns polls of candidates' relative standing. Critics claim that polls undermine both the incentive to vote as well as the vote itself; thus distorting voting decisions. As a consequence the preferences of the populace are warped by the echo chamber of opinion polls. For this reason, many countries have imposed a ban on the publication of pre-election polling results. This ban can range from one day before the election, as in France, to a whole month before the election, as in Luxembourg.[1] Candidates, however, are still allowed to commission opinion polls, even if the general public is not allowed access to the results. In this paper we study the effect of opinion polls on candidates' incentives to invest in their campaign. These investments in turn influence the voters' ballot choice on election day and thus the final election outcome.

Polls are ubiquitous. In addition to traditional providers like Gallup and Rasmussen Reports, newspapers and TV stations conduct their own polls. Poll aggregator websites, such as realclearpolitics.com collect the plethora of poll results and offer a structured overview to the public. In addition, prediction markets such as politicalbetting.com and oddschecker.com offer alternative sources of voters' approval (Berg, Nelson, and Rietz, 2008). Given the relevance of polls, numerous studies have analyzed the various channels through which information about candidates' relative standing might affect voters' behavior. A prominent hypothesis is the existence of a bandwagon effect, which posits that voters increasingly cast their vote for candidates doing well in the polls. This suggests that polls create momentum in the sense that the front-runner improves his position over time. However, there are also studies predicting anti-momentum, for example due to a mobilization effect, with the consequence that the front-runner in the polls experiences a loss of support in the election.

While scholars show a strong interest in the analysis of voters' reaction to polls, the other side of the political market – the politicians and parties – has been virtually neglected. To increase their chances of election, political candidates invest aggressively into their campaigns: Total campaign spending during the 2012 Presidential Campaign in the United States amounted to more than USD 2 billion.[2] Because the fraction of partisan voters has been shrinking in many countries (e.g. Dalton and Wattenberg

---

[1]Source: http://aceproject.org. Accessed: September 10, 2013.

[2]Source: http://www.opensecrets.org/pres12. Accessed: September 10, 2013.

(2001)), there are more swing voters to be swayed in a campaign, making campaigning increasingly important in determining the election outcome. Advances in social media and communication technology offer a plethora of new channels to target specific voter groups more effectively and thus increase the impact of a campaigning dollar.

Voter preferences and voting intentions are not directly observable to the political candidates and typically change over the course of time due to taste or information shocks. The possibility of reliable up-to-date information about candidates' relative standing is thus at the heart of our analysis. We analyze how this information influences the incentives of candidates to spend in the course of their political campaign. In particular, we are interested in how candidates' *relative* incentives to compete are affected and whether polls affect the likelihood of an incumbent winning the election. We construct a model in which candidates for political office may spend resources early and late in the campaign to gain voters' support. If there is no poll, candidates a priori have a common belief about the median voter's candidate ranking and cannot update their beliefs as the election day comes closer. If there is a poll, candidates know the median's candidate ranking when making their investment decisions. Our main results are as follows:

- If candidates are mainly concerned with winning the election, polls give the front-runner an incentive to campaign harder than her opponent and thereby create momentum. Thus, there always exists an equilibrium where the front-runner increases her lead in expectation.

- In an environment where candidates' campaign expenditures are highly effective in influencing voters, polls may also create anti-momentum. The trailing candidate may run a more costly campaign than his adversary, but only if candidates have relatively similar popularity. Equilibria with momentum and anti-momentum co-exist.

- Increasing the precision of polls has an ambiguous effect. In close races more precise polls lead to stronger momentum. In very lopsided races momentum decreases as polls become more precise, because candidates slack off in a mostly decided race.

- Polls can make a campaign both more and less wasteful in terms of aggregate expenditures. If (but not only if) the spending elasticity of campaigning costs decreases polls decrease aggregate spending and thus wastefulness.

- When candidates are sufficiently interested in their vote shares and not only in gaining a majority, unique equilibria exhibiting anti-momentum may occur in a polarized electorate.

To illustrate the intuition behind our results consider two ex-ante equally popular candidates. They campaign over a certain period of time and during this time random (unobservable) shocks to their popularity occur. Without a poll, candidates never learn about their current standing with the voter and thus at any point in time incentives are completely symmetric. With polls on the other hand, candidates receive updates about their current relative standing. This alters their incentives in the following way. A candidate who receives the information that she is ahead, now has an additional incentive to invest. The reason is that any additional investment now affords her an even greater lead in expectation in the next poll. This in turn decreases the expected intensity and thus expected costs of future competition. A trailing candidate on the other hand has a weaker incentive to invest. Any additional unit of investment now brings him closer in expectation to his opponent in the next poll and thus it makes future competition more fierce and costly in expectation. Consequently polls drive a wedge between the investment incentives of the trailing and leading candidate and thereby create momentum.

When campaigning is relatively effective and the candidates are relatively close, this intuition is valid for the trailing candidate as well. A large investment helps him overtake his opponent and at the same time defuse future competition in expectation. In these situations both candidates have an incentive to preempt the other with the objective to save costs in the future. In terms of aggregate expenditures of the candidates, note that initial asymmetries are amplified through a poll. This tends to defuse competition and thus lower investments into the campaigns. In addition though, the candidates also face a less noisy decision environment, which has a differential effect on expenditures. The overall effect of polls on wastefulness is thus ambiguous.

The paper is organized as follows. In the remainder of this section we review the relevant literature. Section 2 sets up and explains the basic model. In Section 3 we explore the effect of providing information through polls in a system of pure majority voting. We discuss the effect of polls on the intensity of political competition in Section 4 and candidates' expected spending profiles in Section 5. Section 6 contains extensions of the basic model. First we show that our technical assumptions are not too restrictive and that in a more general model our findings will be qualitatively preserved. In the second

part of the section we study the effects of polls when candidates care not only about majorities but also about their actual vote shares. In the third part we study imprecise polls. We conclude the section with a discussion of the costs of campaigning. Section 7 concludes. We relegate all formal proofs to the appendix.

Related Literature Other scholars have directed their attention to the effect of polls on election outcomes. The incentives to conduct polls about voters' policy preferences to inform campaign positioning choices are analyzed for example in Bernhardt, Duggan, and Squintani (2009) and Jacobs and Shapiro (1994). In contrast, we focus on the financial aspect of campaigning. In our paper polls inform about candidates' relative standing and we study how polls inform campaign expenditure decisions. Earlier models of how candidates use their money and time during an electoral contest were studied for example by Brams and Davis (1973, 1974), Snyder (1989), Skaperdas and Grofman (1995), Stromberg (2008), Iaryczower and Mattozzi (2013), Denter (2013), and Meirowitz (2008). To be able to study polls and the associated repercussions for electoral outcomes, we study candidate incentives in a dynamic campaigning model. The dynamic nature of the model is essential, since the effects we are interested in can only emerge if candidates can learn over time. The only other paper we are aware of studying dynamics in a strategic model of campaign spending is Klumpp and Polborn (2006). While the authors study incentives in sequential electoral contests during the primaries, we consider spending dynamics and the role of informational feedback within a single contest.

Also the effect of polls on voters' decisions has been investigated. Among others Hong and Konrad (1998), Straffin Jr. (1977), Bikhchandani, Hirshleifer, and Welch (1992), Callander (2007), as well as Ali and Kartik (2012) show how polls can distort voters' decisions at the ballot in favor of the more popular/leading candidate. Rationales for this *bandwagon* effect include information cascades, the desire to vote for the winner, and voters' aversion to uncertainty. Other authors, e.g. Goeree and Großer (2007), argue in the opposite direction, claiming that leading in the polls might actually be harmful, due to a negative mobilization effect. A number of empirical papers have tried to falsify one or the other hypothesis, but by now evidence is at best inconclusive (see for example Blais, Gidengil, and Nevitte (2006); McAllister and Studlar (1991); Nadeau, Cloutier, and Guay (1993); Vowles (2002); Knight and Schiff (2010)). Our paper also addresses the effect of polls on voting decisions and election outcomes, but through an indirect channel: the strategies of political candidates. Candidates vie for voters by spending

time and money on their campaign. We study how candidates' incentives to engage in such persuasive efforts are affected by polls.

Showing that polls may cause momentum in favor of the front-runner, our paper also adds to the literature identifying and explaining an incumbency advantage in political competition. For some early empirical evidence on the existence of an incumbency advantage see for example Erikson (1971) or Gelman and King (1990). A textbook justification for this phenomenon are political business cycles as studied by Nordhaus (1975) or Drazen (2001). A consequence of these theories is that incumbents are likely to have popularity advantages at the outset of a campaign. Our paper shows that this advantage is likely to be amplified through opinion polls.

The momentum effect we identify is closely related to the so-called "discouragement effect" found in dynamic sequential contests (Konrad, 2012). The main intuition here is that the outcome of the present contest or battle affects the perceived value of future contests through its effect on the expected future effort costs. In some cases, such as in Klumpp and Polborn (2006) this effect can also magnify initial advantages. Konrad (2012) illustrates this effect for models of the "tug-o-war" and the "race" class. We show that this effect is also relevant for a different class of models, where competition takes place over a fixed time horizon.

Finally, a literature that is related because of the class of models employed analyzes workers' incentives in labor market tournaments. Aoyagi (2010), Ederer (2010), Gershkov and Perry (2009) as well as Klein and Schmutzler (2013) study the optimal feedback policy of a principal in a dynamic promotion tournament, where the principal's goal is the maximization of aggregate effort. We add to this literature in two ways. First, in order to study momentum in a campaign we depart from the modeling practice of this literature by looking at multiple feedback stages and ex-ante asymmetric candidates. Second, we study the effect of two different "reward schemes:" promotions (concern for winning) and sharing of a bonus (concerns for vote share). Empirical evidence on dynamic tournaments with and without feedback is discussed for example in Fershtman and Gneezy (2011). They find that feedback about relative performance increased the quitting rate among 10th graders in a 60m race. This is in line with our theoretical findings.
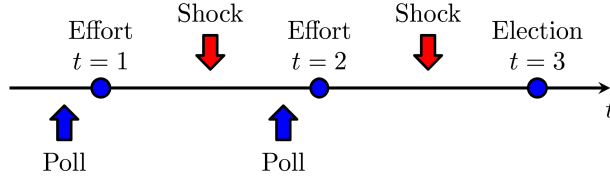
Figure 1: Timing.

2   The model

Two candidates, $i \in \{F, T\}$ compete for voters' approval over two stages, $t = 1$ and $t = 2$. We assume that policy platforms were chosen and committed to before $t = 1$. Campaigning in $t = 1$ and $t = 2$ is thus purely persuasive and candidates sway voters by choosing investments $x_i^t \geq 0$, $i = F, T$, $t = 1, 2$, which can for example be thought of as TV advertisements or public appearances. Throughout the campaign, candidates experience random shocks to their popularity. Polls inform candidates about their relative standing. In stage three, $t = 3$, the election takes place. Figure 1 summarizes the timing of the game.

We assume that a simple majority rule is used ("first past the post"-system, FPTP). This implies that candidates vie for the median voter. The median voter votes for the candidate which affords her the highest expected utility. Formally, we define the (perceived) difference in utilities between candidate F and T on election day ($t = 3$) as[3]

$$d^3 = a + \sum_{t=1}^{2} \left(x_F^t - x_T^t \right) - \sum_{t=1}^{2} e^t \tag{1}$$

The first term, a represents the realized utility difference at the outset of the campaign before any campaign effort was chosen. In case the election were held just before period 1 a would solely determine the election outcome. This parameter can be interpreted as containing the prior decisions on policy

---

[3]In many situations, e.g. with a multi-dimensional policy space, a median voter may fail to exist. In these situations, our so called median voter may also be interpreted simply as a decisive voter. For example, if there are n issues, let preferences of voter v be of the following form:

$$d_v^3 = \sum_{t=1}^{2} \left(x_F^t - x_T^t \right) - \sum_{t=1}^{2} e^t - \sum_{m=1}^{n} w^m \left[(b_v^m - p_F^m)^2 - (b_v^m - p_T^m)^2 \right],$$

where $w^m \geq 0$ are issue weights, p_i^m candidate i's position in issue m, and voters' bliss points b_v^m are distributed on some n-dimensional policy space. Then it is always possible to choose x_i in such a way that the electorate is split evenly between candidates and each receives a popular vote of 50 percent. The type of voter (identified by her bliss point) that is indifferent between candidates in such a situation is the decisive voter that we are referring to as median voter.

platforms, the valence of the candidates, an incumbency advantage or random events prior to $t = 1$. From the candidates' perspective a is the realization of a random variable α, which is normally distributed with mean μ_α and variance σ_α^2. The second term represents the difference in campaigning efforts between F and T in period 1 and 2. Thus, if F outspends T, she will, ceteris paribus, increase her popularity with the median voter. Finally, e^t is the realization of a (macro)-shock after the investment decision in period $t = 1$ and $t = 2$. This shock is common to the whole populace and thus the identity of the median does not change over time.[4] An example for such a shock is a blunder in a publicly broadcasted speech. For example in the midst of economic turmoils 2008 Republican U.S. presidential nominee John McCain claimed that "the fundamentals are strong", making him look extremely out of touch with reality[5]. We denote the respective random variables corresponding to realizations e^1 and e^2 as ϵ^1 and ϵ^2 and assume that they are independently drawn from a normal distribution with mean zero and variance σ_ϵ^2.[6] On election day, the median voter votes for candidate F if $d^3 > 0$, while he votes for candidate T if $d^3 < 0$. If the median voter is indifferent ($d^3 = 0$) so is the electorate as a whole, and both candidates receive a vote share of 50 percent. Denote by δ^3 the random variable belonging to d^3. d^2 and δ^2 are defined analogously as the (realized) difference in perceived utility in $t = 2$. Since $a = d^1$ we will use only the former. Throughout the paper, we shall use the symbols ϕ and Φ to denote the PDF and CDF of the standard normal distribution where needed.

---

[4]The assumption of such a 'macro' shock is not necessary, but significantly simplifies the exposition. However, in general the shocks could as well have any other form that allows inferring the expected ranking of candidates in the next stage. In that case the identity of the median voter changes over time.

[5]http://www.nytimes.com/2008/09/17/world/americas/17iht-mccain.4.16251777.html.

[6]There are many other possible interpretations. For example Gassebner, Jong-A-Pin, and Mierau (2008) show the influence of terrorist attacks on politicians' popularity. While this may still be related to policy or qualification, other studies show that pure random events – from the point of view of a politician – influence election outcomes. For example, Healy, Malhotra, and Mo (2010) study the impact of local college football games just before an election takes place. If the local team wins, the incumbent's chances to win improve significantly. Similarly, natural disasters may have a direct influence of voters' candidate ranking. A recent example for such a shock is the disaster caused by the damaged Fukushima Daiichi nuclear power plant in Japan that followed the earthquake and the thereby caused tsunami of March 11, 2011. In the aftermath, in many countries around the world a shift of voters' preferences in favor of green or anti-nuclear movements was accounted for. In the U.S., support for nuclear energy dropped to a historical low, with numbers even beneath those immediately after the Three Mile Island incident in 1979 (Cooper and Sussman, 2011). In Germany, the Green Party's support surged in the disaster's aftermath and helped them to win the state election in one of Germany's most influential states, Baden-Württemberg, which until then was a traditional stronghold of the conservative party CDU of Angela Merkel (Dempsey, 2011). Similar shifts in public opinion happened in other countries, such as France (Buffery, 2011), Switzerland (Kanter, 2011), and India (Gupta, 2011). Apart from shocks to candidate ranking, ϵ may as well be interpreted as a random shock muting candidates' campaign efforts in a given stage, so that effective effort is not x_i^t but $x_i^t - \epsilon_i^t$. This is a standard assumption in the literature on labor tournaments that has been pioneered by Lazear and Rosen (1981) and Nalebuff and Stiglitz (1983). In the context of political competition it may represent uncertainty about the effects of campaign spending, and that campaigning might backfire as well (e.g. Dukakis in the tank). Finally, it may be interpreted as a poll's error (see Section 6.3).

The candidates' objective functions are equal to

$$E[\pi_F] = Pr\left(\delta^3 > 0\right) - \frac{c}{2}(x_F^1)^2 - \frac{c}{2}(x_F^2)^2,$$

$$E[\pi_T] = Pr\left(\delta^3 < 0\right) - \frac{c}{2}(x_T^1)^2 - \frac{c}{2}(x_T^2)^2.$$

We normalize the utility from holding office to one. By equation (1) we see that expending campaigning effort x_i^t, $t = 1, 2$ increases the chances of winning the median voter's vote and thus the election. Furthermore, two types of factors not under the control of the candidates are relevant for the probability of capturing the median voter on election day. On the one hand there is uncertainty about the exogenous policy taste parameter α of the median voter. Thus, one candidate may start the campaign in $t = 1$ with a popularity advantage. Without loss of generality we focus on situations where $\mu_\alpha > 0$ and thus we refer to candidate F as the "expected front-runner".[7] Similarly, when analyzing the candidates' decisions with polls we focus on cases where $a > 0$ and $d^2 > 0$ such that we refer to candidate F as the current "front-runner". Furthermore, the realizations of the two macro shocks ϵ^1 and ϵ^2 add uncertainty in each stage. Investments come at quadratic costs $C(x_i^t) = \frac{c}{2}(x_i^t)^2$, $c > 0$. The convexity of costs may reflect the fact that the opportunity costs of fundraising go up with every dollar already raised or that each additional unit of investment becomes less effective at a given point in time.[8] For example, if we interpret spending x_i^t as a measure of how many voters a candidate reaches at a given time, convex costs would imply it gets harder and harder for the candidates to reach yet another voter with their messages. Note that the objective function need not be concave in candidates' investments. Thus, in the main part of the paper we focus on situations where c, σ_α, and σ_ϵ are sufficiently high such that pure strategy equilibria exist.

3 The effect of polls

We now start our analysis of the campaign by studying candidate behavior in a FPTP electoral system. A candidate's goal is to win the simple majority of votes and thus, as long as a majority is achieved,

---

[7]F can also be interpreted as the incumbent and μ_α as the expected incumbency advantage.

[8]One might argue that candidates in a campaign are budget constrained. They cannot spend more money than is donated by their supporters. While this is true in the short run, over the course of the campaign, candidates can invest in fundraising and open up new sources of funding. Thus, we believe it is reasonable for our purposes to assume that spending has some opportunity costs which increase with the amount of campaigning effort expended in a given period.

the candidate is indifferent whether he wins by a small or by a large margin. We discuss an extension in Section 6.

In order to identify the effect of polls we consider two polar cases: a situation without any poll, and thus no information about current relative standing, and a situation where polls perfectly reveal current popularity before each investment decision is taken by the candidates. First consider as a benchmark a situation without any poll.

3.1   A campaign without a poll

In both $t = 1$ and $t = 2$ the candidates have the opportunity to buy TV ads for their campaigns. Absent a poll they do not receive a signal as to their current popularity with the electorate in either stage and thus they base their decisions solely on their common prior beliefs $\Phi\left(\frac{\alpha}{\sigma_\alpha}\right)$. The next proposition summarizes our first result:

Proposition 1. *The equilibrium of the campaigning game without polls is unique. Both candidates choose identical investments in each stage, and investments are also identical across stages. Consequently, absent a poll the ex-ante expected popular vote of a candidate in $t = 3$ is identical to his popularity in $t = 1$.*

Proof. See appendix.   □

We find that efforts are completely identical, even if the front-runner F has an expected popularity advantage. To see why it is instructive to look at candidates' marginal incentives. For a realized median position a, note that the marginal increase in popular vote / probability to win of F is exactly the share (density) of currently indifferent voters. Of course, this also holds for T. Taking into account uncertainty about α this must hold also in expectation. Moreover, since candidates have identical costs of investment, they also have identical marginal products and marginal costs of spending. Consequently, for any prior distribution $\Phi\left(\frac{\alpha}{\sigma_\alpha}\right)$ both candidates have identical marginal incentives, and hence the equilibrium must be symmetric. This is similar to a result by Lazear and Rosen (1981) in their analysis of labor market tournaments, when a candidate enjoys a head start.

A direct implication of the proposition is, that the campaigning stage does not have *any* effect on the expected winner of the election, absent polls. Candidates choose identical investments, the net

effect is zero, and the winning probability will be unaffected in equilibrium. Here we clearly see how political campaigns, or contests in general, are similar to the Prisoner's Dilemma: both candidates would be better of by reducing spending by some $k > 0$, or by spending nothing at all. However, individually they have an incentive to outspend their opponent.

3.2 A campaign with polls

In most modern democracies polls are a pervasive element of the political landscape. Information about relative popularity is accessible in abundance. Apart from polls, politicians may learn more about their popularity from prediction markets on the internet. Thus, candidates can react to changes in their popularity, and, at the same time, even if popularity does not change over time (e.g. $a = d^2$), this gives additional information to the candidates. After all, learning about the popularity in $t = 2$ is confounded by less noise than it was in $t = 1$, and for sure by less than absent any feedback. In this section we introduce polls into the political campaign. Before their investment decisions in $t = 1, 2$ both candidates learn their relative popularity with the median voter $d^1 = a$ and $d^2 = a + x_F^1 - x_T^1 - e^1$.

3.2.1 Stage 2

We start with the candidates' investment decisions in stage 2. After learning d^2 through the poll, both candidates decide on x_i^2. This situation is strategically similar to the situation in the absence of polls, with one distinction: the amount of uncertainty until the election in $t = 3$ is reduced. But we have seen from Proposition 1 that the amount of uncertainty was not decisive for *relative* incentives. Hence, although investment levels differ from the situation without poll, candidates in stage 2 also choose identical investments.

Proposition 2. *There is a unique equilibrium in stage 2. Both candidates choose identical investments, no matter what their popularity advantage in $t = 2$.*

Proof. See appendix. □

The intuition for the symmetry of investments is identical to before: by marginally increasing spending, the additional probability to win equals exactly the opponent's loss in winning probability; the marginal benefit of investment is identical for both candidates independent of their popularity.

Because marginal costs are also identical, both must choose identical spending. Hence, we can conclude that the expected popular vote a candidate gets in $t = 3$, $E[\delta^3]$ equals his popularity in stage 2, d^2.

What are the properties of stage 2 investments? Intuitively, higher marginal costs c decrease spending. The effect of increasing the noisiness of the macro shock σ_ϵ is ambiguous. Increasing σ_ϵ implies ϕ goes up at the tails and goes down in the center, overall becoming flatter. Therefore, in case competition is lopsided with one candidate enjoying a popularity advantage, increasing the variance of the shock increases equilibrium spending. Intuitively, if the variance is large the trailing candidate has a realistic chance to catch up without spending overly in the campaign. This also gives him an incentive to increase spending, since, as we have discussed before, ϕ determines the marginal product of investment. To the contrary, if σ_ϵ is relatively low, catching up due to luck is relatively unlikely, and hence the trailing candidate has no incentive to compete. As a result both candidates' investments decrease. In case candidates are similarly popular an increase in variance always decreases the incentives to invest, as the marginal product of investment goes down. Finally consider the effect of the popularity advantage d_2. This comparative static is of special importance for the candidates as they are able to influence their popularity through investing in stage 1. Since the second stage equilibrium is symmetric, the difference in spending, $x_F^2 - x_T^2$, is zero, independent of d^2. Thus, the marginal increase in the winning probability *in equilibrium* is equal to the density of the shock, evaluated at d^2, $\frac{1}{\sigma_\epsilon}\phi(\frac{d_2}{\sigma_\epsilon})$. The normal density is strictly quasi-concave and symmetric at zero and thus $\phi(\frac{d_2}{\sigma_\epsilon})$ is strictly decreasing in $|d_2|$. This implies spending decreases as $|d^2|$ gets larger. Thus, the intensity of competition in $t = 2$ increases as the popularity difference vanishes.

3.2.2 Stage 1

Next turn to stage 1. How does the existence of a poll in period 2 affect candidates' incentives? In order to answer this question we focus, for the moment, on situations where pure strategy equilibria exist.[9] This depends on the cost parameter c and the variance of the macro shock σ_ϵ^2, or more specifically, on $\rho := c\sigma_\epsilon^2$. We can interpret ρ as a measure of how expensive it is to gain a substantial advantage through investing in a close campaign. Obviously higher marginal costs, c, make it more expensive to campaign, regardless of the closeness of the race. A high variance σ_ϵ^2 implies the density in the center

---

[9]We consider an extension in Section 6.

of the shock distribution is low. Thus, on the margin, publishing TV ads in a close race has only a small impact on the probability to win the campaign. First we consider ρ's such that a *unique* pure strategy equilibrium exists. Specifically we require ρ to be sufficiently large or $\rho > (3^{\frac{3}{4}}\sqrt{\pi})^{-1} = \bar{\rho}$.[10]

Consider the decision each candidate has to make in $t = 1$. There are three different channels through which his decision influences a candidate's expected utility. First, campaigning effort increases the probability to win the election. This effect is, as in the previous section, identical for both, and hence cannot be a cause of differences in campaign disbursements. Second, spending has immediate costs. But candidates have identical cost functions, and hence marginal costs are also identical. Therefore, this cannot be a reason for different behavior, either. There is, however, a third channel. By increasing spending in $t = 1$ a candidate changes the expected state of the campaign in $t = 2$ in her favor, because $E[\delta^2] = a + x_F^1 - x_T^1$; the marginal impact of investment in $t = 1$ on $E[\delta^2]$ is equal to one. Changing $E[\delta^2]$ has consequences for expected future spending, and hence for expected future costs. From the discussion of the comparative statics in $t = 2$ we know that costs are highest when the race is tied, $d^2 = 0$, and that costs decrease monotonically if we let $|d^2|$ grow, because equilibrium spending goes down. This implies that the leading candidate, by exerting campaign effort in stage 1, locally increases $|d^2|$ in expectation and hence decreases her expected costs in stage 2, while the opposite holds for the trailing candidate. In more technical terms, this implies the leading candidate's efforts in stage 1 and 2 are strategic substitutes, while the trailing candidates efforts are strategic complements (Bulow, Geanakoplos, and Klemperer (1985)). This implies that the leader has lower (expected) marginal costs of investment than her opponent. As a consequence, she spends more in equilibrium and thereby in expectation increase the difference to her opponent. We have a situation in which the leader acts tough, while her opponent takes a softer stance. We summarize this result in the following proposition:

Proposition 3. *Assume $\rho > \bar{\rho}$. In the unique equilibrium the candidate leading in the poll chooses greater investment in stage 1 than her opponent. If the race is tied both candidates choose identical spending.*

Proof. See appendix. □

Together Propositions 2 and 3 characterize the candidates' spending decisions over the course of

---

[10]A proof of this condition can be found in the appendix together with the proof of the next proposition.

the campaign. While in stage 1 spending is typically asymmetric, with the front-runner outspending the trailing candidate, expenditures are symmetric just before the election in stage 2. Overall, with polls, the front-runner enjoys *momentum* in expectation, i.e. an initial lead grows in expectation over time. In Proposition 1 we characterized spending decisions absent polls. We found that behavior is completely symmetric and thus an initial advantage translates into an equal expected advantage. With these observations we can state our main result:

Corollary 1. *The effect of polls:* *If it is not too cheap to change a close outcome in expectation, $\rho > \bar{\rho}$, polls create momentum in favor of the front-runner. Thus, the introduction of polls increases the likelihood that the front-runner wins the election.*

Proof. This follows immediately from Propositions 1, 2 and 3. $\qquad\square$

We find that the availability of information about relative popularity crucially changes the incentives of the candidates. In fact, a candidate starting into the campaign with an incumbency advantage will make this advantage grow over time in expectation as the campaign progresses.

So far we have assumed that ρ is sufficiently large. When ρ is small campaigning is very effective and cheap. Thus, influencing future popularity through campaigning is attractive. Furthermore, because of the competitive environment, expected costs in period 2 are relatively high, making it attractive to use period one spending to defuse competition and save on future costs. In such a situation "gambling for resurrection" turns out to be a possible equilibrium. However, the popularity difference between candidates must not be too large. To see why this is the case recall the aforementioned logic of strategic substitutes and complements. A candidate's investments are strategic substitutes if and only if he is in expectation more popular than his opponent in stage 2 given equilibrium efforts. If this is not the case, a marginal increase in spending decreases the expected popularity gap and hence increases expected costs. Consequently, the trailing candidate may spend more in equilibrium, but if he does this he needs to spend enough to turn the state in his favor in expectation. Of course, turning the state becomes increasingly expensive as the difference between candidates increases, and there exists a threshold gap determining the maximum lead that the trailing candidate may try to turn by investing heavily in campaigning.

Obviously, if the trailing candidate can take charge and act aggressively in the campaign in equilib-
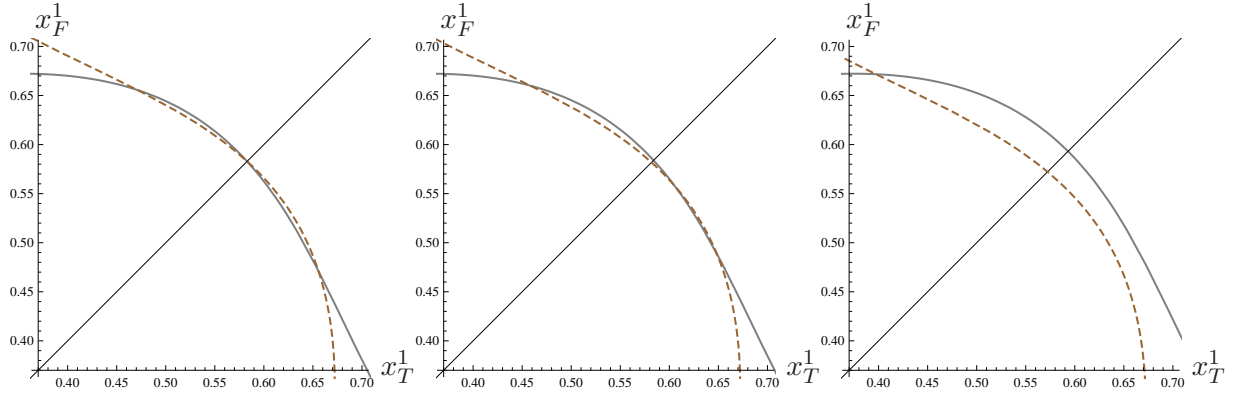
Figure 2: Reaction functions for F (gray) and T (brown, dashed) and stage 1 equilibria for $\rho = 0.242$ and $a \in \{0, 0.04, 0.1\}$.

rium, this is also possible for the front-runner. After all, it is cheaper for her to stay in the lead than for her opponent to turn the electorate's sentiment. Consequently, when $\rho \leq \bar{\rho}$ there exist multiple equilibria.

It remains to discuss what happens when the race is tied. As before, to conjecture that there exists a symmetric equilibrium is appealing. But we have also seen that there are multiple equilibria if no candidate dominates. Those asymmetric equilibria also exist in the tied race. Hence, there exist three equilibria for $a = 0$, and more generally, when a is small. However, note that the symmetric equilibrium at $a = 0$ is asymptotically unstable, while the asymmetric equilibria are stable. In the next proposition we summarize the results formally:

Proposition 4. *If ρ is relatively small, $\rho \leq \bar{\rho}$, there may exist multiple equilibria. If we confine ourselves to asymptotically stable equilibria, there exists no symmetric equilibrium, even for $a = 0$. In close races either candidate may spend more in stage 1 in equilibrium, while if one candidate has a sufficiently large advantage, this candidate outspends her opponent in stage 1 and the equilibrium is unique.*

Proof. See appendix. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ □

Even though the two types of equilibria may look very different at first sight, they follow in fact the same underlying logic. Candidates' find it worthwhile to defuse competition in period 2 by investing asymmetrically in period 1. One candidate takes a soft stance. She will suffer a lower probability of success which will be more than compensated through the saved costs. The other candidate takes

a tough stance and thus incurs high costs in period 1. He is more than compensated through the increased probability of success and the cost savings in period 2. In order to illustrate the multiplicity of equilibria in stage 1 graphically, we simulated the equilibria using the best response functions of the candidates for $\rho < \bar{\rho}$ and $a \in \{0, 0.04, 0.1\}$. They can be found in Figure 2. In panel a) the symmetric case is depicted. In panel b) F has a small advantage, and thus multiple equilibria exist. In panel c) there is a unique equilibrium exhibiting momentum.

Next consider the effect of polls on the election outcome. For the case of $\rho \leq \bar{\rho}$ we need to amend Corollary 1 slightly.

Corollary 2. *The effect of polls: If it is relatively inexpensive to change a close outcome in expectation, $\rho < \bar{\rho}$, polls create momentum if there is a clear front-runner. Thus, polls increase the probability that the front-runner is elected. In close races, momentum and anti-momentum may occur and consequently the outcome of the election relative to a situation without polls is ambiguous.*

Proof. This follows immediately from Propositions 1, 2 and 4. $\square$

A final comment concerning the equilibria when $\rho \leq \bar{\rho}$ is in order. A low value of ρ implies that either marginal costs are low or the campaign is – to borrow from the theory of contests – relatively discriminating (low σ_ϵ), or both. Either makes it unlikely that an interior pure strategy equilibrium exists, since the second order conditions are then likely to be violated. Typically there exists only a small range of values for $\rho \leq \bar{\rho}$ for which the second order conditions hold. Figure 3 illustrates combinations of c and σ_ϵ for which pure-strategy equilibria as discussed above exist.

To sum up, there always exists an equilibrium where the front-runner outspends the trailing candidate and thus creates momentum. Furthermore, in close races when campaigning is relatively effective, also the trailing candidate may outspend the front-runner. While in this section we compared relative spending, in the next section we are interested in the effect of polls on candidates' aggregate campaign spending.

4 Polls and the intensity of political competition

So far we have focused on the effects of polls on the outcome of the election, which is determined through relative candidate spending in the campaign. We have neglected the actual *level* of expenditures.
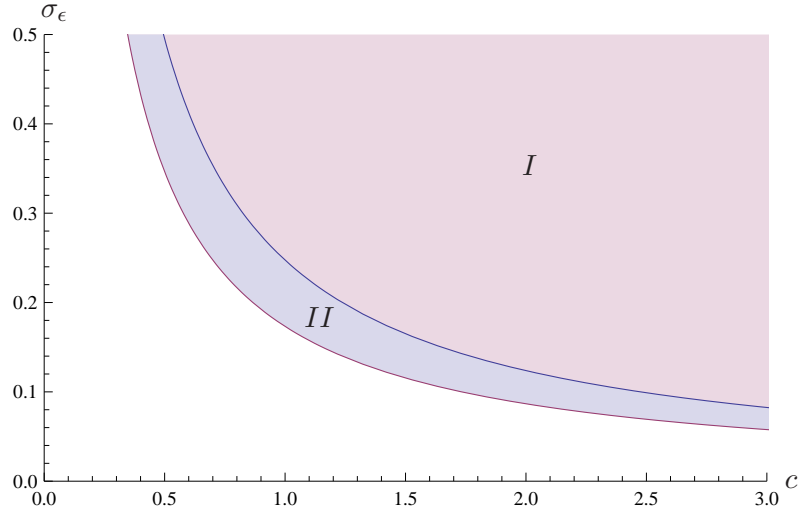
Figure 3: Number of pure strategy equilibria in stage 1 depending on c and σ_ϵ. In Region I there is a unique equilibrium, in region II there are two stable equilibria in close races. In the region below II the second order conditions do not hold for each possible popularity realization a and hence we cannot guarantee existence of pure strategy equilibria.

Campaign expenditures can be quite substantial. For example, they typically exceed one billion U.S. dollars in the case of presidential elections in the U.S. Thus, the "wastefulness of competition" is also an important aspect to study and relate to the existence of polls. We are aware that not all campaign efforts are necessarily wasteful since campaigns also inform voters about the candidates' positions. Nevertheless, we define wastefulness as expected aggregate spending over the course of the campaign. Thus, we implicitly assume that polls do not influence the level of informative campaigning and candidates on the margin invest for example in image building activities. The following proposition shows the result:

Proposition 5. *Expected aggregate campaign expenditures are lower in the presence, than in the absence of polls. Thus, ceteris paribus, polls make competition less wasteful.*

Proof. See appendix. □

Because polls create momentum (or anti-momentum), initial asymmetries are amplified. This decreases the expected intensity of competition in subsequent rounds, and hence expected wastefulness decreases, too.[11]

---

[11]Of course, conducting polls will incur additional costs which need to be taken into consideration in a full-fledged welfare analysis, here we only highlight the direct effect of polls on the intensity of competition.

The effect of increased information on the intensity of competition has been a subject of interest in other studies as well. A recent example is Denter, Morgan, and Sisak (2011), who analyze the effect of mandatory disclosure requirements in competitive environments such as lobbying competition and political campaigns. They study the competitors' incentives to share private information about their characteristics, for example their valuation of winning, and how mandated disclosure affects the investment decision and outcome of competition. Their main result is that mandatory disclosure increases the intensity and decreases allocative efficiency of competition in expectation. In their setting, an uninformed player can use the chance of facing a weak opponent to appease a strong opponent, and at the same time the threat to face a strong opponent to discourage a possible weak opponent. The lack of information about the opponent helps to commit to strategies which would not be credible under complete information. In contrast, in this paper, candidates are identical except for their current popularity. There is no private information held by the candidates, and hence they can not use this lack of knowledge as a commitment.

5 Polls and candidates' spending profiles

In this section we take a closer look at the spending profiles of the candidates. Without polls the result is straightforward, spending decreases as one candidate becomes more and more popular (in expectation). In this section we focus on spending profiles given that polls are conducted. Because comparative statics are only meaningful in settings with unique equilibria, we restrict ourselves to this case. We compare our results with empirical findings and show that the model performs relatively well in predicting spending profiles.

We know already from Section 3 that (given a unique equilibrium) polls always create momentum and the more popular candidate invests more heavily in the campaign. We are now interested in how the candidates' spending varies in absolute terms as we increase the realized popularity advantage $|a|$ of the front-runner from zero. In stage 2 we observed that spending decreases monotonically in $|d^2|$, i.e. when one candidate's becomes more and more popular with the electorate. Also in the studies of Snyder (1989), Erikson and Palfrey (2000), or Klumpp and Polborn (2006) intensity of competition decreases as candidates become increasingly asymmetric. However, in stage 1 we find a non-monotonic spending profile for the front-runner. The reason is that compared to a symmetric race, the front-runner has an

additional incentive to invest in stage 1 in order to reduce future competition. This also explains why the trailing candidate's spending profile is monotonically decreasing. Thus, we find that introducing dynamics may change results significantly. The following proposition states the result:

Proposition 6. *Assume $\rho > \bar{\rho}$. Candidate F's expected total investment increases in a when $a = 0$, while candidate T's expected total investment decreases.*

Proof. See appendix. □

The trailing candidate's campaign spending decreases monotonically as $|a|$ grows larger. His more popular adversary, however, has an incentive to first increase spending when her lead grows larger, before she also cuts down on spending when she becomes more and more advantaged in the campaign. We can now describe the two candidates spending profiles completely:

Corollary 3. *As $|a|$ increases from zero the more popular candidate first increases spending in the campaign and her spending declines only after she reaches a certain popularity advantage. The trailing candidate monotonically decreases spending and spends less than his opponent for all $|a| > 0$.*

How well does the model perform in predicting candidate spending in a campaign? In the left panel of Figure 4 we plotted expected total campaign spending for both candidates for $\sigma_\epsilon = 1$ and $c = \frac{1}{3}$. For a comparison we show in the right panel of Figure 4 the vote-on-spending effects as estimated by Erikson and Palfrey (2000). Although our model is relatively simple it predicts the spending profile surprisingly well. At $a = 0$, which represents a predicted incumbent vote of 50 percent, both candidates choose identical spending. The incumbent's spending increases first and decreases after peaking at around 55 percent until the predicted incumbent vote reaches some 80-85 percent. Then, somewhat surprisingly, spending goes up again.[12] The challenger's spending decreases in the incumbent's popularity advantage and remains below the incumbent's spending at all time.

6   Extending the basic model

In this section we discuss extensions to the basic model. We start off by showing that our results from the previous section are qualitatively robust to more general functional specifications. Second we depart

---

[12]One reason for this upwards turn for a dominating front-runner might lie in her ability to attract funding. In our model costs do not depend on popularity. When increased popularity makes fundraising easier and thus costs of investment lower, an additional benefit of being a front-runner arises.
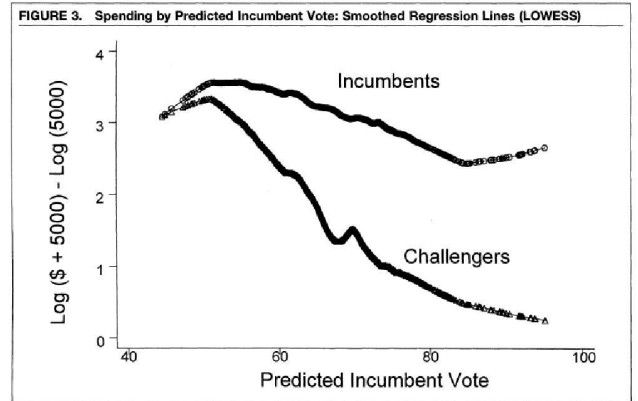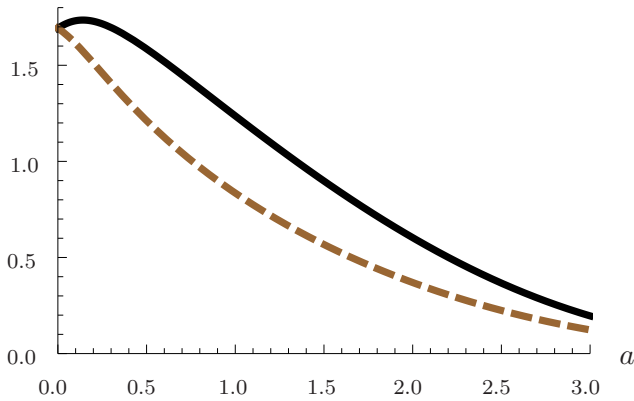
Figure 4: Total spending per candidate, depending on popularity. The left panel shows the predictions of our model (expected total spending) for $\sigma_\epsilon = 1$ and $c = \frac{1}{3}$. The right panel shows Figure 3 from Erikson and Palfrey (2000), which shows candidate spending in the U.S.

from the assumption of pure maximization of the probability to win and consider candidates that are also interested in the popular vote per se, as is for example the case in proportional representation (PR) electoral systems. Third, we consider inaccurate polls. Finally, we discuss some further extensions.

6.1 Functional assumptions

For the baseline model we assumed normally distributed noise and quadratic costs. In this section we show that our results generalize.

Consider the distribution of the macro shock ϵ^t first. While, as we argued above, the normality assumption can be justified – for example by invoking the central limit theorem – it is nevertheless of interest to explore the robustness of our results in more detail. Hence we relax this assumption and assume the following instead:

Assumption 1. ϵ^1 and ϵ^2 are independently distributed on $\mathcal{S}^1 \subseteq \mathbb{R}$ and $\mathcal{S}^2 \subseteq \mathbb{R}$ with densities $g^1(\epsilon^1)$ and $g^2(\epsilon^2)$. $g^t(\epsilon^t)$ is differentiable on $\mathcal{S}^t$, symmetric around zero and $g^{t\prime}(|\epsilon^t|) \leq 0$.

By Assumption 1 each shock has a quasi-concave density, implying that a large shock is less likely to occur than a small shock. Furthermore, noise is unbiased and thus it is unlikely to have a series of

random events only in support of one of the two candidates, i.e. each candidate gains from some events and loses from others in expectation. Distributions fulfilling Assumption 1 are for example the uniform distribution or a symmetrically truncated normal distribution.

Now turn to the cost function. Instead of quadratic costs we assume the following:

Assumption 2. *Spending on TV ads x implies costs $C(x)$, where $C(0) = 0$, $C'(0) = 0$, $C' > 0$ for all $x > 0$, $C'' > 0$ and $|C'''|$ is finite.*

This cost function is strictly convex and hence may reflect both increasing marginal costs of fund raising as well as diminishing marginal returns of campaigning. In any other respect, the campaign game is as before.

Proposition 7. *All results qualitatively resemble those of Section 3. If there is a unique equilibrium (sufficient noise, $C''(x)$ sufficiently large) polls always create momentum.*

Proof. See appendix. $\square$

As before, if the macro shock's variance is sufficiently large or if marginal costs increase sufficiently fast, there is a unique equilibrium and polls create momentum. Otherwise, multiple equilibria may exist and polls may create anti-momentum in close races.

An interesting feature of the more general model is worth noting: candidates may drop out of the race by spending nothing. This could not happen in the baseline model since marginal costs of the first TV ad published were zero and the marginal benefit of publishing TV ads was strictly positive due to the infinite support of the normal distribution. This possibility can be best illustrated by letting $E[(\epsilon^t)^2] \to 0$ and assuming the quadratic cost function from the previous section.[13] The campaign is then a perfectly discriminating contest or – using the language of auction theory – an All-pay auction.[14] Absent a poll both candidates publish an identical number of TV ads in stage $t = 1, 2$. If there is a poll, there exists a unique equilibrium in stage 2. If $|d^2| < \sqrt{\frac{2}{c}}$ this equilibrium is in mixed strategies and both candidates choose in expectation identical spending. Otherwise, the campaign is decided already and both candidates spend nothing. In stage 1, for $|a| < \sqrt{\frac{2}{c}}$, there always exists a pure strategy

---

[13]The formal discussion can be found in Section G of the appendix.

[14]This form of contests has been studied extensively with applications ranging from lobbying contest (Hillman and Riley, 1989; Baye, Kovenock, and de Vries, 1993) to political campaigns (Meirowitz, 2008), revolution or war (Polborn, 2006) and internal labor markets (Nalebuff and Stiglitz, 1983; Moldovanu and Sela, 2001).

equilibrium in which the leading candidate publishes TV ads and the trailing candidate stays passive (for $a \geq \sqrt{\frac{2}{c}}$ the front-runner is so far ahead, that it is strictly dominated for the trailing candidate to spend anything, and thus there is no competition). In close races also the opposite – an equilibrium in which the trailing candidate acts tough and spends a positive amount, while the front-runner remains passive – exists. Hence, there are multiple pure strategy equilibria if the campaign is close in $t = 1$. Consequently, in stage 1 the trailing candidates always drops out if his opponent is sufficiently popular. In analogy to before, in close campaigns also the front-runner might drop out in equilibrium.

This discussion also confirms our previous findings for $\rho < \bar{\rho}$ for the limit as $\rho = 0$. In close races, multiple equilibria emerge, exhibiting momentum and anti-momentum. In general though, it is likely to see the front-runner outspending the trailing candidate and thus the front-runner is more likely to improve her position over time.

While we have shown that momentum is a robust characteristic of polls, can we say the same about wastefulness? In Section 4, using our benchmark model, we have shown that polls make political competition less wasteful because they introduce more asymmetries between the candidates. Next we generalize Proposition 5 for more general noise and cost functions, but focusing on situations where a unique equilibrium exists:

Proposition 8. *Given assumptions 1 and 2 hold and a unique equilibrium exists. When $C'''(x) \geq 0$ $\forall x$, aggregate campaign expenditures decrease in expectation with the introduction of polls, strictly so when $\mid \mu_\alpha \mid > 0$. For $C'''(x) < 0$ $\forall x$ wastefulness may increase.*

Proof. See appendix. ☐

Thus, we find that the shape of the marginal cost function determines the effect of polls on wastefulness of competition. Note that if (but not only if) the spending elasticity of costs is decreasing in x wastefulness always declines.[15] This is in line with earlier findings by Aoyagi (2010), Ederer (2010) as well as Gershkov and Perry (2009) regarding the effect of feedback in labor market tournaments. To see why the shape of the cost function matters, consider second period investments. Investments are symmetric and decreasing in popularity $|d^2|$. The revelation of current relative standing reduces the noisiness of the decision environment. Thus, when standing $|d^2|$ is close to zero, candidates' incentives to invest are increased by the reduction in noise. On the other hand, when relative popularity

---

[15]If this elasticity is decreasing $C'''(x)$ must be positive.

$|d^2|$ is large, candidates' incentives are diminished with a reduction in noise as the trailing candidate then faces little chance of a "lucky break". When marginal costs increase faster with higher spending, or $C'''(x) > 0$, the latter effect is dominant and polls decrease wastefulness. In this case the difference in spending for relatively equal popularity is negligible because for high expenditures the effect of increasing marginal costs kicks in the strongest. In contrast to Aoyagi (2010), Ederer (2010) and Gershkov and Perry (2009) polls decrease wastefulness also for $C''' = 0$ and even $C''' < 0$ if sufficiently close to zero because of the effect described in Section 4: Polls cause momentum for the front-runner and thus amplify early asymmetries and defuse competition.

6.2 Concerns for vote share

In some situations candidates do not only care about winning the election, but also about the vote share they receive. For example, a government that has only slightly more than 50 percent of the vote share, might have difficulties in passing laws because some delegates might refuse to toe the party line. Similarly, in proportional representation electoral systems the vote share is important.

Assume a candidate or party values winning the majority by $\lambda \in [0, 1]$. We can interpret this as a plurality premium as in Iaryczower and Mattozzi (2013). In addition, candidates get utility $s_i(1 - \lambda)$ from gaining a share of s_i of the total popular vote. $(1 - \lambda)$ measures the relative importance of the vote share. For example, in many countries political parties receive financial subsidies in proportion to their vote share. Increasing those subsidies would imply a higher $(1 - \lambda)$. Similarly, the influence of a normal member of parliament that is not part of the government may determine $(1 - \lambda)$, too. The benefit of candidate i in the election is hence

$$
b_i = \begin{cases} \lambda + s_i(1 - \lambda) & \text{if } s_i > \frac{1}{2}, \\ s_i(1 - \lambda) & \text{else.} \end{cases}
$$

As before, candidates maximize their expected benefit subject to costs of campaigning effort. Otherwise we leave the model unchanged relative to Section 3. Candidates now care about the whole distribution of voters, and not only the median voter. We assume that voters vote for the candidate that affords them the highest expected utility as defined in equation (1), and candidates' spending affects all voters in the same way. Analogous to the median's relative ex-ante candidate assessment α, α_i is voter i's

relative assessment of the candidates. We assume the distribution of α_i follows:

Assumption 3. *For a given realization of the median $'a'$, the distribution of voters' candidate assessments has density*

$$p(\alpha_i, a) = \frac{\phi(\frac{-\mu_p - a + \alpha_i}{\sigma_p}) + \phi(\frac{\mu_p - a + \alpha_i}{\sigma_p})}{2},$$

where $\mu_p \geq 0$ and $\sigma_p > 0$.

$p(\alpha_i, a)$ is a mixture distribution of two normals with identical standard deviation σ_p and means $-\mu_p$ and μ_p. Hence, it is symmetric around a. It is easy to generalize the model to more flexible distributions, but the additional insights gained from such a generalization are only marginal.[16]

There are two different effects at work that may lead to differential incentives: the effect of a marginal change in spending on the winner-take-all part and the effect on the vote share part. We already know how the first effect looks like: investments are decreasing in $|d^2|$ and thus there always exist momentum equilibria. Hence, turn to the latter effect, caused by the proportionality of the obtained rent and the vote share. In the second stage competition will be fiercest when the mass of marginal/swing voters to be won over is largest, i.e. the distribution peaks at current popularity d^2. Thus, in stage 1 a candidate's incentives to publish TV ads depend on whether this will bring him closer in expectation to a position of fierce future competition. If the electorate is polarized, i.e. there is a large mass of voters at the tails of the distribution, while only a small mass in the center, competition is fiercest when candidates are not equally popular. Thus, in these cases the trailing candidate may have an incentive to publish more ads in stage 1 than the front-runner as this will enable him to defuse competition in the future.

Proposition 9. *If candidates care sufficiently about vote share (λ is small) and the electorate is sufficiently polarized (μ_p is large), there exist unique anti-momentum equilibria in close races.*

Proof. See appendix. □

Of course, even if the electorate is polarized, this does not yet mean that there is this kind of anti-momentum, because – as before – also the probability to win the majority counts. Hence, if λ, the

---

[16]We could also allow for different variances of the two normal distributions. Then the distribution $p(\alpha_i, a)$ is not symmetric anymore. This would have an asymmetric effect on candidates that is not related to their standing in the race, but to the particular distribution. Assuming symmetry allows us to identify effects due to relative standing.

plurality premium, is relatively large, the results will be identical to those in the previous section no matter what the distribution looks like:

Corollary 4. *There always exists a $\lambda < 1$ such that all results remain as in Section 3. If $\mu_p \leq \sqrt{2\sigma_\epsilon^2 + \sigma_p^2}$ this is generally the case, independent of λ.*

Proof. See appendix. □

What can we conclude from this? In FPTP systems it is unlikely that candidates value the vote share more than the probability to win, and hence λ is likely to be large. But as we have seen, in this case the results from Section 3 remain valid and momentum is likely to occur. However, in other electoral systems results may differ. For example, in proportional representation electoral systems λ is likely to be lower than under FPTP. Hence, if in such a system the electorate is polarized, the conditions from Proposition 9 are more likely to apply and anti-momentum might be the unique outcome.

6.3 The precision of polls

We assumed that without a poll candidates receive no feedback about their relative standing with the populace, while in the presence of polls they learn current popularity for sure. These were interesting benchmark cases and served well to illustrate the relevant intuitions. We now extend our baseline model to analyze imprecise polls to show that our results hold more generally.

Assume that instead of fully revealing current popularity, polls offer a noisy, though unbiased signal. Let d^t denote the result of the poll at time t. Note that in this extended model we have $E\left[d^1\right] = a$ instead of $d^1 = a$ as well as $E\left[d^2\right] = a + x_F^1 - x_T^1$ instead of $d^2 = a + x_F^1 - x_T^1$ due to the imprecision of the poll. Denote the error of the poll at t by $\eta^t \sim N(0, \sigma_\eta^2)$ for some $\sigma_\eta \geq 0$.[17]

In order to focus on the effect of poll precision, assume now that $\sigma_\epsilon \to 0$ and thus there are no popularity shocks. Hence, any change in candidate evaluation is due to campaigning. Given that candidates have the belief that the median is normally distributed with mean μ_α and variance σ_α^2, and if the polling error is distributed as described above, the posterior belief after observing the poll at

---

[17]Note that as before we need 'sufficient' noise to guarantee existence of a pure strategy equilibrium. We assume this to be the case. But note that this extended model nests the all-pay case discussed in section 6.1 as a special case ($\sigma_\eta^2 = 0$).

$t = 1$ is still normal with updated mean

$$\tilde{\mu}_\alpha(d^1) = \frac{\sigma_\alpha^2 d^1 + \sigma_\eta^2 \mu_\alpha}{\sigma_\alpha^2 + \sigma_\eta^2}$$

and variance[18]

$$\tilde{\sigma}_\alpha^2 = \frac{\sigma_\alpha^2 \sigma_\eta^2}{\sigma_\alpha^2 + \sigma_\eta^2}.$$

When the poll becomes very imprecise, $\sigma_\eta^2 \to +\infty$, candidates ignore the poll and beliefs remain unchanged. This would correspond to our no-poll benchmark. On the other hand, when $\sigma_\eta^2 \to 0$ the prior is ignored and beliefs solely reflect the poll result. This corresponds to the All-pay auction discussed in Section 6.1. Hence, we bridge the two models by allowing for imprecise polls. For $t = 2$ we get a posterior belief with mean

$$\tilde{\tilde{\mu}}_\alpha(d^2) = \frac{\tilde{\sigma}_\alpha^2 d^2 + \sigma_\eta^2 \tilde{\mu}_\alpha}{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}$$

and variance

$$\tilde{\tilde{\sigma}}_\alpha^2 = \frac{\tilde{\sigma}_\alpha^2 \sigma_\eta^2}{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}.$$

Considering the candidates' investments in the second stage, we can simply replace the density of the shock ϵ^t with the density of the posterior popularity belief in $t = 2$ with variance $\tilde{\tilde{\sigma}}_\alpha^2$. The reason is that even though ϵ^t and η^t represent different types of uncertainty, they have the same effect with respect to expected popularity in $t = 3$. ϵ^t represents uncertainty as to shifts in popularity between t and $t + 1$, while η^t represents uncertainty as to actual popularity measured at t. Both translate into uncertainty about popularity on election day and thus both have the same effect on behavior in $t = 2$.

The effect of imprecise polls on the spending decisions in stage 1 is slightly more complex, as before, and works through an expected shift of the stage 2 belief distribution.

Proposition 10. *Consider the extended model and assume marginal costs are sufficiently high such that a unique pure strategy equilibrium exists. The first-period front-runner always outspends the trailing candidate. Furthermore, when the first-period advantage of the front-runner, d^1 is sufficiently small, an increase in poll precision increases the spending gap between front-runner and trailing candidate. If*

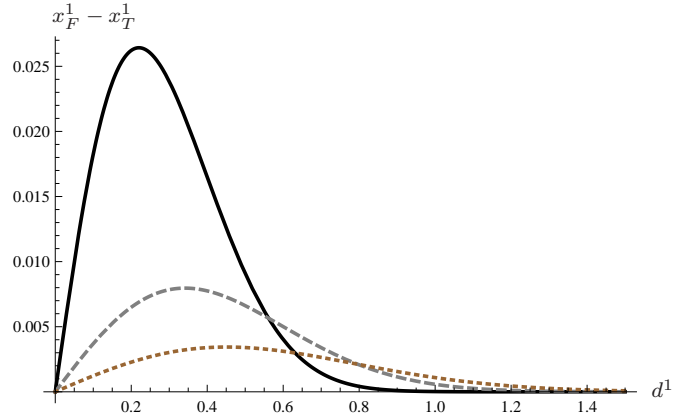[18]See, for example, Casella and Berger (2002).

Figure 5: The effect of a decrease of σ_η on $x_F^1 - x_T^1$ in the extended model for $c = 2$ and $\sigma_\eta = 2/3$ (brown dotted line), $\sigma_\eta = 1/2$ (gray dashed line), and $\sigma_\eta = 1/3$ (solid black line). In close races more precise polls increase the spending difference, while in very lopsided races the spending difference decreases as the poll becomes more precise.

the front-runner's advantage is quite large, increasing the precision of the poll decreases momentum as the spending differential goes to zero.

Proof. See appendix. □

Thus, we find that increasing the precision of a poll increases the difference in incentives between front-runner and trailing candidate and thus amplifies the difference in spending $x_F^1 - x_T^1$ in close races. In lopsided races the opposite is true. Figure 5 illustrates the result by example.

Intuitively, a front-runner in the poll at $t = 1$ can be more sure that she is indeed preferred by the median voter, ceteris paribus, when the precision of polls increases. This has two effects. On the one hand, she will be more likely to benefit from reduced expenditures in the future, giving her an extra incentive to invest. On the other hand, when $|d^1|$ is large and thus the race seems very lopsided, both candidates will be further discouraged from investing as they become more sure about their relative standing. This latter effect can dominate and lead to a decrease in spending differences despite an increase in poll precision. This is seen easiest by looking at the All-pay auction case discussed earlier. When the front-runner's advantage becomes too large no candidate invests in the campaign anymore and hence the race is over before it starts.

6.4 Costs of campaigning

We assumed that costs of campaigning effort are additive across periods. This implies that the candidates' expected costs of spending in the future do not (directly) depend on current expenditures and vice versa. Thinking about costs of fund raising, one can easily imagine though that high expenditures in early periods make it harder to raise money later (ceteris paribus). This would be the case if first the most willing donors are targeted for donations. Then, in later periods, new, less willing donors have to be convinced to contribute. We expect this to have a mitigating effect on the magnitude of momentum. When thinking about whether to increase spending, the front-runner will trade off a decrease in future competition with higher expected costs of raising funds.

On the other hand, it may also seem plausible that early success lowers (marginal) costs of raising funds in the future. This would be the case when donors condition their donation on expected success in the election, maybe because they are expecting the implementation of favorable policies from the winner. Then investing early has an additional advantage in that it improves popularity in expectation and thus lowers costs of fundraising. In such a case momentum will be amplified. Which effect is more relevant is an empirical question.

Finally, we assumed that candidates are similar in their costs of raising funds. Thus, we abstract away from any differences between candidates not related to their relative popularity. Adding asymmetries, for example in ability or costs, will naturally lead to more asymmetric campaigning behavior. The general intuition though should be preserved also in such a more general setting. The benefit of our modeling assumption is being able to disentangle different effects that play a role in creating momentum more clearly.

7 Conclusion

In this paper we explore the effect of opinion polls on candidates' incentives to campaign and the final election outcome. We find that when candidates mostly care about winning the election, polls generally cause momentum for the front-runner. The reason is that the front-runner has an additional incentive to invest in her campaign. By investing early it becomes more likely that she comes out ahead in the poll in the future. This in turn defuses competition and thus she can save on campaigning expenditures

closer to the election. In close campaigns, when campaigning is very effective, also the trailing candidate may adopt a tough stance and outspend the leader. We show that those results are robust to functional form assumptions or the extent to which random events influence the election outcome.

Whether polls also make competition more wasteful, on the other hand, depends on the shape of the marginal cost function of campaigning expenditures. If raising additional money becomes more expensive with the amount raised, polls generally lead to a decrease in expected spending in the campaign and so decrease the expected intensity of the electoral competition.

As discussed, we show that momentum arises quite generally when candidates care mostly about winning the election. In case candidates are mainly interested in their vote share and the electorate is relatively polarized, the trailing candidate may have stronger incentives to campaign in relatively close elections. Anti-momentum may be the unique equilibrium. Thus, we expect the effect of polls to differ between countries with a FPTP system and countries with proportional representation and that in the latter no general statement about momentum can be made.

An implication of the model is that given an incumbency advantage at the outset of the campaign, polls are likely to foster this advantage and thus decrease the rate of turnover in political offices. Indeed, the spending profile predicted by our model fits well with the estimated profiles in Erikson and Palfrey (2000), who show that incumbents with an early popularity edge tend to improve their chances by spending more than their opponents in the competition.

Our model shows that a ban on the publication of opinion polls, as many countries impose in the pre-election period, will likely not eliminate all effects on voters. In particular, we have shown that there exists an indirect effect of polls on candidates' campaigning investments, which in turn influence voters' decisions at the ballot, and which is still in place because candidates are still allowed to commission polls in that period. Thus, a ban on the publication of polling results needs to be complemented with a ban on campaigning if potential causes of momentum are to be eliminated.

In our model voters react only to expenditures of the candidates, not to their relative standing. Other papers have argued that voters' preferences may be influenced also by relative standing (e.g. Hong and Konrad (1998) or Callander (2007)). Voters receive additional utility by voting for a candidate who is ahead in the polls. Thus, for given popularity levels and equal campaigning effort, the advantage of the front-runner grows. In terms of our model this can be interpreted as an additional

popularity dependent shift of the voter distribution in favor of the front-runner. Under this interpretation of the "bandwagon" effect momentum still results. Thus, accounting for the bandwagon effect only introduces further stickiness into our model and potentially intensifies momentum.

For future research it is interesting to empirically validate the results proposed by the theory. In contrast to theories that focus on the direct effect of polls on voters, our theory has predictions regarding candidates' campaign expenditures. Another interesting direction for future research is to study how polls influence donors' decisions to contribute to the candidates' campaigns. As Figure 4 shows, clear front-runners have a significant spending advantage over trailing candidates which is not completely captured by our model, but which we believe to be related to donors' behavior.

Mathematical appendix

A   Proof of Proposition 1

Proof. We prove the proposition assuming the more general model discussed in Section 6.1, which as a special case nests the model discussed in Section 3.1. If there is no poll, candidates do not observe the realization of ϵ^1 nor investments and hence their information does not change between periods 1 and 2. Hence, they maximize

$$\max_{(x_F^1, x_F^2) \in \mathbb{R}_+^2} \quad \int_{-\infty}^{\infty} \tilde{G}(a + x_F^1 + x_F^2 - x_T^1 - x_T^2) \frac{1}{\sigma_\alpha} \phi\left(\frac{a - \mu_\alpha}{\sigma_\alpha}\right) da - C(x_F^1) - C(x_F^2),$$

$$\max_{(x_T^1, x_T^2) \in \mathbb{R}_+^2} \quad 1 - \int_{-\infty}^{\infty} \tilde{G}(a + x_F^1 + x_F^2 - x_T^1 - x_T^2) \frac{1}{\sigma_\alpha} \phi\left(\frac{a - \mu_\alpha}{\sigma_\alpha}\right) da - C(x_T^1) - C(x_T^2),$$

where $\tilde{G}$ is the cdf of the convolution $g^1 * g^2$. The system of first order conditions is

$$\int_{-\infty}^{\infty} \tilde{g}(a + x_F^1 + x_F^2 - x_T^1 - x_T^2) \frac{1}{\sigma_\alpha} \phi\left(\frac{a - \mu_\alpha}{\sigma_\alpha}\right) da - C'(x_i^t) = 0,$$

$i = F, T$ and $t = 1, 2$. It is easily observed that in any interior pure strategy equilibrium it must hold that $x_F^1 = x_F^2 = x_T^1 = x_T^2 = x^*$. In particular,

$$x^* = C'^{-1}\left(\int_{-\infty}^{\infty} \tilde{g}(a) \frac{1}{\sigma_\alpha} \phi\left(\frac{a - \mu_\alpha}{\sigma_\alpha}\right) da\right).$$

Note that corner solutions with any $x_i^t = 0$ cannot be an equilibrium for unbounded densities g^i such as the normal due to our assumption that $C'(0) = 0$ and the fact that $g^i(a) > 0$ for all a. The second order conditions evaluated at x^* are

$$\int_{-\infty}^{\infty} \tilde{g}'(a) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) d\,a - C'' \left(C'^{-1} \left(\int_{-\infty}^{\infty} \tilde{g}(a) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) d\,a \right) \right) < 0$$

for F and

$$-\int_{-\infty}^{\infty} \tilde{g}'(a) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) d\,a - C'' \left(C'^{-1} \left(\int_{-\infty}^{\infty} \tilde{g}(a) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) d\,a \right) \right) < 0$$

for T. If $C''(.)$ is sufficiently large relative to $\tilde{g}'(.)$ the second order conditions hold for all a and hence this is an equilibrium. We assume this to be the case. Note that if the variance of g^1 and g^2 vanishes, the second order conditions always hold, as we show later in Appendix G. If the variance becomes very large, the slope of g^1 and g^2 decreases and hence the second order conditions also always hold. For the parameterized model in Section 3.1 the second order conditions hold if

$$c > | \int_{-\infty}^{\infty} \tilde{g}'(a) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) d\,a |.$$

For the remainder of the paper we assume C'' to be sufficiently large to guarantee the second order conditions. Hence, the proof is complete and this also contributes to the proof of Proposition 7. $\square$

B Proof of Proposition 2

Proof. We again prove the proposition directly in the general version of Proposition 7.

Consider the subgame in stage 2 when there is a poll. In this case both candidates know the median's exact position in stage 2, d^2. Candidates maximize

$$\max_{x_F^2 \in \mathbb{R}_+} \quad G^2(d^2 + x_F^2 - x_T^2) - C(x_F^2),$$

$$\max_{x_T^2 \in \mathbb{R}_+} \quad 1 - G^2(d^2 + x_F^2 - x_T^2) - C(x_T^2).$$

First order conditions read

$$g^2(d^2 + x_F^2 - x_T^2) - C'(x_i^2) = 0,$$

where $i = F, T$, and the assumption on C' guarantee that there exists a unique x_i^2 fulfilling these. It is immediately observed that – as without a poll – both candidates choose identical spending, $x_F^2 = x_T^2$. In particular,

$$x_F^2 = x_T^2 = x^{**}(d^2) = C'^{-1}(g^2(d^2)).$$

The second order condition for F is

$$g^{2\prime}(d^2) - C''(C'^{-1}(g^2(d^2))) < 0.$$

Since it must hold for all d^2, if this holds for F it also holds for T. We assume this holds for all d^2. The problem each candidate faces is then continuous and concave and hence a pure strategy equilibrium exists. In our parameterized example, if $\sigma_{\epsilon_2} = 1$ it must hold that $c > 0.242$. For the remainder of the paper we focus on situations where C'' is sufficiently large to guarantee the second order conditions are fulfilled.

This proves Proposition 2 and also contributes to the proof of Proposition 7. $\square$

C Proofs of Propositions 3, 4, and 7

Proof. The expected utility of F and T, conditional on being in state d^2 in the second stage, is

$$
\begin{aligned}
EU_F^*(d^2) &= G^2(d^2) - C\left(C'^{-1}\left(g^2(d^2)\right)\right), \\
EU_T^*(d^2) &= (1 - G^2(d^2)) - C\left(C'^{-1}\left(g^2(d^2)\right)\right).
\end{aligned}
$$

Note that $d^2 = a + x_F^1 - x_T^1 - e^1$. Then we can write the optimization problem of the candidates as:

$$
\max_{x_F^1 \geq 0} \quad \int_{\mathcal{S}^1} G^2(a + x_F^1 - x_T^1 - e^1) - C\left(C'^{-1}\left(g^2(a + x_F^1 - x_T^1 - e^1)\right)\right) g^1(e^1) de^1 - C(x_F^1),
$$

$$
\max_{x_T^1 \geq 0} \quad \int_{\mathcal{S}^1} (1 - G^2(a + x_F^1 - x_T^1 - e^1)) - C\left(C'^{-1}\left(g^2(a + x_F^1 - x_T^1 - e^1)\right)\right) g^1(e^1) de^1 - C(x_T^1).
$$

32

Taking the derivative with respect to the respective own strategy yields

$$\int_{\mathcal{S}^1} \left(g^2(\cdot) - \left[C'(C'^{-1}(\cdot)) \frac{\partial C'^{-1}(\cdot)}{\partial g^2(\cdot)} \frac{\partial g^2(\cdot)}{\partial x_F^1} \right] \right) g^1(e^1)\, de^1 - C'(x_F^1),$$

$$\int_{\mathcal{S}^1} \left(g^2(\cdot) - \left[C'(C'^{-1}(\cdot)) \frac{\partial C'^{-1}(\cdot)}{\partial g^2(\cdot)} \frac{\partial g^2(\cdot)}{\partial x_T^1} \right] \right) g^1(e^1)\, de^1 - C'(x_T^1).$$

Making use of the fact that $C'(C'^{-1}(z)) = z$ and $\partial g^2(\cdot)/\partial x_F^1 = \partial g^2(\cdot)/\partial a = -\partial g^2(\cdot)/\partial x_T^1$ these equations simplify to

$$\int_{\mathcal{S}^1} \left(g^2(\cdot) - \left[g^2(\cdot) \frac{\partial C'^{-1}(\cdot)}{\partial g^2(\cdot)} \frac{\partial g^2(\cdot)}{\partial a} \right] \right) g^1(e^1)\, de^1 - C'(x_F^1)$$

$$\int_{\mathcal{S}^1} \left(g^2(\cdot) + \left[g^2(\cdot) \frac{\partial C'^{-1}(\cdot)}{\partial g^2(\cdot)} \frac{\partial g^2(\cdot)}{\partial a} \right] \right) g^1(e^1)\, de^1 - C'(x_T^1)$$

Using $\partial C'^{-1}(g^2(\cdot))/\partial g^2(\cdot) = 1/(C''(C'^{-1}(g^2(\cdot))))$ and $\partial g^2(\cdot)/\partial a = g^{2\prime}(\cdot)$, and letting $a + x_F^1 - x_T^1 =: \kappa$ we get

$$\int_{\mathcal{S}^1} \left(g^2(\kappa - e^1) - \left[\frac{g^2(\kappa - e^1) g^{2\prime}(\kappa - e^1)}{C''(C'^{-1}(g^2(\kappa - e^1)))} \right] \right) g^1(e^1)\, de^1 - C'(x_F^1), \tag{C.1}$$

$$\int_{\mathcal{S}^1} \left(g^2(\kappa - e^1) + \left[\frac{g^2(\kappa - e^1) g^{2\prime}(\kappa - e^1)}{C''(C'^{-1}(g^2(\kappa - e^1)))} \right] \right) g^1(e^1)\, de^1 - C'(x_T^1). \tag{C.2}$$

Before we analyze the equilibrium in detail we now need to show that an equilibrium exists.

Lemma 1. *If $C''(x) > \Gamma(g^1, g^2)$ (defined below) a pure strategy equilibrium exists for all realizations a.*

Proof. The second derivative of F's payoff function is

$$\int_{\mathcal{S}^1} \left[g^{2\prime}(\kappa - e^1) - \left(\frac{g^{2\prime}(\kappa - e^1)}{C''(x^{**}(\cdot))} \right)^2 \left(C''(x^{**}(\cdot)) - \frac{g^{2\prime}(\kappa - e^1) C'''(x^{**}(\cdot))}{C''(x^{**}(\cdot))} \right) - \frac{g^2(\kappa - e^1) g^{2\prime\prime}(\kappa - e^1)}{C''(x^{**}(\cdot))} \right] dG(e^1) - C''(x_F^1).$$

To show strict concavity of the payoff function we need to show that this is strictly negative for all κ and x_F^1. Consider

$$\left(\frac{g^{2\prime}(\kappa - e^1)}{C''(x^{**}(\kappa - e^1))} \right)^2 \left(C''(x^{**}(\kappa - e^1)) - \frac{g^{2\prime}(\kappa - e^1) C'''(x^{**}(\kappa - e^1))}{C''(x^{**}(\kappa - e^1))} \right). \tag{C.3}$$

Let the $\gamma_1 > 0$ be the smallest value of C'' for which

$$\gamma_1^2 > g^{2\prime}(\kappa - e^1)C'''(x^{**}(\kappa - e^1))$$

for all $\kappa - e^1$. If the infimum of C'' is at least γ_1 (C.3) is non-negative. Since C''' is bounded by Assumption 2 and because it follows from differentiability and quasi-concavity that $g^{2\prime}(\kappa - e^1)$ is bounded, too, there exists finite γ_1 for which this is the case. Hence assume (C.3) is zero. If this is the case and we can show that the second derivative is negative, this is even more so the case when (C.3) is positive. Therefore, we are left with

$$\int_{\mathcal{S}^1} \left[g^{2\prime}(\kappa - e^1) - \frac{g^2(\kappa - e^1)g^{2\prime\prime}(\kappa - e^1)}{C''(x^{**}(\cdot))} \right] dG(e^1) - C''(x_F^1). \tag{C.4}$$

From strict quasi-concavity of g^2 it follows that the expectation is bounded. Hence, there exists $\gamma_2 > 0$ such that if $C''(x) > \gamma_2$ for all x, (C.4) is strictly negative. It follows that if $C''(x) > \Gamma(g^1, g^2) := \max\{\gamma_1, \gamma_2\}$ the second derivative is strictly negative and hence the problem is strictly concave. Thus, assuming this holds, and since payoffs are continuous in x_F^1 and x_T^2, existence of a pure strategy Nash equilibrium follows from Theorem 1.2 in Fudenberg and Tirole (1991), which is due to Debreu (1952), Fan (1952), and Glicksberg (1952). $\square$

Now let us go more into the details of the first order conditions to determine the properties an equilibrium must have. For this purpose we need the following lemma:

Lemma 2. *Let*
$$\xi(\kappa) := E_{\epsilon^1} \left[\frac{g^2(\kappa - \epsilon^1)g^{2\prime}(\kappa - \epsilon^1)}{C''(C'^{-1}(g^2(\kappa - \epsilon^1)))} \right].$$

Then $\xi(0) = 0$, $\xi(+) < 0$ and $\xi(-) > 0$.

Proof. Define
$$\omega(\kappa) := \frac{g^2(\kappa - \epsilon^1)g^{2\prime}(\kappa - \epsilon^1)}{C''(C'^{-1}(g^2(\kappa - \epsilon^1)))},$$

which is the function we want to take the expectation of. Now remember that by adding an arbitrary constant k to the argument of a function, the graph of the function is shifted horizontally by $-k$. Therefore, if $g^2(-\epsilon^1)$ is axis-symmetric across zero, $g^2(\kappa - \epsilon^1)$ is axis-symmetric across $-\kappa$. As a

consequence, the same holds true also for functions of this function, like $C'^{-1}(g^2(\kappa-\epsilon^1))$, $C''(C'^{-1}(g^2(\kappa-\epsilon^1)))$, and $\frac{g^2(\kappa-\epsilon^1)}{C''(C'^{-1}(g^2(\kappa-\epsilon^1)))}$. Because $g^2(-\epsilon^1)$ is axis-symmetric across zero, its derivative is point-symmetric across zero. By a similar argument as above it also holds that $g^{2\prime}(\kappa-\epsilon^1)$ is point-symmetric across $-\kappa$. Therefore, it holds true that also the product of an axis-symmetric function across $-\kappa$ and a point-symmetric function across this point, in our case this function is $\omega(\kappa)$, is point-symmetric across $-\kappa$.

We first show that $\xi(0) = 0$. Let $f(z)$ be a function which is axis-symmetric across zero and let $h(z)$ be another function which is point-symmetric across zero. Both functions share the same support $\mathcal{K}$. Then, if we want to find $\int_{\mathcal{K}} f(z)h(z)dz$, we can split the integral into two parts:

$$\int_{\mathcal{K}} f(z)h(z)dz = \int_{\{z \in \mathcal{K}:z \leq 0\}} f(z)h(z)dz + \int_{\{z \in \mathcal{K}:z > 0\}} f(z)h(z)dz.$$

Because of the symmetry properties $f(z) = f(-z)$ and $h(z) = -h(-z)$ we can rewrite the second term as

$$\int_{\{z \in \mathcal{K}:z > 0\}} f(z)h(z)dz = -\int_{\{z \in \mathcal{K}:z \leq 0\}} f(z)h(z)dz.$$

Using this substitution it is easily verified that

$$\int_{\mathcal{K}} f(z)h(z)dz = \int_{\{z \in \mathcal{K}:z \leq 0\}} f(z)h(z)dz - \int_{\{z \in \mathcal{K}:z \leq 0\}} f(z)h(z)dz = 0.$$

Now let $f(z) = \omega(\cdot)$ and $h(z) = g^1(\epsilon^1)$ and observe that the integral we want to calculate is the expectation of $\omega(\cdot)$ and therefore equal to $\xi(0)$ to complete this part of the proof.

Next, consider $\kappa > 0$. ω is shifted to the left and is point-symmetric across $-\kappa$. Now note two things: First, to the left of $-\kappa$ the values of ω are positive, to the rights the values are negative. Second, for any shock e leading to a realization $\omega(\kappa - e) = m$ there exists exactly one other shock e', which leads to a realization $\omega(\kappa - e') = -m$ and is an inversion of the former point at $(-\kappa, 0)$. Moreover, this holds true for any point in the graph of ω. Accordingly, we can define the whole graph as pairs of inversion points. Now observe, that the probability of an outcome $-m$ is always weakly larger than the probability of outcome m for all $m \geq 0$. To see this note that a shock generating m must be of size $-\kappa - c$, while the shock generating $-m$ must be $-\kappa + c$, for some constant $c \geq 0$. But then the

shock e that produces outcome m is in absolute value weakly larger than e'. As a consequence, because shocks are distributed symmetrically around zero, the density at e' is weakly larger than the density at e, $g^2(e') \geq g^2(e)$. Note that this must hold for all m, e and e', and accordingly the expectation of ω must be negative. From a similar argument it follows that the converse must hold if we assume $\kappa < 0$. Hence the proof is complete. $\qquad\square$

Knowing now that a pure strategy equilibrium exists we focus on interior equilibria henceforth. In Section 6.1 we also discuss the case when noise vanishes and the contest takes the form of a fully discriminating All-pay auction. In this case there are only corner equilibria and the results are qualitatively identical.

From (C.1) and (C.2) it follows that the first order conditions in an interior equilibrium are

$$\int_{\mathcal{S}^1} \left(g^2(\kappa - e^1) - \left[\frac{g^2(\kappa - e^1)g^{2\prime}(\kappa - e^1)}{C''(C'^{-1}(g^2(\kappa - e^1)))} \right] \right) g^1(e^1)\, de^1 - C'(x_F^1) = 0$$

and

$$\int_{\mathcal{S}^1} \left(g^2(\kappa - e^1) + \left[\frac{g^2(\kappa - e^1)g^{2\prime}(\kappa - e^1)}{C''(C'^{-1}(g^2(\kappa - e^1)))} \right] \right) g^1(e^1)\, de^1 - C'(x_T^1) = 0.$$

Simple manipulations reveal that equivalently the following must hold:

$$x_F^1 = C'^{-1}\left(\int_{\mathcal{S}^1} \left(g^2(\kappa - e^1) - \left[\frac{g^2(\kappa - e^1)g^{2\prime}(\kappa - e^1)}{C''(C'^{-1}(g^2(\kappa - e^1)))} \right] \right) g^1(e^1)\, de^1 \right)$$

$$x_T^1 = C'^{-1}\left(\int_{\mathcal{S}^1} \left(g^2(\kappa - e^1) + \left[\frac{g^2(\kappa - e^1)g^{2\prime}(\kappa - e^1)}{C''(C'^{-1}(g^2(\kappa - e^1)))} \right] \right) g^1(e^1)\, de^1 \right).$$

Using $\Delta^1 = x_F^1 - x_T^1$, it follows that in equilibrium it must hold that

$$\begin{aligned}
\Delta^1 &= \Sigma(\Delta^1 + a) := C'^{-1}\left(\int_{\mathcal{S}^1} \left(g^2(\Delta^1 + a - e^1) - \left[\frac{g^2(\Delta^1 + a - e^1)g^{2\prime}(\Delta^1 + a - e^1)}{C''(C'^{-1}(g^2(\Delta^1 + a - e^1)))} \right] \right) g^1(e^1)\, de^1 \right) \\
&\quad - C'^{-1}\left(\int_{\mathcal{S}^1} \left(g^2(\Delta^1 + a - e^1) + \left[\frac{g^2(\Delta^1 + a - e^1)g^{2\prime}(\Delta^1 + a - e^1)}{C''(C'^{-1}(g^2(\Delta^1 + a - e^1)))} \right] \right) g^1(e^1)\, de^1 \right).
\end{aligned} \tag{C.5}$$

The shape of this function is now important to determine equilibrium behavior. We now establish a few lemmata that help us to characterize equilibria.

Lemma 3. *Sign $\left[\Sigma(\Delta^1 + a) \right]$ = Sign$[\Delta^1 + a]$. Moreover, $\Sigma(\Delta^1 + a)$ is continuous, bounded, point symmetric at $-a$ in Δ^1, and $\lim_{|\Delta^1| \to \infty} \Sigma(\Delta^1 + a) = 0$.*

36

Proof. From $C'' > 0$ it follows that the inverse C'^{-1} is increasing. This together with Lemma 2 directly implies $\Sigma(0) = 0$, $\Sigma(+) = (+)$, and $\Sigma(-) = (-)$. For the symmetry properties look at $a = 0$ first. Then we have that

$$C'^{-1} \left(\int_{\mathcal{S}^1} \left(g^2(\Delta^1 - e^1) - \left[\frac{g^2(\Delta^1 - e^1)g^{2\prime}(\Delta^1 - e^1)}{C''(C'^{-1}(g^2(\Delta^1 - e^1)))} \right] \right) g^1(e^1) \, de^1 \right)$$

and

$$C'^{-1} \left(\int_{\mathcal{S}^1} \left(g^2(\Delta^1 - e^1) + \left[\frac{g^2(\Delta^1 - e^1)g^{2\prime}(\Delta^1 - e^1)}{C''(C'^{-1}(g^2(\Delta^1 - e^1)))} \right] \right) g^1(e^1) \, de^1 \right)$$

are mirror images of each other (in Δ^1) with the reflection axis being the vertical through zero. This follows from the first argument of C'^{-1} being axis symmetric at zero and the second being point symmetric at zero. Hence, the difference must be point symmetric at zero. Now note that by adding an arbitrary constant – for example a – to the argument of a function, the function is shifted horizontally by $-a$. Hence, Σ must be point symmetric at $-a$. Continuity and boundedness follow directly from all terms and C'^{-1} being continuous and bounded. That the limit vanishes follows from $\lim_{|x| \to \infty} g^2(x)g^{2\prime}(x) = 0$, which follows from quasi-concavity of g^2. This proves the lemma. $\square$

Lemma 4. *Assume $a \neq 0$ and let the investment of the more popular candidate in stage 1 be x_F^1 and the investment of his opponent be x_T^1. In any equilibrium it holds that*

$$x_T^1 \notin (x_F^1, x_F^1 + |a|).$$

Proof. To see this look at the first order conditions. Without loss of generality assume $a > 0$ and also assume $x_T^1 \in (x_F^1, x_F^1 + a)$. This implies $\kappa > 0$ and thus, by Lemma 2, $\xi(\kappa) < 0$. Hence, T's investments are strategic complements, and F's strategic substitutes. If F's first order condition holds, T's must be strictly negative and he hence would like to decrease investment. If T's first order condition holds, F's must be strictly positive and he would like to increase investment. Hence, this cannot be an equilibrium. $\square$

Lemma 5. *Assume $a \neq 0$ and let the investment of the more popular candidate in stage 1 be x_F^1 and*

the investment of his opponent be x_T^1. There exists $\overline{d} \geq 0$ such that if $a > \overline{d}$ the equilibrium in stage 1 is unique and $x_F^1 \geq x_T^1$.

Proof. This follows from Lemma 4 and the fact, that any investment greater than $\bar{x} := C^{-1}(1)$ is strictly dominated. Thus, when a becomes larger and larger, outspending the leading candidate becomes too expensive. The inequality is weak because we did not assume ϵ^2 has full support and hence the race might be decided if a is sufficiently large. If we assume ϵ^2 has full support $- \mathcal{S}^2 = \mathbb{R}$ – the inequality is strict because the leading candidate always spends positive amounts. $\qquad\square$

The derivative of Σ with respect to Δ^1 is

$$
\begin{aligned}
\frac{\partial \Sigma}{\partial \Delta^1} &= \frac{E_{\epsilon^1}\left[g^{2\prime}(\cdot)\right] - E_{\epsilon^1}\left[\frac{(g^{2\prime}(\cdot))^2}{C''(C'^{-1}(g(\cdot)))} + \frac{g^2(\cdot)g^{2\prime\prime}(\cdot)}{C''(C'^{-1}(g(\cdot)))} - \frac{g^2(\cdot)(g^{2\prime}(\cdot))^2 C'''(C'^{-1}(g(\cdot)))}{(C''(C'^{-1}(g(\cdot))))^3}\right]}{C''\left(C'^{-1}\left(E_{\epsilon^1}\left[\left(g^2(\cdot) - \left[\frac{g^2(\cdot)g^{2\prime}(\cdot)}{C''(C'^{-1}(g^2(\cdot)))}\right]\right)\right]\right)\right)} \\
&\quad - \frac{E_{\epsilon^1}\left[g^{2\prime}(\cdot)\right] + E_{\epsilon^1}\left[\frac{(g^{2\prime}(\cdot))^2}{C''(C'^{-1}(g(\cdot)))} + \frac{g^2(\cdot)g^{2\prime\prime}(\cdot)}{C''(C'^{-1}(g(\cdot)))} - \frac{g^2(\cdot)(g^{2\prime}(\cdot))^2 C'''(C'^{-1}(g(\cdot)))}{(C''(C'^{-1}(g(\cdot))))^3}\right]}{C''\left(C'^{-1}\left(E_{\epsilon^1}\left[\left(g^2(\cdot) + \left[\frac{g^2(\cdot)g^{2\prime}(\cdot)}{C''(C'^{-1}(g^2(\cdot)))}\right]\right)\right]\right)\right)}
\end{aligned}
\tag{C.6}
$$

The proposition we want to prove states that in close games there might be both equilibria in which the leading candidate spends more and some in which the trailing candidate spends more, depending on the distributions of ϵ^1 and ϵ^2 and the shape of the cost function. If one candidate has a sufficiently large advantage, in all equilibria this candidate will spend weakly more. If the equilibrium is unique for all a, in this equilibrium the leading candidate will always spend weakly more. A necessary and sufficient condition for a unique equilibrium for all a is that $\Sigma'(\Delta^1) < 1$ for all Δ^1. To see this note that if there are to be multiple equilibria, that is $\Delta^1 = \Sigma(\Delta^1)$ intersect more than once, then Σ must be steeper than Δ^1 somewhere. Starting from an intersection of the two functions, if the slope is strictly smaller than 1 to the right of the intersection Σ is strictly smaller than Δ^1, and to the left strictly larger, and hence there cannot be another equilibrium. If, however, there is some region in which the slope is larger than 1, there exists a shifting Σ in a way such that there are multiple equilibria. Hence,

if and only if

$$
\begin{aligned}
\frac{\partial \Sigma}{\partial \Delta^1} &= \frac{E_{\epsilon^1}\left[g^{2\prime}(\cdot)\right] - E_{\epsilon^1}\left[\frac{(g^{2\prime}(\cdot))^2}{C''(C'^{-1}(g(\cdot)))} + \frac{g^2(\cdot)g^{2\prime\prime}(\cdot)}{C''(C'^{-1}(g(\cdot)))} - \frac{g^2(\cdot)(g^{2\prime}(\cdot))^2 C'''(C'^{-1}(g(\cdot)))}{(C''(C'^{-1}(g(\cdot))))^3}\right]}{C''\left(C'^{-1}\left(E_{\epsilon^1}\left[\left(g^2(\cdot) - \left[\frac{g^2(\cdot)g^{2\prime}(\cdot)}{C''(C'^{-1}(g^2(\cdot)))}\right]\right)\right]\right)\right)} \\[2mm]
&\quad - \frac{E_{\epsilon^1}\left[g^{2\prime}(\cdot)\right] + E_{\epsilon^1}\left[\frac{(g^{2\prime}(\cdot))^2}{C''(C'^{-1}(g(\cdot)))} + \frac{g^2(\cdot)g^{2\prime\prime}(\cdot)}{C''(C'^{-1}(g(\cdot)))} - \frac{g^2(\cdot)(g^{2\prime}(\cdot))^2 C'''(C'^{-1}(g(\cdot)))}{(C''(C'^{-1}(g(\cdot))))^3}\right]}{C''\left(C'^{-1}\left(E_{\epsilon^1}\left[\left(g^2(\cdot) + \left[\frac{g^2(\cdot)g^{2\prime}(\cdot)}{C''(C'^{-1}(g^2(\cdot)))}\right]\right)\right]\right)\right)} < 1
\end{aligned} \tag{C.7}
$$

for all Δ^1, there is a unique equilibrium. Note that the absolute value of the slope is strictly decreasing in C'' (evaluated at the equilibrium). C'' enters only in the denominator of the condition and all the other terms are bounded by assumption. Hence, if $C''(x)$ is sufficiently large for all x the equilibrium is unique for all a. For $a = 0$ the equilibrium is symmetric and $\Delta^1 = 0$. If we now increase a we thereby shift Σ to the left, which, because of the fact that $\Sigma(+) = (+)$ (see Lemma 3), implies that the intersection is now where $\Delta^1 > 0$. This remains true for all $a > 0$, and the opposite is similarly true for $a < 0$.

If
$$
\frac{\partial \Sigma}{\partial \Delta^1} > 1
$$

for some Δ^1 there are multiple equilibria for some a. This follows from the discussion above. If this is true for small $|a|$ it is likely that either candidate might spend more in equilibrium. A sufficient condition for such equilibria is that

$$
\frac{\partial \Sigma}{\partial \Delta^1}\Big|_{a=0} = \frac{-2E_{\epsilon^1}\left[\frac{(g^{2\prime}(-\epsilon^1))^2}{C''(C'^{-1}(g(-\epsilon^1)))} + \frac{g^2(-\epsilon^1)g^{2\prime\prime}(-\epsilon^1)}{C''(C'^{-1}(g(-\epsilon^1)))} - \frac{g^2(\cdot)(g^{2\prime}(\cdot))^2 C'''(C'^{-1}(g(\cdot)))}{(C''(C'^{-1}(g(\cdot))))^3}\right]}{C''\left(C'^{-1}\left(E_{\epsilon^1}\left[(g^2(-\epsilon^1))\right]\right)\right)} > 1.
$$

This derivative is strictly positive (follows from Lemma 2). If C'' is sufficiently small the derivative gets larger than 1. Then there is one equilibrium $\Delta^1 = 0$. Moreover, because Σ vanishes as $|\Delta^1| \to \infty$ (see Lemma 3) and Σ is continuous, it follows from the intermediate value theorem that there are at least two more equilibria, one with $\Delta^1 > 0$ and one with $\Delta^1 < 0$. Because Σ is point symmetric in Δ^1 at zero (see Lemma 3) the asymmetric equilibria are symmetric to each other. Note that this analysis relies on the validity of the first-order conditions. We have shown though that when C'' gets too small the second order conditions are violated. In the parameterized version below we show that equilibria

with this property can exist.

From Lemma 5 it follows that in races with one dominant candidate this candidate will always expend more in the campaign, or both spend zero.

Now go the specific example of a normal distribution with variance σ_ϵ^2 and zero mean and a cost function $C(x) = \frac{c}{2}x^2$. The marginal cost function is then linear and the second derivative of the cost function is c. Hence, Σ simplifies significantly:

$$\Sigma(\Delta^1 + a) = -E_{\epsilon^1}\left[\frac{2\phi(\frac{\Delta^1+a-\epsilon^1}{\sigma_\epsilon})\phi'(\frac{\Delta^1+a-\epsilon^1}{\sigma_\epsilon})}{c^2}\right] = \frac{(\Delta^1+a)e^{\frac{-(\Delta^1+a)^2}{3\sigma_\epsilon^2}}}{\sqrt{27}\pi c^2\sigma_\epsilon^4}. \tag{C.8}$$

Now look at the shape of Σ. Since a only shifts the function horizontally, we assume $a = 0$ for now. The derivative with respect to Δ^1 is

$$\frac{(3\sigma_\epsilon^2 - 2(\Delta^1)^2)e^{\frac{-(\Delta^1)^2}{3\sigma_\epsilon^2}}}{9\sqrt{3}\pi c^2\sigma_\epsilon^6}. \tag{C.9}$$

This is strictly positive for $|\Delta^1| < \sqrt{\frac{3}{2}}\sigma_\epsilon$, negative for $|\Delta^1| > \sqrt{\frac{3}{2}}\sigma_\epsilon$, and zero for $|\Delta^1| = \sqrt{\frac{3}{2}}\sigma_\epsilon$. The maximum of $\overline{\Sigma} = \frac{1}{3\sqrt{2e}\pi c^2\sigma_\epsilon^3}$ is attained at $\Delta^1 = \sqrt{\frac{3}{2}}\sigma_\epsilon$ and the minimum of $\underline{\Sigma} = -\frac{1}{3\sqrt{2e}\pi c^2\sigma_\epsilon^3}$ is attained at $\Delta^1 = -\sqrt{\frac{3}{2}}\sigma_\epsilon$. Now look at the second derivative of Σ,

$$\frac{2\Delta^1(2\Delta^1 - 9\sigma_\epsilon^2))e^{-\frac{(\Delta^1)^2}{3\sigma_\epsilon^2}}}{27\sqrt{3}\pi c^2\sigma_\epsilon^8}.$$

This is strictly negative on $[-\infty, -\frac{3}{\sqrt{2}}\sigma_\epsilon) \cup (0, \frac{3}{\sqrt{2}}\sigma_\epsilon)$, and hence the function is strictly concave in this region, which also must include (and does) the maximum. Hence, Σ is strictly concave between zero and the maximum, and decreases monotonically thereafter. Hence, if $\Sigma'(\Delta^1) \leq 1$, the slope is strictly smaller than 1 (the slope of Δ^1) for all $\Delta^1 \geq 0$. Therefore, there exists a unique Δ^1 fulfilling $\Delta^1 = \Sigma(\Delta^1)$ not only for $a = 0$, but for all $a \in \mathbb{R}$. Using the condition $\Sigma'(0) \leq 1$ we can derive the lower bound on ρ for which a unique equilibrium exists in the benchmark model, $\rho > \bar{\rho} = \frac{1}{3^{3/4}\sqrt{\pi}}$. Thus, for given σ_ϵ the derivative becomes arbitrarily small at $\Delta^1 = 0$ if we increase c, and hence there exists $\tilde{c}$ such that for all $c > \tilde{c}$ the slope is less than 1, and larger than 1 else. In Figure 6 we show two examples. If the derivative at $\Delta^1 = 0$ is larger than 1 there exist multiple equilibria. Because Σ is strictly concave
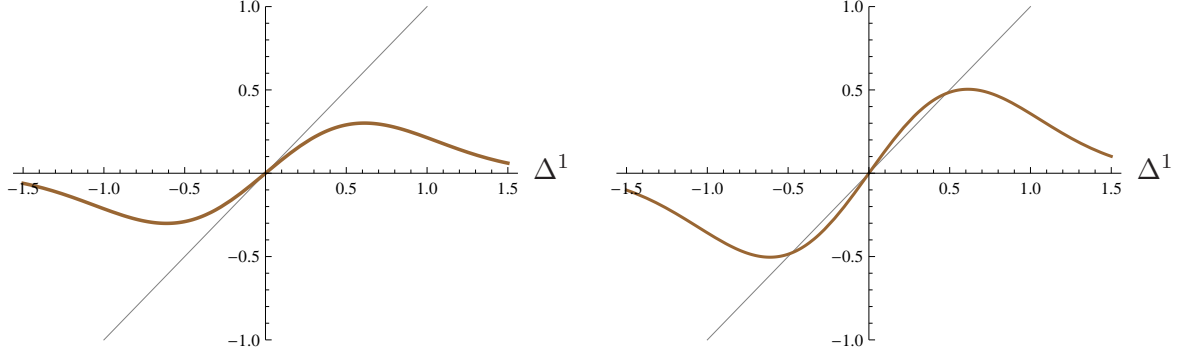
Figure 6: Δ^1 (gray) and $\Sigma(\Delta^1 + a)$ (brown) for $\sigma_\epsilon = .5$ and $c \in \{1.1, 0.97\}$.

on $[0, \sqrt{\frac{3}{2}}\sigma_\epsilon]$, and the derivative is zero at the end of this interval, it must become equal to 1 at some $\Delta^1 \in [0, \sqrt{\frac{3}{2}}\sigma_\epsilon]$. Denote this by $\tilde{\Delta}$. If we increase a now from zero we thereby shift Σ to the left by a. Hence, the two outer intersections of Δ^1 and Σ move to the right (Δ^1 increases), while the inner intersection moves to the left. Hence, there are two intersections converging to each other, the ones where $\Delta^1 < 0$. At $\tilde{d}$ they converge to $\tilde{\Delta}$, and hence there are only two equilibria left. If we increase a now further this equilibrium vanishes and only one equilibrium remains, in which $\Delta^1 > 0$.

To complete the proof we now show by example that the second order condition can hold in both stages when there are multiple equilibria in stage 1. The second order condition in stage 2 for F is

$$\phi'(\frac{d^2}{\sigma_\epsilon}) - c < 0,$$

and this must hold for all d^2. The second derivative has a maximum if $d^2 = -\sigma_\epsilon$, and if $c > \frac{1}{\sqrt{2e\pi}\sigma_\epsilon^2} \Leftrightarrow \rho > \frac{1}{\sqrt{2e\pi}}$ this maximum is strictly negative and F's second order condition holds for all d^2. Note that this is then also guarantees T's second order condition since candidates are symmetric in d^2. Moreover, note that $\frac{1}{\sqrt{2e\pi}} < \bar{\rho} = \frac{1}{3^{3/4}\sqrt{\pi}}$. Hence, if the second order condition holds marginally in stage 2, there are multiple equilibria in stage 1.

Now look at stage 1. Assuming the second order condition in stage 2 holds marginally, the second order condition in stage 1 is

$$-\frac{\kappa e^{-\frac{\kappa^2}{4\sigma_\epsilon^2}}}{4\sqrt{\pi}\sigma_\epsilon^3} - \frac{1}{\sqrt{2e\pi}\sigma_\epsilon^2} + \frac{(3\sigma_\epsilon^2 - 2\kappa^2)e^{\frac{1}{2}-\frac{\kappa^2}{3\sigma_\epsilon^2}}}{9\sqrt{6\pi}\sigma_\epsilon^4},$$
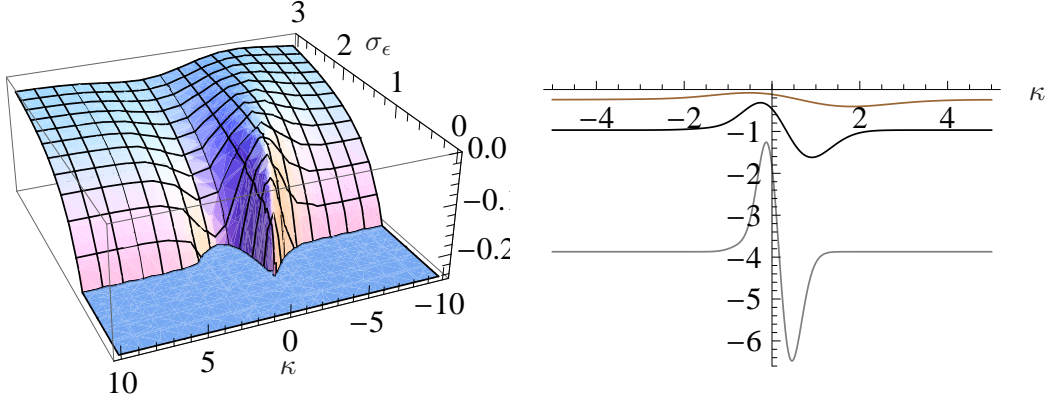
41

Figure 7: Numerical calculations of F's second derivative in stage 1, assuming second order conditions in stage 2 hold marginally, $c > \frac{1}{\sqrt{2e\pi\sigma_\epsilon^2}}$. Right panel: $\sigma_\epsilon \in \{1/4, 1/2, 1\}$.

where, as above, $\kappa = \Delta^1 + a$. This seems to be strictly negative for all $(\sigma_\epsilon, \kappa)$, and is strictly negative in all our numerical calculation, see Figure 7.

What is left to prove Proposition 4 is to show that no stable symmetric equilibrium exists when $\rho < \bar{\rho}$. In the following we show that the slope of each candidate's best response function for $a = 0$, evaluated at the symmetric intersection, is less than minus one.

Each candidate's best response is implicitly defined by

$$BR_F^1(x_T^1) = \max\left\{\{x_F^1 : \int_{\mathcal{S}^1}\left(g^2(x_F^1 - x_T^1 - e^1) - \left[\frac{g^2(x_F^1 - x_T^1 - e^1)g^{2\prime}(x_F^1 - x_T^1 - e^1)}{C''(C'^{-1}(g^2(x_F^1 - x_T^1 - e^1)))}\right]\right) g^1(e^1)\, de^1 - C'(x_F^1) = 0\}, 0\right\},$$

$$BR_T^1(x_F^1) = \max\left\{\{x_T^1 : \int_{\mathcal{S}^1}\left(g^2(x_F^1 - x_T^1 - e^1) + \left[\frac{g^2(x_F^1 - x_T^1 - e^1)g^{2\prime}(x_F^1 - x_T^1 - e^1)}{C''(C'^{-1}(g^2(x_F^1 - x_T^1 - e^1)))}\right]\right) g^1(e^1)\, de^1 - C'(x_T^1) = 0\}, 0\right\}.$$

Assuming a symmetric equilibrium with $x_F^1 = x_T^1$, it follows from Lemma 2 that the indirect effect is zero for both and hence

$$x_F^1 = x_T^1 = C'^{-1}(E_{\epsilon^1}[g^2(\epsilon^1)])$$

is an equilibrium. From the implicit function theorem it follows that the slope of the best responses is

$$\frac{\partial BR_i^1(x_j^1)}{\partial x_j^1} = \frac{E_{\epsilon^1}\left[\left(\frac{g^{2\prime}(\epsilon^1)}{C''(x^{**}(\epsilon^1))}\right)^2\left(C''(x^{**}(\epsilon^1)) - \frac{g^2(\epsilon^1)C'''(x^{**}(\epsilon^1))}{C''(x^{**}(\epsilon^1))}\right) + \frac{g^2(\epsilon^1)g^{2\prime\prime}(\epsilon^1)}{C''(x^{**}(\epsilon^1))}\right]}{E_{\epsilon^1}\left[\left(\frac{g^{2\prime}(\epsilon^1)}{C''(x^{**}(\epsilon^1))}\right)^2\left(C''(x^{**}(\epsilon^1)) - \frac{g^2(\epsilon^1)C'''(x^{**}(\epsilon^1))}{C''(x^{**}(\epsilon^1))}\right) + \frac{g^2(\epsilon^1)g^{2\prime\prime}(\epsilon^1)}{C''(x^{**}(\epsilon^1))}\right] + C''(C'^{-1}(E_{\epsilon^1}g^2(\epsilon^1)))}.$$

If this is smaller than minus 1 the equilibrium is unstable. Note that the denominator must be positive,

because it is the negative of the second derivative in equilibrium, and this has to be negative in equilibrium. This is the case whenever

$$-2E_{\epsilon^1}\left[\left(\frac{g^{2\prime}(\epsilon^1)}{C''(x^{**}(\epsilon^1))}\right)^2\left(C''(x^{**}(\epsilon^1)) - \frac{g^2(\epsilon^1)C'''(x^{**}(\epsilon^1))}{C''(x^{**}(\epsilon^1))}\right) + \frac{g^2(\epsilon^1)g^{2\prime\prime}(\epsilon^1)}{C''(x^{**}(\epsilon^1))}\right] > C''(C'^{-1}(E_{\epsilon^1}g^2(\epsilon^1))).$$

Hence, the cost function must be sufficiently convex but not too convex. Using the functional forms assumed this simplifies to

$$-2E_{\epsilon^1}\left[\left(\phi'\left(\frac{\epsilon^1}{\sigma_\epsilon}\right)\right)^2 + \phi\left(\frac{\epsilon^1}{\sigma_\epsilon}\right)\phi''\left(\frac{\epsilon^1}{\sigma_\epsilon}\right)\right] > c^2,$$

which further simplifies to

$$\rho = c\sigma_\epsilon^2 < (3^{3/4}\sqrt{\pi})^{-1} = \bar{\rho}.$$

This is the condition for the existence of multiple equilibria in the game with quadratic costs and normal shocks. Hence, whenever there are multiple equilibria in this game, there is no stable symmetric equilibrium. This completes the proof.

The proofs in Appendix A and B complete the proof of Proposition 7.

$\square$

D   Proof of Propositions 5 and 8

Proof. From the proof of Proposition 1 we know that aggregate expenditures in the general model and without a poll are equal to

$$4C'^{-1}\left(\int_{-\infty}^{\infty}\tilde{g}(a)\frac{1}{\sigma_\alpha}\phi\left(\frac{a-\mu_\alpha}{\sigma_\alpha}\right)da\right).$$

With a poll, second period expected aggregate expenditures are also symmetric and equal to

$$2\int_{-\infty}^{\infty}\left(\int_{\mathcal{S}^1}C'^{-1}(g^2(d^2))g^1(e^1)de^1\right)\frac{1}{\sigma_\alpha}\phi\left(\frac{a-\mu_\alpha}{\sigma_\alpha}\right)da.$$

as the proof of Proposition 2 shows. In period 1 we unfortunately cannot generally get a closed form solution for equilibrium investments with poll. So we need to take an indirect approach. Recall from

the proof of Proposition 3 that the first-order conditions are given by

$$\int_{\mathcal{S}^1} \left(g^2(d^2) - \left[g^2(d^2) \frac{\partial C'^{-1}(g^2(d^2))}{\partial g^2(d^2)} \frac{\partial g^2(d^2)}{\partial a} \right] \right) g^1(e^1)\, de^1 - C'\left(x_F^1\right) = 0$$

$$\int_{\mathcal{S}^1} \left(g^2(d^2) + \left[g^2(d^2) \frac{\partial C'^{-1}(g^2(d^2))}{\partial g^2(d^2)} \frac{\partial g^2(d^2)}{\partial a} \right] \right) g^1(e^1)\, de^1 - C'\left(x_T^1\right) = 0$$

where $d^2 = a + x_F^1 - x_T^1 - e^1$. Aggregate spending in period 1 is thus implicitly defined by

$$X^1(a) = \quad C'^{-1} \left(\int_{\mathcal{S}^1} \left(g^2(d^2) - \left[g^2(d^2) \frac{\partial C'^{-1}(g^2(d^2))}{\partial g^2(d^2)} \frac{\partial g^2(d^2)}{\partial a} \right] \right) g^1(e^1)\, de^1 \right)$$
$$+ C'^{-1} \left(\int_{\mathcal{S}^1} \left(g^2(d^2) + \left[g^2(d^2) \frac{\partial C'^{-1}(g^2(d^2))}{\partial g^2(d^2)} \frac{\partial g^2(d^2)}{\partial a} \right] \right) g^1(e^1)\, de^1 \right),$$

and expected aggregate stage 1 spending is

$$\int_{-\infty}^{\infty} X^1(a) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da.$$

By Jensen's inequality we know that if $C'''(x) \geq 0$

$$C'^{-1} \left(\int_{-\infty}^{\infty} \tilde{g}(a) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) d\,a \right)$$

$$\geq \quad \int_{-\infty}^{\infty} C'^{-1} \left(\int_{\mathcal{S}^1} g^2(a - e^1) g^1(e^1)\, de^1 \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da$$

$$\geq \quad \int_{-\infty}^{\infty} \left(\int_{\mathcal{S}^1} C'^{-1}(g^2(a - e^1)) g^1(e^1)\, de^1 \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da.$$

This inequality is strict for $C'''(x) > 0$ and non-degenerate distribution functions. Furthermore, we used the fact that $\tilde{g}(a)$ is density of the sum of of ϵ^1 and ϵ^2 and thus the convolution of g^1 and g^2. The

difference between expected aggregate investment with and without poll is equal to

$$
\int_{-\infty}^{\infty} \left(C'^{-1} \left(\int_{\mathcal{S}^1} \left(g^2(d^2) - \Upsilon(d^2) \right) g^1(e^1) de^1 \right) + C'^{-1} \left(\int_{\mathcal{S}^1} \left(g^2(d^2) + \Upsilon(d^2) \right) g^1(e^1) de^1 \right) \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da
$$

$$
+ 2 \int_{-\infty}^{\infty} \left(\int_{\mathcal{S}^1} C'^{-1}(g^2(d^2)) g^1(e^1) de^1 \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da - 4 C'^{-1} \left(\int_{-\infty}^{\infty} \tilde{g}(a) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da \right)
$$

$$
\leq \int_{-\infty}^{\infty} \left(C'^{-1} \left(\int_{\mathcal{S}^1} \left(g^2(d^2) - \Upsilon(d^2) \right) g^1(e^1) de^1 \right) + C'^{-1} \left(\int_{\mathcal{S}^1} \left(g^2(d^2) + \Upsilon(d^2) \right) g^1(e^1) de^1 \right) \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da
$$

$$
- 2 \int_{-\infty}^{\infty} C'^{-1} \left(\int_{\mathcal{S}^1} g^2(a - e^1) g^1(e^1) de^1 \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da
$$

$$
+ 2 \int_{-\infty}^{\infty} \left(\int_{\mathcal{S}^1} \left(C'^{-1}(g^2(d^2)) - C'^{-1}(g^2(a - e^1)) \right) g^1(e^1) de^1 \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da
$$

$$
\leq 2 \int_{-\infty}^{\infty} \left(C'^{-1} \left(\int_{\mathcal{S}^1} \left(g^2(d^2) \right) g^1(e^1) de^1 \right) - C'^{-1} \left(\int_{\mathcal{S}^1} g^2(a - e^1) g^1(e^1) de^1 \right) \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da
$$

$$
+ 2 \int_{-\infty}^{\infty} \left(\int_{\mathcal{S}^1} \left(C'^{-1}(g^2(d^2)) - C'^{-1}(g^2(a - e^1)) \right) g^1(e^1) de^1 \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da
$$

$$
< 0
$$

where we denote $\Upsilon(d^2) = \left[g^2(d^2) \frac{\partial C'^{-1}(g^2(d^2))}{\partial g^2(d^2)} \frac{\partial g^2(d^2)}{\partial a} \right]$ for ease of exposition. The first two inequalities follow from Jensen's inequality and the fact that C'^{-1} is a concave function. The last inequality is strict because of the fact that $Sign[a] = Sign[x_F^1 - x_T^1]$ (momentum) and thus for all $a \neq 0$, $|g^2(d^2)| < |g^2(a - e^1)|$ due to quasi-concavity and symmetry around zero of g^2. Hence, for $C'''(x) \geq 0$ polls always decrease expected aggregate spending.

Now consider again the benchmark model and a situation where there is anti-momentum. If in an anti-momentum equilibrium it holds that $|\Delta^1| > |2a|$ we are done because then the above proof establishes the result also for this case. From the first order conditions we know the difference in investments is Σ as defined above in (C.8) and equals

$$
\Sigma(\Delta^1 + a) = \frac{(\Delta^1 + a) e^{\frac{-(\Delta^1 + a)^2}{3\sigma_\epsilon^2}}}{\sqrt{27} \pi c^2 \sigma_\epsilon^4}.
$$

To find the maximum a for which multiple equilibria exist we take the derivative with respect to Δ^1 and set it equal to 1:

$$
\Sigma'(\Delta^1 + a) = \frac{e^{-\frac{(\Delta^1 + a)^2}{3\sigma_\epsilon^2}} \left(3\sigma_\epsilon^2 - 2(\Delta^1 + a)^2 \right)}{9\sqrt{3} \pi c^2 \sigma_\epsilon^6} \overset{!}{=} 1.
$$

The solution is not to be obtained in closed form, but can be expressed in terms of the Lambert W

function. Let $z = \frac{3}{2}\sqrt{3e}\pi c^2\sigma_\epsilon^4$ and let $W(z)$ be the solution of the Lambert equation

$$z = W(z)e^{W(z)}.$$

Because $z > 0$ there exists a unique solution to this problem and $W'(z) > 0$. Hence, $W(z)$ is bounded from below by $\underline{W}$ and from above by $\overline{W}$. The maximum value that z can take on is achieved exactly when $\rho = \bar{\rho}$, in which case $z = \sqrt{e}/2$ and hence $\overline{W} = 1/2$. As the lower bound we can take the minimum value of ρ for which the second order conditions in stage 2 hold, which is $\underline{\rho} = 1/(\sqrt{2e\pi})$. This yields $z = \frac{3}{4}\sqrt{\frac{3}{e}}$, in which case then $\underline{W} \approx 0.485075$. The solution to the above equation is

$$\Delta^1 + a = \pm\sigma_\epsilon\sqrt{\frac{3}{2} - 3W(z)}.$$

Note that since $W(z) \leq \frac{1}{2}$ the square root has a real solution. The value of Σ evaluated at this point is

$$\Sigma(\bar{a}) = \pm\frac{\sigma_\epsilon\sqrt{\frac{3}{2} - 3W(z)}}{2W(z)},$$

which by definition is the difference in investments at this point. We focus on $a \geq 0$ without loss of generality, implying that in both expressions above we have to take the one with a minus. Hence, we can determine the maximum a for which anti-momentum can exist in this case:

$$\bar{a} = -\sigma_\epsilon\sqrt{\frac{3}{2} - 3W(z)} + \frac{\sigma_\epsilon\sqrt{\frac{3}{2} - 3W(z)}}{2W(z)} = \sigma_\epsilon\sqrt{\frac{3}{2} - 3W(z)}\left(\frac{1}{2W(z)} - 1\right) \geq 0.$$

$\bar{a}$ is zero iff $\rho = \bar{\rho}$, which must of course be the case. Conditional on ρ, $\bar{a}$ is linear in σ_ϵ.

So what we need to show is that

$$\Delta^1 : \left\{ \Delta^1 = \frac{(\Delta^1 + a)e^{\frac{-(\Delta^1+a)^2}{3\sigma_\epsilon^2}}}{\sqrt{27}\pi c^2\sigma_\epsilon^4} \mid a \in [0,\bar{a}] \wedge \rho \in [\underline{\rho},\overline{\rho}] \right\} \leq -2a.$$

It is easy to show that at $\overline{a}$ this always holds,

$$
\begin{aligned}
\Sigma(\overline{a}) + 2\overline{a} &= -\frac{\sigma_\epsilon\sqrt{\frac{3}{2} - 3W(z)}}{2W(z)} + 2\left(-\sigma_\epsilon\sqrt{\frac{3}{2} - 3W(z)} + \frac{\sigma_\epsilon\sqrt{\frac{3}{2} - 3W(z)}}{2W(z)}\right) \\
&= \sigma_\epsilon\left(\frac{\sqrt{\frac{3}{2} - 3W(z)}}{2W(z)} - 2\sqrt{\frac{3}{2} - 3W(z)}\right) \\
&= \sigma_\epsilon\sqrt{\frac{3}{2} - 3W(z)}\left(\frac{1}{2W(z)} - 2\right) < 0
\end{aligned}
$$

Remember that at this point the slope of Σ with respect to a is positive, it is one. Also remember that the slope of Σ is strictly negative for all finite $a \notin \left[-\sqrt{\frac{3}{2}}\sigma_\epsilon, \sqrt{\frac{3}{2}}\sigma_\epsilon\right]$, zero at the boundaries of this interval, and positive within. In other words, the set of values of a for which the slope is positive is convex (see equation (C.9) and paragraph that follows it). If we can show that in the equilibrium to which the anti-momentum converges as $a \to 0$ the slope is also positive, we have proven that, starting at $a = 0$, in all anti-momentum equilibria, the larger a grows, the smaller becomes $|\Delta^1|$. We know already that for the largest a that admits anti-momentum equilibria $|\Delta^1|$ is large enough. Hence, if the slope at the equilibrium at $a = 0$ is positive we proved the proposition (see also Figure 6). It is easily verified that in this equilibrium $|\Delta^1| = \sqrt{3}\sqrt{-\sigma_\epsilon^2\log\left(3\sqrt{3}\pi c^2\sigma_\epsilon^4\right)}$, which is well defined since $\rho \le \overline{\rho}$. Using this expression in the derivative of Σ above yields

$$
\frac{\partial\Sigma}{\partial\Delta}\Big|_{a=0 \wedge \Delta^1=-\sqrt{3}\sqrt{-\sigma_\epsilon^2\log\left(3\sqrt{3}\pi c^2\sigma_\epsilon^4\right)}} = 2\log\left(3\sqrt{3}\pi c^2\sigma_\epsilon^4\right) + 1,
$$

which is positive for all $\rho \in [\underline{\rho}, \overline{\rho}]$. This proves the proposition. $\qquad\square$

E   Proof of Proposition 6

Recall the candidates' first order conditions in stage 1 are

$$
\begin{aligned}
\frac{\partial\pi_F^1}{\partial x_F^1} &= E_{\epsilon^1}\left[\frac{\phi\left(\frac{a+\Delta^1-\epsilon^1}{\sigma_\epsilon}\right)}{\sigma_\epsilon}\right] - E_{\epsilon^1}\left[\frac{\phi\left(\frac{a+\Delta^1-\epsilon^1}{\sigma_\epsilon}\right)\phi'\left(\frac{a+\Delta^1-\epsilon^1}{\sigma_\epsilon}\right)}{c\,\sigma_\epsilon^3}\right] - c\,x_F^1 \overset{!}{=} 0 \\
\frac{\partial\pi_T^1}{\partial x_T^1} &= E_{\epsilon^1}\left[\frac{\phi\left(\frac{a+\Delta^1-\epsilon^1}{\sigma_\epsilon}\right)}{\sigma_\epsilon}\right] + E_{\epsilon^1}\left[\frac{\phi\left(\frac{a+\Delta^1-\epsilon^1}{\sigma_\epsilon}\right)\phi'\left(\frac{a+\Delta^1-\epsilon^1}{\sigma_\epsilon}\right)}{c\,\sigma_\epsilon^3}\right] - c\,x_T^1 \overset{!}{=} 0
\end{aligned}
$$

in symmetric equilibrium, where $a = 0$, it is then easily verified that $x_F^1 = x_T^1 = \frac{1}{2\sqrt{\pi}c\sigma_\epsilon}$. To see this note that from Lemma 2 the second term in the first order conditions is zero. The first is simply the expectation of the normal density and hence the convolution of two normal densities, and thus equals $E_{\epsilon^1}\left[\frac{\phi(\frac{a+\Delta^1-\epsilon^1}{\sigma_\epsilon})}{\sigma_\epsilon}\right] = \frac{1}{2\sqrt{\pi}\sigma_\epsilon}$. Investment follows immediately.

Letting $\kappa = a + \Delta^1 - \epsilon^1$, second order conditions are

$$
\begin{aligned}
\frac{\partial^2 \pi_F^1}{\partial(x_F^1)^2} &= E_{\epsilon^1}\left[\frac{\phi'(\frac{\kappa}{\sigma_\epsilon})}{\sigma_\epsilon^2}\right] - E_{\epsilon^1}\left[\frac{\phi'(\frac{\kappa}{\sigma_\epsilon})^2 + \phi(\frac{\kappa}{\sigma_\epsilon})\phi''(\frac{\kappa}{\sigma_\epsilon})}{c\,\sigma_\epsilon^4}\right] - c \overset{!}{<} 0 \\
\frac{\partial^2 \pi_T^1}{\partial(x_T^1)^2} &= -E_{\epsilon^1}\left[\frac{\phi'(\frac{\kappa}{\sigma_\epsilon})}{\sigma_\epsilon^2}\right] - E_{\epsilon^1}\left[\frac{\phi'(\frac{\kappa}{\sigma_\epsilon})^2 + \phi(\frac{\kappa}{\sigma_\epsilon})\phi''(\frac{\kappa}{\sigma_\epsilon})}{c\,\sigma_\epsilon^4}\right] - c \overset{!}{<} 0
\end{aligned}
$$

For comparative statics we totally differentiate the candidates' FOCs:

$$
\begin{aligned}
\left(\frac{\partial^2 \pi_F^1}{\partial(x_F^1)^2}\right) dx_F^1 + \left(\frac{\partial^2 \pi_F^1}{\partial x_F^1 \partial x_T^1}\right) dx_T^1 &= -\left(\frac{\partial^2 \pi_F^1}{\partial x_F^1 \partial a}\right) da \\
\left(\frac{\partial^2 \pi_T^1}{\partial x_F^1 \partial x_T^1}\right) dx_F^1 + \left(\frac{\partial^2 \pi_T^1}{\partial(x_T^1)^2}\right) dx_T^1 &= -\left(\frac{\partial^2 \pi_T^1}{\partial x_T^1 \partial a}\right) da
\end{aligned}
$$

Evaluating this in the symmetric equilibrium, where $a = 0$ and $x_F^1 = x_T^1 = \frac{1}{2\sqrt{\pi}c\sigma_\epsilon}$, we can express the total differential in matrix notation. Defining

$$
\mathbf{M} = \begin{pmatrix} \frac{\partial^2 \pi_F^1}{\partial(x_F^1)^2} & \frac{\partial^2 \pi_F^1}{\partial x_F^1 \partial x_T^1} \\ \frac{\partial^2 \pi_T^1}{\partial x_F^1 \partial x_T^1} & \frac{\partial^2 \pi_T^1}{\partial(x_T^1)^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} - c & -\frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} \\ -\frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} & \frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} - c \end{pmatrix}
$$

$$
\mathbf{P} = \begin{pmatrix} -\frac{\partial^2 \pi_F^1}{\partial x_F^1 \partial a} \\ -\frac{\partial^2 \pi_T^1}{\partial x_T^1 \partial a} \end{pmatrix} = \begin{pmatrix} -\frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} \\ \frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} \end{pmatrix}
$$

$$
\mathbf{dx} = \begin{pmatrix} dx_F^1 \\ dx_T^1 \end{pmatrix}
$$

we can rewrite the total differential as

$$
\mathbf{M} \cdot \mathbf{dx} = \mathbf{P}\,da
$$

We apply Cramer's rule for comparative statics. Let

$$M_F = \begin{pmatrix} -\frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} & -\frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} \\ \frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} & \frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} - c \end{pmatrix}$$

and

$$M_T = \begin{pmatrix} \frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} - c & -\frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} \\ -\frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} & \frac{1}{6\sqrt{3}c\pi\sigma_\epsilon^4} \end{pmatrix}.$$

The derivative we are searching for is then $\frac{\partial x_F^1}{\partial a} = \frac{|M_F|}{|M|}$ and $\frac{\partial x_T^1}{\partial a} = \frac{|M_T|}{|M|}$, or

$$\frac{\partial x_F^1}{\partial a}\Big|_{a=0} = \frac{1}{6\sqrt{3}\pi c^2\sigma_\epsilon^4 - 2} \tag{E.1}$$

$$\frac{\partial x_T^1}{\partial a}\Big|_{a=0} = -\frac{1}{6\sqrt{3}\pi c^2\sigma_\epsilon^4 - 2} \tag{E.2}$$

$\frac{\partial x_F^1}{\partial a}$ is positive iff $c\sigma_\epsilon^2 > (3^{3/4}\sqrt{\pi})^{-1} \Leftrightarrow \rho > \bar{\rho}$, and $\frac{\partial x_T^1}{\partial a}$ is then negative. Expected investment in stage 2 is

$$Ex^2(\kappa) = E_{\epsilon^1}\left[\frac{1}{\sigma_\epsilon c}\phi\left(\frac{a + x_F^1 - x_T^1 - \epsilon^1}{\sigma_\epsilon}\right)\right] = \frac{e^{-\frac{(a+x_F^1-x_T^1)^2}{4\sigma_\epsilon^2}}}{2\sqrt{\pi}c\sigma_\epsilon} = \frac{e^{-\frac{\kappa^2}{4\sigma_\epsilon^2}}}{2\sqrt{\pi}c\sigma_\epsilon}.$$

The derivative of this expected investment with respect to κ is

$$\frac{\partial Ex^2}{\partial\kappa} = -\frac{\kappa e^{-\frac{\kappa^2}{4\sigma_\epsilon^2}}}{4\sqrt{\pi}c\sigma_\epsilon^3}$$

which, if we evaluate this at the symmetric equilibrium at $a = 0$, is equal to zero. Hence, marginally changing a only has an effect on stage 1 spending but no effect on spending in stage 2. Hence, expected total investment of F is increasing in a when $a = 0$ and T's expected total investment decreases. This proves the proposition.

F Proof of Corollary 3

The derivative of the candidates' equilibrium spending, derived using Cramer's rule as above, is

$$
\frac{\partial x_F^1}{\partial a} = \frac{27\sqrt{\pi}c\sigma_\epsilon^3(a+x_F^1-x_T^1)e^{\frac{6x_T^1(a+x_F^1)+(a+x_F^1)^2+x_T^{1\,2}}{12\sigma_\epsilon^2}} + 2\sqrt{3}e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(2(a+x_F^1-x_T^1)^2 - 3\sigma_\epsilon^2\right)}{4\sqrt{3}e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right) - 108\pi c^2\sigma_\epsilon^6 e^{\frac{(a+x_F^1)^2+x_T^{1\,2}}{3\sigma_\epsilon^2}}}
$$

$$
\frac{\partial x_T^1}{\partial a} = \frac{27\sqrt{\pi}c\sigma_\epsilon^3(a+x_F^1-x_T^1)e^{\frac{6x_T^1(a+x_F^1)+(a+x_F^1)^2+x_T^{1\,2}}{12\sigma_\epsilon^2}} + 2\sqrt{3}e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right)}{4\sqrt{3}e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right) - 108\pi c^2\sigma_\epsilon^6 e^{\frac{(a+x_F^1)^2+x_T^{1\,2}}{3\sigma_\epsilon^2}}}
$$

We proceed as follows. First we show that the denominator is strictly negative. Then we show that for T the numerator is strictly positive. Finally, we show that for F the numerator is negative for small a and positive thereafter. To do all this we make the following transformations: $c = \frac{t}{3^{3/4}\sqrt{\pi}\sigma_\epsilon^2}$ for some $t > 1$ and we let $\kappa = a + x_F^1 + x_T^1 > 0$ for all $a > 0$ in the following where applicable:

$$
\frac{\partial x_F^1}{\partial a} = \frac{9\sqrt[4]{3}st(a+x_F^1-x_T^1)e^{\frac{6x_T^1(a+x_F^1)+(a+x_F^1)^2+x_T^{1\,2}}{12\sigma_\epsilon^2}} + 2\sqrt{3}e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right)}{4\sqrt{3}\left(e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right) - 3\sigma_\epsilon^2 t^2 e^{\frac{(a+x_F^1)^2+x_T^{1\,2}}{3\sigma_\epsilon^2}}\right)}
$$

$$
\frac{\partial x_T^1}{\partial a} = \frac{9\sqrt[4]{3}st(a+x_F^1-x_T^1)e^{\frac{6x_T^1(a+x_F^1)+(a+x_F^1)^2+x_T^{1\,2}}{12\sigma_\epsilon^2}} + 2\sqrt{3}e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right)}{4\sqrt{3}\left(e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right) - 3\sigma_\epsilon^2 t^2 e^{\frac{(a+x_F^1)^2+x_T^{1\,2}}{3\sigma_\epsilon^2}}\right)}
$$

Look at the denominator, which is identical for both, first and assume it is negative:

$$
4\sqrt{3}\left(e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right) - 3\sigma_\epsilon^2 t^2 e^{\frac{(a+x_F^1)^2+x_T^{1\,2}}{3\sigma_\epsilon^2}}\right) < 0
$$

$$
\Leftrightarrow 4\sqrt{3}e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right) < 12\sqrt{3}\sigma_\epsilon^2 t^2 e^{\frac{(a+x_F^1)^2+x_T^{1\,2}}{3\sigma_\epsilon^2}}
$$

$$
\Leftrightarrow e^{-\frac{(a+x_F^1-x_T^1)^2}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a+x_F^1-x_T^1)^2\right) < 3\sigma_\epsilon^2 t^2 \Leftrightarrow e^{-\frac{\kappa^2}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2\kappa^2\right) < 3\sigma_\epsilon^2 t^2
$$

It is easy to see that the last line must be true. The left side is decreasing in κ. At $\kappa = 0$ in the symmetric equilibrium the inequality also holds. Hence, it holds generally.

Now look at the numerator of T and assume it is positive:

$$9st(a + x_F^1 - x_T^1)e^{\frac{(a+x_F^1+x_T^1)^2}{12\sigma_\epsilon^2}} + 2\sqrt[4]{3}e^{\frac{x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a + x_F^1 - x_T^1)^2\right) > 0$$

$$\Leftrightarrow 2\sqrt[4]{3}e^{-\frac{(a+x_F^1-x_T^1)^2}{12\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2(a + x_F^1 - x_T^1)^2\right) + 9st(a + x_F^1 - x_T^1) > 0$$

$$\Leftrightarrow 2\sqrt[4]{3}e^{-\frac{\kappa^2}{12\sigma_\epsilon^2}}\left(3\sigma_\epsilon^2 - 2\kappa^2\right) + 9\kappa st > 0$$

Note that the last line always holds when $\sqrt{\frac{3}{2}}\sigma_\epsilon > \kappa$. Hence, let $\kappa = \sqrt{\frac{3}{2}}\sigma_\epsilon m$ for $m \in [0,1]$. We get then

$$3\sqrt{2}\sqrt[4]{3}mt - 4e^{-\frac{m^2}{8}}\left(m^2 - 1\right) > 0.$$

$m \leq 1$ implies $-4(m^2 - 1) \geq 0$, and hence this must be strictly positive and the inequality must holds. Together with the negative denominator this consequently implies T's investment decreases monotonically in a, proving one part of the corollary.

Finally, consider the numerator of F.

$$27\sqrt{\pi}c\sigma_\epsilon^3(a + x_F^1 - x_T^1)e^{\frac{6x_T^1(a+x_F^1)+(a+x_F^1)^2+x_T^{1\,2}}{12\sigma_\epsilon^2}} + 2\sqrt{3}e^{\frac{2x_T^1(a+x_F^1)}{3\sigma_\epsilon^2}}\left(2(a + x_F^1 - x_T^1)^2 - 3\sigma_\epsilon^2\right) > 0$$

$$\Leftrightarrow \frac{9}{2}\sqrt{3\pi}c\sigma_\epsilon^3(a + x_F^1 - x_T^1)e^{\frac{(a+x_F^1-x_T^1)^2}{12\sigma_\epsilon^2}} + 2(a + x_F^1 - x_T^1)^2 > 3\sigma_\epsilon^2$$

$$\Leftrightarrow \frac{9}{2}\sqrt{3\pi}c\kappa\sigma_\epsilon^3 e^{\frac{\kappa^2}{12\sigma_\epsilon^2}} + 2\kappa^2 > 3\sigma_\epsilon^2$$

It is immediately observed that the LHS is zero for $\kappa = 0$, monotonically increasing in κ, and reaching infinity as $\kappa \to +\infty$. Hence, iff κ is large the inequality holds. In addition, there exists a unique $\bar{\kappa}$ for which the LHS equals the RHS and for all smaller values the LHS is smaller. Because the denominator is negative this yields that F's spending increases first, reaches a maximum and declines then monotonically. Together with Proposition 6 this completes the proof.

G   All-pay auction

Consider the example when $E[(\epsilon^t)^2] \to 0$. Also consider the situation without a poll first. In this situation the candidates do not know the median's exact ranking but have beliefs $F(\alpha)$. For candidate

F the probability to win the race is

$$Pr[x_F^1 + x_F^2 + a - x_T^1 - x_T^2 > 0] = Pr[a > x_T^1 + x_T^2 - x_F^1 - x_F^2] = 1 - F(x_T^1 + x_T^2 - x_F^1 - x_F^2),$$

and similarly for T we get $F(x_T^1 + x_T^2 - x_F^1 - x_F^2)$. Hence, we may write the candidates objectives as

$$\max_{(x_F^1, x_F^2) \in \mathbb{R}_+^2} \quad 1 - F(x_T^1 + x_T^2 - x_F^1 - x_F^2) - C(x_F^1) - C(x_F^2),$$
$$\max_{(x_T^1, x_T^2) \in \mathbb{R}_+^2} \quad F(x_T^1 + x_T^2 - x_F^1 - x_F^2) - C(x_T^1) - C(x_T^2).$$

The corresponding first order condition for i in t reads

$$f(x_T^1 + x_T^2 - x_F^1 - x_F^2) - C'(x_i^t) \stackrel{!}{=} 0.$$

This is identical for all $t = 1, 2$ and $i = F, T$. Hence,

$$x_F^t = C'^{-1}(f(0)).$$

For this to be an equilibrium the second order conditions need to hold:

$$f'(0) - C''(C'^{-1}(f(0))) \leq 0.$$

If costs are sufficiently convex and / or the variance of beliefs is sufficiently large this inequality holds generally.

Now look at the case of a campaign with polls. Candidates have perfect knowledge of the median's ranking, and since there is no shock, there is no exogenous noise left. Consider stage 2 first, and assume without loss of generality $d^2 \geq 0$. It is easily shown that there cannot be a pure strategy equilibrium in this stage (see e.g. Nalebuff and Stiglitz (1983) or Hillman and Riley (1989)). The stage game is similar to the game analyzed in Meirowitz (2008) with the difference that we have strictly convex costs instead of linear cost functions. Let $\bar{x} := C^{-1}(1)$, the maximum investment that is not strictly dominated. Moreover, let $Q_i(x_i^2)$ be candidate i's mixed strategy with support S_i, $i = F, T$. Then stage 2 equilibrium is given by
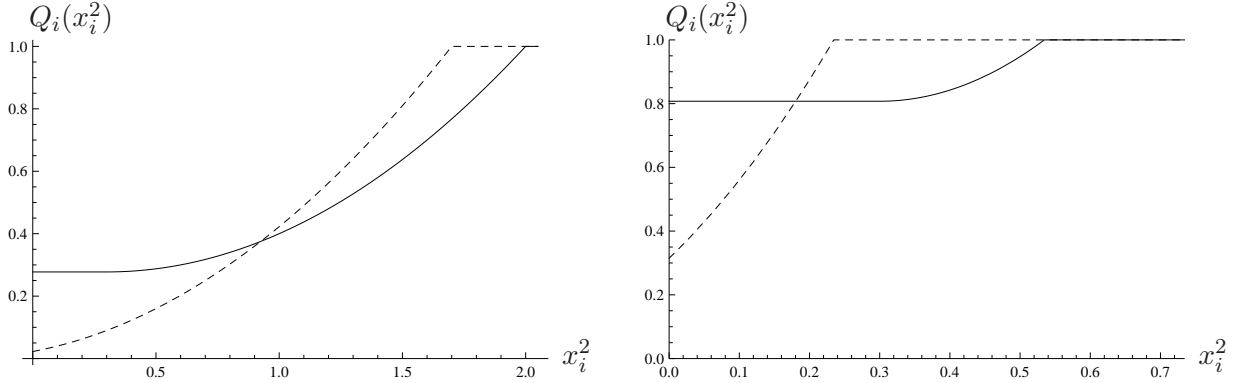
Figure 8: Investment distribution functions for the leader (dashed) and his opponent (solid). We assumed $d^2 = \frac{3}{10}$ and in the left panel $c = 1$, while in the right $c = 7$.

Proposition 11. *Without loss of generality, let $d^2 \geq 0$. There is a unique equilibrium in mixed strategies if $d^2 < \bar{x}$ with $S_F = [0, \bar{x} - d^2]$ and $S_T = \{0\} \cup [d^2, \bar{x}]$. Candidate F randomizes according to*

$$
Q_F(x_F^2) = \begin{cases} 0 & \text{if } x_F < 0, \\ C(d^2 + x_F) & \text{if } x_F \in [0, \bar{x} - d^2], \\ 1 & \text{if } x_F > \bar{x} - d^2, \end{cases}
$$

while T's mixed strategy is given by

$$
Q_T(x_T^2) = \begin{cases} 0 & \text{if } x_T < 0, \\ 1 - C(\bar{x} - d^2) & \text{if } x_T \in [0, d^2], \\ 1 - C(\bar{x} - d^2) + C(x_T - d^2) & \text{if } x_T \in [d^2, \bar{x}], \\ 1 & \text{if } x_T > \bar{x}. \end{cases}
$$

Expected utilities are $1 - C(\bar{x} - d^2)$ for F and 0 for T. If $d^2 \geq \bar{x}$ both candidates expend zero, $x_F^2 = x_T^2 = 0$, and expected utilities are 1 and 0 respectively.

Proof. First we show that this is an equilibrium. Consider F choosing investment $x \in [0, \bar{x} - d^2]$. The

53

associated payoff is then

$$
\begin{aligned}
Pr[x + d^2 > x_T^2] - C(x) &= Q_T(x + d^2) - C(x) = 1 - C(\bar{x} - d^2) + C(x + d^2 - d^2) - C(x) \\
&= 1 - C(\bar{x} - d^2) + C(x) - C(x) = 1 - C(\bar{x} - d^2).
\end{aligned}
$$

He is hence indifferent between all pure strategies in S_F given Q_T. Note also, that $x > \bar{x} - d^2$ is strictly dominated since lowering investment to $\bar{x} - d^2$ saves costs without altering the probability to win. Hence, given Q_T F is indifferent between all pure strategies in S_F and cannot improve on the expected payoff by deviating. Now consider T. Given Q_F, choosing $x \in S_T$ gives the following payoff:

$$
\begin{aligned}
Pr[x > x_F^2 + d^2] - C(x) &= Pr[x - d^2 > x_F^2] - C(x) = Q_F(x - d^2) - C(x) \\
&= C(d^2 + x - d^2) - C(x) = C(x) - C(x) = 0.
\end{aligned}
$$

Investment $x = 0$ never wins but does not imply costs, either. All investments $x \in [0, d^2]$ are strictly dominated since they cannot win but imply costs. Choosing investment $x \in [d^2, \bar{x}]$ also yields zero payoff, since the probability to win exactly equals the associated investment costs. Hence, all $x \in S_B$ yield zero expected payoff. Note that deviating to $x > \bar{x}$ is strictly dominated, since it would win for sure but imply costs that are greater than the value of winning. Hence, there is no beneficial deviation for T as well and thus Q_F and Q_T are an equilibrium.

Now consider the support S_F. By elimination of strictly dominated strategies it is apparent that F will never choose investment greater than $\bar{x} - d^2$, because T will never choose investment greater than $\bar{x}$. The maximum bid neither can be lower, because than T would have a winning strategy and could simply overbid F and win for sure. That F's minimum bid must be zero and that the support cannot have holes, or that it must be convex, follows directly from the proof in Hillman and Riley (1989). Hence, S_F is the only support a mixed strategy of F can have in equilibrium. Now consider S_T. It is apparent that all $x \in (0, d^2)$ can never win but imply positive costs, and are hence strictly dominated. The maximum possible bid is naturally $\bar{x}$, and must also be $\bar{x}$. Otherwise F could overbid x_T^2 and win for sure. Convexity of S_T follows again from Hillman and Riley (1989).

It is now easy to see that given S_F and S_T there do not exist other investment distribution making the opponent indifferent between all investment in his mixed strategy. Hence, the proof is complete. $\square$

To see which candidate has stronger incentives to invest we now need to calculate the expected investment of each. We resort to the quadratic form employed before, where $C(x) = \frac{c}{2}x^2$. Candidate F's expected investment in the campaign is

$$
\begin{aligned}
Ex_F^2 &= \int_{S_F} x\,C'(x+d^2)dx = \int_{S_F} xc(x+d^2)dx = \left[\frac{c}{3}x^3 + \frac{cd^2}{2}x^2\right]_0^{\bar{x}-d^2} = \frac{c}{3}(\bar{x}-d^2)^3 + \frac{cd^2}{2}(\bar{x}-d^2)^2 \\
&= \frac{c}{3}\left(\sqrt{\frac{2}{c}}-d^2\right)^3 + \frac{cd^2}{2}\left(\sqrt{\frac{2}{c}}-d^2\right)^2 \\
&= \sqrt{\frac{8}{9c}} - d^2 + \frac{c}{6}(d^2)^3
\end{aligned}
$$

Similarly, for candidate T we get

$$
\begin{aligned}
Ex_T^2 &= \int_{S_T} x\,C'(x-d^2)dx = \int_{S_T} xc(x-d^2)dx = \left[\frac{c}{3}x^3 - \frac{cd^2}{2}x^2\right]_{d^2}^{\bar{x}} \\
&= \frac{c}{3}(\bar{x})^3 - \frac{cd^2}{2}(\bar{x})^2 - \frac{c}{3}(d^2)^3 + \frac{c}{2}(d^2)^3 \\
&= \sqrt{\frac{8}{9c}} - d^2 + \frac{c}{6}(d^2)^3
\end{aligned}
$$

Hence, the expected difference in investments is zero:

Corollary 5. *The difference in expected investments in stage 2 is equal to zero.*

This resembles our findings from a campaign with exogenous noise.

Now look at stage 1. We focus on pure strategy equilibria. Assume $a \geq 0$ and $x_T^1 = 0$. F chooses investment to maximize

$$
1 - C(\bar{x}-d^2) - C(x_F^1) = 1 - C(\bar{x}-a-x_F^1) - C(x_F^1).
$$

His first order condition reads

$$
C'(\bar{x}-a-x_F^1) - C'(x_F^1) = 0 \Leftrightarrow \bar{x}-a = 2x_F^1.
$$

The equivalence in the last step follows from the strict convexity of $C(.)$, due to which $C'(.)$ is strictly

increasing. Because the second derivative is strictly negative,

$$-C''(\bar{x} - a - x_F^1) - C''(x_F^1) < 0,$$

we can use the first order condition to find the global maximum. F's optimal reaction to $x_T^1 = 0$ is

$$x_F^* = \frac{\bar{x} - a}{2}.$$

(Note that for this to make sense we need that $|a| < \bar{x}$. Otherwise no candidate has any incentive to invest in their campaign.) Now turn to T. He either spends zero and gets zero in expectation or maximizes

$$1 - C\left(\frac{3}{2}\bar{x} + \frac{1}{2}a - x_T^1\right) - C(x_T^1).$$

The first order condition reads

$$C'(\frac{3}{2}\bar{x} + \frac{1}{2}a - \tilde{x}_T^1) = C'(\tilde{x}_T^1) \Leftrightarrow 2\tilde{x}_T = \frac{3}{2}\bar{x} + \frac{1}{2}a,$$

and hence

$$\tilde{x}_T = \frac{3}{4}\bar{x} + \frac{a}{4}.$$

The second order condition holds, but note that we only looked at positive investment. For this to be optimal it needs to hold in addition that expected utility is weakly positive:

$$1 - 2C\left(\frac{3}{4}\bar{x} + \frac{a}{4}\right) \geq 0.$$

This inequality could in principle hold for small a, depending on the curvature of the costs function and when a is relatively small. For concreteness, however, we stick to the functional employed mostly in the paper, $C(x) = \frac{c}{2}x^2$. Then

$$1 - 2C\left(\frac{3}{4}\bar{x} + \frac{a}{4}\right) = 1 - c\left(\frac{3}{4}\sqrt{\frac{2}{c}} + \frac{a}{4}\right)^2 < 0.$$

To see this note that the second term is strictly increasing in a. Hence, if utility is negative for $a = 0$
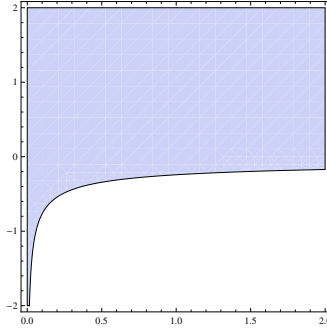
Figure 9: Combinations of c (horizontal) and a (vertical) admitting a pure strategy Nash equilibrium in which F spends more.

this is true also for all $a > 0$. Using $a = 0$ we get

$$1 - c\left(\frac{3}{4}\sqrt{\frac{2}{c}} + \frac{a}{4}\right)^2 = 1 - c\left(\frac{3}{4}\sqrt{\frac{2}{c}}\right)^2 = 1 - \frac{9}{8} = -\frac{1}{8} < 0.$$

Hence, $\tilde{x}_T$ cannot be an optimal choice and T chooses zero investment instead. Therefore, we established that this is an equilibrium. Moreover, note that as long as the utility from spending $\tilde{x}_T^1$ is negative, $(x_F^*, 0)$ is an equilibrium. This is the case for all

$$a \geq \frac{4 - 3\sqrt{2}}{\sqrt{c}} < 0.$$

Hence, as before when $E[(\epsilon)^2] \neq 0$, it may also be the case that the trailing candidate spends more. In Figure 9 we show combinations of c and a for which there exists a pure strategy Nash equilibrium in which F spends $x_F^1 > 0$ while T stays passive.

Note that by symmetry the whole analysis also applies for T if $a \leq 0$. Hence, there always exists a pure strategy Nash equilibrium in stage 1 if we assume costs to be quadratic, and there are multiple pure strategy equilibria in close enough games:

Proposition 12. *In stage 1 of the competition with quadratic costs there always exists a pure strategy Nash equilibrium. In a close race there may also exist multiple equilibria, and either candidate may adopt a tough stance and spend more. If one candidate has a relatively large advantage, this candidate will also choose larger investment.*

Proof. This follows immediately from the discussion above. $\square$

57

It follows from Corollary 5 and Proposition 12 that the campaign game qualitatively perfectly resembles what we have seen before in the case of a noisy competition. This proves the example.

H   Proof of Proposition 9

Proof. To show existence of equilibrium we can proceed as before. To save on space we do not show the proof (available upon request) but note that, as before, if c is large enough in any stage the candidates' problems are strictly concave with continuous payoff functions and compact strategy spaces $[0, \sqrt{\frac{2}{c}}]$, and hence equilibrium existence follows from standard proofs.

First we replicate the result of the benchmark model without polls. Candidates maximize

$$\int_{-\infty}^{\infty} \left(\lambda \Phi \left(\frac{a + \Delta^1 + \Delta^2}{\sqrt{2}\sigma_\epsilon} \right) + (1-\lambda)E_{\epsilon^1,\epsilon^2} \left[1 - P(a + \Delta^1 + \Delta^2 - \epsilon^1 - \epsilon^2) \right] \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da - \frac{c}{2}(x_F^1)^2 - \frac{c}{2}(x_F^2)^2$$

$$\int_{-\infty}^{\infty} \left(\lambda (1 - \Phi \left(\frac{a + \Delta^1 + \Delta^2}{\sqrt{2}\sigma_\epsilon} \right)) + (1-\lambda)E_{\epsilon^1,\epsilon^2} \left[P(a + \Delta^1 + \Delta^2 - \epsilon^1 - \epsilon^2) \right] \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da - \frac{c}{2}(x_T^1)^2 - \frac{c}{2}(x_T^2)^2$$

respectively. Taking first derivatives with respect to spending in t leads to

$$\int_{-\infty}^{\infty} \left(\frac{\lambda}{\sqrt{2}\sigma_\epsilon} \phi \left(\frac{a + \Delta^1 + \Delta^2}{\sqrt{2}\sigma_\epsilon} \right) + (1-\lambda)E_{\epsilon^1,\epsilon^2} \left[p(a + \Delta^1 + \Delta^2 - \epsilon^1 - \epsilon^2) \right] \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da - cx_F^t \overset{!}{=} 0$$

$$\int_{-\infty}^{\infty} \left(\frac{\lambda}{\sqrt{2}\sigma_\epsilon} \phi \left(\frac{a + \Delta^1 + \Delta^2}{\sqrt{2}\sigma_\epsilon} \right) + (1-\lambda)E_{\epsilon^1,\epsilon^2} \left[p(a + \Delta^1 + \Delta^2 - \epsilon^1 - \epsilon^2) \right] \right) \frac{1}{\sigma_\alpha} \phi \left(\frac{a - \mu_\alpha}{\sigma_\alpha} \right) da - cx_T^t \overset{!}{=} 0$$

It is apparent that in any interior equilibrium both choose identical investments in both stages. Moreover, since investments perfectly cancel out, the equilibrium must be unique.

Now consider the situation with polls. For the stage 1 analysis we need an expression of the second stage investment. Noticing that the noise structure changes since only one random term is remaining, and that candidates now know d^2 precisely, from the FOCs is follows that

$$x_F^2 = x_T^2 = x^2 = \frac{\frac{\lambda}{\sigma_\epsilon} \phi \left(\frac{d^2}{\sigma_\epsilon} \right) + (1-\lambda)E_{\epsilon^2} \left[p(d^2 - \epsilon^2) \right]}{c}.$$

Making use of Assumption 3, we find

$$x^2 = \frac{\frac{\lambda}{\sigma_\epsilon}\phi\left(\frac{d^2}{\sigma_\epsilon}\right) + \frac{(1-\lambda)}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\frac{\phi\left(\frac{-\mu_p+d^2}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)+\phi\left(\frac{\mu_p+d^2}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)}{2}}{c}.$$

Differentiation with respect to d^2 gives

$$\frac{\partial x^2}{\partial d^2} = \frac{\frac{\lambda}{\sigma_\epsilon^2}\phi'\left(\frac{d^2}{\sigma_\epsilon}\right) + \frac{(1-\lambda)}{\sigma_\epsilon^2+\sigma_p^2}\frac{\phi'\left(\frac{-\mu_p+d^2}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)+\phi'\left(\frac{\mu_p+d^2}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)}{2}}{c}.$$

At $d^2 = 0$ the derivative is clearly zero. We hence take the second derivative to see the curvature:

$$\frac{\partial^2 x^2}{\partial (d^2)^2} = \frac{\frac{\lambda}{\sigma_\epsilon^3}\phi''\left(\frac{d^2}{\sigma_\epsilon}\right) + \frac{(1-\lambda)}{(\sigma_\epsilon^2+\sigma_p^2)^{3/2}}\frac{\phi''\left(\frac{-\mu_p+d^2}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)+\phi''\left(\frac{\mu_p+d^2}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)}{2}}{c}.$$

At $d^2 = 0$ we get

$$\frac{\partial^2 x^2}{\partial (d^2)^2}\Big|_{d^2=0} = \frac{\frac{\lambda}{\sigma_\epsilon^3}\phi''(0) + \frac{(1-\lambda)}{(\sigma_\epsilon^2+\sigma_p^2)^{3/2}}\frac{\phi''\left(\frac{-\mu_p}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)+\phi''\left(\frac{\mu_p}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)}{2}}{c}$$

$$= \frac{\frac{\lambda}{\sigma_\epsilon^3}\phi''(0) + \frac{(1-\lambda)}{(\sigma_\epsilon^2+\sigma_p^2)^{3/2}}\phi''\left(\frac{\mu_p}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)}{c}$$

The first term is strictly negative, implying the derivative $\frac{\partial x^2}{\partial d^2}$ decreases and so does investment as we increase $|d^2|$ from zero. Hence look at the second term. It is straightforward to verify that

$$Sign\left[\phi''\left(\frac{\mu_p}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)\right] = Sign\left[\mu_p - \sqrt{\sigma_\epsilon^2 + \sigma_p^2}\right].$$

It follows that $\mu_p > \sqrt{\sigma_\epsilon^2 + \sigma_p^2}$ is necessary for investment to be increasing in d^2. It is, however, not sufficient. It is clear that when λ becomes large, the first term dominates and the sign of the second derivative must be positive. However, if λ is small, that is if the plurality premium is very important,

the second term dominates.

It is apparent from the discussion that for too small μ_p incentives are as under FPTP. However, if $\mu_p > \sqrt{\sigma_\epsilon^2 + \sigma_p^2}$ stage 2 investments increase over some range in $|d^2|$. It is easily verified that when $\mu_p = \overline{\mu}_p = \sqrt{3}\sqrt{\sigma_\epsilon^2 + \sigma_p^2}$ the second derivative of second stage investment with respect to d^2 has its maximum. It is easily shown that comparative statics $\partial x_F^1/\partial a$ at $a = 0$, in case a unique equilibrium exists, are negative if $\mu_p = \overline{\mu}_p$ and if λ is small (available upon request). Hence, in this region there is always anti-momentum in close games. $\qquad\square$

I   Proof of Corollary 4

Proof. We know already from before that the comparative statics in stage 2 drive stage 1 behavior. To prove the corollary it is hence sufficient to show that second stage investment is weakly decreasing in $|d^2|$. Note that there are two parts that determine the slope of investment, the marginal winning probability and the marginal share. The first is always decreasing. Hence, if the second is also decreasing the result holds. It is easily shown that for $\mu_p \leq \sqrt{\sigma_\epsilon^2 + \sigma_p^2}$ this must always be true. Note that this does not mean the electorate is not divided. There are still two peaks of the preference distribution, but in expectation the distribution in stage 3 becomes single peaked. Reasoning from stage 1 the condition modifies to $\mu_p \leq \sqrt{2\sigma_\epsilon^2 + \sigma_p^2}$, because the variance due to noise exactly doubles. That this is sufficient for momentum follows basically from Proposition 7. If this condition does not hold comparative statics may be different and anti-momentum might exist. However, there always exist a $\lambda < 1$ for which there is strict momentum. To see this differentiate the investment in the second stage with respect to d^2:

$$\frac{\partial x^2}{\partial d^2} = \frac{\frac{\lambda}{\sigma_\epsilon^2}\phi'\left(\frac{d^2}{\sigma_\epsilon}\right) + \frac{(1-\lambda)}{\sigma_\epsilon^2+\sigma_p^2}\frac{\phi'\left(\frac{-\mu_p+d^2}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)+\phi'\left(\frac{\mu_p+d^2}{\sqrt{\sigma_\epsilon^2+\sigma_p^2}}\right)}{2}}{c}.$$

As discussed above the first term is strictly negative, but the second may be positive as well. However, since all functions are bounded for all d^2 there exist a $\lambda < 1$ for which this is strictly negative, also in expectation. $\qquad\square$

J Proof of Proposition 10

Proof. We start solving the game from the end. Expected utility of F and T before exerting effort in stage 2 equals

$$
\begin{aligned}
\pi_F^2 &= \Phi\left(\frac{\tilde{\tilde{\mu}}_\alpha + x_F^2 - x_T^2}{\tilde{\tilde{\sigma}}_\alpha}\right) - \frac{c}{2}(x_F^2)^2 \\
\pi_T^2 &= 1 - \Phi\left(\frac{\tilde{\tilde{\mu}}_\alpha + x_F^2 - x_T^2}{\tilde{\tilde{\sigma}}_\alpha}\right) - \frac{c}{2}(x_T^2)^2
\end{aligned}
$$

From a similar reasoning as before we can show that equilibrium efforts are identical for both and equal to

$$
x^{**} = \frac{\phi\left(\frac{\tilde{\tilde{\mu}}_\alpha}{\tilde{\tilde{\sigma}}_\alpha}\right)}{c}.
$$

Effort depends on the belief about the relative standing, which was influenced by the earlier polls as well as the prior belief. Expected utility in equilibrium is

$$
\begin{aligned}
\overline{\pi}_F^2 &= \Phi\left(\frac{\tilde{\tilde{\mu}}_\alpha}{\tilde{\tilde{\sigma}}_\alpha}\right) - \frac{\left(\phi\left(\frac{\tilde{\tilde{\mu}}_\alpha}{\tilde{\tilde{\sigma}}_\alpha}\right)\right)^2}{2c} \\
\overline{\pi}_T^2 &= 1 - \Phi\left(\frac{\tilde{\tilde{\mu}}_\alpha}{\tilde{\tilde{\sigma}}_\alpha}\right) - \frac{\left(\phi\left(\frac{\tilde{\tilde{\mu}}_\alpha}{\tilde{\tilde{\sigma}}_\alpha}\right)\right)^2}{2c}
\end{aligned}
$$

Substituting

$$
\tilde{\tilde{\mu}}_\alpha = \frac{\tilde{\mu}_\alpha \sigma_\eta^2 + d^2 \tilde{\sigma}_\alpha^2}{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}
$$

and

$$
\tilde{\tilde{\sigma}}_\alpha = \frac{\sigma_\eta^2 \tilde{\sigma}_\alpha^2}{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}
$$

and taking the expectation over d^2 yields the expectation over stage 2 expected utility from the point of view of stage 1:[19]

$$
\begin{aligned}
\overline{\overline{\pi}}_F^2 &= \int_{-\infty}^{\infty} \left(\Phi\left(\frac{\tilde{\mu}_\alpha \sigma_\eta^2 + d^2 \tilde{\sigma}_\alpha^2}{\sigma_\eta^2 \tilde{\sigma}_\alpha^2}\right) - \frac{\left(\phi\left(\frac{\tilde{\mu}_\alpha \sigma_\eta^2 + d^2 \tilde{\sigma}_\alpha^2}{\sigma_\eta^2 \tilde{\sigma}_\alpha^2}\right)\right)^2}{2c} \right) \phi\left(\frac{\tilde{\mu}_\alpha + x_F^1 - x_T^1 - d^2}{\tilde{\sigma}_\alpha}\right) dp^2 \\
\overline{\overline{\pi}}_T^2 &= \int_{-\infty}^{\infty} \left(1 - \Phi\left(\frac{\tilde{\mu}_\alpha \sigma_\eta^2 + d^2 \tilde{\sigma}_\alpha^2}{\sigma_\eta^2 \tilde{\sigma}_\alpha^2}\right) - \frac{\left(\phi\left(\frac{\tilde{\mu}_\alpha \sigma_\eta^2 + d^2 \tilde{\sigma}_\alpha^2}{\sigma_\eta^2 \tilde{\sigma}_\alpha^2}\right)\right)^2}{2c} \right) \phi\left(\frac{\tilde{\mu}_\alpha + x_F^1 - x_T^1 - d^2}{\tilde{\sigma}_\alpha}\right) dp^2
\end{aligned}
$$

---

[19]As in the baseline model we assume that investment is not observable to the candidates and it is thus not used to update the prior beliefs. On the other hand, the candidates do anticipate the effect of their investments on the distribution of the poll signal. We chose to present this specification because it encompasses the no-poll benchmark and the All-Pay auction case as the limiting cases. Our results are robust to assuming that investments are observable and are thus used by the candidates to update their prior beliefs. An appendix with details is available from the authors upon request.

This can be shown to be

$$
\begin{aligned}
\pi_F^1 &= \Phi\left(\frac{\tilde{\mu}_\alpha\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)+\tilde{\sigma}_\alpha^2(x_F^1-x_T^1)}{\tilde{\sigma}_\alpha\sigma_\eta\sqrt{2\tilde{\sigma}_\alpha^2+\sigma_\eta^2}}\right) - \frac{1}{2} - \frac{e^{-\frac{\frac{\tilde{\mu}_\alpha^2\left(\tilde{\sigma}_\alpha^4+\tilde{\sigma}_\alpha^2\sigma_\eta^2+2\sigma_\eta^4\right)}{\tilde{\sigma}_\alpha^2\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)}+2\tilde{\mu}_\alpha(x_F^1-x_T^1)+(x_F^1-x_T^1)^2}{2\sigma_\eta^2}}}{4\pi c\tilde{\sigma}_\alpha^2\sigma_\eta^2} \\
&\quad \times \left(\frac{\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)^{3/2}e^{\frac{\sigma_\eta^2\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)\left(\frac{\tilde{\mu}_\alpha+x_F^2-x_T^1}{\sigma_\eta^2}-\frac{2\tilde{\mu}_\alpha}{\tilde{\sigma}_\alpha^2+\sigma_\eta^2}\right)^2}{2\left(3\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)}}}{\sqrt{3\tilde{\sigma}_\alpha^2+\sigma_\eta^2}} - 2\pi c\tilde{\sigma}_\alpha^2\sigma_\eta^2 e^{\frac{\frac{\tilde{\mu}_\alpha^2\left(\tilde{\sigma}_\alpha^4+\tilde{\sigma}_\alpha^2\sigma_\eta^2+2\sigma_\eta^4\right)}{\tilde{\sigma}_\alpha^2\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)}+2\tilde{\mu}_\alpha(x_F^1-x_T^1)+(x_F^1-x_T^1)^2}{2\sigma_\eta^2}}\right) \\
&\quad - \frac{c}{2}(x_F^1)^2 \\[2em]
\pi_T^1 &= -\Phi\left(\frac{\tilde{\mu}_\alpha\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)+\tilde{\sigma}_\alpha^2(x_F^1-x_T^1)}{\tilde{\sigma}_\alpha\sigma_\eta\sqrt{2\tilde{\sigma}_\alpha^2+\sigma_\eta^2}}\right) - \frac{e^{-\frac{\frac{\tilde{\mu}_\alpha^2\left(\tilde{\sigma}_\alpha^4+\tilde{\sigma}_\alpha^2\sigma_\eta^2+2\sigma_\eta^4\right)}{\tilde{\sigma}_\alpha^2\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)}+2\tilde{\mu}_\alpha(x_F^1-x_T^1)+(x_F^1-x_T^1)^2}{2\sigma_\eta^2}}}{4\pi c\tilde{\sigma}_\alpha^2\sigma_\eta^2} \\
&\quad \times \left(\frac{\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)^{3/2}e^{\frac{\sigma_\eta^2\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)\left(\frac{\tilde{\mu}_\alpha+x_F^1-x_T^1}{\sigma_\eta^2}-\frac{2\tilde{\mu}_\alpha}{\tilde{\sigma}_\alpha^2+\sigma_\eta^2}\right)^2}{2\left(3\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)}}}{\sqrt{3\tilde{\sigma}_\alpha^2+\sigma_\eta^2}} - 4\pi c\tilde{\sigma}_\alpha^2\sigma_\eta^2 e^{\frac{\frac{\tilde{\mu}_\alpha^2\left(\tilde{\sigma}_\alpha^4+\tilde{\sigma}_\alpha^2\sigma_\eta^2+2\sigma_\eta^4\right)}{\tilde{\sigma}_\alpha^2\left(\tilde{\sigma}_\alpha^2+\sigma_\eta^2\right)}+2\tilde{\mu}_\alpha(x_F^1-x_T^1)+(x_F^1-x_T^1)^2}{2\sigma_\eta^2}}\right) \\
&\quad - \frac{c}{2}(x_T^1)^2
\end{aligned}
$$

The first order conditions evaluated at the symmetric equilibrium at $\tilde{\mu}_\alpha = 0$ are

$$
\begin{aligned}
\frac{\partial \pi_F^1}{\partial x_F^1}\Big|_{\tilde{\mu}_\alpha=0} &= \frac{\tilde{\sigma}_\alpha}{\sqrt{2\pi}\sigma_\eta\sqrt{2\tilde{\sigma}_\alpha^2+\sigma_\eta^2}} - cx_F^1 \overset{!}{=} 0 \\
\frac{\partial \pi_T^1}{\partial x_T^1}\Big|_{\tilde{\mu}_\alpha=0} &= \frac{\tilde{\sigma}_\alpha}{\sqrt{2\pi}\sigma_\eta\sqrt{2\tilde{\sigma}_\alpha^2+\sigma_\eta^2}} - cx_T^1 \overset{!}{=} 0
\end{aligned}
$$

and equilibrium efforts at $\tilde{\mu}_\alpha = 0$ are thus

$$
x_F^1 = x_T^1 = \frac{\tilde{\sigma}_\alpha}{\sqrt{2\pi}\sigma_\eta\sqrt{2\tilde{\sigma}_\alpha^2+\sigma_\eta^2}c}.
$$

It is easily shown that iff the slope at the symmetric equilibrium at $\tilde{\mu}_\alpha = 0$ is less than one there is a unique equilibrium and there are multiple equilibria else.

As before, for comparative statics we totally differentiate the candidates' FOCs (which are omitted

here; they can be easily reproduced for example with Mathematica):

$$\left(\frac{\partial^2 \pi_F^1}{\partial(x_F^1)^2}\right) dx_F^1 + \left(\frac{\partial^2 \pi_F^1}{\partial x_F^1 \partial x_T^1}\right) dx_T^1 = -\left(\frac{\partial^2 \pi_F^1}{\partial x_F^1 \partial \tilde{\mu}_\alpha}\right) d\tilde{\mu}_\alpha$$

$$\left(\frac{\partial^2 \pi_T^1}{\partial x_F^1 \partial x_T^1}\right) dx_F^1 + \left(\frac{\partial^2 \pi_T^1}{\partial(x_T^1)^2}\right) dx_T^1 = -\left(\frac{\partial^2 \pi_T^1}{\partial x_T^1 \partial \tilde{\mu}_\alpha}\right) d\tilde{\mu}_\alpha$$

We express the total differential in matrix notation. Defining

$$\mathbf{M} = \begin{pmatrix} \frac{\partial^2 \pi_F^1}{\partial(x_F^1)^2} & \frac{\partial^2 \pi_F^1}{\partial x_F^1 \partial x_T^1} \\ \frac{\partial^2 \pi_T^1}{\partial x_F^1 \partial x_T^1} & \frac{\partial^2 \pi_T^1}{\partial(x_T^1)^2} \end{pmatrix},$$

where

$$\frac{\partial^2 \pi_F^1}{\partial(x_F^1)^2} = \frac{\left(\sqrt{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2} - 6c^2\pi\sigma_\eta^4\sqrt{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right)\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\left(\sqrt{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2} - 2c^2\pi\sigma_\eta^4\sqrt{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right)}{2c\pi\sigma_\eta^4\left(3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\right)^{3/2}},$$

$$\frac{\partial^2 \pi_F^1}{\partial x_F^1 \partial x_T^1} = -\frac{\left(\frac{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right)^{3/2}}{2c\pi\sigma_\eta^4},$$

$$\frac{\partial^2 \pi_T^1}{\partial x_F^1 \partial x_T^1} = -\frac{\left(\frac{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right)^{3/2}}{2c\pi\sigma_\eta^4},$$

$$\frac{\partial^2 \pi_T^1}{\partial(x_T^1)^2} = \frac{\left(\sqrt{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2} - 6c^2\pi\sigma_\eta^4\sqrt{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right)\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\left(\sqrt{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2} - 2c^2\pi\sigma_\eta^4\sqrt{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right)}{2c\pi\sigma_\eta^4\left(3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\right)^{3/2}},$$

all evaluated at the symmetric equilibrium at $\tilde{\mu}_\alpha = 0$, and letting

$$\mathbf{P} = \begin{pmatrix} -\frac{\partial^2 \pi_F^1}{\partial x_F^1 \partial \tilde{\mu}_\alpha} \\ -\frac{\partial^2 \pi_T^1}{\partial x_T^1 \partial \tilde{\mu}_\alpha} \end{pmatrix} = \begin{pmatrix} -\frac{\left(\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\right)^{5/2}}{2\pi c\tilde{\sigma}_\alpha^2\sigma_\eta^4\left(3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\right)^{3/2}} \\ \frac{\left(\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\right)^{5/2}}{2\pi c\tilde{\sigma}_\alpha^2\sigma_\eta^4\left(3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\right)^{3/2}} \end{pmatrix}$$

$$\mathbf{dx} = \begin{pmatrix} dx_F^1 \\ dx_T^1 \end{pmatrix}$$

we rewrite the total differential as

$$\mathbf{M} \cdot \mathbf{dx} = \mathbf{P}d\tilde{\mu}_\alpha$$

To apply Cramer's rule we substitute $\mathbf{P}$ in the respective columns of $\mathbf{M}$ as before. The derivative we

are searching for is

$$\frac{\partial x_F^1}{\partial \tilde{\mu}_\alpha}|_{\tilde{\mu}_\alpha=0} = \frac{\left(\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\right)^{5/2}}{2\tilde{\sigma}_\alpha^2 \left(\pi c^2 \sigma_\eta^6 \sqrt{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2} - \tilde{\sigma}_\alpha^2 \left(\sqrt{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2} - 3\pi c^2 \sigma_\eta^4 \sqrt{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right) - \sigma_\eta^2 \sqrt{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right)}$$

$$\frac{\partial x_T^1}{\partial \tilde{\mu}_\alpha}|_{\tilde{\mu}_\alpha=0} = \frac{-\left(\tilde{\sigma}_\alpha^2 + \sigma_\eta^2\right)^{5/2}}{2\tilde{\sigma}_\alpha^2 \left(\pi c^2 \sigma_\eta^6 \sqrt{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2} - \tilde{\sigma}_\alpha^2 \left(\sqrt{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2} - 3\pi c^2 \sigma_\eta^4 \sqrt{3\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right) - \sigma_\eta^2 \sqrt{\tilde{\sigma}_\alpha^2 + \sigma_\eta^2}\right)}$$

or, if we substitute $\tilde{\sigma}_\alpha = \sqrt{\frac{\sigma_\eta^2 \sigma_\alpha^2}{\sigma_\eta^2 + \sigma_\alpha^2}}$,

$$\frac{\partial x_F^1}{\partial \tilde{\mu}_\alpha}|_{\tilde{\mu}_\alpha=0} = \frac{\left(2\sigma_\alpha^2 + \sigma_\eta^2\right)^{5/2}}{\sigma_\alpha^2 \left(\sigma_\alpha^2 + \sigma_\eta^2\right)^{3/2} \left(\pi c^2 \sigma_\eta^4 \sqrt{\frac{3\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^2} + 1} - \sigma_\alpha^2 \left(\sqrt{\frac{2\sigma_\alpha^2 + \sigma_\eta^2}{\left(\sigma_\alpha^2 + \sigma_\eta^2\right)^3}} - 3\pi c^2 \sigma_\eta^4 \sqrt{\frac{4\sigma_\alpha^2 + \sigma_\eta^2}{\left(\sigma_\alpha^2 + \sigma_\eta^2\right)^3}}\right) - \sqrt{\frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^2} + 1}\right)}$$

$$\frac{\partial x_T^1}{\partial \tilde{\mu}_\alpha}|_{\tilde{\mu}_\alpha=0} = \frac{-\left(2\sigma_\alpha^2 + \sigma_\eta^2\right)^{5/2}}{\sigma_\alpha^2 \left(\sigma_\alpha^2 + \sigma_\eta^2\right)^{3/2} \left(\pi c^2 \sigma_\eta^4 \sqrt{\frac{3\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^2} + 1} - \sigma_\alpha^2 \left(\sqrt{\frac{2\sigma_\alpha^2 + \sigma_\eta^2}{\left(\sigma_\alpha^2 + \sigma_\eta^2\right)^3}} - 3\pi c^2 \sigma_\eta^4 \sqrt{\frac{4\sigma_\alpha^2 + \sigma_\eta^2}{\left(\sigma_\alpha^2 + \sigma_\eta^2\right)^3}}\right) - \sqrt{\frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^2} + 1}\right)}$$

This derivative is strictly positive for F (and thus also strictly negative for T) iff

$$c > \sqrt{\frac{\sigma_\alpha^2 \sqrt{\frac{2\sigma_\alpha^2 + \sigma_\eta^2}{\left(\sigma_\alpha^2 + \sigma_\eta^2\right)^3}} + \sqrt{\frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^2} + 1}}{3\pi \sigma_\alpha^2 \sigma_\eta^4 \sqrt{\frac{4\sigma_\alpha^2 + \sigma_\eta^2}{\left(\sigma_\alpha^2 + \sigma_\eta^2\right)^3}} + \pi \sigma_\eta^4 \sqrt{\frac{3\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^2} + 1}}}$$

and hence if this condition is met momentum exists. For example, if $\sigma_\eta = \sigma_\alpha = 1$ we need $c > \frac{\left(\frac{3}{5}\right)^{3/4}}{\sqrt{\pi}} \approx$ 0.3846. This parallels our previous finding because for large enough c a unique equilibrium exists and there is momentum.

Finally, the derivative of $\frac{\partial x_F^1}{\partial \tilde{\mu}_\alpha}$ with respect to poll (im)precision σ_η at $\tilde{\mu}_\alpha = 0$ is

$$\frac{-2\sigma_\eta \left(2\sigma_\alpha^2 + \sigma_\eta^2\right)^{3/2}}{\sigma_\alpha^2 \left(\sigma_\alpha^2 + \sigma_\eta^2\right)^3 \left(\sqrt{\frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^2} + 1} - \pi c^2 \sigma_\eta^4 \sqrt{\frac{3\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\eta^2} + 1} + \sigma_\alpha^2 \left(\sqrt{\frac{2\sigma_\alpha^2 + \sigma_\eta^2}{\left(\sigma_\alpha^2 + \sigma_\eta^2\right)^3}} - 3\pi c^2 \sigma_\eta^4 \sqrt{\frac{4\sigma_\alpha^2 + \sigma_\eta^2}{\left(\sigma_\alpha^2 + \sigma_\eta^2\right)^3}}\right)\right)^2}$$
$$\times \quad [16\pi c^2 \sigma_\alpha^4 \sigma_\eta^2 \sqrt{4\sigma_\alpha^2 + \sigma_\eta^2} + \sigma_\alpha^2 \left(2\sqrt{2\sigma_\alpha^2 + \sigma_\eta^2} + 5\pi c^2 \sigma_\eta^4 \sqrt{4\sigma_\alpha^2 + \sigma_\eta^2}\right) + \sigma_\eta^2 \left(\sqrt{2\sigma_\alpha^2 + \sigma_\eta^2} + \pi c^2 \sigma_\eta^4 \sqrt{4\sigma_\alpha^2 + \sigma_\eta^2}\right)] < 0$$

and is strictly negative. Hence, when a unique equilibrium with momentum exists, as the poll becomes more imprecise, momentum decreases.
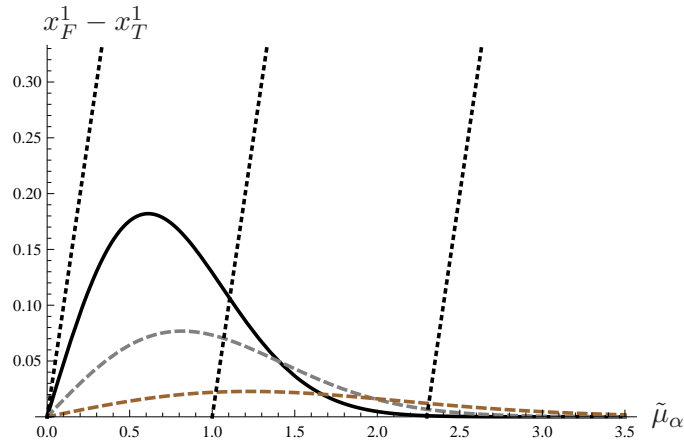
Figure 10: Difference in spending for $c = 2$ and $\sigma_\epsilon \in \{1/2, 3/4, 1\}$. The dotted straight lines represent $\tilde{\mu}_\alpha \in \{0, 1, 5/2\}$ and intersections with the spending difference curve mark the respective equilibria. We see that at $\tilde{\mu}_\alpha = 0$ the spending difference is zero independent of σ_ϵ. As we increase $\tilde{\mu}_\alpha$ the spending difference increases in both $\tilde{\mu}_\alpha$ and σ_ϵ. However, as we continue increasing $\tilde{\mu}_\alpha$ this changes and for example at $\tilde{\mu}_\alpha = 2.3$ momentum is largest when precision is smallest.

If $\tilde{\mu}_\alpha$ becomes large in absolute value, momentum may decrease. This is easiest seen by considering the All-Pay action as a polar case of this model, featuring perfect poll precision. For $\mid \tilde{\mu}_\alpha \mid > \sqrt{\frac{2}{c}}$ both candidates expend zero and thus momentum vanishes. Figure 10 illustrates how poll precision is related to momentum using a parameterized version of the model. That proves the proposition. $\qquad \square$

References

ALI, S. N., AND N. KARTIK (2012): "Herding with Collective Preferences," *Economic Theory*, 51(3), 601–626.

AOYAGI, M. (2010): "Information Feedback in a Dynamic Tournament," *Games and Economic Behavior*, 70(2), 242 – 260.

BAYE, M. R., D. KOVENOCK, AND C. G. DE VRIES (1993): "Rigging the Lobbying Process: An Application of the All-Pay Auction," *American Economic Review*, 83(1), 289–94.

BERG, J. E., F. D. NELSON, AND T. A. RIETZ (2008): "Prediction Market Accuracy in the Long Run," *International Journal of Forecasting*, 24(2), 285–300.

BERNHARDT, D., J. DUGGAN, AND F. SQUINTANI (2009): "Private Polling in Elections and Voter Welfare," *Journal of Economic Theory*, 144(5), 2021–2056.

BIKHCHANDANI, S., D. HIRSHLEIFER, AND I. WELCH (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100(5), 992–1026.

BLAIS, A., E. GIDENGIL, AND N. NEVITTE (2006): "Do Polls Influence the Vote?," in *Capturing campaign effects*, ed. by H. E. Brady, and R. Johnston, p. 263–279. Ann Arbor: University of Michigan Press.

BRAMS, S. J., AND M. D. DAVIS (1973): "Resource-Allocation Models in Presidential Campaigning: Implications for Democratic Representation," *Annals of the New York Academy of Sciences*, 219, 105–123.

BRAMS, S. J., AND M. D. DAVIS (1974): "The 3/2's Rule in Presidential Campaigning," *The American Political Science Review*, 68(1), pp. 113–134.

BUFFERY, V. (2011): "Majority of French want to Drop Nuclear Energy-Poll," *Reuters*, April 13.

BULOW, J. I., J. D. GEANAKOPLOS, AND P. D. KLEMPERER (1985): "Multimarket Oligopoly: Strategic Substitutes and Complements," *Journal of Political Economy*, 93(3), 488–511.

CALLANDER, S. (2007): "Bandwagons and Momentum in Sequential Voting," *Review of Economic Studies*, 74(3), 653–684.

CASELLA, G., AND R. BERGER (2002): *Statistical Inference*. Duxbury Thomson Learning.

COOPER, M., AND D. SUSSMAN (2011): "Nuclear Power Loses Support in New Poll," *The New York Times*, March 22.

DALTON, R. J., AND M. P. WATTENBERG (2001): *Parties Without Partisans: Political Change in Advanced Industrial Democracies*. New York, Oxford University Press.

DEMPSEY, J. (2011): "Merkel Loses Key German State on Nuclear Fears," *The New York Times*, March 27.

DENTER, P. (2013): "A Theory of Communication in Political Campaigns," Discussion Paper 1302, University of St. Gallen, School of Economics and Political Science.

DENTER, P., J. MORGAN, AND D. SISAK (2011): ""Where ignorance is bliss, 'tis folly to be wise": Transparency in Contests," Discussion Paper 2011-28, University of St. Gallen, School of Economics and Political Science.

DRAZEN, A. (2001): "The Political Business Cycle After 25 Years," in *NBER Macroeconomics Annual 2000, Volume 15*, NBER Chapters, pp. 75–138. National Bureau of Economic Research, Inc.

EDERER, F. (2010): "Feedback and Motivation in Dynamic Tournaments," *Journal of Economics & Management Strategy*, 19(3), 733–769.

ERIKSON, R. S. (1971): "The Advantage of Incumbency in Congressional Elections," *Polity*, 3(3), pp. 395–405.

ERIKSON, R. S., AND T. R. PALFREY (2000): "Equilibria In Campaign Spending Games: Theory And Data," *American Political Science Review*, 94, 595–609.

FERSHTMAN, C., AND U. GNEEZY (2011): "The Trade-off between Performance and Quitting in High-Power Tournaments," *Journal of the European Economic Association*, 9(2), 318–336.

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. The MIT Press.

GASSEBNER, M., R. JONG-A-PIN, AND J. O. MIERAU (2008): "Terrorism and Electoral Accountability: One Strike, You're Out!," *Economics Letters*, 100(1), 126–129.

GELMAN, A., AND G. KING (1990): "Estimating Incumbency Advantage without Bias," *American Journal of Political Science*, 34(4), pp. 1142–1164.

GERSHKOV, A., AND M. PERRY (2009): "Tournaments with Midterm Reviews," *Games and Economic Behavior*, 66(1), 162–190.

GOEREE, J., AND J. GROSSER (2007): "Welfare Reducing Polls," *Economic Theory*, 31(1), 51–68.

GUPTA, J. (2011): "The View from Jaitapur," *Climate Spectator*, March 29.

HEALY, A. J., N. MALHOTRA, AND C. H. MO (2010): "Irrelevant Events Affect Voters' Evaluations of Government Performance," *Proceedings of the National Academy of Sciences*, 107(29), 12804–12809.

HILLMAN, A., AND J. RILEY (1989): "Politically Contestable Rents and Transfers," *Economics and Politics*, 1, 17–40.

HONG, C. S., AND K. KONRAD (1998): "Bandwagon Effects and Two-Party Majority Voting," *Journal of Risk and Uncertainty*, 16(2), 165–172.

IARYCZOWER, M., AND A. MATTOZZI (2013): "On the Nature of Competition in Alternative Electoral Systems," *The Journal of Politics*, 75, 743–756.

JACOBS, L. R., AND R. Y. SHAPIRO (1994): "Issues, Candidate Image, and Priming: The Use of Private Polls in Kennedy's 1960 Presidential Campaign," *The American Political Science Review*, 88(3), pp. 527–540.

KANTER, J. (2011): "Switzerland Decides on Nuclear Phase-Out," *The New York Times*, March 25.

KLEIN, A., AND A. SCHMUTZLER (2013): "Intertemporal Effort Provision," Working paper.

KLUMPP, T., AND M. K. POLBORN (2006): "Primaries and the New Hampshire Effect," *Journal of Public Economics*, 90(6-7), 1073–1114.

KNIGHT, B., AND N. SCHIFF (2010): "Momentum and Social Learning in Presidential Primaries," *Journal of Political Economy*, 118(6), 1110 – 1150.

KONRAD, K. A. (2012): "Dynamic Contests and the Discouragement Effect," *Revue d'économie politique*, 112(2), 233–256.

LAZEAR, E., AND S. ROSEN (1981): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89(5), 841–864.

MCALLISTER, I., AND D. T. STUDLAR (1991): "Bandwagon, Underdog, or Projection? Opinion Polls and Electoral Choice in Britain, 1979-1987," *The Journal of Politics*, 53(3), pp. 720–741.

MEIROWITZ, A. (2008): "Electoral Contests, Incumbency Advantages, and Campaign Finance," *The Journal of Politics*, 70(03), 681–699.

MOLDOVANU, B., AND A. SELA (2001): "The Optimal Allocation of Prizes in Contests," *American Economic Review*, 91(3), 542–558.

NADEAU, R., E. CLOUTIER, AND J.-H. GUAY (1993): "New Evidence about the Existence of a Bandwagon Effect in the Opinion Formation Process," *International Political Science Review*, 14(2), pp. 203–213.

NALEBUFF, B. J., AND J. E. STIGLITZ (1983): "Prices and Incentives: Towards a General Theory of Compensation and Competition," *Bell Journal of Economics*, 14(1), 21–43.

NORDHAUS, W. D. (1975): "The Political Business Cycle," *Review of Economic Studies*, 42(2), 169–90.

POLBORN, M. (2006): "Investment under Uncertainty in Dynamic Conflicts," *Review of Economic Studies*, 73(2), 505–529.

SKAPERDAS, S., AND B. GROFMAN (1995): "Modeling Negative Campaigning," *The American Political Science Review*, 89(1), 49–61.

SNYDER, J. M. (1989): "Election Goals and the Allocation of Campaign Resources," *Econometrica*, 57(3), pp. 637–660.

STRAFFIN JR., P. D. (1977): "The Bandwagon Curve," *American Journal of Political Science*, 21(4), pp. 695–709.

STROMBERG, D. (2008): "How the Electoral College Influences Campaigns and Policy: The Probability of Being Florida," *American Economic Review*, 98(3), 769–807.

VOWLES, J. (2002): "Did the Polls Influence the Vote? A Case Study of the 1999 New Zealand General Election," *Political Science*, 54(1), 67–78.