Dynamic Campaign Spending[*]

Avidit Acharya, Edoardo Grillo, Takuo Sugaya, Eray Turkel[†]

February 19, 2019

Abstract

We build a model of electoral campaigning in which two purely office-motivated candidates allocate money over time to control the movement of their relative popularity, which evolves as a (mean-reverting) Brownian motion. We establish a key result: the ratio of spending by each candidate equals the ratio of their available budgets. This result enables us to characterize the path of spending over time as a function of the parameters of the popularity process. We then use this relationship to recover estimates of the decay rate in the popularity process for U.S. elections from 2000-2014 and find substantial weekly decay rates well above 50%, consistent with other approaches in the literature on political advertising.

Key words: campaigns, dynamic allocation problems, contests

JEL codes: C72

1 Introduction

Electoral campaigns are dynamic contests in which the candidates allocate their resources over time to increase their relative popularity prior to the election.

In U.S. House elections, for example, the mean level of spending on TV ads by candidates in the most competitive races from 2000-2014 was $2.5 million, and for the most competitive Senate and gubernatorial races in this period, the corresponding figures were $11.4 million and $11.5 million. Figure 1 shows the pattern of spending over time for various candidates in these races, corresponding to different percentiles of total spending, as well as the average spending paths for Democrats and Republicans. For the most part, candidates tend to increase their spending over time ahead of the election date, ramping it up in the final weeks, especially for candidates in the most competitive financial contests who spend higher amounts. Overall, the average spending patterns by Democrats and Republicans in these races are nearly identical.

These patterns are the consequence of deliberate and strategic choices made by the campaigns—a fact that raises the question of what the underlying calculations are that drive these decisions. We answer this question through a simple dynamic allocation model in which two candidates allocate their stock of available resources across a finite number of periods to influence the movement of their relative popularity.

In our model, the candidates begin with one being possibly more popular than the other. At each moment in time, relative popularity may go up, meaning that candidate 1 increases his lead in the polls; or it may go down, meaning that candidate 2 increases her lead. Relative popularity evolves between periods according to a modified Brownian motion so that the next period's starting level of relative popularity is normally distributed with a fixed variance and a mean that depends only on the current level of relative popularity and the ratio of candidate 1's spending to candidate 2's spending that period. At the final date, an election takes place and the more popular candidate wins office. Money left over has no value, so the game is zero-sum.

The solution to the spending decision in our model rests on a key result: at every history, the equilibrium ratio of candidate 1's spending to candidate 2's spending equals the ratio of their available resources. This result, enables us to provide a complete characterization of the equilibrium path of spending over time as a function of the popularity process. For example, in the case where the drift of the popularity process is affected only by relative spending, the two candidates spread their resources evenly over periods independent of the current level of relative popularity. Therefore, along the
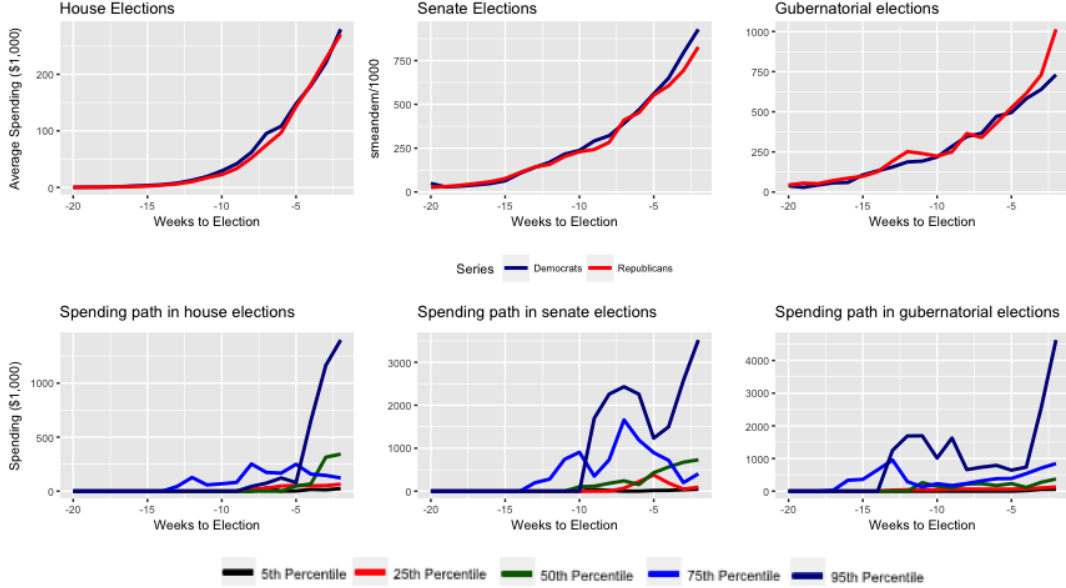
Figure 1: Upper figures are average spending paths by Democrats and Republicans on TV ads in "competitive" House, Senate and gubernatorial races in the period 2000-2014. These are elections in which both candidates spent a positive amount; see Section 4.1 for the source of these data, and more details. Bottom figures are spending paths for 5th, 25th, 50th, 75th and 95th percentile candidates in terms of total money spent in the corresponding elections of the upper panel.

equilibrium path relative popularity follows a constant-drift Brownian motion, where the drift depends only on the ratio of starting budgets.

Alternatively, if any gains in a candidate's relative popularity tend to decay over time, then maintaining an advantage in popularity is harder. As a result, the candidates increase their spending over time, and relative popularity follows a mean-reverting Brownian motion (the Ornstein-Uhlenbeck process), with long-term mean determined by the ratio of starting budgets. Moreover, the rate at which spending increases over time is higher when the speed of reversion of the stochastic process is greater. This case is salient because it rationalizes the spending patterns depicted in Figure 1.

We test the theoretical predictions of our model using spending data from the elections that we used to generate Figure 1. The key prediction of our model is that the ratio of candidates' spending is constant over time. We find that this pattern is borne out in the data. We then use the equilibrium relationship between the rate of spending over time and the degree of mean reversion in the polls to recover estimates of the weekly decay rate of a polling lead. Our estimates imply a substantial level of decay in popularity leads. In House elections, for example, our point estimate for the average

weekly decay rate of a polling lead implied by the candidates' spending patterns is 88%. In Senate and gubernatorial elections, the point estimates are 74% and 73%.

Despite polling data being very sparse, we compare these estimates to direct estimates of the decay rate from polling averages and find that the decay rate in polling is very close to our estimates from spending, but typically higher. We also find that our estimates are roughly in line with estimates of the decay rate coming from prior work on political advertising, which uses very different methods such as field experiments.

Related Literature— We build on a long tradition of using contest theory to model electoral campaigning. Prior work in this literature has studied static models. Erikson and Palfrey (1993, 2000), for example, use a contest model to reveal the empirical challenges in estimating the electoral effects of incumbent spending. Meirowitz (2008) studies a related model to show how asymmetries in the cost of effort can explain the incumbency advantage. Polborn and David (2004) and Skaperdas and Grofman (1995) examine static campaigning models in which candidates must choose between positive or negative advertising.[1]

A few papers use dynamic models to study campaigning. de Roos and Sarafidis (2018) explain how candidates that have won past races may enjoy "momentum," which results from a complementarity between prior electoral success and current spending.[2] Gul and Pesendorfer (2012) study a model of campaigning in which candidates provide information to voters in continuous time, and face the strategic timing problem of when to optimally stop. Kawai and Sunada (2015) estimate a model of fund-raising and campaigning in which the inter-temporal resource allocation decisions that candidates make are across different elections rather than across periods in the run-up to a particular election. Iaryczower et al. (2017) estimate a model in which campaign spending weakens electoral accountability, but in which the opportunity cost of spending is exogenous. The key difference between our work and these past papers is that we study a strategic allocation problem in which resources must be allocated over time subject to an inter-temporal budget constraint.

---

[1]Other static models of campaigning include Prat (2002) and Coate (2004), who investigate how one-shot campaign advertising financed by interest groups can affect elections and voter welfare, and Krasa and Polborn (2010) who study a model in which candidates compete on the level of effort that they apply to different policy areas.

[2]Other dynamic models of electoral campaigns in which candidates enjoy momentum—such as Callander (2007), Knight and Schiff (2010), Ali and Kartik (2012)—are models of *sequential voting*.

Our paper also connects to other empirical work on the relationship between campaign spending and electoral success. As Jacobson (1978) notes, a key challenge is how to tackle simultaneity bias: campaign spending decisions affect the outcome of the election, but these decisions are themselves affected by expectations about how the campaign will unfold, who will win, and with what margin (see Jacobson, 2015, for a review of the literature). Erikson and Wlezien (2012) address these concerns with a time-series analysis, and produce evidence that campaign spending in U.S. presidential elections influences polling numbers. Other authors (e.g. Green and Krasno, 1988, Gerber, 1998, Cox and Thies, 2000) address the challenge by estimating the effects of campaign spending using instrumental variables. Spenkuch and Toniatti (2018) leverage a natural experiment to show that TV ads affect vote shares without affecting aggregate turnout. Martin (2014) develops a structural model to estimate the persuasive and informative channels of TV ads, and finds that the persuasive channel is twice as large.

In a famous paper, Gerber et al. (2011) conduct a field experiment to show that exposure to campaign advertising influences voters, though the effects are ephemeral. In line with their finding, and the findings of the marketing literature on the decay of advertising effects more generally (e.g., Dubé et al., 2005, Leone, 1995, Tellis et al., 2005), our model assumes that relative campaign overspending has a positive effect on relative popularity, but the effect fades over time. By exploiting the one-to-one link between the decay rate and the candidates' optimal spending paths, our paper is the first to our knowledge to estimate decay rates across a large set of elections using only spending data. This is particularly valuable because of the cost of field experiments, and the lack of rich time-series polling data across many elections.

Finally, our paper relates to the literature on dynamic contests (see Konrad et al., 2009, and Vojnović, 2016, for reviews of this literature). In this literature, Gross and Wagner (1950) study a continuous Blotto game; Harris and Vickers (1985, 1987), Klumpp and Polborn (2006) and Konrad and Kovenock (2009) study models of races; and Glazer and Hassin (2000) and Hinnosaar (2018) study sequential contests. Ours is the first paper, to our knowledge, that studies a dynamic strategic allocation problem.

2 Model

Consider the following complete information *dynamic campaigning game* between two candidates, $i = 1, 2$, ahead of an election. Time runs continuously from 0 to T and

candidates take actions at times in $\mathcal{T} := \{0, \Delta, 2\Delta, ..., (N-1)\Delta\}$, with $\Delta := T/N$ being the time interval between consecutive actions. We identify these times with N discrete periods indexed by $n \in \{0, ..., N-1\}$. For all $t \in [0, T]$, we use $\underline{t} := \max\{\tau \in \mathcal{T} : \tau \leq t\}$ to denote the last time that the candidates took actions.

At the start of the game the candidates are endowed with positive resource stocks, $X_0 \geq 0$ and $Y_0 \geq 0$ respectively for candidates 1 and 2.[3] They allocate their resources across periods to influence changes in their relative *popularity*. Relative popularity at time t is measured by a continuous random variable $Z_t \in \mathbb{R}$ whose realization at time t is denoted by z_t. We will interpret this as a measure of candidate 1's lead in the polls. If $z_t > 0$, then candidate 1 is ahead of candidate 2. If $z_t < 0$, then candidate 2 is ahead; and if $z_t = 0$, it is a dead heat. We assume that at the beginning of the game, relative popularity is equal to $z_0 \in \mathbb{R}$.

At any time $t \in \mathcal{T}$, the candidates simultaneously decide how much of their resource stock to invest in influencing their future relative popularity. Candidate 1's investment is denoted x_t while candidate 2's is denoted y_t. The size of the resource stock that is available to candidate 1 at time $t \in \mathcal{T}$ is denoted $X_t = X_0 - \sum_{\tau \in \{t' \in \mathcal{T} : t' < t\}} x_\tau$ and that available to candidate 2 is $Y_t = Y_0 - \sum_{\tau \in \{t' \in \mathcal{T} : t' < t\}} y_\tau$.

Throughout, we will maintain the assumption that for all times t, the evolution of popularity is governed by the following modified Brownian motion:

$$dZ_t = \left(q\left(x_{\underline{t}}/y_{\underline{t}}\right) - \lambda Z_t\right) dt + \sigma dW_t \tag{1}$$

where $\lambda \geq 0$ and $\sigma > 0$ are parameters and $q(\cdot)$ is a strictly increasing, strictly concave function on $[0, \infty)$. Thus, the drift of popularity depends on the ratio of investments through the function $q(\cdot)$, and it may be mean-reverting if $\lambda > 0$.[4]

Finally, we assume that the *winner* of the election collects a payoff of 1 while the *loser* collects a payoff of 0. For analytical convenience, we make the assumption that if either candidate $i = 1, 2$ invests an amount equal to 0 at any time in $\mathcal{T}$, then the

---

[3]Although candidates raise funds over time, our assumption that they start with a fixed stock is tantamount to assuming that they can forecast how much will be available to them. In fact, some large donors make pledges early on and disburse their funds as it is needed in the campaign. In Appendix C we consider an extension of the model in which the candidates' resources evolve over time.

[4]If $\lambda = 0$ the process governing the evolution of popularity in the interval between two consecutive times in $\mathcal{T}$ is a standard Brownian motion— the continuous time limit of the random walk in which popularity goes up with probability probability $\frac{1}{2} + q(x_{\underline{t}}/y_{\underline{t}})\sqrt{\Delta}$ and goes down with complementary probability. If $\lambda > 0$, instead, popularity evolves in this interval according to the Ornstein-Uhlenbeck process, under which the leading candidate's lead has a tendency to decay.

game ends immediately. If $j \neq i$ invested a positive amount at that time, then j is the winner while if j also invested 0 at that time, then each candidate wins with equal probability.[5] If both candidates invest a positive amount at every time $t \in \mathcal{T}$, then the game only ends at time T, with candidate 1 winning if $z_T > 0$, losing if $z_T < 0$, and both candidates winning with equal probability if $z_T = 0$. In other words, if the game does not end before time T, then the winner is the candidate that is more popular at time T, and if they are equally popular they win with equal probability.

3   Analysis

Since the game is in continuous time, strategies must be measurable with respect to the filtration generated by W_t. However, since candidates take actions only at discrete times, we will forgo this additional formalism and treat the game as a game in discrete time. To that end, define the function

$$p\left(x/y\right) = \begin{cases} q(x/y) & \text{if } \lambda = 0 \\ q(x/y)/\lambda & \text{if } \lambda > 0 \end{cases}$$

By our assumption about the popularity process in (1), the distribution of $Z_{t+\Delta}$ at any time $t \in \mathcal{T}$, conditional on (x_t, y_t, z_t), is normal with constant variance and a mean that is a weighted sum of $p(x_t/y_t)$ and z_t; specifically,

$$Z_{t+\Delta} \mid (x_t, y_t, z_t) \sim \begin{cases} \mathcal{N}\left(p(x_t/y_t)\Delta + z_t, \sigma^2\Delta\right) & \text{if } \lambda = 0 \\ \mathcal{N}\left(p(x_t/y_t)(1 - e^{-\lambda\Delta}) + z_t e^{-\lambda\Delta}, \sigma^2(1 - e^{-2\lambda\Delta})/2\lambda\right) & \text{if } \lambda > 0 \end{cases}$$

$$(2)$$

where $\mathcal{N}(\cdot, \cdot)$ denotes the normal distribution whose first component is mean and second is variance. Note that the mean and variance of $Z_{t+\Delta}$ in the $\lambda = 0$ case correspond to the limits as $\lambda \to 0$ of the mean and variance in the $\lambda > 0$ case.

The model is therefore strategically equivalent to a discrete time model in which relative popularity is a state variable that transitions over discrete periods, and in each

---

[5]These assumptions close the model since the function q, which depends on spending *ratios*, is undefined if the denominator in the ratio is 0. The assumptions also guarantee that relative popularity, Z_t, follows an Itô process at every history. This model can be considered the limiting case of two different models. One is a model in which the marginal return to investing an ϵ amount of resources starting at 0 goes to infinity. The other is a model in which candidates must spend a minimum amount ϵ in each period to sustain the campaign, and ϵ goes to 0.

period it is normally distributed with a constant variance and a mean that depends on the popularity in the last period and on the ratio of candidates' spending.

With this, our equilibrium concept is subgame perfect Nash equilibrium (SPE) in pure strategies. We will refer to this concept succinctly as "equilibrium."

In the remainder of this section, we establish results on the paths of spending and popularity over time. We begin with a key observation, established in Section 3.1 below, that facilitates the analysis: on the equilibrium path of play, the ratio of the candidates' spending, x_t/y_t, is constant across all periods $t \in \mathcal{T}$.

3.1 Equal Spending Ratios

We refer to the ratio of a candidate's current spending to current budget as that candidate's *spending ratio*. For candidate 1 this is x_t/X_t and for candidate 2 it is y_t/Y_t. We will show that on the equilibrium path, these two ratios equal each other at every time t that the candidates make spending decisions.

Consider any time $t \in \mathcal{T}$ at which the game has not ended and candidates have to make their investment decisions. If $t = (N-1)\Delta$, then both candidates will spend their remaining budgets, i.e. $x_{(N-1)\Delta} = X_{(N-1)\Delta}$ and $y_{(N-1)\Delta} = Y_{(N-1)\Delta}$. Therefore, both candidates' spending ratios equal 1.

Now suppose that $t < (N-1)\Delta$ and assume that the stock of resources available to the two candidates are $X_t, Y_t > 0$.[6] Also, suppose that after the candidates choose their spending levels x_t and y_t, the probability that candidate 1 will win the election at time T when evaluated at time $t+\Delta$ depends on $X_{t+\Delta} = X_t - x_t$ and $Y_{t+\Delta} = Y_t - y_t$ only through the ratio $(X_t - x_t)/(Y_t - y_t)$. Denote this probability by $\pi_t\left((X_t - x_t)/(Y_t - y_t), z_{t+\Delta}\right)$. Further, let $F(z_{t+\Delta}|x_t/y_t, z_t)$ denote the c.d.f. of $Z_{t+\Delta}$ conditional on (x_t, y_t, z_t), and let $f(z_{t+\Delta}|x_t/y_t, z_t)$ denote the associated p.d.f. (Recall that these are normal distributions that depend on x_t and y_t only through the ratio x_t/y_t.)

If both candidates spend a positive amount in every period, candidate 1's expected payoff at time t is given by

$$\Pi_t(x_t, y_t | X_t, Y_t, z_t) = \int \pi_t \left(\frac{X_t - x_t}{Y_t - y_t}, z_{t+\Delta} \right) dF\left(z_{t+\Delta}|x_t/y_t, z_t\right)$$

---

[6]Recall that if either X_t or Y_t equal 0, the game will end at time t: either both candidates have no money to spend, or the one with a positive budget will spend any positive amount and win.

and candidate 2's expected payoff is $1 - \Pi_t(x_t, y_t | X_t, Y_t, z_t)$. The pair of necessary first order conditions for interior equilibrium values of x_t and y_t are

$$\frac{1}{y_t} \int \pi_t\left(\frac{X_t - x_t}{Y_t - y_t}, z_{t+\Delta}\right) \frac{\partial f\left(z_{t+\Delta} | x_t/y_t, z_t\right)}{\partial(x_t/y_t)} dz_{t+\Delta} =$$
$$= \frac{1}{Y_t - y_t} \int \frac{\partial \pi_t\left(\frac{X_t - x_t}{Y_t - y_t}, z_{t+\Delta}\right)}{\partial\left(\frac{X_t - x_t}{Y_t - y_t}\right)} dF\left(z_{t+\Delta} | x_t/y_t, z_t\right); \qquad (3)$$

$$\frac{x_t}{(y_t)^2} \int \pi_t\left(\frac{X_t - x_t}{Y_t - y_t}, z_{t+\Delta}\right) \frac{\partial f\left(z_{t+\Delta} | x_t/y_t, z_t\right)}{\partial(x_t/y_t)} dz_{t+\Delta} =$$
$$= \frac{X_t - x_t}{(Y_t - y_t)^2} \int \frac{\partial \pi_t\left(\frac{X_t - x_t}{Y_t - y_t}, z_{t+\Delta}\right)}{\partial\left(\frac{X_t - x_t}{Y_t - y_t}\right)} dF\left(z_{t+\Delta} | x_t/y_t, z_t\right). \qquad (4)$$

Taking the ratios of the respective left and right hand sides of these equations implies that $x_t/y_t = (X_t - x_t)/(Y_t - y_t)$, or $x_t/y_t = X_t/Y_t$. This observation suggests that our supposition that the remaining budgets $X_t - x_t$ and $Y_t - y_t$ affect continuation payoffs only through their ratio can be established by induction provided that the second order conditions are satisfied. The main steps in the proof of the following proposition involve establishing these facts. This and all other proofs appear in the Appendix.[7]

Proposition 1. *There exists an essentially unique equilibrium. If $X_t, Y_t > 0$ are the remaining budgets of candidates 1 and 2 at any time $t \in \mathcal{T}$, then in all equilibria,*

$$x_t/X_t = y_t/Y_t,$$

and, x_t/X_t and y_t/Y_t are independent of the past history $(z_\tau)_{\tau \leq t}$ of relative popularity.

The equal spending ratio result of Proposition 1 relies on the fact the distribution of Z_T given $((x_\tau, y_\tau, z_\tau)_{\tau \leq t-\Delta}, z_t)$ depends on $(x_\tau, y_\tau)_{\tau \geq t}$ only through ratios $(x_\tau/y_\tau)_{\tau \geq t}$. Since this is the case, if $(x_\tau^*, y_\tau^*)_{\tau \geq t}$ is an equilibrium in the continuation game in which the candidates' remaining budgets are (X_t, Y_t), then $(\beta x_\tau^*, \beta y_\tau^*)_{\tau \geq t}$ must be an equilibrium when budgets are $(\beta X_t, \beta Y_t)$, for all $\beta > 0$.[8] The independence of spending on the history

---

[7]The word "essentially" appears in the proposition below only because the equilibrium is not unique at histories at which either $X_t = 0 < Y_t$ or $X_t > 0 = Y_t$ — histories that do not arise on the path of play. In these cases, the candidate with a positive resource stock may spend any amount in period t and win. Apart from this trivial source of multiplicity, the equilibrium is unique.

[8] If this was not the case, we could find $(\tilde{x}_\tau)_{\tau \geq t}$ that gives a higher probability of winning to candidate 1 given $(\beta y_\tau^*)_{\tau \geq t}$. Because Z_T is determined by $(x_\tau/y_\tau)_{\tau \geq t}$, this would imply that the distribution of Z_T given $(\tilde{x}_\tau/\beta y_\tau^*)_{\tau \geq t}$ is more favorable to candidate 1 than the distribution given $(\beta x_\tau^*/\beta y_\tau^*)_{\tau \geq t} = (x_\tau^*/y_\tau^*)_{\tau \geq t}$. Because $(\tilde{x}_\tau/\beta)_{\tau \geq t}$ is a feasible continuation spending path when the budget profile is (X_t, Y_t), this contradicts the optimality of $(x_\tau^*)_{\tau \geq t}$ when candidate 2 plays $(y_\tau^*)_{\tau \geq t}$.

of popularity $(z_\tau)_{\tau \le t}$ relies on the fact that the law of motion of popularity depends in an additively separable way on a function $h(x_t, y_t)$ of candidates' spending, and on the current popularity level.[9] Under this assumption, we have

$$Z_T = \left(1 - e^{-\lambda\Delta}\right) \sum_{\tau=0}^{T-1} e^{-\lambda\Delta(T-1-\tau)} h(x_\tau, y_\tau) + z_0 e^{-\lambda T\Delta} + \sum_{\tau=0}^{T-1} e^{-\lambda\Delta(T-1-\tau)} \varepsilon_\tau,$$

where $(\varepsilon_\tau)_{\tau \ge 0}$ are i.i.d. normal shocks with mean 0 and variance $\sigma^2(1 - e^{-2\lambda\Delta})/2\lambda$. Hence, a pure strategy equilibrium exists if $h(\cdot, y)$ is concave in all y and $h(x, \cdot)$ is convex in all x; and the equilibrium spending profile (x_t, y_t) is notably independent of z_t.

In the appendix, we show that the equal spending ratio result of Proposition 1 carries over if we extend the model in two different directions. In Appendix C we study the case in which the candidates' budgets evolve over time in response to movements in relative popularity. This extension is motivated by the fact that campaigns raise funds over time and, contrary to the baseline model, they may not be able to perfectly forecast ex ante how much money they will have raised by election day.

In Appendix D.1 we study an electoral contest that takes place over multiple districts, each with its own relative popularity process, and in which the winner is determined by aggregating performance across districts. The extension covers both the U.S. electoral college, and competition by two parties for majority seats in a legislature composed of representatives from winner-take-all single-member districts. We show that the equal spending ratio result generalizes district-by-district.

3.2 Equilibrium Spending and Popularity Paths

An immediate corollary of Proposition 1 is a characterization of the process governing the evolution of relative popularity on the equilibrium path.

Corollary 1. *On the equilibrium path, relative popularity follows the process*

$$dZ_t = (q(X_0/Y_0) - \lambda Z_t)\, dt + \sigma dW_t \tag{5}$$

If $\lambda = 0$, this is a Brownian motion with constant drift $p(X_0/Y_0)$. If $\lambda > 0$, it is the Ornstein-Uhlenbeck process with long-term mean $p(X_0/Y_0)$ and speed of reversion λ.

---

[9]See Karatzas and Shreve (1998) equation (6.30). Using this result, we can write sufficient conditions to obtain this separability. Details are available upon request.

Therefore, when $\lambda > 0$ popularity leads have a tendency to decay towards zero. The instantaneous volatility of the process is σ and the stationary variance is $\sigma^2/2\lambda$.

Proposition 1 also enables us to solve, in closed form, for the equilibrium spending ratio at each history.

Proposition 2. *Let $t \in \mathcal{T}$ be a time at which $X_t, Y_t > 0$. Then, in equilibrium, spending ratios depend only on calendar time and on the speed of reversion λ. In particular,*

$$\frac{x_t}{X_t} = \frac{y_t}{Y_t} = \begin{cases} \Delta/(T-t) & \text{if } \lambda = 0 \\ \frac{e^{-\lambda(T-t-\Delta)}-e^{-\lambda(T-t)}}{1-e^{-\lambda(T-t)}} & \text{if } \lambda > 0 \end{cases}$$

which is continuous at $\lambda = 0$.

Proposition 2 implies that the fraction of their initial budget that each candidate spends in each period $n\Delta$ is the same for both candidates, and so is the ratio of spending in consecutive periods $n\Delta$ and $(n+1)\Delta$; we define these quantities as dependent on n and λ to be, respectively,

$$\gamma_\lambda(n) := \frac{x_{n\Delta}}{X_0} = \frac{y_{n\Delta}}{Y_0} \quad \text{and} \quad r_n(\lambda) := \frac{x_{(n+1)\Delta}}{x_{n\Delta}} = \frac{y_{(n+1)\Delta}}{y_{n\Delta}} \tag{6}$$

If $\lambda = 0$, then Proposition 2 implies that the candidates will spend a fraction $\gamma_0(n) = 1/N$ of their available resources in each period $n\Delta$, and the ratio of spending in consecutive periods is $r_n(0) = 1$. The $\lambda > 0$ case is handled in the following proposition.

Proposition 3. *Fix the number of periods N, total time $T = N\Delta$, and consider the case in which $\lambda > 0$. Then, for all n,*

$$\gamma_\lambda(n) = \frac{e^{\lambda\Delta}-1}{e^{\lambda N\Delta}-1}e^{\lambda\Delta n} \quad \text{and} \quad r_n(\lambda) = e^{\lambda\Delta}.$$

Since $r_n(\lambda)$ is increasing in λ, the shape of $\gamma_\lambda(n)$ is clear: it is increasing in n, and as λ grows it becomes higher for higher values of n and lower for lower values. Figure 2 depicts these properties by plotting $\gamma_\lambda(n)$ for different values of λ. The key property is that as the speed of reversion increases, candidates save even more of their resources for the final stages of the campaign.

The intuition behind these results is straightforward. When $\lambda = 0$, popularity advantages do not decay at all, and candidates equate the marginal benefit of spending against the marginal (opportunity) cost by spending evenly over time. As λ increases, then the
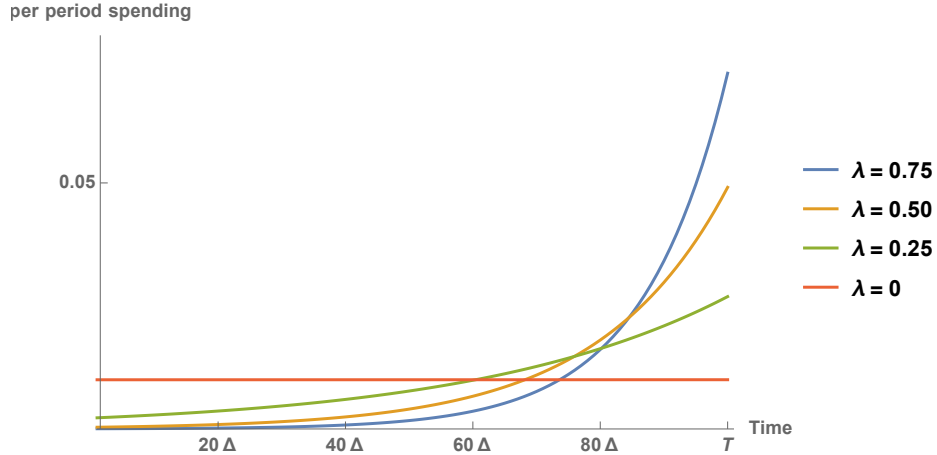
Figure 2: The fraction $\gamma_n(\lambda)$ of initial budget that the candidates spend over time, for $N = 100$ and various values of λ.

marginal benefit of spending early drops since any popularity advantage produced by an early investment has a tendency to decay, where this tendency is greater the greater is λ. In particular, if λ is high then any advantage in popularity that a candidate builds early on is harder to grow or even maintain. This means that candidates have an incentive to invest less in the early stages and more in the later stages of the campaign.

Finally, we can write a clean closed-form expression for the fraction of a candidate's initial budget cumulatively spent at time t by taking the continuous time limit as $\Delta \to 0$, fixing T. We have

$$\lim_{\Delta \to 0} \sum_{n\Delta \leq t} \gamma_\lambda(n) = \frac{e^{\lambda t} - 1}{e^{\lambda T} - 1}.$$

4 Empirical Analysis

As Proposition 3 implies and Figure 2 reveals, the speed of reversion λ can be recovered by fitting the actual pattern of spending to the predicted pattern of spending exploiting the fact that the predicted pattern of spending is uniquely determined by λ. In Section 4.2 below, we establish an identification result, introduce an estimator for λ, apply it to estimate λ from elections data, and compare the implied decay rates to estimates of the decay rate for TV ads from past studies. We first describe the data for the elections

12

we study (which include U.S. House, Senate, and gubernatorial elections) in Section 4.1 below, and examine some of the model's predictions.

4.1 Data

While spending in our model refers to all spending (e.g., TV ads, calls, mailers, door-to-door canvasing visits) that directly affects the candidates' relative popularity, it is not straightforward to separate out this kind of spending from other campaign spending (e.g. fixed costs, or administrative costs) that does not influence relative popularity. So we collect data only on TV ad spending and proceed under the assumption that any residual spending on the type of campaign activities that directly affect relative popularity is proportional to spending on TV ads.

We collect TV ad spending data from Wesleyan Media Project and Wisconsin Advertising Database. For each election in which TV ads were bought, the database contains information about the candidate each ad supports, the date it was aired, and the estimated cost. For the year 2000, the data covers only the 75 largest Designated Market Areas (DMAs), and for years 2002-2004, it covers only the 100 largest DMAs. The data from 2006 onwards covers all of the 210 DMAs. For 2006, where ad price data are missing, we estimate prices using ad prices in 2008.

We aggregate ad spending made on behalf of the two major parties' candidates by week and focus on the 20 weeks leading to election day, though we will drop the final week which is typically incomplete since elections are held on Tuesdays.[10] We get 1918 unique House, Senate and gubernatorial elections between 2000 and 2014. We then drop all elections that are clearly not genuine contests to which our model does not apply—i.e., elections in which one of the candidates did not spend anything for at least 18 weeks. This leaves us with 600 House, 167 Senate, and 161 gubernatorial elections. We focus on the last 20 weeks of the race both because TV ad spending is usually zero prior to this period, and because we want to restrict attention to the general election campaign. Nevertheless, there are still some states where primaries are held after the last week of June. So, whenever possible, we restrict attention to ads bought for the

---

[10]Election day is defined by law as "the first Tuesday after November 1," so candidates do not have a full week to spend on the last calendar week of the cycle.

Table 1: Time trend of $x_t/X_t - y_t/Y_t$

	(All)	(All)	(House)	(Senate)	(Governor)
	Dependent variable: $x_t/X_t - y_t/Y_t$				
Constant	0.0005	0.0005	0.001	−0.0008	0.0002
	(0.0009)	(0.001)	(0.001)	(0.002)	(0.0002)
Time	−0.0001	−0.0001	0.0001	−0.0002	−0.0007**
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)
Time × Dem Victory		0.0001	−0.0003	0.0003	0.0010**
		(0.0002)	(0.0002)	(0.0003)	(0.0003)
Observations	11,780	11,780	6,137	2,888	2,755
R^2	0.0001	0.0002	0.001	0.0005	0.006

Note:	*p<0.1; **p<0.05; ***p<0.01

Note: Standard errors are robust to heteroskedasticity and autocorrelation. The sample includes 620 elections (out of a total of 928) for which we can identify the winner from our data.

general election campaign.[11] Figure 1 in the introduction plots weekly spending averages from these races, showing that spending over time is generally increasing.

We also investigate the main robust prediction of our model that $x_t/X_t - y_t/Y_t$ is constant over time. In the data, we define $x_t/X_t - y_t/Y_t$ as the difference between the weekly spending of the Democratic candidate and the Republican candidate, as a percentage of their remaining budget. We regress this difference on a time trend. Column (1) of Table 1 presents the result, showing a very small shift in the difference in favor of the Republican candidate.

It is also possible that while the coefficient on the time trend is insignificant, the ratio x_t/y_t tilts towards the winning candidate especially if donors adjust their contributions to how candidates are doing in comparison to expectations, which might be more likely in state-wide elections than House races due to the greater importance of state-wide offices.[12] To investigate this possibility, Column (2) adds an interaction of the time-trend with Democratic victory, and Columns (3) - (5) disaggregate the data to House,

---

[11] The dataset allows us to do this for the elections in 2000, 2012 and 2014. As a robustness exercise, we also conduct the same analysis using data from only the last 12 weeks of campaigns and find that the results are similar (available upon request).

[12] The extension in Appendix C in which the candidates' budgets evolve over time in response to the realizations of their relative popularity is motivated in part by the possibility that the amount of money the candidates raise could be sensitive to popularity path, especially in statewide contests.
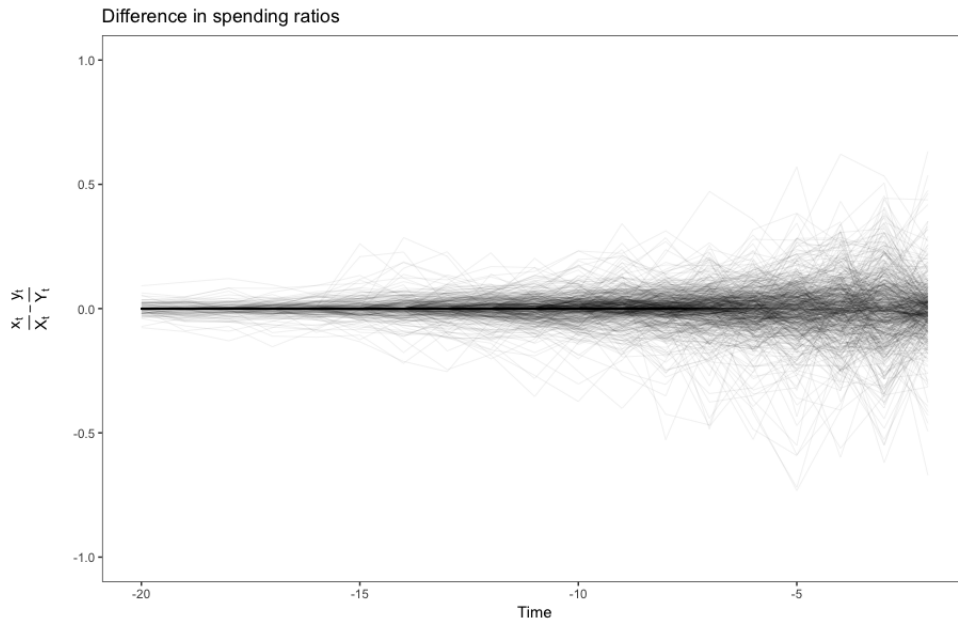
Figure 3: Difference between the TV ad spending of the Democratic candidate and the Republican candidate as a percentage of their remaining budget, in the last 20 weeks of the election.

Senate and gubernatorial elections. Although the time trend coefficients are statistically significant in the gubernatorial sample at the 5% level, the substantive magnitudes are very small. Figure 3 plots $x_t/X_t - y_t/Y_t$ over time for all elections.

4.2   Estimating Decay Rates from Spending Data

We begin by establishing an identification result that shows that we can empirically identify λ for an arbitrary choice of Δ.

Proposition 4. *Fix any integer $\kappa \geq 1$, and for $t \in \mathcal{T}$ let $\bar{x}_t = \sum_{j=0}^{\kappa-1} x_{t+\Delta j}$ be aggregate equilibrium spending for candidate 1 over a unit of time $\bar{\Delta} := \kappa\Delta$. Then, given the remaining budget X_t at time t,*

$$\frac{\bar{x}_t}{X_t} = \begin{cases} \bar{\Delta}/(T-t) & \text{if } \lambda = 0 \\ \frac{e^{-\lambda(T-t-\bar{\Delta})}-e^{-\lambda(T-t)}}{1-e^{-\lambda(T-t)}} & \text{if } \lambda > 0 \end{cases} \tag{7}$$

and the same result also holds for candidate 2.

The key implication of this proposition is that λ and Δ cannot be separately identified from the data; only their product $\lambda\Delta$ can be identified.

Therefore, for our analysis of spending in the final twenty weeks of each election, we fix $\Delta = 1$ week, set $T = 19$ (recall that we drop the final incomplete week), and estimate λ for these values of Δ and T. We also report results on the implied decay rate, where

$$\text{decay rate} = 1 - e^{-\lambda}$$

is the percentage decay in a polling lead absent any financial influence of the candidates on the path of relative popularity. We transform the 95% confidence intervals for our estimates of λ to get the exact 95% confidence intervals for the decay rate.

To estimate λ we use a truncated maximum likelihood estimator. Let $\{x_{n\Delta}\}$ denote a path of spending, and assume that we observe in the data $\{\tilde{x}_{n\Delta}\}$, where

$$\log \tilde{x}_{n\Delta} = \max\{0, \log x_{n\Delta} + \varepsilon_{n\Delta}\}$$

where $\varepsilon_{n\Delta}$ is i.i.d. mean zero normal measurement error. Proposition 3 shows that $\log x_{n\Delta} = \log \gamma_\lambda(n) X_0$ and Proposition 4 allows us to take $\Delta = 1$, allowing λ to vary, so we can write the likelihood function as

$$L(\lambda, \mu, \sigma) := \prod_{n:\tilde{x}_{n\Delta}=0} \Phi\left(\frac{-\mu - \lambda n}{\sigma}\right) \prod_{n:\tilde{x}_{n\Delta}>0} \phi\left(\frac{\log \tilde{x}_{n\Delta} - \mu - \lambda n}{\sigma}\right)$$

where

$$\mu = \log(e^\lambda - 1) - \log(e^{\lambda T} - 1) + \log X_0$$

The estimator for (λ, μ, σ) maximizes the log of this likelihood function. It is well known that under regularity conditions this estimator is consistent and asymptotically normal, which gives us an estimator for the standard error of λ (see Amemiya, 1973).

Figure 4 presents the estimated values of λ across the House, Senate and gubernatorial elections in our sample, as well as the implied decay rates along with 95% confidence intervals. The median estimated λ across House elections is 2.02 (95% CI = [1.57, 2.47]), corresponding to a weekly decay rate of 86% ([78%,91%]). The median estimated λ in Senate elections is 1.23 ([1.00, 1.47]) corresponding to a decay rate of 70% ([63%, 77%]) while the median estimated λ in gubernatorial elections is 1.29 ([0.99, 1.46]) corresponding to a decay rate of 72%([62%, 76%]).

The densities of our point estimates for λ values and decay rates across all three settings, House, Senate, and gubernatorial elections are depicted in Figure 5. The figure
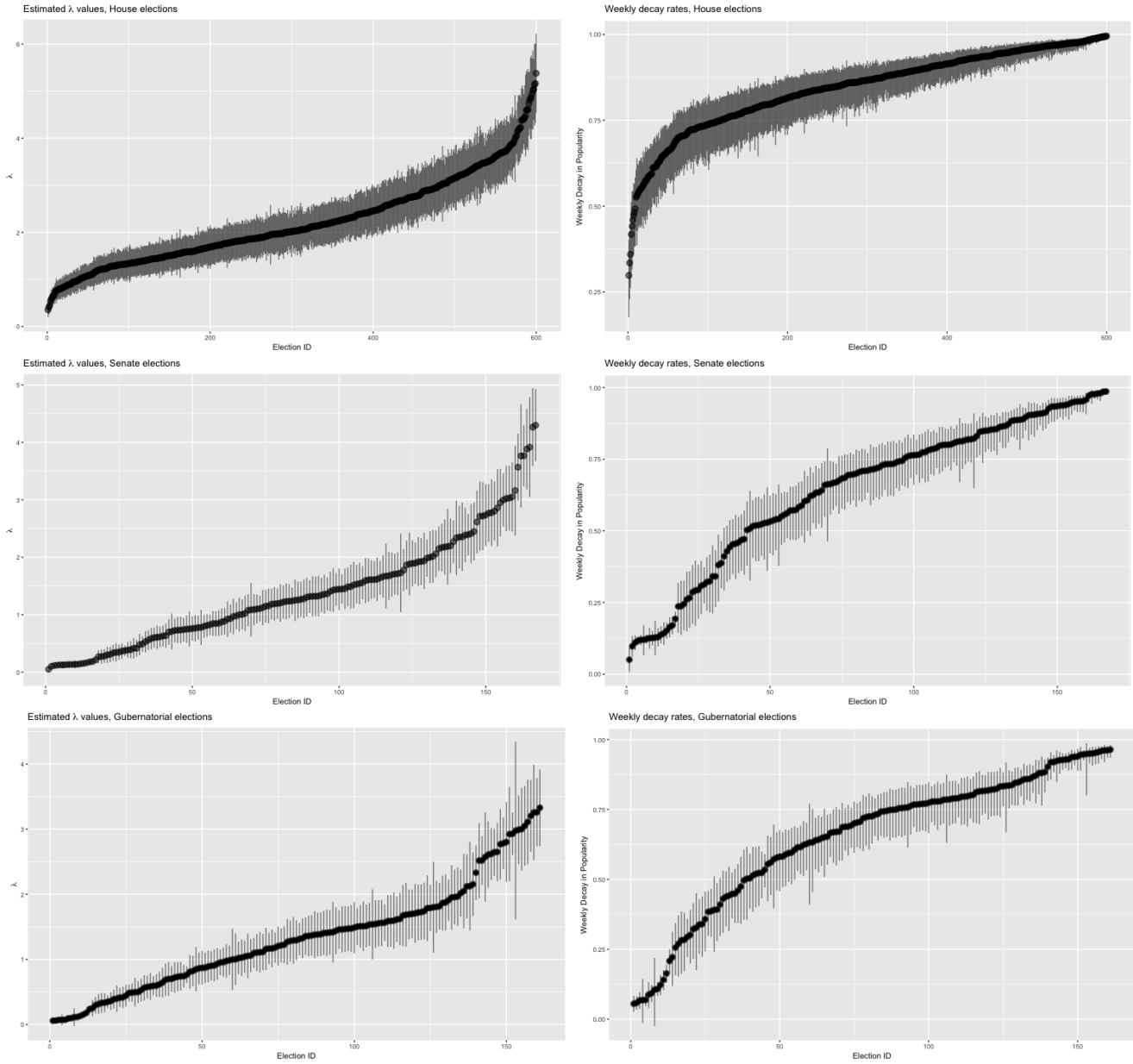
16

Figure 4: Estimated λ values for House, Senate and gubernatorial elections and implied weekly decay rates along with 95% confidence intervals.
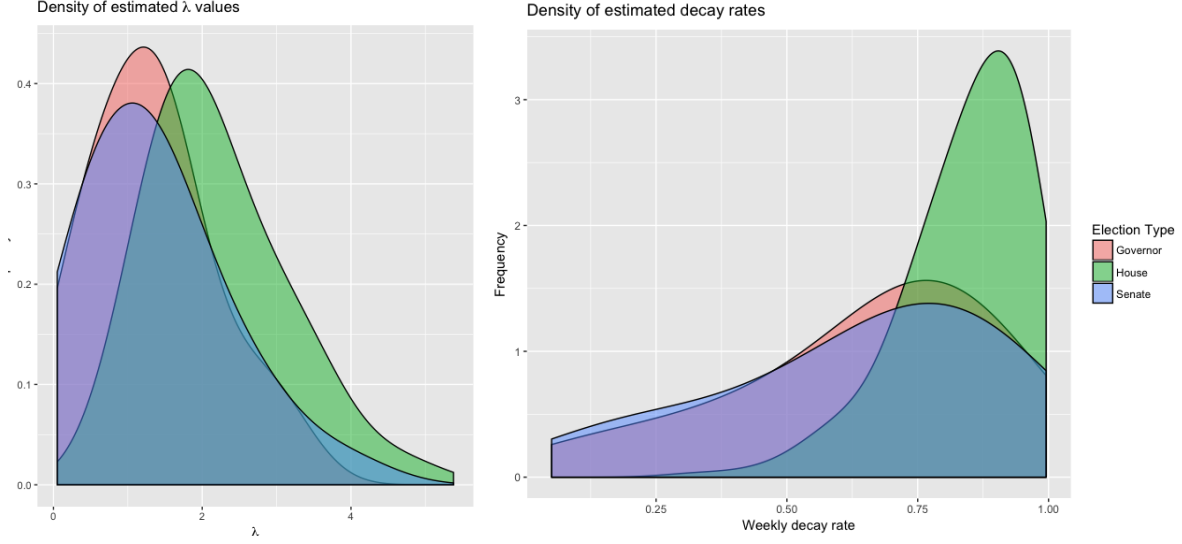
17

Figure 5: Densities of estimated λ and decay rates for House, Senate and gubernatorial elections.

shows that while the distribution of decay rates is remarkably similar across Senate and gubernatorial elections, decay rates for House elections are typically higher.

4.3 Comparisons

How do these estimates compare to other studies and estimation techniques? One alternative approach is to estimate the decay rate directly from polling data. To investigate this approach, we collect polling data from the public version of FiveThirtyEight's polls database and from HuffPost's Pollster database. We find, not surprisingly, that polling data for these elections are very sparse; so our estimates are likely to be very noisy, precluding us from doing any meaningful inference.[13] Nevertheless, we implement the approach to compare point estimates across the two methodologies.

Given equation (2), our model implies that for $\lambda > 0$, relative popularity evolves according to a simple AR(1) process:

$$Z_{(n+1)\Delta} = \beta_0 + \beta_1 Z_{n\Delta} + \varepsilon \tag{8}$$

where

$$\beta_0 = (1 - e^{-\lambda})q(X_0/Y_0)/\lambda \quad \text{and} \quad \beta_1 = e^{-\lambda} \tag{9}$$

---

[13]The sparsity of polling data is an additional reason for why our model's ability to indirectly recover estimates of the decay rate from spending data is particularly valuable.
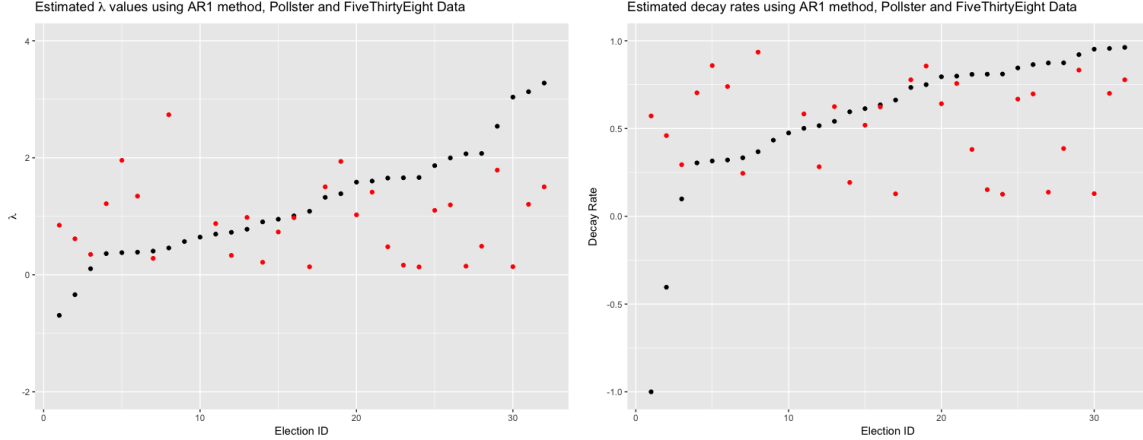
Figure 6: AR(1) model point estimates of the λ values and decay rates from polling data in blue, and maximum likelihood point estimates of the λ values and decay rates from spending data in red.

since again we set $\Delta = 1$ week. Therefore, the weekly decay rate is simply $1 - \beta_1$. For this estimation to work, however, we need at least three weeks of consecutive polling data. Applying this criteria, we get 27 elections from Pollster's database and 68 elections from FiveThirtyEight's database, three of which are overlapping. In this case, we use Pollster's data since Pollster's polling data are richer for these elections. This gives us a total of 92 elections, all of which are statewide elections. For 60 of these, however, we get point estimates of β_1 that are negative, implying that consecutive period polling is negatively correlated. We drop these since $-\log \beta_1$ is undefined for these elections, meaning that it not possible to recover estimates of λ. Decay rates and corresponding estimates of λ for the remaining 32 elections are depicted in blue in Figure 6. The median decay rate is 65%, which is close to but lower than the median estimated decay rates across the House, Senate and gubernatorial elections using spending data.

For 30 of the statewide elections, we have both weekly spending data and weekly polling data, so we can do an election-by-election comparison of the estimated decay rates using the two different methodologies. The plots in Figure 6 also depict in red the point estimates for λ and the decay rate from spending data. The figure shows that point estimates of the decay rate from polling data are generally higher than estimates of the decay rate from spending data, with the average difference in λ being $+0.34$ and the average difference in decay rates being 4.5 percentage points.

We can also compare our decay rates to decay rates found by other studies. If we look at estimates of the decay in advertising effects in general, our estimates are substantially higher, though these studies usually examine advertising of well-known

brands of perishable consumer goods.[14] For example, in a study similar to ours, Dubé et al. (2005) study advertising carry-over in the frozen food industry, where firms build a capital of "goodwill" through ads, which decays over time. They report a half life of 6 weeks, which corresponds to a weekly decay rate of about 12%.

However, our estimates of the decay rates are broadly consistent with estimates of the decay rates found in studies of political advertising more specifically. One study by Hill et al. (2013) finds the weekly decay rate to be between 70% and 95% for subnational U.S. elections in 2006, which is consistent with our estimates, though higher than our median. The famous Gerber et al. (2011) study conducts a field experiment during the 2006 Texas gubernatorial election, about eleven months prior to election day, and depending on the econometric specification finds the decay rate to be between 25% and 94%.[15] For this specific election, we get a point estimate for λ of 3.11, ($[2.46, 3.75]$), corresponding to a weekly decay rate of 95% ($[91\%, 97\%]$), which is in the ballpark—though toward the higher end—of their estimates.

5   Extensions and Other Analyses

In the appendix, we report various other theoretical and empirical results. For example, Proposition 3 shows that the ratio of spending in consecutive periods $r_\lambda(n)$ is increasing in λ. We show in Appendix B that this prediction holds in the data. In the same appendix, we also use polling data along with our estimates of λ to recover point estimates of $q(X_0/Y_0)/\lambda$. Given that this quantity measures the long run mean of the popularity process, we use the estimates to study the effect of financial advantage on the long-term mean of popularity, as well as to study the incumbency advantage in campaigning.

Despite the noisiness of the direct estimates of the decay rate from polling data, our finding that the point estimates of the decay rate recovered from spending data are typically higher than those estimated from polling data might be interpreted as meaning that candidates are over-spending in the early weeks. This leads us to reexamine the model's key assumption that the candidates' budgets are fixed (or they can perfectly forecast how much money they will raise for the campaign). If access-seeking donors are likely to watch the early polling results of elections to important state-wide offices, and

---

[14]See Leone (1995) for an extensive survey of this literature.

[15]For example, their polynomial distributed lag model estimates show that the standing of the advertising candidate increases by 5.4 percentage points in the week that the ad is aired, and the effect goes down to 2.19 percentage points the following week (a 60% decay). In another specification, the first week effect is 5.78, and goes down to 1.86 (a 68% decay).

they adjust their donations in response—say, to pick the winner—then early spending may be more valuable than we have assumed.

Motivated by this possibility we study an extension of the model in Appendix C in which budgets evolve over time in response to movements in popularity. The extension captures a form of momentum, suggested by Aldrich (1980), that arises from donors flocking to the candidate that is leading in the polls (but it is general enough to cover the opposite case in which they channel their support to the underdog). If donors tend to channel their support to the leading candidate, then there is a new tradeoff between spending early versus saving for later. Spending early now has greater benefit: although its effects on popularity are short-lived, short-term increases in popularity can attract more money that can be put in the war chest for later use.

Finally, we noted that it is difficult to compare estimates of the decay rate from spending data to those from polling data because of the sparsity of polling data. This is not a concern though for presidential contests, where polling is significantly richer. However, one complication in studying presidential contests is that our baseline model does not directly apply because presidential elections are multi-district contests across many winner-take-all districts, whose outcomes are aggregated to produce a final winner according to the electoral college. Given this, in Appendix D we present an extension of the model in which election to office depends on how the candidates performs across many electoral districts, each with its own popularity process. This extension is general enough to cover the electoral college, but also other settings such as competition between two parties seeking to control a majoritarian legislature, etc.

In the remainder of Appendix D, we estimate parameters of the popularity process in the 2012 Obama-Romney presidential race and find that the decay rate estimated from spending data is precisely 11% per week, whereas the decay rate estimated from polling data alone is 66.5% ([30.9%, 100%)) per week. Since the latter number is more in line with our estimates from the other races, we interpret this as evidence that there is significant overspending early on given the assumptions of our baseline model. To explain the discrepancy, we speculate that perhaps the extension in Appendix C is particularly relevant to presidential contests if candidates in these races have a substantial incentive to spend early and perform well to influence the inflow of future resources.

6   Conclusion

We have written a new model of dynamic campaigning, and used it to propose a novel methodology for recovering estimates of the decay rate in the popularity process using spending data alone. This is particularly valuable given the current sparsity of polling data for subnational elections, and the cost of running field experiments to estimate the decay rate of political advertising. We then implemented the methodology to recover decay rate estimates from almost a thousand elections, and compared these estimates to those using different approaches.

Our theoretical contribution raises new questions, however. Since we focused on the strategic choices made by the campaigns, we abstracted away from some important considerations. For example, we left unmodeled the behavior of the voters that generates over-time fluctuations in relative popularity. In addition, we abstracted away from the motivations and choices of the donors, and the effort decisions of the candidates in how much time to allocate to campaigning versus fundraising. These abstractions leave open questions about how to micro-found the behavior of voters and donors, and effort allocation decision for the candidates. We leave these questions to future work.[16]

Moreover, we have abstracted from the fact that in real life, campaigns may not know what the return to spending is at the various stages of the campaign, and what the decay rate is, as this may be specific to the personal characteristics of their respective candidates, and changes in the political environment, including the "mood" of voters. Real-life campaigns face an optimal experimentation problem whereby they try to learn about their environment through early spending. Our model also abstracted away from the question of how early spending may benefit campaigns by providing them with information about what kinds of campaign strategies seem to work well for their candidate. This is a considerably difficult problem, especially in the face of a fixed election deadline, and the endogeneity of donor interest and available resources. But there is no doubt that well-run campaigns spend to acquire valuable information about how voters are engaging with and responding to the candidates over the course of the campaign. These are interesting and important questions that ought to be addressed by future work.

---

[16]Bouton et al. (2018) address some of these questions in a static model. They study the strategic choices of donors who try to affect the electoral outcome and show that donor behavior depends on the competitiveness of the election. Similarly, Mattozzi and Michelucci (2017) analyze a two-period dynamic model in which donors decide how much to contribute to each of two possible candidates without knowing ex-ante who is the more likely winner.

Appendix

A   Proof of Results in the Main Text

A.1   Proof of Proposition 1

We consider the case of $\lambda > 0$. The $\lambda = 0$ case must be handled separately, but is very similar, so we omit the details.[17]

We prove by induction that, in any equilibrium, if $X_t, Y_t > 0$, then for all $t \in \mathcal{T}$,

(i) $x_\tau / y_\tau = X_t / Y_t$ at all times $\tau \geq t$ at which the candidates take actions;

(ii) if $t < (N-1)\Delta$, then the distribution of Z_T computed at time $t + \Delta \in \mathcal{T}$ given $z_{t+\Delta}$ is

$$\mathcal{N}\left(p\left(\frac{X_t - x_t}{Y_t - y_t} \right) (1 - e^{-\lambda(T - t - \Delta)}) + z_{t+\Delta} e^{-\lambda(T - t - \Delta)}, \frac{\sigma^2 (1 - e^{-2\lambda(T - t - \Delta)})}{2\lambda} \right)$$

The claim is obviously true at $t = (N-1)\Delta$, since in any equilibrium the candidates' payoffs depend only on z_T and in the final period they must spend the remainder of their budget.

Suppose, for the inductive step, that for all $\tau \geq t + \Delta$, both statements (i) and (ii) above hold. The distribution of $Z_{t+\Delta}$ at time $t \in \mathcal{T}$ given (x_t, y_t, z_t) is

$$\mathcal{N}\left(p\left(\frac{x_t}{y_t} \right) (1 - e^{-\lambda\Delta}) + z_t e^{-\lambda\Delta}, \frac{\sigma^2 (1 - e^{-2\lambda\Delta})}{2\lambda} \right)$$

By this hypothesis, the distribution of Z_T computed at time $t + \Delta \in \mathcal{T}$ given $z_{t+\Delta}$ is

$$\mathcal{N}\left(p\left(\frac{X_t - x_t}{Y_t - y_t} \right) (1 - e^{-\lambda(T - t - \Delta)}) + z_{t+\Delta} e^{-\lambda(T - t - \Delta)}, \frac{\sigma^2 (1 - e^{-2\lambda(T - t - \Delta)})}{2\lambda} \right)$$

---

[17] We have continuity at the limit: all of the results for the $\lambda = 0$ case hold as the limits of the $\lambda > 0$ case as $\lambda \to 0$.

The compound of normal distributions is also a normal distribution. Therefore, the distribution of Z_T at time t, given (x_t, y_t, z_t) is normal with mean and variance:

$$\mu_{Z_T|t} = p\left(\frac{X_t - x_t}{Y_t - y_t}\right)(1 - e^{-\lambda(T-t-\Delta)}) + p\left(\frac{x_t}{y_t}\right)(e^{-\lambda(T-t-\Delta)} - e^{-\lambda(T-t)}) + z_t e^{-\lambda(T-t)}$$

$$\sigma_{Z_T|t}^2 = \frac{\sigma^2(1 - e^{-2\lambda(T-t)})}{2\lambda}.$$

These expressions follow from the law of iterated expectation, $\mu_{Z_T|t} = E_t[E_{t+1}[Z_T]]$, and the law of iterated variance, $\sigma_{Z_T|t}^2 = E_t[Var_{t+1}[Z_T]] + Var_t[E_{t+1}[Z_T]]$.

Now, define the standardized random variable

$$\tilde{Z}_T = \frac{Z_T - \mu_{Z_T|t}}{\sigma_{Z_T|t}}.$$

Candidate 1 wins if $Z_T > 0$ or, equivalently, if

$$\tilde{Z}_T > -\frac{\mu_{Z_T|t}}{\sigma_{Z_T|t}} =: \tilde{z}_T^*$$

Therefore, taking y_t as given, candidate 1's objective is to maximizes his probability of winning, which is given by

$$\pi_t(x_t, y_t | X_t, Y_t, z_t) := \int_{\tilde{z}_T^*}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-s/2} ds.$$

Factoring common constants, the first order condition for this optimization problem is satisfied if and only if $0 = \partial\mu_{Z_T|t}/\partial x_t$, i.e.,

$$0 = p'\left(\frac{x_t}{y_t}\right)\frac{e^{-\lambda(T-t-\Delta)} - e^{-\lambda(T-t)}}{y_t} - p'\left(\frac{X_t - x_t}{Y_t - y_t}\right) \cdot \frac{1 - e^{-\lambda(T-t-\Delta)}}{Y_t - y_t} \tag{10}$$

Moreover, substituting the first order condition in the second order condition and rearranging terms, we get that the second order expression is given by a positive constant that multiplies

$$\frac{\partial^2\mu_{Z_T|t}}{\partial(x_t)^2} = p''\left(\frac{x_t}{y_t}\right)\frac{(e^{-\lambda(T-t-\Delta)} - e^{-\lambda(T-t)})}{(y_t)^2} + p''\left(\frac{X_t - x_t}{Y_t - y_t}\right) \cdot \frac{1 - e^{-\lambda(T-t-\Delta)}}{(Y_t - y_t)^2}$$

Because the function q is strictly concave, p is strictly concave as well. Hence, the second order condition is always satisfied and the objective function is strictly quasi-concave

24

in x_t. By an analogous argument, we can show that candidate 2's problem is strictly quasi-concave in y_t.

Therefore, the first order approach in the main text of Section 3.1 is valid, and we have $x_t/y_t = X_t/Y_t$ for all $\tau \geq t$. This implies $(X_t - x_t)/(Y_t - y_t) = X_t/Y_t$. Therefore, we can conclude that the distribution of Z_T computed at time t is given by a normal distribution with mean and variance:

$$\mu_{Z_T|t} = p\left(\frac{x_t}{y_t}\right)(1 - e^{-\lambda(T-t)}) + z_t e^{-\lambda(T-t)},$$

$$\sigma^2_{Z_T|t} = \frac{\sigma^2(1 - e^{-2\lambda(T-t)})}{2\lambda}.$$

A.2 Proof of Proposition 2

Suppose that $\lambda > 0$. Then, the first order condition for x_t from (10) is

$$p'\left(\frac{x_t}{y_t}\right)\frac{e^{-\lambda(T-t-\Delta)} - e^{-\lambda(T-t)}}{y_t} = p'\left(\frac{X_t - x_t}{Y_t - y_t}\right)\cdot\frac{1 - e^{-\lambda(T-t-\Delta)}}{Y_t - y_t}$$

The analogous first order condition for y_t is

$$p'\left(\frac{x_t}{y_t}\right)\frac{x_t\left(e^{-\lambda(T-t-\Delta)} - e^{-\lambda(T-t)}\right)}{(y_t)^2} = p'\left(\frac{X_t - x_t}{Y_t - y_t}\right)\frac{(X_t - x_t)\left(1 - e^{-\lambda(T-t-\Delta)}\right)}{(Y_t - y_t)^2}$$

These two equations (together with the fact that $x_t/y_t = (X_t - x_t)/(Y_t - y_t)$) imply

$$\frac{x_t}{X_t} = \frac{y_t}{Y_t} = \frac{e^{-\lambda(T-t-\Delta)} - e^{-\lambda(T-t)}}{1 - e^{-\lambda(T-t)}}.$$

Now consider the $\lambda = 0$ case. The corresponding first order conditions for x_t and y_t are, respectively,

$$p'\left(\frac{x_t}{y_t}\right)\frac{\Delta}{y_t} = p'\left(\frac{X_t - x_t}{Y_t - y_t}\right)\cdot\frac{T - t}{Y_t - y_t},$$

$$p'\left(\frac{x_t}{y_t}\right)\frac{x_t\Delta}{(y_t)^2} = p'\left(\frac{X_t - x_t}{Y_t - y_t}\right)\frac{(X_t - x_t)(T - t)}{(Y_t - y_t)^2}.$$

Therefore, we have $x_t/X_t = y_t/Y_t = \Delta/(T - t)$.

25

A.3   Proof of Proposition 3

Since spending ratios are equal for the two candidates, we can focus without loss of generality on candidate 1. From Proposition 2, we have

$$\frac{x_{n\Delta}}{X_{n\Delta}} = \frac{e^{-\lambda(T-(n+1)\Delta)} - e^{-\lambda(T-n\Delta)}}{1 - e^{-\lambda(T-n\Delta)}} = \frac{e^{\lambda\Delta} - 1}{e^{\lambda(T-n\Delta)} - 1}$$

Then since

$$\frac{e^{\lambda(T-(n+1)\Delta)} - 1}{e^{\lambda(T-n\Delta)} - 1} = \frac{x_{n\Delta}/X_{n\Delta}}{x_{(n+1)\Delta}/X_{(n+1)\Delta}} = \frac{x_{n\Delta}}{x_{(n+1)\Delta}} \frac{X_{(n+1)\Delta}}{X_{n\Delta}} = \frac{x_{n\Delta}}{x_{(n+1)\Delta}} \frac{X_{n\Delta} - x_{n\Delta}}{X_{n\Delta}}$$

we have

$$r_n(\lambda) = \frac{x_{(n+1)\Delta}}{x_{n\Delta}} = \left(1 - \frac{x_{n\Delta}}{X_{n\Delta}}\right) \frac{e^{\lambda(T-n\Delta)} - 1}{e^{\lambda(T-(n+1)\Delta)} - 1}$$

$$= \left(1 - \frac{e^{\lambda\Delta} - 1}{e^{\lambda(T-n\Delta)} - 1}\right) \frac{e^{\lambda(T-n\Delta)} - 1}{e^{\lambda(T-(n+1)\Delta)} - 1} = e^{\lambda\Delta}$$

This gives us

$$x_{n\Delta} = e^{\lambda\Delta n} x_0 \quad \text{and} \quad X_0 = \sum_{n=0}^{N-1} x_{n\Delta} = \sum_{n=0}^{N-1} e^{\lambda\Delta n} x_0 = \frac{e^{\lambda\Delta N} - 1}{e^{\lambda\Delta} - 1} x_0$$

Therefore, we have

$$\gamma_\lambda(n) = \frac{x_{n\Delta}}{X_0} = \frac{e^{\lambda\Delta} - 1}{e^{\lambda\Delta N} - 1} e^{\lambda\Delta n}.$$

A.4   Proof of Proposition 4

For $\kappa = 1$ the result reduces to the formula reported in Proposition 2. Suppose $\kappa > 0$ and let $\lambda > 0$. Then,

$$
\begin{aligned}
\frac{\bar{x}_t}{X_t} &= \frac{x_t + x_{t+\Delta} + \cdots + x_{t+(\kappa-1)\Delta}}{X_t} \\
&= \frac{x_t}{X_t} + \frac{x_{t+\Delta}}{X_{t+\Delta}} \frac{X_{t+\Delta}}{X_t} + \cdots + \frac{x_{t+(\kappa-1)\Delta}}{X_{t+(\kappa-1)\Delta}} \frac{X_{t+(\kappa-1)\Delta}}{X_{t+(\kappa-2)\Delta}} \cdots \frac{X_{t+\Delta}}{X_t} \\
&= \frac{e^{-\lambda(T-t-\Delta)} - e^{-\lambda(T-t)}}{1 - e^{-\lambda(T-t)}} + \frac{e^{-\lambda(T-t-2\Delta)} - e^{-\lambda(T-t-\Delta)}}{1 - e^{-\lambda(T-t-\Delta)}} \left(1 - \frac{e^{-\lambda(T-t-\Delta)} - e^{-\lambda(T-t)}}{1 - e^{-\lambda(T-t)}}\right) + \cdots \\
&= e^{-\lambda(T-t)} \frac{e^{\lambda k\Delta} - 1}{1 - e^{-\lambda(T-t)}} = \frac{e^{-\lambda(T-t-\overline{\Delta})} - e^{-\lambda(T-t)}}{1 - e^{-\lambda(T-t)}}
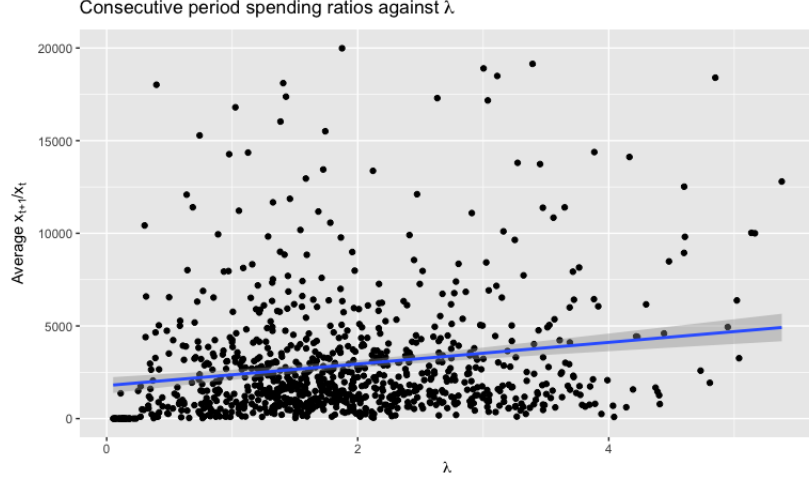\end{aligned}
$$

26

Figure B.1: Consecutive period spending ratios against estimates of λ from spending data.

which is the formula stated in the proposition. The proof of the case in which $\lambda = 0$ is similar and omitted.

B   Additional Analyses

A prediction of our model is that the ratio of spending in consecutive periods $r_\lambda(n)$ is increasing in λ. We test this prediction by plotting the average consecutive period spending against the maximum likelihood estimates of λ from spending data in Figure B.1. The figure shows an overall positive relationship.

In addition, we can use polling data to calibrate values of the long-term popularity mean, which in the equilibrium of our model is $q(X_0/Y_0)/\lambda$. Suppose that relative popularity data is collected N times at fixed intervals over a period T, with Δ normalized to 1. Equation (2) in the main text implies that we can calibrate $q(X_0/Y_0)/\lambda$ to minimize the sum of squared deviations of Z_{n+1} from its mean; i.e., to solve

$$\min_{q(X_0/Y_0)/\lambda} \sum_{n=0}^{N-1} \left| z_{n+1} - \left[(1 - e^{-\lambda})q(X_0/Y_0)/\lambda + e^{-\lambda}z_n \right] \right|^2, \tag{11}$$

which gives

$$q(X_0/Y_0)/\lambda = \frac{\frac{1}{N}\sum_{n=0}^{N-1} z_{n+1} - e^{-\lambda}z_n}{1 - e^{-\lambda}}.$$
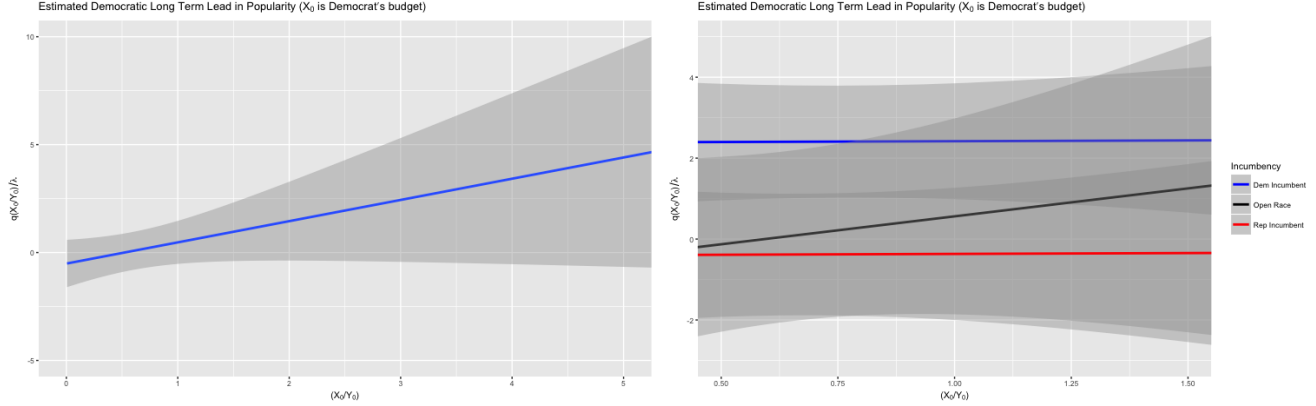
27

Figure B.2: Estimated $q(X_0/Y_0)/\lambda$ against budget ratios. The left figure shows the long term lead of the Democratic candidate. The right figure disaggregates races based on incumbency versus open seat.

To calibrate $q(X_0/Y_0)/\lambda$ values, we use the polling data from FiveThirtyEight's polls database. We need at least two weeks of polling data for each race to be able to calculate $q(X_0/Y_0)/\lambda$. Applying this criteria, we can match 389 elections in the FiveThirtyEight database to elections for which we have spending data. The polling is sparse: most races in our sample have two to four weeks of data. We take weekly averages of these polls to construct z_t as the average lead of the Democratic candidate in polls in week t. We use the maximum likelihood estimated λ values recovered from spending data in calculating $q(X_0/Y_0)/\lambda$, and we fit a line through the calculated $q(X_0/Y_0)/\lambda$ values.

We define the X_0/Y_0 ratio to be the ratio of the Democratic candidate's budget to the republican candidate's budget. So, $q(X_0/Y_0)/\lambda$ is interpreted as the long-run lead of the democratic candidate in polls. We restrict attention races with budget ratios in the $(0.1, 10)$ interval to focus on genuine contests, and plot the results in Figure B.2.

The left figure shows that the effect of financial advantage on the long term polling lead: increasing the ratio X_0/Y_0 has an increasing but statistically indiscernible effect on the long term lead. The right figure shows the effect of incumbency. For a given ratio of budgets X_0/Y_0, we can compare $q(X_0/Y_0)/\lambda$ in an open race with $q(X_0/Y_0)/\lambda$ in an election with an incumbent to understand the incumbency advantage in campaigning in terms of long term lead in popularity.

In an open race (black line), a Democratic candidate who spends the same amount as a Republican candidate seems to have no significant long term lead in the polls. Our point estimate is a 0.55 point lead in favor of the Democrat, with the 95% confidence interval spanning $[-1.85, +2.97]$. A Democratic incumbent (blue line) who spends the

same amount as a Republican challenger enjoys a long term lead in the polls by 2.41 points (95% CI =[0.97, 3.85]). Similarly, a Republican incumbent has a long term lead of 0.3 points ([2.01,-1.28]).

C   Extension on Evolving Budgets

C.1   Model

Instead of assuming that candidates are endowed with a fixed budget at the start of the game that they must allocate over time, we assume that the size of the resources stock also evolves in a way that depends on the evolution of popularity. In particular, we retain all the features of the baseline model except the ones described below.

Candidates start with exogenous budgets X_0 and Y_0 as in the baseline model. However, the budgets now evolve according to the following geometric Brownian motions

$$\frac{dX_t}{X_t} = az_t dt + \sigma_X dW_t^X$$
$$\frac{dY_t}{Y_t} = bz_t dt + \sigma_Y dW_t^Y$$

where a, b, σ_X and σ_Y are constants, and W_t^X and W_t^Y are Wiener processes. Under this assumption, the evolution of the budget is stochastic because it depends on the evolution of relative popularity, which is itself stochastic, and because there also an exogenous source of randomness. In addition, even though none of our results hinge on it, we make the assumption for simplicity that dW_t is independent of dW_t^X and of dW_t^Y, while dW_t^X and dW_t^Y have covariance $\rho \geq 0$.

This setting is general enough to allow for several possibilities. For example, donors may raise their support for candidate that is leading in the polls and withdraw support from the one that is trailing. This is the case where $b < 0 < a$. Alternatively, donors may channel their resources to the underdog, which is the case where $a < 0 < b$. Popularity therefore feeds back into the budget process. If the difference $a - b$ is positive, then the feedback is positive and if it is negative then the feedback is negative. We refer to a and b as the feedback parameters.[18]

---

[18]Also, note that dX_t and dY_t may be negative. One interpretation is that X_t and Y_t are expected total budgets available for the remainder of the campaign, where the expectation is formed at time t. Depending on the level of relative popularity, the candidates revise their expected future inflow of funds and adjust their spending choices accordingly.

All other features of the model are exactly the same as in the baseline model, including the process (1) governing the evolution of popularity, though we now assume for analytical tractability that

$$q(x/y) = \log(x/y).$$

C.2 Characterizing the Evolution of Spending Over Time

The baseline model suggests that the ratio of budgets, X_t/Y_t, plays an important role in characterizing the equilibrium. This is also the case for this extension. In fact, the main finding for this extension is that even though candidates' budgets now evolve stochastically over time, the main feature of equilibrium in the baseline model still holds here: the equilibrium spending ratios, x_t/X_t and y_t/Y_t at time t are equal; thus x_t/y_t is pinned down by the ratio of available budgets, X_t/Y_t, at time t.

Applying Itô's lemma, we can write the process governing the evolution of this ratio for this model as:

$$\frac{d(X_t/Y_t)}{X_t/Y_t} = \mu_{XY}(z_t)dt + \sigma_X dW_t^X - \sigma_Y dW_t^Y, \tag{12}$$

where

$$\mu_{XY}(z_t) = (a - b)z_t + \sigma_Y^2 - \rho\sigma_X\sigma_Y.$$

Hence, the instantaneous volatility of this process is simply $\sigma_{XY} = \sqrt{\sigma_X^2 + \sigma_Y^2 - \rho\sigma_X\sigma_Y}$. Therefore, if at time $t \in \mathcal{T}$ the candidates have an amount of available resources equal to X_t and Y_t and spend x_t and y_t, then $Z_{t+\Delta}$ conditional on all information, $\mathcal{I}_t$, available at time time t is a normal random variable:

$$Z_{t+\Delta} \mid \mathcal{I}_t \sim \mathcal{N}\left(\log\left(\frac{x_t}{y_t}\right)\frac{1 - e^{-\lambda\Delta}}{\lambda} + z_t e^{-\lambda\Delta}, \frac{\sigma^2(1 - e^{-2\lambda\Delta})}{2\lambda}\right),$$

and Itô's lemma implies that

$$\log\left(\frac{X_{t+\Delta}}{Y_{t+\Delta}}\right) \mid \mathcal{I}_t \sim \mathcal{N}\left(\log\left(\frac{X_t - x_t}{Y_t - y_t}\right) + \mu_{XY}(z_t)\Delta, \sigma_{XY}^2\Delta\right).$$

Then we have the following result, whose proof (along with all other proofs for this appendix) is given in Section C.4 below.

Proposition C.1. *In the model with evolving budgets, for every N, T, and $\lambda > 0$, there exists $-\eta < 0$ such that whenever $a - b \geq -\eta$, there is an essentially unique equilibrium.*

For all $t \in \mathcal{T}$, if $X_t, Y_t > 0$, then in all equilibria,

$$x_t/X_t = y_t/Y_t.$$

To understand the condition $a - b \geq -\eta$, note that when $a < 0 < b$, there is a negative feedback between popularity and the budget flow: a candidate's budget increases when she is less popular than her opponent. The condition $a - b \geq -\eta$ puts a bound on how negative this feedback can be. If this condition is not satisfied, candidates may want to *reduce* their popularity as much as they can in the early stages of the campaign to accumulate a larger war chest to use in the later stages. This could undermine the existence of an SPE in pure strategies.

As in the case of the baseline model, we can characterize the stochastic process of relative campaign spending for this model with evolving budgets as well.

To characterize the stochastic process of relative campaign spending, fix the number of periods N that the candidates take actions. Let $g_1(0) = 1$ and $g_2(0) = 0$, and define recursively for every $m \in \{1, ..., N-1\}$,

$$\begin{pmatrix} g_1(m) \\ g_2(m) \end{pmatrix} = \begin{pmatrix} e^{-\lambda\Delta} & a - b \\ \frac{1-e^{-\lambda\Delta}}{\lambda} & 1 \end{pmatrix} \begin{pmatrix} g_1(m-1) \\ g_2(m-1) \end{pmatrix} \tag{13}$$

Proposition C.2. *Let $t = (N-m)\Delta \in \mathcal{T}$ be a time at which $X_t, Y_t > 0$. Then, in the essentially unique equilibrium, spending ratios are equal to*

$$\frac{x_t}{X_t} = \frac{y_t}{Y_t} = \frac{g_1(m-1)}{g_1(m-1) + g_2(m-1)\frac{\lambda}{1-e^{-\lambda\Delta}}}. \tag{14}$$

Moreover, in equilibrium, $(\log(x_{t+n\Delta}/y_{t+n\Delta}), z_{t+n\Delta}) \mid \mathcal{I}_t$ follows a bivariate normal distribution with mean

$$\begin{pmatrix} 1 & (a-b)\Delta \\ \frac{1-e^{-\lambda\Delta}}{\lambda} & e^{-\lambda\Delta} \end{pmatrix}^n \begin{pmatrix} \log\left(\frac{X_t}{Y_t}\right) - \frac{\lambda\Delta(\sigma_Y^2 - \rho\sigma_X\sigma_Y)}{a-b}\frac{e^{-\lambda\Delta}}{1-e^{-\lambda\Delta}} \\ z_t + \frac{(\sigma_Y^2 - \rho\sigma_X\sigma_Y)}{a-b} \end{pmatrix} + \begin{pmatrix} \frac{\lambda(\sigma_Y^2 - \rho\sigma_X\sigma_Y)}{a-b}\frac{e^{-\lambda\Delta}}{1-e^{-\lambda\Delta}} \\ \frac{\rho\sigma_X\sigma_Y - \sigma_Y^2}{a-b} \end{pmatrix}$$

and variance

$$\begin{pmatrix} 1 & (a-b)\Delta \\ \frac{1-e^{-\lambda\Delta}}{\lambda} & e^{-\lambda\Delta} \end{pmatrix}^n \begin{pmatrix} \sigma_{XY}^2\Delta & 0 \\ 0 & \frac{\sigma^2(1-e^{-2\lambda\Delta})}{2\lambda} \end{pmatrix} \begin{pmatrix} 1 & \frac{1-e^{-\lambda\Delta}}{\lambda} \\ (a-b)\Delta & e^{-\lambda\Delta} \end{pmatrix}^n.$$

C.3 The Added Incentive to Spend Early

One question that we can ask of this extension is how the distribution of spending over time varies with the feedback parameters a and b that determine the rate of flow of candidates' budget in response to shifts in relative popularity. In the baseline model, when $\lambda > 0$ the difficulty in maintaining an early lead means that there is a disincentive to spend resources early on. This produces the result, reported in Proposition 3 and depicted in Figure 2, that spending is increasing over time. However, in this extension, if $b < 0 < a$ then there is a force working in the other direction: spending money to build early leads may advantageous because it results in faster growth of the war chest, which is valuable for the future. The disincentive to spend early is mitigated by this opposing force, and may even be overturned if a is much larger than b, i.e., if donors have a greater tendency to flock to the leading candidate.

We can establish this intuition formally. Recall that $r_n(\lambda)$ defined in the main text gave the ratio of equilibrium spending in consecutive periods, n and $n + 1$. For this extension with evolving budgets, we define the analogous ratio, $\tilde{r}_n$, which we show in the appendix depends on a and b only through the difference $a - b$ and is the same for both candidates. Specifically,

$$\tilde{r}_n(\lambda, a - b) = \frac{x_{(n+1)\Delta}/X_{(n+1)\Delta}}{x_{n\Delta}/X_{n\Delta}} = \frac{y_{(n+1)\Delta}/Y_{(n+1)\Delta}}{y_{n\Delta}/Y_{n\Delta}}$$

The following proposition establishes the key properties of this ratio, particularly its dependence on the feedback parameters, a and b.

Proposition C.3. *Fix the number of periods N, total time $T = N\Delta$, and consider the case in which $\lambda > 0$. Then, for all n, if $a - b$ is sufficiently small then the ratio $\tilde{r}_n(\lambda, a - b)$ of spending in consecutive periods n and $n + 1$ conditional on the history up to period n is (i) greater than 1, (ii) increasing in λ, and (iii) decreasing in $a - b$.*

The baseline model is the special case of the model with evolving budgets in which there is no budget volatility: $a = b = \sigma_X = \sigma_Y = 0$. What Proposition C.3 establishes is that starting with this special case, as we increase the difference $a - b$ from zero, spending plans becomes more balanced over time: there is a greater incentive to spend in earlier periods of the race than there is if $a = b$.

However, it is worth noting that Proposition C.3 does not necessarily hold when $a - b$ is very large. Indeed, we have examples in which $\tilde{r}_n(\lambda, a - b)$ is *increasing* in $a - b$ for

large λ, n, and $a - b$.[19] The intuition behind these examples rests on the fact that when the degree of mean reversion is high, then it is important for candidates to build up a large war chest that they can deploy in the final stages of the race. If the election date is distant and $a - b$ is large, then early spending is mostly for the purpose of building up these resources. But how should the candidates allocate their spending across the very early periods? Spending too much in any one period is risky: if the resource stock does not grow (or even if it grows but insufficiently) then there is less money, and hence less opportunity, to grow it in the subsequent periods. Since q is concave, the candidates would like to have many attempts to grow the war chest early on, and this is even more the case as the importance of the relative feedback $a - b$ gets large.

C.4 Proofs

Proof of Proposition C.1. Consider time $t = n\Delta \in \mathcal{T}$ and suppose that the game has arrived at time t with both candidates having a positive amount of resources still available, $X_t, Y_t > 0$. We will prove the proposition by induction on the times at which candidates take actions, $t = (N - m)\Delta \in \mathcal{T}$, $m = 1, 2, ..., N$. To simplify notation, it is convenient to define recursively the following expressions. Let

$$g_1(1) = e^{-\lambda\Delta}, \quad g_2(1) = \frac{1 - e^{-\lambda\Delta}}{\lambda}, \quad g_3(1) = 0, \quad g_4(1) = \frac{\sigma^2(1 - e^{-2\lambda\Delta})}{2\lambda}.$$

and define recursively for any $m > 1$

$$g_1(m) = e^{-\lambda\Delta}g_1(m - 1) + g_2(m - 1)(a - b);$$

$$g_2(m) = g_1(m - 1)\frac{1 - e^{-\lambda\Delta}}{\lambda} + g_2(m - 1);$$

$$g_3(m) = g_2(m - 1) + g_3(m - 1).$$

$$g_4(m) = (g_1(m - 1))^2\frac{\sigma^2(1 - e^{-2\lambda\Delta})}{2\lambda} + (g_2(m - 1))^2\sigma_{XY}^2\Delta + g_4(m - 1)$$

Note that this gives us the following relationship in (13), so by diagonalizing the matrix

$$\begin{pmatrix} e^{-\lambda\Delta} & a - b \\ \frac{1 - e^{-\lambda\Delta}}{\lambda} & 1 \end{pmatrix}$$

---

[19] One example is $\lambda = 0.8$, $\Delta = 0.9$, and $n = a - b = 10$.

and solving for $(g_1(m), g_2(m))'$ with initial conditions $g_0(1) = 1$ and $g_2(0) = 0$, we can conclude that, for each $N \in \mathbb{N}$ and $\lambda, \Delta > 0$, there exists $-\eta < 0$ such that, for each $a - b \geq -\eta$, both $g_1(m)$ and $g_2(m)$ are non-negative for each $m = 1, ..., N$. For the rest of the proof, we assume $g_1(m) \geq 0$ and $g_2(m) \geq 0$.

Consider the following inductive hypothesis: for every $s = (N - m)\Delta \in \mathcal{T}$, $m \in \{1, ..., N\}$, if $X_s, Y_s > 0$, then

(i) the continuation payoff of each candidate is a function of current popularity z_s, current budget ratio X_s/Y_s and calendar time s;

(ii) the distribution of Z_T given z_s and X_s/Y_s is $\mathcal{N}\left(\hat{\mu}_{(N-m)\Delta}(z_s), \hat{\sigma}^2_{(N-m)\Delta}\right)$, where

$$\hat{\mu}_{(N-m)\Delta}(z_{(N-m)\Delta}) = g_1(m) z_{(N-m)\Delta} + g_2(m) \log\left(\frac{X_{(N-m)\Delta}}{Y_{(N-m)\Delta}}\right) + g_3(m)(\sigma_Y^2 - \rho\sigma_X\sigma_Y),$$

and $\hat{\sigma}^2_{(N-m)\Delta} = g_4(m)$.

Suppose the game reaches period $t = (N - 1)\Delta$ and both candidates still have a positive amount of resources, $X_{(N-1)\Delta}, Y_{(N-1)\Delta} > 0$. Because money left at time T is useless, candidates spend all their remaining resources at t, $x_{(N-1)\Delta} = X_{(N-1)\Delta}$ and $y_{(N-1)\Delta} = Y_{(N-1)\Delta}$. Hence, $x_{(N-1)\Delta}/y_{(N-1)\Delta} = X_{(N-1)\Delta}/Y_{(N-1)\Delta}$ and

$$Z_T \mid \mathcal{I}_{(N-1)\Delta} \sim \mathcal{N}\left(\log\left(\frac{X_{(N-1)\Delta}}{Y_{(N-1)\Delta}}\right)\frac{1 - e^{-\lambda\Delta}}{\lambda} + z_{(N-1)\Delta}e^{-\lambda\Delta}, \frac{\sigma^2(1 - e^{-2\lambda\Delta})}{2\lambda}\right).$$

Because Z_T fully determines candidates' payoffs, the continuation payoff of candidates is a function of current popularity $z_{(N-1)\Delta}$, the ratio $X_{(N-1)\Delta}/Y_{(N-1)\Delta}$ and calendar time. Furthermore, at $t = (N - 1)\Delta$ the distribution of Z_T given $z_{t+\Delta}$ is a degenerate distribution on $z_{t+\Delta}$. Given the recursive expressions defined above, we can conclude that the second part of the inductive hypothesis also holds at $t = (N - 1)\Delta$. This concludes the base step.

Suppose the inductive hypothesis holds true at any time $s = (N - m)\Delta \in \mathcal{T}$ with $m \in \{1, 2, ..., m^* - 1\}$, $m^* < N$. We want to show that at time $t = (N - m^*)\Delta \in \mathcal{T}$, if $X_t, Y_t > 0$, then (i) an equilibrium exists, and in all equilibria, $x_t/y_t = X_t/Y_t$ and the continuation payoffs of both candidates are functions of relative popularity z_t, the ratio X_t/Y_t, and calendar time t, and (ii) Z_T given period t information is distributed according to $\mathcal{N}\left(\hat{\mu}_{(N-m^*)\Delta}(z_t), \hat{\sigma}^2_{(N-m^*)\Delta}\right)$.

Consider period $N - m^*$ and let $x, y > 0$ be the candidate's spending levels in this period. Exploiting the inductive hypothesis, we can compound the normal distributions and conclude that $Z_T \mid \mathcal{I}_t \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$ where

$$
\begin{aligned}
\tilde{\mu} = {} & g_1(m^* - 1) \left[\log\left(\frac{x}{y}\right) \frac{1 - e^{-\lambda\Delta}}{\lambda} + z_t e^{-\lambda\Delta} \right] \\
& + g_2(m^* - 1) \left[\log\left(\frac{X_{(N-m^*)\Delta} - x}{Y_{(N-m^*)\Delta} - y}\right) + \mu_{XY}(z_t)\Delta \right] + g_3(m^* - 1)(\sigma_Y^2 - \rho\sigma_X\sigma_Y) \\
\tilde{\sigma}^2 = {} & (g_1(m^* - 1))^2 \frac{\sigma^2(1 - e^{-2\lambda\Delta})}{2\lambda} + (g_2(m^* - 1))^2 \sigma_{XY}^2 \Delta + g_4(m^* - 1)
\end{aligned}
$$

Then, note that we can write

$$
\tilde{\mu} = \hat{\mu}_t(x, y) := G_1 \log\left(\frac{x}{y}\right) + G_2 \log\left(\frac{X_{(N-m^*)\Delta} - x}{Y_{(N-m^*)\Delta} - y}\right) + G_3,
$$

where

$$
\begin{aligned}
G_1 &= g_1(m^* - 1) \frac{1 - e^{-\lambda\Delta}}{\lambda} \\
G_2 &= g_2(m^* - 1) \\
\text{and } G_3 &= g_1(m^* - 1) z_t e^{-\lambda\Delta} + g_2(m^* - 1)\mu_{XY}(z_t)\Delta + g_3(m^* - 1)(\sigma_Y^2 - \rho\sigma_X\sigma_Y) \quad (15)
\end{aligned}
$$

Furthermore $\tilde{\sigma}^2$ is independent of x and y.

Candidate 1 wins the election if $Z_T > 0$. Thus, in equilibrium he chooses x to maximize his winning probability

$$
\int_{-\frac{\hat{\mu}_t(x,y)}{\tilde{\sigma}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-s/2} ds.
$$

The first order necessary condition for x is given by

$$
\frac{1}{\sqrt{2\pi}} e^{\frac{\hat{\mu}_t(x,y)}{2\tilde{\sigma}_t}} \frac{\hat{\mu}_t'(x,y)}{\tilde{\sigma}} = \frac{1}{\sqrt{2\pi}\tilde{\sigma}} e^{\frac{\hat{\mu}_t(x,y)}{2\tilde{\sigma}}} \left[\frac{G_1(X_t - x) - G_2 x}{x(X_t - x)} \right].
$$

Furthermore, when the first order necessary condition holds, the second order condition is given by

$$
\frac{1}{\sqrt{2\pi}} e^{\frac{\hat{\mu}_t(x,y)}{2\tilde{\sigma}_t}} \frac{\hat{\mu}''_t(x,y)}{\tilde{\sigma}} = \frac{-1}{\sqrt{2\pi}} e^{\frac{\hat{\mu}_t(x,y)}{2\tilde{\sigma}}} \left[\frac{G_1(X_t - x)^2 + G_2 x^2}{x^2(X_t - x)^2} \right] < 0.
$$

Hence, the problem is strictly quasi-concave for candidate 1 for each y. A symmetric argument shows that the corresponding problem for candidate 2 is strictly quasi-concave for each x. Hence an equilibrium exists and the optimal investment of the two candidates is pinned down by the first order necessary condition. Solving the first order condition, we get

$$\frac{x_t}{X_t} = \frac{y_t}{Y_t} = \frac{G_1}{G_1 + G_2}. \tag{16}$$

∎

Proof of Proposition C.2. From the equilibrium condition in (16), we know that the distribution of

$$Z_T \mid \mathcal{I}_t,$$

and hence the continuation payoffs of candidates at times $t = (N - m)\Delta$, are functions only of current popularity z_t, the current budget ratio X_t/Y_t, and calendar time t. Furthermore, substituting $\mu_{XY}(z_t) = (a - b)z_t + \sigma_Y^2 - \rho\sigma_X\sigma_Y$ in the distribution of

$$Z_T \mid \mathcal{I}_{(N-m^*)\Delta},$$

the equilibrium condition also implies that the mean of the distribution is

$$
\begin{aligned}
\tilde{\mu} = {}& [g_1(m^* - 1)e^{-\lambda\Delta} + g_2(m^* - 1)(a - b)]z_t \\
& + \left[g_1(m^* - 1)\frac{1 - e^{-\lambda\Delta}}{\lambda} + g_2(m^* - 1)\right]\log\left(\frac{X_{(N-m^*)\Delta}}{Y_{(N-m^*)\Delta}}\right) \\
& + [g_2(m^* - 1) + g_3(m^* - 1)](\sigma_Y^2 - \rho\sigma_X\sigma_Y).
\end{aligned}
$$

Given the expressions recursively defined above, we conclude that

$$Z_T \mid \mathcal{I}_{(N-m^*)\Delta} \sim \mathcal{N}(\hat{\mu}_{(N-m^*)\Delta}, \hat{\sigma}^2_{(N-m^*)\Delta})$$

where

$$\hat{\mu}_{(N-m^*)\Delta}(z_{(N-m^*)\Delta}) = g_1(m^*)z_{(N-m^*)\Delta} + g_2(m^*)\log\left(\frac{X_{(N-m)\Delta}}{Y_{(N-m)\Delta}}\right) + g_3(m^*)(\sigma_Y^2 - \rho\sigma_X\sigma_Y),$$

$$\hat{\sigma}^2_{(N-m^*)\Delta} = g_4(m^*).$$

36

The expression for x_t/X_t and y_t/Y_t in the proposition thus follows from (13), (15), and (16). To derive the distribution, we start by using the findings from our proof of Proposition C.1 above to derive the distribution of $x_{t+j\Delta}/y_{t+j\Delta}$ and $z_{t+j\Delta}$ given x_t/y_t and z_t. Note that, since $X_t/Y_t = x_t/y_t$ for each t, we can recursively write that

$$\left(\begin{array}{c} \log\left(\frac{x_{t+n\Delta}}{y_{t+n\Delta}}\right) \\ z_{t+n\Delta} \end{array} \right) \Bigg| \left(\begin{array}{c} \frac{x_{t+(n-1)\Delta}}{y_{t+(n-1)\Delta}} \\ z_{t+(n-1)\Delta} \end{array} \right)$$

follows the multivariate normal distribution

$$\mathcal{N}\left(\left(\begin{array}{c} \log\left(\frac{x_{t+(n-1)\Delta}}{y_{t+(n-1)\Delta}}\right) + \mu_{XY}(z_{t+(n-1)\Delta})\Delta \\ \log\left(\frac{x_{t+(n-1)\Delta}}{y_{t+(n-1)\Delta}}\right)\frac{1-e^{-\lambda\Delta}}{\lambda} + z_{t+(n-1)\Delta}e^{-\lambda\Delta} \end{array} \right), \left(\begin{array}{cc} \sigma_{XY}^2\Delta & 0 \\ 0 & \frac{\sigma^2(1-e^{-2\lambda\Delta})}{2\lambda} \end{array} \right) \right)$$

Therefore,

$$\left(\begin{array}{c} \log\left(\frac{x_{t+n\Delta}}{y_{t+n\Delta}}\right) - \frac{\lambda(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b}\frac{e^{-\lambda\Delta}}{1-e^{-\lambda\Delta}} \\ z_{t+n\Delta} + \frac{(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b} \end{array} \right) \Bigg| \left(\begin{array}{c} \log\left(\frac{x_{t+(n-1)\Delta}}{y_{t+(n-1)\Delta}}\right) - \frac{\lambda(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b}\frac{e^{-\Delta\lambda}}{1-e^{-\lambda\Delta}} \\ z_{t+(n-1)\Delta} + \frac{(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b} \end{array} \right)$$

follows the multivariate normal distribution

$$\mathcal{N}\left(\left(\begin{array}{cc} 1 & (a-b)\Delta \\ \frac{1-e^{-\lambda\Delta}}{\lambda} & e^{-\lambda\Delta} \end{array} \right) \left(\begin{array}{c} \log\left(\frac{x_{t+(n-1)\Delta}}{y_{t+(n-1)\Delta}}\right) - \frac{\lambda\Delta(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b}\frac{e^{-\lambda\Delta}}{1-e^{-\lambda\Delta}} \\ z_{t+(n-1)\Delta} + \frac{(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b} \end{array} \right), \left(\begin{array}{cc} \sigma_{XY}^2\Delta & 0 \\ 0 & \frac{\sigma^2(1-e^{-2\lambda\Delta})}{2\lambda} \end{array} \right) \right).$$

Therefore,

$$\left(\begin{array}{c} \log\left(\frac{x_{t+n\Delta}}{y_{t+n\Delta}}\right) - \frac{\lambda(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b}\frac{e^{-\lambda\Delta}}{1-e^{-\lambda\Delta}} \\ z_{t+n\Delta} + \frac{(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b} \end{array} \right) \Bigg| \left(\begin{array}{c} \log\left(\frac{x_t}{y_t}\right) - \frac{\lambda\Delta(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b}\frac{e^{-\lambda\Delta}}{1-e^{-\lambda\Delta}} \\ z_t + \frac{(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b} \end{array} \right)$$

follows the multivariate normal distribution with mean

$$\left(\begin{array}{cc} 1 & (a-b)\Delta \\ \frac{1-e^{-\lambda\Delta}}{\lambda} & e^{-\lambda\Delta} \end{array} \right)^n \left(\begin{array}{c} \log\left(\frac{X_t}{Y_t}\right) - \frac{\lambda\Delta(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b}\frac{e^{-\lambda\Delta}}{1-e^{-\lambda\Delta}} \\ z_t + \frac{(\sigma_Y^2-\rho\sigma_X\sigma_Y)}{a-b} \end{array} \right)$$

and variance

$$\left(\begin{array}{cc} 1 & (a-b)\Delta \\ \frac{1-e^{-\lambda\Delta}}{\lambda} & e^{-\lambda\Delta} \end{array} \right)^n \left(\begin{array}{cc} \sigma_{XY}^2\Delta & 0 \\ 0 & \frac{\sigma^2(1-e^{-2\lambda\Delta})}{2\lambda} \end{array} \right) \left(\begin{array}{cc} 1 & \frac{1-e^{-\lambda\Delta}}{\lambda} \\ (a-b)\Delta & e^{-\lambda\Delta} \end{array} \right)^n.$$

37

$\blacksquare$

Proof of Proposition C.3. Fix λ and Δ and let $N - m = n$. We must show that for all $m \in \{1, ..., N - 1\}$,

$$\hat{r}_m(a - b) = \frac{x_{m\Delta}}{X_{m\Delta}} \Big/ \frac{x_{(m+1)\Delta}}{X_{(m+1)\Delta}}$$

is decreasing in $\alpha := a - b$ around $\alpha = 0$. (Note that $\hat{r}_m$ is the same as $\tilde{r}_{N-m}$ but $\hat{r}_m$ counts time backwards.)

Since we can write

$$\hat{r}_m(\alpha) = \frac{\dfrac{g_1\left(m - 1\right)}{g_1\left(m - 1\right) + g_2\left(m - 1\right)\frac{\lambda}{1 - e^{-\lambda\Delta}}}}{\dfrac{g_1\left(m\right)}{g_1\left(m\right) + g_2\left(m\right)\frac{\lambda}{1 - e^{-\lambda\Delta}}}} = \frac{g_1\left(m - 1\right)}{g_1\left(m\right)}\frac{g_2\left(m + 1\right)}{g_2\left(m\right)},$$

we first relate $\left(g_1\left(m\right), g_2\left(m + 1\right)\right)$ to $\left(g_1\left(m - 1\right), g_2\left(m\right)\right)$. From

$$\left(\begin{array}{c} g_1\left(m\right) \\ g_2\left(m\right) \end{array}\right) = \left(\begin{array}{c} e^{-\lambda\Delta}g_1\left(m - 1\right) + \alpha g_2\left(m - 1\right) \\ \frac{1 - e^{-\lambda\Delta}}{\lambda}g_1\left(m - 1\right) + g_2\left(m - 1\right) \end{array}\right),$$

we derive

$$g_1\left(m\right) = \frac{\left(\lambda + \alpha\right)e^{-\lambda\Delta} - \alpha}{\lambda}g_1\left(m - 1\right) + \alpha g_2\left(m\right)$$

and

$$g_2\left(m + 1\right) = \frac{1 - e^{-\lambda\Delta}}{\lambda}\frac{\left(\lambda + \alpha\right)e^{-\lambda\Delta} - \alpha}{\lambda}g_1\left(m - 1\right) + \frac{\alpha - \alpha e^{-\lambda\Delta} + \lambda}{\lambda}g_2\left(m\right).$$

Hence,

$$\left(\begin{array}{c} g_1\left(m\right) \\ g_2\left(m + 1\right) \end{array}\right) = \left(\begin{array}{c} \frac{\left(\lambda+\alpha\right)e^{-\lambda\Delta}-\alpha}{\lambda}g_1\left(m - 1\right) + \alpha g_2\left(m\right) \\ \frac{1-e^{-\lambda\Delta}}{\lambda}\frac{\left(\lambda+\alpha\right)e^{-\lambda\Delta}-\alpha}{\lambda}g_1\left(m - 1\right) + \frac{\alpha-\alpha e^{-\lambda\Delta}+\lambda}{\lambda}g_2\left(m\right) \end{array}\right). \qquad (17)$$

Substituting in the expression for $\hat{r}_m(\alpha)$ and simplifying, we get

$$\hat{r}_m(\alpha) = \frac{1}{\frac{\left(\lambda+\alpha\right)e^{-\lambda\Delta}-\alpha}{\lambda} + \alpha g_m}\left(\frac{1 - e^{-\lambda\Delta}}{\lambda}\frac{\left(\lambda + \alpha\right)e^{-\lambda\Delta} - \alpha}{\lambda}\frac{1}{g_m} + \frac{\alpha + \lambda - \alpha e^{-\lambda\Delta}}{\lambda}\right) \qquad (18)$$

38

where $g_m := g_2(m)/g_1(m-1)$. From (17), we further have

$$g_{m+1} = \frac{\frac{1-e^{-\lambda\Delta}}{\lambda}\frac{(\lambda+\alpha)e^{-\lambda\Delta}-\alpha}{\lambda} + \frac{\alpha+\lambda-\alpha e^{-\lambda\Delta}}{\lambda}g_m}{\frac{(\lambda+\alpha)e^{-\lambda\Delta}-\alpha}{\lambda} + \alpha g_m}. \tag{19}$$

Computing (18) one step forward and substituting for g_{m+1} as obtained from (19) and, subsequently, for g_m as obtained from (18), we get $\hat{r}_{m+1}$ as a function of α and $\hat{r}_m$:

$$\hat{r}_{m+1}(\alpha,\hat{r}_m)$$

$$= \frac{\left(\begin{array}{c}\alpha\left(-2e^{\lambda\Delta}+(\hat{r}_m-3)\left(-e^{2\Delta\lambda}\right)+\hat{r}_m-1\right)+\lambda\left(e^{\lambda\Delta}\left(e^{\lambda\Delta}+\hat{r}_m-1\right)+\hat{r}_m\right)+ \\ +(1+e^{\lambda\Delta})\sqrt{(\alpha(\hat{r}_m+1)(e^{\Delta\lambda}-1)+\lambda(e^{\lambda\Delta}-\hat{r}_m))^2-4\alpha\hat{r}_m(e^{\lambda\Delta}-1)(\alpha(e^{\lambda\Delta}-1)-\lambda)}\end{array}\right)}{\left(\begin{array}{c}\alpha\left(e^{\lambda\Delta}-1\right)\left((\hat{r}_m+1)e^{\Delta\lambda}-\hat{r}_m+1\right)+\lambda\left(e^{\lambda\Delta}\left(e^{\lambda\Delta}-\hat{r}_m+1\right)+\hat{r}_m\right)+ \\ +(1+e^{\lambda\Delta})\sqrt{(\alpha(\hat{r}_m+1)(e^{\Delta\lambda}-1)+\lambda(e^{\lambda\Delta}-\hat{r}_m))^2-4\alpha\hat{r}_m(e^{\lambda\Delta}-1)(\alpha(e^{\lambda\Delta}-1)-\lambda)}\end{array}\right)}.$$

We first show that $\hat{r}_m > e^{\lambda\Delta} > 1$ for each m around $\alpha = 0$. For $m = 1$, since we have $\frac{x_{(N-1)\Delta}}{X_{(N-1)\Delta}} = 1$ and $\frac{x_{(N-2)\Delta}}{X_{(N-2)\Delta}} = \frac{g_1(1)+g_2(1)\frac{\lambda}{1-e^{-\lambda\Delta}}}{g_1(1)}$, we have

$$\hat{r}_1 - e^{\lambda\Delta} = 1 + \frac{\frac{1-e^{-\lambda\Delta}}{\lambda}}{e^{-\lambda\Delta}}\frac{\lambda}{1-e^{-\lambda\Delta}} - e^{\lambda\Delta} = 1,$$

so $\hat{r}_1 > 1$. Given $\hat{r}_m > e^{\lambda\Delta}$, subtracting $e^{\lambda\Delta}$ from the right hand side of the expression of $\hat{r}_{m+1}$ and setting $\alpha = 0$, we get

$$\hat{r}_{m+1} - e^{\lambda\Delta} = 1 - \frac{e^{\lambda\Delta}}{\hat{r}_m} > 0.$$

Hence, if $\hat{r}_m > e^{\lambda\Delta}$, then $\hat{r}_{m+1} > e^{\lambda\Delta}$. Therefore, we have $\hat{r}_m > e^{\lambda\Delta}$ for each m in a neighborhood of $\alpha = 0$.

Now we prove that $\hat{r}'_m(\alpha) < 0$ around $\alpha = 0$. To this end, observe that $\hat{r}_{m+1}(\alpha,\hat{r}_m)$ is decreasing in α and increasing in $\hat{r}_m$:

$$\frac{\partial \hat{r}_{m+1}(\alpha,\hat{r}_m)}{\partial \alpha}\bigg|_{\alpha=0} = -\frac{(\hat{r}_m-1)e^{\lambda\Delta}\left(e^{2\lambda\Delta}-1\right)}{\hat{r}_m(\hat{r}_m-e^{\lambda\Delta})} < 0;$$

$$\frac{\partial \hat{r}_{m+1}(\alpha,\hat{r}_m)}{\partial \hat{r}_m}\bigg|_{\alpha=0} = \frac{e^{\lambda\Delta}}{(\hat{r}_m)^2} > 0.$$

39

Hence, for each m, $\hat{r}_m(\alpha)$ is decreasing in α inductively.

Finally, now viewing $\hat{r}_m$ as a function of λ as well, we have

$$\left.\frac{\partial \hat{r}_{m+1}(\alpha, \hat{r}_m, \lambda)}{\partial \lambda}\right|_{\alpha=0} = \frac{e^{\lambda\Delta}(\hat{r}_m - 1)\Delta}{\hat{r}_m} > 0 \text{ for each } \lambda > 0.$$

Hence, for each m, $\hat{r}_m$ is increasing in λ inductively near $\alpha = 0$. ∎

D   The 2012 Obama-Romney Presidential Race

In this appendix we demonstrate how our model can speak to presidential races and use it to study the 2012 Obama-Romney presidential race.

One advantage of studying a presidential contest like this is the richness of polling data enables us to more precisely estimate the decay rate from a time series analysis of the relative popularity numbers. In addition, this election is particularly salient because it is, to date, the most expensive electoral contest in history and was an extremely competitive race, both financially and in terms of the candidates' polling performance. Each candidate raised approximately $1 billion and spent close to their entire war chests by election day: Obama had spent $986 million and had only $28 million cash on hand by the end of November (when most, but not all, payments to campaign staff were made), while Romney had spent $992 million and had $29 million cash on hand.

Figure D.1 reveals how both candidates increased their spending over time, ramping it up in the final months, and also shows how spending in each month is roughly equal for the two candidates. Figure D.2 shows the competitiveness in polling for the two candidates as well, though it suggests an overall polling advantage for Obama.

To study this election, we start by presenting a framework that extends our baseline model to multi-district contests that covers the electoral college. We then present an estimation strategy, along with some additional assumptions that are required to study aggregate spending data across these multiple districts. We end by presenting the estimates of the popularity process, and comment on the findings.

D.1 Multidistrict Contests

We now provide an extension to address the possibility that the candidates compete in S winner-take-all districts (rather than a single district) and each must win a certain subset of these to win the electoral contest.[20]

Relative popularity in each district s is the random variable Z_t^s with realizations z_t^s, and we assume that the joint distribution of the vector $(Z_{t+1}^s)_{s=1}^S$ depends on $(x_t^s/y_t^s, z_t^s)_{s=1}^S$ only. This allows for correlation of relative popularity across districts.

All other structural features are the same as in the baseline model. In particular, to close this version of the model, we assume that if a candidate stop spending money in a particular district, then she loses the election right away if the other candidate is spending a positive amount in all districts and she wins the election with probability $1/2$ if the other candidate does not campaign in at least one district as well.

Proposition D.1. *In any equilibrium of this extension, if $X_t, Y_t > 0$ are the remaining budgets of candidates 1 and 2 at any time $t \in \mathcal{T}$, then for all districts s,*

$$x_t^s/X_t = y_t^s/Y_t.$$

The key implication of this result is that the total spending of each of the two candidates across all districts at a given time also respects the equal spending ratio result: if $x_s := \sum_s x_t^s$ is candidate 1's total spending at time t and $y_t := \sum_s y_t^s$ is candidate 2's then the proposition above implies $x_t/X_t = y_t/Y_t$. We can use this fact to study the aggregate spending pattern across U.S. states of each candidate in presidential contests despite the contest being subject to the electoral college.

Proof of Proposition D.1. Note that the game ends in a immediate defeat for any candidate that spends 0 in any district in any period. Therefore, in equilibrium we must have an interior spending for any district and any period.

Given this, we will prove the proposition by induction. Consider the final period as the basis case. Fix $(z_{T-\Delta}^s)_{s=1}^S$ arbirarily. Suppose candidates 1 and 2 have budgets X and Y, respectively in the last period. Fix an equilibrium strategy profile $(x_{T-\Delta}^{s,*}, y_{T-\Delta}^{s,*})_{s=1}^S$. We show that, if they have budget of βX and βY, then $(\beta x_{T-\Delta}^{s,*}, \beta y_{T-\Delta}^{s,*})_{s=1}^S$ is an equi-

---

[20]For example, if the set of districts is $\mathcal{S} = \{1, ..., S\}$ then consider any electoral rule such that for all partitions of $\mathcal{S}$ of the form $\{\mathcal{S}_1, \mathcal{S}_2\}$, either candidate 1 wins if she wins all the districts in $\mathcal{S}_1$ or candidate 2 wins if she wins all the districts in $\mathcal{S}_2$. The rule should be monotonic in the sense that for any partitions $\{\mathcal{S}_1, \mathcal{S}_2\}$ and $\{\mathcal{S}_1', \mathcal{S}_2'\}$ if candidate i wins by winning $\mathcal{S}_i$ then i wins by winning $\mathcal{S}_i' \supseteq \mathcal{S}_i$.

librium. This implies that the equilibrium payoff in the last period is determined by $\left(z_{T-\Delta}^s\right)_{s=1}^S$ and $X_{t-\Delta}/Y_{t-\Delta}$.

Suppose otherwise. Without loss, assume that there is $\left(\tilde{x}_{T-\Delta}^{s,*}\right)_{s=1}^S$ such that it gives a higher probability of winning to candidate 1 given $\left(z_{T-\Delta}^s\right)_{s=1}^S$ and $\beta y_{T-\Delta}^{s,*}$, satisfying $\sum_{s=1}^S \tilde{x}_{T-\Delta}^{s,*} \le \beta X$. Since the distribution of $(Z_T^s)_{s=1}^S$ is determined by $\left(z_{T-\Delta}^s\right)_{s=1}^S$ and $\left(x_{t-\Delta}^s/y_{t-\Delta}^s\right)_{s=1}^S$, this means that the distribution of $(Z_T^s)_{s=1}^S$ given $\left(z_{T-\Delta}^s\right)_{s=1}^S$ and $\left(\tilde{x}_{t-\Delta}^s/\beta y_{t-\Delta}^{*,s}\right)_{s=1}^S$ is more favorable to candidate 1 than that given $\left(z_{T-\Delta}^s\right)_{s=1}^S$ and

$$\left(\beta x_{t-\Delta}^{*,s}/\beta y_{t-\Delta}^{*,s}\right)_{s=1}^S = \left(x_{t-\Delta}^{*,s}/y_{t-\Delta}^{*,s}\right)_{s=1}^S.$$

On the other hand, candidate 1 could take $\left(\frac{1}{\beta}\tilde{x}_{T-\Delta}^{s,*}\right)_{s=1}^S$ when the budgets are (X,Y). Since $(x_{T-\Delta}^{s,*}, y_{T-\Delta}^{s,*})_{s=1}^S$ is an equilibrium, the distribution of $(Z_T^s)_{s=1}^S$ given $(z_{T-\Delta}^s)_{s=1}^S$ and $\left(\frac{1}{\beta}\tilde{x}_{t-\Delta}^s/y_{t-\Delta}^{*,s}\right)_{s=1}^S = \left(\tilde{x}_{t-\Delta}^s/\beta y_{t-\Delta}^{*,s}\right)_{s=1}^S$ is no more favorable to candidate 1 than that given $(z_{T-\Delta}^s)_{s=1}^S$ and $(x_{t-\Delta}^{*,s}/y_{t-\Delta}^{*,s})_{s=1}^S$. This is a contradiction.

Now consider the inductive step. Take the inductive hypothesis to be that the continuation payoff for either candidate in period $t \in \mathcal{T}$ can be written as a function of only the budget ratio X_{t+1}/Y_{t+1} and vector $\left(z_{t+1}^s\right)_{s=1}^S$ and candidates spend a positive amount in each district and in each following period. We have to show that $x_t^s/X_t = y_t^s/Y_t$.

Denote the continuation payoff of candidate 1 in period t with $W_{t+1}\left(\frac{X_{t+1}}{Y_{t+1}}, (z_{t+1}^s)_{s=1}^S\right)$. Candidate 1's objective is

$$\max_{x_t} \int W_{t+1}\left(\frac{X_t - \sum_{s=1}^S x_t^s}{Y_t - \sum_{s=1}^S y_t^s}, (z_{t+1}^s)_{s=1}^S\right) f_t\left((z_{t+1}^s)_{s=1}^S \mid \left(\frac{x_t^s}{y_t^s}\right)_{s=1}^S, (z_{t+1}^s)_{s=1}^S\right) dz_{t+1}.$$

The first order condition for an interior optimum is

$$\frac{1}{Y_t - \sum_{s=1}^S y_t^s} \int \frac{\partial W_{t+1}\left((X_t - x_t)/(Y_t - y_t), z_{t+1}\right)}{\partial(x_t^s/y_t^s)} f_t\left(z_{t+1} \mid \frac{x_t}{y_t}, z_t\right) dz_{t+1} =$$

$$= \frac{1}{y_t^s} \int W_{t+1}\left(\frac{X_t - x_t}{Y_t - y_t}, z_{t+1}\right) \frac{\partial f_t\left(z_{t+1} \mid (x_t^s/y_t^s)_{s=1}^S, z_t\right)}{\partial(x_t^s/y_t^s)} dz_{t+1}.$$

Similarly, the objective for candidate 2 is

$$\min_{y_t} \int W_{t+1}\left(\frac{X_t - x_t}{Y_t - y_t}, z_{t+1}\right) f_t\left(z_{t+1} \mid \frac{x_t}{y_t}, z_t\right) dz_{t+1}.$$

42

and the corresponding first order condition is

$$
\frac{X_t - \sum_{s=1}^{S} x_t^s}{\left(Y_t - \sum_{s=1}^{S} y_t^s\right)^2} \int \frac{\partial W_{t+1}\left((X_t - x_t)/(Y_t - y_t), z_{t+1}\right)}{\partial (x_t^s/y_t^s)} f_t\left(z_{t+1} \mid \frac{x_t}{y_t}, z_t\right) dz_{t+1}
$$

$$
= \frac{x_t^s}{(y_t^s)^2} \int W_{t+1}\left(\frac{X_t - x_t}{Y_t - y_t}, z_{t+1}\right) \frac{\partial f_t\left(\left(z_{t+1}^s\right)_{s=1}^{S} \mid (x_t^s/y_t^s)_{s=1}^{S}, (z_t^s)_{s=1}^{S}\right)}{\partial (x_t^s/y_t^s)} dz_{t+1}.
$$

Dividing the candidate 1's first order condition by candidate 2's, we have

$$
\frac{X_t - \sum_{s=1}^{S} x_t^s}{Y_t - \sum_{s=1}^{S} y_t^s} = \frac{x_t^s}{y_t^s}.
$$

Hence there exists β such that $x_t^s = \beta y_t^s$ for each s, and so

$$
\beta = \frac{X_t - \beta \sum_{s=1}^{S} y_t^s}{Y_t - \sum_{s=1}^{S} y_t^s},
$$

which implies $\beta = X_t/Y_t$. Therefore, $x_t^s/y_t^s = X_t/Y_t$ for each s. ∎

D.2   Estimation Strategy

The extension shows how under mild conditions the equal spending ratio result holds district-by-district. To be able to analyze district-aggregated spending data, we make additional assumptions on the popularity process that enable us to estimate its parameters. In particular, we assume that for each district s, relative popularity for that district evolves according to the process

$$
dZ_t^s = \left(q(x_{\underline{t}}^s/y_{\underline{t}}^s) - \lambda^s Z_t^s\right)dt + \sigma^s dW_t^s
$$

where λ^s represents the speed of reversion in each district s; W_t^s is a district-specific Wiener process; and σ^s is the overall district-specific volatility. The notation $\underline{t} := \max\{\tau \in \mathcal{T} : \tau \leq t\}$ is the same as in the baseline model. We also assume for that

1. process W_t^s is independent of process $W_t^{s'}$ for all $s \neq s'$

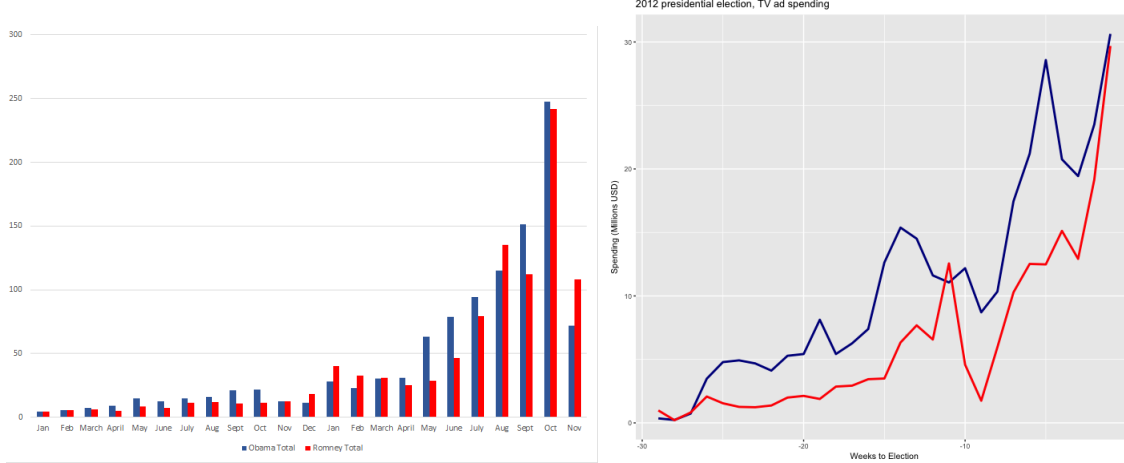2. $\lambda^s = \lambda^{s'}$ for all s, s'

Figure D.1: Monthly spending by Obama and Romney in millions of dollars for the 2012 presidential race. The left figure is total spending data from the Federal Election Commission as reported by the *New York Times*. Obama in blue and Romney in red. The right figure is TV ad spending data collected from the Wesleyan Advertising Database.

and we let λ denote the common value of the speed of reversion in all districts. These are strong assumptions that we make for convenience.

We let $Z_t := \sum_s \pi^s Z_t^s$ be a weighted average of relative popularity using the vector of nonnegative weights $(\pi^s)_{s=1}^S$ that sum to 1. In our estimations we use national polling average data rather than state-specific data, so we observe realizations of a version of Z_t under weights $(\pi^s)_{s=1}^S$ rather than realization the vector $(Z_t^s)_{s=1}^S$. Then, we have

$$dZ_t = d\left(\sum_s \pi^s Z_t^s\right) = \left(\sum_s \pi^s q(x_{\underline{t}}^s/y_{\underline{t}}^s) - \sum_s \pi^s \lambda^s Z_t^s\right) dt + \sum_s \pi^s \sigma^s dW_t^s$$
$$= (q(X_0/Y_0) - \lambda Z_t) \, dt + \sigma dW_t$$

for some σ and process W_t, where the first line follows from linearity and the second follows from Proposition D.1 and the assumptions above.

We recover the parameters λ and $q(X_0/Y_0)/\lambda$ using the same methods we used in the main text and Appendix B. In particular, when we use polling data, we take $\sum_s \pi^s Z_t^s$ to represent national polling averages.

D.3 Estimates of the Popularity Process

We now report our estimates of the parameters of the popularity process using TV ad spending data for the 2012 presidential contest between Obama (candidate 1) and
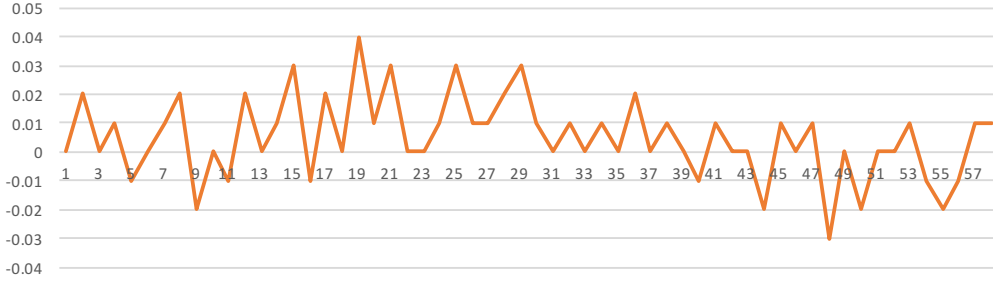
Figure D.2: Obama's weekly average support in polls minus Romney's weekly average for the 58 weeks prior to the election. Data are from HuffPost Pollster.

Romney (candidate 2), and the weekly averages of polling data for these candidates as reported by Pollster. These data are depicted in Figures D.1 and D.2.

We use the same maximum likelihood procedure described in the main text for our estimation, and we use spending data from the last 20 weeks of the election, after the completion of the Republican primaries. We estimate the common λ across districts to be 0.1200 ([0.1198, 0.1202]), which corresponds to a weekly decay rate of 11%.

Given the extensive polling data for this election, we can also estimate decay rates using equation the AR(1) procedure described in the main text. This gives us a weekly decay rate of 66.5% ([30.9%, 100%)) corresponding to an estimated λ of 1.09 ([0.37, $+\infty$)). The comparison of this estimate to the estimate from spending data suggests that candidates are over-spending in early weeks of the campaign. This is a large discrepancy, but as we suggested for the Senate and gubernatorial elections in our evolving budgets extension in Appendix C, there might be substantial fundraising benefits from early spending not captured in our baseline model.

We also compute Obama's long term polling lead, $q(X_0/Y_0)/\lambda$, using equation (11) and Pollster data. We compute $q(X_0/Y_0)/\lambda$ to be 0.007. The richness of polling data for this election also affords us an alternative approach to estimate Obama's long term polling lead that relies only on polling data, which is particularly useful given the large discrepancy between the spending estimates of λ and polling estimates. Note that by equations (8) and (9) in the main text, $q(X_0/Y_0)/\lambda$ is estimated by $\beta_0/(1 - \beta_1)$. From our AR(1) estimates of β_0 and β_1, this gives us an estimate of the long term polling lead of 0.007, which is exactly the same as the equation (11) estimate. In addition, the variance of $\beta_0/(1 - \beta_1)$ can be estimated using the following relationship, which holds

asymptotically by the Delta method:

$$\text{Var}\left[\frac{\beta_0}{1-\beta_1}\right] = \left(\frac{1}{1-\beta_1}\right)^2 \text{Var}[\beta_0] + \frac{\beta_0}{(1-\beta_1)^3}\left[\frac{\beta_0}{1-\beta_1} - 2\bar{z}\right]\text{Var}[\beta_1]$$

where $\bar{z} = 0.0071$ is the mean relative popularity in the data. Using this relationship and our estimates of β_0, $\text{Var}[\beta_0]$, β_1 and $\text{Var}[\beta_1]$ of 0.004, 0.00007, 0.3348, and 0.1813, we get a 95% confidence interval for the estimated long term polling lead of $[0.0003, 0.013]$.

References

ALDRICH, J. H. (1980): "A Dynamic Model of Presidential Nomination Campaigns," *American Political Science Review*, 74, 651–669.

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

AMEMIYA, T. (1973): "Regression analysis when the dependent variable is truncated normal," *Econometrica: Journal of the Econometric Society*, 997–1016.

BOUTON, L., M. CASTANHEIRA, AND A. DRAZEN (2018): "A Theory of Small Campaign Contributions," *NBER Working Paper No. 24413*.

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

COATE, S. (2004): "Political Competition with Campaign Contributions and Informative Advertising," *Journal of the European Economic Association*, 2, 772–804.

COX, G. W. AND M. F. THIES (2000): "How Much Does Money Matter? Buying Votes in Japan, 1967-1990," *Comparative Political Studies*, 33, 37–57.

DE ROOS, N. AND Y. SARAFIDIS (2018): "Momentum in dynamic contests," *Economic Modelling*, 70, 401–416.

DUBÉ, J.-P., G. J. HITSCH, AND P. MANCHANDA (2005): "An empirical model of advertising dynamics," *Quantitative marketing and economics*, 3, 107–144.

ERIKSON, R. S. AND T. R. PALFREY (1993): "The Spending Game: Money, Votes, and Incumbency in Congressional Elections," .

——— (2000): "Equilibria in campaign spending games: Theory and data," *American Political Science Review*, 94, 595–609.

ERIKSON, R. S. AND C. WLEZIEN (2012): *The timeline of presidential elections: How campaigns do (and do not) matter*, University of Chicago Press.

GERBER, A. (1998): "Estimating the effect of campaign spending on senate election outcomes using instrumental variables," *American Political Science Review*, 92, 401–411.

GERBER, A. S., J. G. GIMPEL, D. P. GREEN, AND D. R. SHAW (2011): "How Large and Long-lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment," *American Political Science Review*, 105, 135–150.

GLAZER, A. AND R. HASSIN (2000): "Sequential rent seeking," *Public Choice*, 102, 219–228.

GREEN, D. P. AND J. KRASNO (1988): "Salvation for the spendthrift incumbent: Reestimating the effects of campaign spending in House elections," *American Journal of Political Science*, 884–907.

GROSS, O. AND R. WAGNER (1950): "A Continuous Colonel Blotto game," *Manuscript*.

GUL, F. AND W. PESENDORFER (2012): "The war of information," *The Review of Economic Studies*, 79, 707–734.

HARRIS, C. AND J. VICKERS (1985): "Perfect Equilibrium in a Model of a Race," *The Review of Economic Studies*, 52, 193–209.

——— (1987): "Racing with uncertainty," *The Review of Economic Studies*, 54, 1–21.

HILL, S. J., J. LO, L. VAVRECK, AND J. ZALLER (2013): "How quickly we forget: The duration of persuasion effects from mass communication," *Political Communication*, 30, 521–547.

HINNOSAAR, T. (2018): "Optimal sequential contests," *Manuscript*.

IARYCZOWER, M., G. L. MOCTEZUMA, AND A. MEIROWITZ (2017): "Career Concerns and the Dynamics of Electoral Accountability," *Manuscript*.

JACOBSON, G. (1978): "The effects of campaign spending in congressional elections," *American Political Science Review*, 72, 469–491.

JACOBSON, G. C. (2015): "How Do Campaigns Matter?" *Annual Review of Political Science*, 18, 31–47.

KARATZAS, I. AND S. E. SHREVE (1998): "Brownian motion," in *Brownian Motion and Stochastic Calculus*, Springer, 47–127.

KAWAI, K. AND T. SUNADA (2015): "Campaign finance in us house elections," *Manuscript.*

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

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

KONRAD, K. A. AND D. KOVENOCK (2009): "Multi-battle contests," *Games and Economic Behavior*, 66, 256–274.

KONRAD, K. A. ET AL. (2009): *Strategy and dynamics in contests*, Oxford University Press.

KRASA, S. AND M. POLBORN (2010): "Competition between specialized candidates," *American Political Science Review*, 104, 745–765.

LEONE, R. P. (1995): "Generalizing what is known about temporal aggregation and advertising carryover," *Marketing Science*, 14, G141–G150.

MARTIN, G. J. (2014): "The Informational Content of Campaign Advertising," *Mimeo.*

MATTOZZI, A. AND F. MICHELUCCI (2017): "Electoral Contests with Dynamic Campaign Contributions," *CERGE-EI Working Paper Series No. 599.*

MEIROWITZ, A. (2008): "Electoral contests, incumbency advantages, and campaign finance," *The Journal of Politics*, 70, 681–699.

POLBORN, M. K. AND T. Y. DAVID (2004): "A Rational Choice Model of Informative Positive and Negative Campaigning," *Quarterly Journal of Political Science*, 1, 351–372.

PRAT, A. (2002): "Campaign Advertising and Voter Welfare," *The Review of Economic Studies*, 69, 999–1017.

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

SPENKUCH, J. L. AND D. TONIATTI (2018): "Political advertising and election outcomes," *The Quarterly Journal of Economics*, 133.

TELLIS, G. J., R. K. CHANDY, D. MACINNIS, AND P. THAIVANICH (2005): "Modeling the microeffects of television advertising: Which ad works, when, where, for how long, and why?" *Marketing Science*, 359–366.

VOJNOVIĆ, M. (2016): *Contest theory: Incentive mechanisms and ranking methods*, Cambridge University Press.