Elections with Incomplete Information

by

Scott Ashworth

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2001
(Signed 2001)
© Scott Ashworth, MMI. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly
paper and electronic copies of this thesis document in whole or in part.

Author

Department of Economics
May 15, 2001

Certified by

James M. Snyder, Jr.
Professor of Economics and Political Science
Thesis Supervisor

Accepted by

Richard S. Eckaus
Chairman, Department Committee on Graduate Students
Elections with Incomplete Information

by

Scott Ashworth

Submitted to the Department of Economics
on May 15, 2001, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This dissertation consists of three chapters exploring the role of incomplete information and learning in elections. The first chapter examines the dynamics of voter learning about candidate ability in repeated elections. The dynamic process of belief revision gives rise to incentives that vary strongly over a politician's career. In particular, candidates become entrenched over time, so, even though they exert little effort, the voter cannot commit to throw incumbents out of office. I embed the basic model in a common agency framework to study seniority norms in legislative organization. The model organizes many of the stylized facts about the U.S. Congress, including the incumbency advantage, the dynamics of effort allocation over a career, the importance of constituency service, and seniority norms in committee assignments.

In chapter 2, I study a simple model of campaign finance with possibly asymmetric candidates. Each candidate has the option of promising favors to interest groups in exchange for the funds they need to reveal information to the voters. When the incumbent has a sufficiently large \textit{ex-ante} advantage, the challenger will be unable to raise funds at all. In this case, incumbent spending is unambiguously too high from the perspective of voter welfare. In fact, if the value of a good candidate is high relative to the value of favors a winner can promise, it will be socially optimal to simultaneously restrict spending by the incumbent and encourage spending by the challenger.

In chapter 3, (joint with Aaron Hantman) we propose a simple model of rational learning in elections. A linear approximation to the model is used to justify a version of the Gelman-King (1990) estimator of the incumbency advantage. Restricting the model to elections for which the linear approximation should be valid produce different estimates for the incumbency advantage than those found by either Gelman and King or more traditional studies bases on the "slurge".

Thesis Supervisor: James M. Snyder, Jr.
Title: Professor of Economics and Political Science
Acknowledgments

I could not have written this dissertation without help from many people. First, I owe an enormous debt of gratitude to my advisors, Jim Snyder, Susan Athey, and Abhijit Banerjee. All three of them provided crucial advice and guidance on the models in all three chapters.

In addition, Jim made sure I was aware of the political science literature to which I was contributing, and he made sure it was possible for me to apply for the job in the Harvard Government Department one month after their application deadline. Susan gave me extensive advise on how to write in addition to the help she gave on content. She always managed to be encouraging about the quality of the model while telling me the writing was unacceptable. If this thesis is readable, she gets a lot of the credit. Abhijit gave me crucial advice during the modeling stages of chapter 2, and he gave me important guidance about what to emphasize in writing.

I also learned a great deal from discussion with my fellow graduate students. Mitali Das taught me all of the econometrics that I know, sometimes after becoming very frightened that I was going to fail an exam the next day. My knowledge of game theory and the economics of information grew by leaps and bounds during many discussions with Jon Levin. I also got a lot out of conversations with Mark Aguiar, Max Amarante, Fernando Broner, Jon Dvorak, Nathan Larson, and Justin Johnson. More important than the learning, I also had a lot of fun with all of them.

Special thanks go to my coauthor on chapter 3, Aaron Hantman. In addition to belonging on the list in the previous paragraph, I couldn’t have written this thesis without him, as I can barely turn on a computer, much less make it do econometrics. He also made sure that the negative productivity shocks during the last leg of the writing were quite enjoyable.

Most of the credit for my being sane at the end of this process should go to Georgia Libares. She has made the past few years much better than they could have been otherwise, and I’m sure she will continue to do so for many years to come. She also deserves special credit for putting up with someone who is always speaking some strange language and who is too absent-minded to pay the rent without being reminded.

My greatest thanks go to my parents. Their love and support has been crucial, both during my time at M.I.T. and for all of the years before. I couldn’t have done it without them. Thank you.
Contents

1 Reputation Effects in Electoral Competition ........................................ 9
   1.1 Introduction .................................................................................. 9
   1.2 Model .......................................................................................... 14
       1.2.1 Description of the Model ....................................................... 14
       1.2.2 Information and Beliefs ......................................................... 14
       1.2.3 Definition of Equilibrium ...................................................... 15
   1.3 Main Results .................................................................................. 16
       1.3.1 Discussion of Empirical Evidence ........................................ 21
       1.3.2 Comparison with Market Incentives ...................................... 22
   1.4 Service Activities and Term Limitation ....................................... 24
   1.5 Committee Assignments ............................................................... 26
   1.6 Conclusion .................................................................................... 29
   1.7 Appendix ...................................................................................... 30
       1.7.1 Proof of Proposition 1 ............................................................ 30
       1.7.2 Proof of Proposition 2 ............................................................ 32
       1.7.3 Proof of Proposition 3 ............................................................ 33

2 An Equilibrium Model of Campaign Finance .................................... 35
   2.1 Introduction .................................................................................. 35
       2.1.1 Related Literature ................................................................. 38
   2.2 Model ........................................................................................... 39
   2.3 Preliminary Results ......................................................................... 42
   2.4 Who Advertises? ............................................................................ 45
   2.5 Social Welfare ............................................................................... 51
2.6 Conclusion ........................................... 54

3 Rational Learning and the Incumbency Advantage in Congressional Elections ........... 57
   3.1 Introduction ....................................... 57
   3.2 A Simple Model ................................... 59
   3.3 Data ............................................... 63
   3.4 Econometric model ................................. 64
   3.5 Results ............................................ 66
   3.6 Conclusion ......................................... 69

A References ............................................. 75
Chapter 1

Reputation Effects in Electoral Competition

1.1 Introduction

I present a tractable model of electoral control with both moral hazard and symmetric learning about politician quality. Although simple, the model explains many stylized facts about the U.S. Congress. In elections, it predicts that voters will use retrospective voting rules, incumbents will do better than non-incumbents, and that the incumbency advantage will be associated with extensive provision of constituency service. The model is also able to rationalize laws imposing term limits on a state’s Representatives, even when no other state has term limits. Inside the Congress, the model matches the observed dynamics of shirking by legislators, and it shows that the floor will have incentives to adopt a seniority norm in committee assignments.

The model is based on Holmström’s (1999) career concerns model.\(^1\) In each of three periods, an incumbent politician exerts effort which (stochastically) increases the utility of a representative voter. The voter uses a simple retrospective voting rule, reelecting the incumbent only if he believes that the incumbent has higher ability than a (randomly selected) challenger. The politician works to “fool” the voter about her ability, since high ability politicians are more attractive in the next election. In equilibrium, the voter correctly

---

\(^1\)Loehmann (1998) has used a two period version of the career concerns model to study the policy bias toward special interests.
forecasts the politician's action, so he makes correct inferences about the unknown ability. After analyzing this basic case, I extend the model so I can study the interaction between electoral incentives and the incentives provided by party leaders or by interest groups.

The main results of the basic model are best understood in the context of what a statistician would observe in a cross-sectional analysis of many electoral districts, each of which had a voter learning about a sequence of incumbents. The statistician would observe two things: (i) the probability that a candidate wins an election is increasing with her tenure in office, and (ii) that new politicians work harder than veteran politicians, even when both face the same future of reelection contests.

The fact that the probability of reelection increases with tenure is a form of incumbency advantage. This advantage arises in this model because the voter is selecting the best candidate at each election, so the expected ability of a candidate increases with each electoral victory. It is important to recognize that the voter is not using incumbency as a voting cue; he observes the entire history of signals about the candidate, so he learns nothing from incumbency status. The statistician, on the other hand, does not observe the history of voter utilities, so she does revise her beliefs when she conditions on reelection.

The decline of effort with tenure is a bit more subtle, as that effect results from the interaction of two different comparative statics results: (i) incumbents who are believed to have higher ability take less effort, and (ii) the expected ability of politicians is increasing with their tenure in office. Since effort incentives come from the possibility of fooling the voter into having more favorable beliefs, high ability legislators do not need to work too much. Their constituents already think their incumbent is good, so the marginal benefit of effort is low. Because the expected ability of an incumbent increases with each reelection, this decrease in effort with ability implies a decrease (on average) with tenure.

These effort dynamics can be thought of as a formal model of entrenchment. Once a politician establishes a good reputation, she works less. The voter would like her to exert more effort, but he cannot threaten to throw her out of office if she shirks; because he believes that the incumbent has high ability, such a threat is not credible.

These results match many of the stylized facts on Congressional elections. Fiorina (1981) finds that an individual voter is more likely to vote for an incumbent if that voter's recent, personal financial situation has improved. At the state level, Lowry, Alt, and Ferree (1998) find that Republican incumbents are penalized by voters for increased spending, while
Democrats are penalized for decreased spending. Furthermore, the incumbency advantage is one of the best documented facts in the literature on Congressional elections. This advantage has grown throughout the post-war period, and is now worth 7 to 10 percent of the two-party vote. (See, for example, Gelman and King [1990] and Levitt and Wolfram [1997].)

The model’s predictions also match the empirical evidence on effort choices over Congressional careers. Levitt (1996) finds that Senators place more weight on constituency preferences in their first term than in later terms, and that the weight on constituent preferences is higher for Senators who eventually lose than it is for those who eventually retire. Both of these facts indicate that Senators place less weight on constituent preferences when they are relatively confident of their reelection prospects.

Previous agency-based models of reelection do not provide a unified explanation of these stylized facts. Austen-Smith and Banks (1989) and Ferejohn (1986) study pure moral hazard models of reelection. A critical assumption in these models is that all politicians are identical. This means that the voters can commit to arbitrary reelection rules. The analyst uses this freedom to specify a retrospective voting rule which gives the incumbent maximal incentive to work. Since these rules are stationary, incumbent behavior is time invariant, and there are no dynamics of effort choices or of the incumbency advantage.

Banks and Sundaram (1998) consider a model with both adverse selection and moral hazard. In their model, each politician is term-limited, and can serve only two terms. They also find that effort declines with tenure, but in their model this is due entirely to the last period effect of the term limit. This two-period political lifetime also makes it impossible for their model to explain the empirically observed incumbency advantage, namely that incumbents do better than otherwise identical candidates.

With a unified explanation of effort dynamics and the incumbency advantage in hand, I turn to a model in which the incumbent has several tasks, so that I can explore the implications of the learning-based incumbency advantage for the division of effort between policymaking and constituent service. A traditional explanation for the incumbency advantage is that members of Congress help their constituents deal with the bureaucracy and

\[\text{\footnotesize 2They also (1993) study an infinite horizon version, but in that paper they study stationary strategies in which the incumbents do not condition their behavior on the public beliefs. Thus that version has the same limitations as the pure moral hazard models.}\]
they provide pork barrel spending, so the voters are more likely to vote for them. (See, for example, Fiorina [1977].) It is left unclear, however, why voters don’t expect challengers to provide the same services.

I show that, in a version of my model in which incumbents can divide their time between several tasks, they will devote excessive time to the more easily monitored tasks. This is because that is the most efficient way to try to influence the voter’s posterior beliefs. The intuition can be seen in a simple example: a member’s performance in making sure that constituents receive their Social Security checks is a better signal of ability than her work trying to fix the Social Security system. Too many other factors influence the ambitious task, and information about how well it was performed comes too late for the voter to learn much from it.

Along with placing the constituency service story on a firm footing, the formal model also provides an explanation for why voters in a state would impose term limits on themselves unilaterally. The voters of more than twenty states have passed term limits for their federal representatives. In previous models explaining term limits (see, for example, Dick and Lott [1993]), this would have been a bad idea. Those models assume that voters in the various states are playing a prisoner’s dilemma: veteran congressmen are bad for all states, but having the most senior representatives in Congress is good for an individual state. This is an explanation for nation-wide term limits, not unilateral ones. In my model, term limits can be optimal for a district if incumbents will take useful actions without reelection incentives.

Since the power of electoral incentives decreases with tenure, my model predicts that constituency service activities will be more important for new members and for members who are electorally vulnerable. This is consistent with the data in Cain, Ferejohn, and Fiorina (1987), who find that freshmen and electorally vulnerable Congressmen engage in more casework. Furthermore, this story is reminiscent of Fenno’s (1978) description of the “career in the district”. Congressmen begin their careers in an “expansionist” phase, devoting substantial effort to building a base of support in their district. Over time, the representatives move into a “protectionist” phase, in which they no longer attempt to seriously expand their base. In the context of my model, this is just what we expect if “home style” is used to convince the voters that a candidate is high quality.

The Supreme Court struck down these laws.
An exclusive focus on the reelection contest paints an oversimplified picture of the incentives politicians face. They also get incentives from their party leadership, legislative leaders, and interest groups. An important effect of entrenchment in my model is that politicians with longer tenure are more responsive to incentives from interest group lobbying and party leadership. I study this effect by embedding the reelection model in a common agency framework, so I can discuss the role of seniority in the design of informative committees.

Starting with Gilligan and Krehbiel (1987, 1990), a large literature has developed around the idea that legislatures organize themselves to facilitate information collection about the effects of potential policies. A typical assumption is that the floor of the legislature is unsure of the effect of some policy change, as that effect depends on an unknown state of the world. Since preferences are over outcomes and not directly over policy, there is a demand for information. Information gathering is costly, so some members must be given incentives find out the state. One way to do this is to assign members to committees which specialize in gathering information, and reward the members when they produce information.

My results imply that in a legislature with relatively homogeneous policy preferences, the floor gets a higher expected payoff if it delegates information gathering to a more senior member. This is because the senior member is likely to be relatively immune to electoral pressure and consequently will be cheaper for the floor to influence. The intuition is simple. A freshman faces overwhelming incentives to engage in casework, and this work (partially) crowds out information gathering. Veterans, on the other hand, can afford to shift time and staff resources to information gathering. This provides a possible explanation for the seniority system in committee leadership in the U.S. Congress.

McKelvey and Riezman (1992) also study a model of a legislature with both an electoral stage and a committee system. In that paper, a legislature designs a seniority-based committee system so that the voters will suffer a capital loss if they turn out an incumbent. That model is based on the distributive view of committees, while mine is based on the informative view. Since both approaches to committees are needed to understand actual legislatures, the approaches are complementary. Finally, since McKelvey’s and Riezman’s model has perfect information, it cannot address the effort allocation issues I consider.

\footnote{By homogeneity, I have in mind the dimension of the issue space. For the “floor” to have well-defined preferences, we want to appeal to a median voter theorem.}
1.2 Model

1.2.1 Description of the Model

At each of three dates, \( t = 0, 1, 2 \), an incumbent politician must choose an action \( a \geq 0 \). This action might be effort directed to constituent services, or it might be using time or political capital to fight for better fiscal policy, as suggested by Lowry et al. (1998). The action costs the incumbent \( c(a) \), where \( c \) is increasing and strictly convex, and \( c(0) = 0 \). Incumbents also get a rent equal to \( R > 0 \) each period that they hold office. Any politician not in office gets a period payoff of 0. The rent \( R \) could be an “ego rent,” that is, a direct, psychic benefit from holding office, or it could represent a monetary benefit from bribes or from the diversion of public funds.

At the end of periods 0 and 1, a single voter chooses to reelect the incumbent or to elect a fresh candidate. The initial incumbent is called candidate 0, while the challenger at date \( i (i = 1, 2) \) is called candidate \( i \). Each politician has an unknown ability or talent, \( \theta \). The prior beliefs are that \( \theta \sim N(0, \sigma^2_\theta) \), for both the date 0 incumbent and for any challenger. Uncertainty about \( \theta \) will be symmetric throughout the game. The assumption of symmetric learning considerably simplifies the analysis, as there is no possibility of signaling. For many aspects of politician quality, this is a reasonable assumption. For example, a new legislator is unlikely to know how good she will be at negotiating with lobbyists and party leaders. The symmetry assumption is a bad one, however, for the candidate’s ideology. This is why I focus on ability rather than policy preferences.

Each period, the voter has utility \( u_t = \theta + a_t + \epsilon_t \), where \( \epsilon \sim N(0, \sigma^2_\epsilon) \). The noise term, \( \epsilon \) contains the effects on the voter’s utility of all factors not affected by the policymaker. This payoff is also the voter’s only utility of all factors not affected by the policymaker. The \( \theta \)s and \( \epsilon \)s are mutually independent, and there is no discounting.

1.2.2 Information and Beliefs

Denote by \( h^i_t \) the history of candidate \( i \) up to date \( t \), where a history includes the actions believed to be taken by \( i \) and the payoff to the voter while \( i \) held office. Write \( \overline{\theta}_t \) for \( \mathbb{E}(\theta_t | h^i_t) \). When there is no chance of confusion, I omit the \( i \) and \( t \).

The standard formula for updating in the normal learning process (see DeGroot 1970, pp. 166-168) gives us a recursion for the posterior mean \( \overline{\theta}_t \), as a function of the signal and
the expected action $a^*$:

$$
\tilde{\theta}_t = \lambda (u_{t-1} - a^*_{t-1}) + (1 - \lambda)\tilde{\theta}_{t-1},
$$

(1.1)

where $\lambda = \sigma^2_0/(\sigma^2_0 + \sigma^2_\epsilon)$. This equation says that the (public) estimate of the expected ability of an incumbent is a weighted average of the old estimate, $\tilde{\theta}_{t-1}$, and the new observation,

$$
u_{t-1} - a^*_{t-1} = \theta_{t-1} + \epsilon_{t-1}.$$

The weights in this average depend on the variances in an intuitive way: the more informative is the new signal (low $\sigma^2_\epsilon$), the greater is the weight on the new observation, and the more precise is the old estimate of ability (low $\sigma^2_0$), the greater is the weight on the old belief.

I also assume that the true ability of an incumbent follows a random walk, whose innovations, $\eta$, have variance

$$
\sigma^2_\eta = \frac{(\sigma^2_0)^2}{\sigma^2_0 + \sigma^2_\epsilon}.
$$

This ensures that the variance of beliefs about ability does not decline with tenure. This will highlight the fact that the comparative statics results about effort and tenure are entirely due to the entrenchment effect, and not due to increasing precision of the voter’s beliefs.

1.2.3 Definition of Equilibrium

Any perfect Bayesian equilibrium of the game will be symmetric among the candidates and Markovian, so we can write the incumbent’s effort choices as a sequence $\{a^*_0, a^*_1(\tilde{\theta}_1), a^*_2(\tilde{\theta}_2)\}$. Similarly, the voter’s strategy is given by two functions, $\rho_0$ and $\rho_1$, where $\rho_t : \tilde{\theta}_{t+1} \to \{0, 1\}$.

The incumbent is reelected if and only if $\rho_t(\tilde{\theta}_{t+1}) = 1$.

The incumbent’s decisions must maximize her payoff given the voter’s equilibrium election rules. Formally, this means that the effort decisions must satisfy:

$$
a^*_0 = \arg \max_a \mathbb{E}[\rho_0^*(\tilde{\theta})V(\tilde{\theta})|a^*_0, a] - c(a),
$$

$$
a^*_1(\theta) = \arg \max_a \mathbb{E}[\rho^*_1(\tilde{\theta})[R - c(a^*_1(\tilde{\theta}))]|a^*_1(\theta), \theta, a] - c(a),
$$

$$
a^*_2(\theta) = \arg \max_a -c(a),
$$
where

\[ V(\theta) = \max_a \mathbb{E}[\rho^*_{1}(\hat{\theta})[R - c(a^*_2(\hat{\theta}))][a^*_1(\theta), \theta, a] - c(a)]. \]

Each of these expressions says that the incumbent politician maximizes the probability of getting reelected times the expected reward from coming back to office, minus the cost of effort. At date 2, there is no chance of getting reelected, so this simplifies to minus the cost of effort.

Given these effort choices, the reelection rules must maximize voter utility, so \( \rho^*_1(\theta) = 1 \) if and only if

\[ \mathbb{E}[\theta + a^*_2(\theta)] \geq \mathbb{E}[a^*_2(0)], \]

and \( \rho^*_0(\theta) = 1 \) if and only if \( W(\theta) \geq W(0) \), where

\[ W(\theta) = \theta + a^*_1(\theta) + \mathbb{E}[\rho^*_1(\hat{\theta})[\hat{\theta} + a^*_2(\hat{\theta})] + (1 - \rho^*_1(\hat{\theta}))[a^*_2(0)]]. \]

Both of these expressions say that the voter chooses whichever candidate offers the most (expected) output. The function \( W \) accounts for the option values of candidates in the date 0 election; the voter knows he will get another chance to review whomever he elects.

Say that the voter uses a cutoff strategy if each \( \rho_t \) is monotone, so \( \theta > \theta' \) and \( \rho_t(\theta') = 1 \) imply that \( \rho_t(\theta) = 1 \).

### 1.3 Main Results

We begin by showing that there is an equilibrium in which the voter uses a simple retrospective voting rule. This is a crucial step in the development, as the voting rule will induce a strongly nonlinear incentive scheme.

Because the incumbent’s objective function will contain the cdf of a normal distribution, which is a convex function for negative values of its argument, we need to make an assumption to ensure that the incumbent’s first-order conditions characterize optimal behavior. The basic idea of the assumption is that the per-period payoff to holding office is not too large relative to the variance of the voter’s utility. Let \( \sigma^2 \) be the variance of the
prior distribution of the voter's posterior mean belief.

Assumption S \( R \leq \frac{\sigma^2}{\lambda} \sqrt{(\pi e)/2} \) and \( c'' \geq 1 \).

Proposition 1 Under assumption S, there is a pure strategy equilibrium of the game, in which the voter uses a cutoff strategy, and, along the equilibrium path, effort is a decreasing function of the posterior beliefs about ability.

The details of the proof are in the appendix, but it will be instructive to go over the main points here. Since period 2 ends the game, the last incumbent has no incentive to exert effort and will set \( a_2^* \) to zero no matter what beliefs are. Given this, the voter will choose whichever candidate has higher \( \bar{\theta} \) in the date 1 election.

Next, we can characterize the incumbent's effort in period 1. She will choose a level of effort which maximizes her utility, given the voter's reelection rule. We first derive an expression for the probability of reelection given the incumbent's effort, which we then use to find a simple form for the incumbent's first-order condition.

Given the prior mean of ability, \( \bar{\theta} \), the action expected by the voter, \( a^* \), and the actual action taken by the incumbent, \( a \), the posterior beliefs will be normal with mean

\[
\lambda(\theta + a + \epsilon_1 - a^*) + (1 - \lambda)\bar{\theta}.
\]

By the law of iterated expectations, the expected value of this posterior mean is

\[
\bar{\theta} + \lambda(a - a^*). \tag{1.2}
\]

The prior distribution of the posterior mean is normal, with mean given by (1.2) and variance \( \sigma^2 \). Since the incumbent is reelected if and only if the posterior mean greater than 0, the reelection probability is

\[
1 - \Phi \left( \frac{-\bar{\theta} - \lambda(a - a^*)}{\sigma} \right).
\]

This probability is increasing in \( a \). Intuitively, the incumbent can substitute effort for ability. Since the voter thinks he has corrected for this by subtracting \( c^* \), increases in effort fool the voter into thinking ability is high.
The incumbent solves

$$\max_a R \left[ 1 - \Phi \left( \frac{-\bar{\theta} - \lambda (a - a^*)}{\sigma} \right) \right] - c(a).$$

The first-order condition for this problem is

$$-R\phi \left( \frac{-\bar{\theta} - \lambda (a - a^*)}{\sigma} \right) \left( -\frac{\lambda}{\sigma} \right) - c'(a) = 0.$$ 

In equilibrium, the voter’s forecast of the incumbent’s action is correct, so (using the symmetry of the normal density), we have the equilibrium condition

$$\frac{\lambda R}{\sigma} \phi \left( \frac{\bar{\theta}}{\sigma} \right) = c'(a^*_1).$$

The appendix proves that the global second-order condition holds, so the above analysis in fact describes equilibrium behavior.

An important step in the proof, which is also useful in the later results, is described in the following lemma.

**Lemma 1** The equilibrium effort choice of the incumbent is maximal at $\bar{\theta} = 0$, and it is decreasing in the absolute value of the posterior mean.

**Proof** Since the cost function is strictly convex, the marginal cost is invertible and we can write

$$a^*_1(\bar{\theta}) = (c')^{-1} \left[ \frac{\lambda R}{\sigma} \phi \left( \frac{\bar{\theta}}{\sigma} \right) \right].$$

Since $(c')^{-1}$ is increasing, $a^*_1$ is maximal, increasing, or decreasing as $\phi \left( \frac{\bar{\theta}}{\sigma} \right)$ is maximal, increasing, or decreasing. But now the result is clear from the properties of the normal distribution. □

This comparative statics result is quite intuitive; the change in the probability of re-election due to a small, unexpected change in $a$ is approximately equal to the change in $a$ times the height of the density function for ability, evaluated at zero. As the expected ability moves away from zero, the density at zero decreases. Thus the marginal benefit to effort decreases.
It turns out that the voter’s value of a candidate in the date 0 election is an increasing function of the candidate’s expected ability. This means that the voter again finds it optimal to reelect the date 0 incumbent if and only if her expected ability is greater than zero. Given this cutoff rule, an analysis just like the one above holds for the date 0 incumbent.

Now that we have a description of the equilibrium, we can present the main results of this section. The first result describes the dynamics of effort choices.

**Proposition 2** (i) A date 1 incumbent who has won reelection exerts less effort than someone elected as a challenger. (ii) Efforts are declining over time: the date 0 incumbent takes a higher effort than any date 1 incumbent, and every date 1 incumbent takes a higher effort than any date 2 incumbent.

**Proof** (i) Since the voter will never reelect an incumbent whose ability is less than zero, every reelected candidate will have expected ability at least zero, and almost every one will have ability strictly greater than zero. Thus lemma 1 implies that almost every reelected incumbent will exert less effort than a new candidate.

Proof of (ii) is in the appendix. □

The fact that efforts decline over time is a consequence of the fixed time horizon. More interesting is the decline with tenure of the date 1 effort. This says that, even though they both face one reelection contest, veteran politicians shirk more on average. This is due to two effects. First, for a fixed cutoff point, an incumbent has a smaller marginal benefit of effort. Second, the voter sets a lower cut-point for veterans. Unlike the decline of all efforts over time, the decline of effort with tenure is robust to the time horizon. In an infinite-horizon model, efforts will decrease with tenure so long as the voter uses a cutoff reelection rule in every election.

This result is robust to the assumed normality of the beliefs about ability and of the noise. As long as signals and abilities satisfy the monotone likelihood ratio property, better incumbents will face a lower hurdle for reelection. Furthermore, increases in the expected ability lead to lower marginal benefits of effort (for fixed reelection hurdles) under the relatively mild assumption of logconcavity.\textsuperscript{5}

\textsuperscript{5}If both the ability and the noise have logconcave densities, then the distribution of the posterior mean will also have a logconcave density. Since logconcavity implies single-peakedness, we will again get an eventually declining effort choice, as marginal cost is proportional to the unconditional signal density.
Although effort declines with tenure, the voter is willing to retain the incumbents because the increase in talent more than compensates them. The next proposition formalizes this result. The particular formalization we choose is useful both for discussing the incumbency advantage in reelection and for interpreting some of the empirical work discussed below. Say that the probability distribution $F$ is more favorable than distribution $G$ if, $F$ dominates $G$ in the monotone likelihood ratio (MLR) order.\footnote{See Milgrom (1981) for technical details and several applications in economics.} Formally, $F$ is more favorable than $G$ if, for all $x > y$,

\[
\frac{f(x)}{g(x)} > \frac{f(y)}{g(y)},
\]

where $f$ is the density of $F$ and $g$ is the density of $G$. This is a strong notion of “larger” for distributions; in particular, it implies that the distributions conditional on any interval are ordered by first-order stochastic dominance.

**Proposition 3** (i) The distribution of ability conditional on winning reelection once is more favorable than the prior distribution of ability. (ii) The distribution of date 1 ability conditional on winning the next election is more favorable than the distribution conditional on losing the next election.

**Proof** See the appendix. □

**Corollary 1** A two-term incumbent has a higher probability of winning reelection than a one-term incumbent.

**Proof** The more favorable than relation is just the likelihood ratio order, so proposition 3 implies that the ability distribution of a two-term incumbent first-order stochastically dominates that of a one-term incumbent. Since the reelection rule is a monotone function of the voter’s beliefs, the result follows. □

This incumbency advantage arises because the voter censors the distribution of abilities every time he selects the best candidate. Without this selection, the mean value of expected ability in a cross-section of many districts would be 0. (This is just the law of iterated
expectations.) With selection, however, this mean will drift up over time. This is because the voter is more likely to remove a candidate the less is her true ability.

The result about the incumbency advantage fails to match one fact about House elections, namely the “sophomore surge” (see Gelman and King [1990]). This refers to the fact that part (but not all) of the incumbency advantage is shared by members in their first reelection effort. This effect is missing from my model because there is no learning in campaigns. Using the law of iterated expectations, we expect that the first signal about an incumbent leads to a posterior mean greater than 0 exactly half the time. This implies that politicians who have served one term will win reelection with probability 1/2. If the model were extended so that the prior mean of challenger ability were drawn from some distribution, and if that prior mean was observed by the voter before the election, then the same selection effect identified in proposition 3 would apply to freshmen. I leave this out of the formal model to preserve tractability.

The best way to understand how these results interact is to think about the differences we would observe between a cross-sectional and a longitudinal data analysis. In each case actual effort would decline with tenure. Of course, effort is less likely to be observed than outcomes. When only outcomes are observed, we would see a difference between the cross-section and the panel. Because the increase in expected ability is greater than the decrease in effort, we should see a positive relationship between outcomes and tenure in cross-sectional data. Indeed, this is what leads to reelection. On the other hand, in a panel setting, we see a fixed talent added to declining effort, so the outcomes produced by a fixed candidate will decrease with tenure. Finally, a dynamic analysis which ignored the fact that voters are selecting which incumbents to retain would find that outcomes were increasing with tenure.

1.3.1 Discussion of Empirical Evidence

Ansolabehere, Snyder, and Stewart (1999) present some evidence that suggests the learning view of the incumbency advantage is empirically important. They use post-census redistricting to compare an incumbent's advantage in counties new to his district to the advantage in counties which were in his old district. They find that the advantage is larger in the old counties and that the incumbency advantage in the new counties is an increasing, concave function of the number of years since redistricting. This finding makes perfect sense in a learning context; the new voters have less information about the incumbent. It is not con-
sistent, however, with the idea that the advantage is solely due to committee assignments or perks of office: these factors are the same for the new and the old districts.

Levitt’s (1996) analysis of Senate roll-call votes contains two facts which strongly support the above analysis. Under the assumption that votes in the House of Representatives represent state interests, Levitt is able to estimate the weights that Senators assign to their home constituencies, party, and personal preferences in voting decisions. Although my model is not one of position taking by legislators, it does have implications for ideological location thanks to a paper by Groseclose (1997). He studies a Hotelling-type model of competition in platforms when one candidate has a “personal advantage”. The candidate with an advantage uses this to move away from the median voter’s position toward her own ideal point. The divergence from the median is an increasing function of the personal advantage, so, using my model for the personal advantage, we predict that the distance from the median voter’s ideal point increases (on average) with tenure.

The first result relevant to my analysis is that Senators in their first term place more weight on constituencies than do other senators. Second, Levitt partitions the Senators into those who eventually lost an election and those who eventually retired. The ones who went on to lose placed higher weight on their voters throughout their careers. To see why the second fact is an implication of my model, consider an incumbent who has won reelection once. Proposition 3 tells us that two-term incumbents who lose the election were likely to have had lower expected ability at date 1. Consequently, we expect that they took high actions. This is intuitive: candidates take high actions when the election is expected to be close. This means that high effort and subsequent poor electoral showings should be associated in the data, which is what Levitt finds.

1.3.2 Comparison with Market Incentives

The comparative statics results provide a formal model of entrenchment. As the beliefs about the incumbent become more favorable, the reelection contest provides weaker incentives. She is insulated from competitive pressures.

In this regard, it is instructive to compare the incentives provided by reelection with those in a labor market context. To highlight the comparison between elections and wages,

---

7 If House votes are influenced by parties or interest groups, then the point estimates of the weights are not meaningful. This does not affect the differences between weights that I cite, however.
I consider a model with the same technology and information structure as above. The only difference is that the worker gets a wage equal to her marginal product, \( \bar{\theta} \). (This is essentially the model from Holmström [1999].) Along with pointing out the importance of the nonlinearity in the reelection incentive scheme, this comparison shows what voters lose from their inability to make bids for the services of politicians from other districts.

Recall that the equilibrium condition in the election context is

\[
\frac{\lambda R}{\sigma} \phi \left( \frac{\bar{\theta}}{\sigma} \right) = c'(a^*).
\]

In the labor market case, the worker supplies effort to affect the beliefs about her ability because her period two wage is given by \( \mathbb{E}(\theta | s, a^*) \), where

\[
s = \theta + a + \epsilon
\]

is the labor market's signal about date 1 output. If all random variables are normal, this conditional expectation is given by

\[
\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\epsilon^2} (s - a^*) + \frac{\sigma_\epsilon^2}{\sigma_\theta^2 + \sigma_\epsilon^2} \bar{\theta}.
\]

Since \( s = \theta + a + \epsilon \), the first order condition for the first period action choice is

\[
c'(a) = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\epsilon^2}.
\]

We see that, in this case, the action is constant in the prior mean on ability. This is because the labor market provides \textit{constant} incentives.\(^8\)

The key difference between these two cases is that reelection provides highly nonlinear incentives, while the labor market provides linear incentives. In the labor market context, an increase in \( a \) by \( \Delta a \) leads to an expected increase in wages of \( [(\sigma_\theta^2)/(\sigma_\theta^2 + \sigma_\epsilon^2)] \Delta a \). In contrast, a small increase in \( a \) by an incumbent politician leads to an increase in expected reward equal to \( R \) times the change in the probability that the posterior mean of ability is higher than zero. Because the prior distribution of the posterior mean is normal with mean \( \bar{\theta} + [(\sigma_\theta^2)/(\sigma_\theta^2 + \sigma_\epsilon^2)] \Delta a \), the amount of probability mass carried across zero by \( \Delta a \) is

\(^8\)The fact that the action is exactly constant also depends on the absence of wealth effects in the cost of effort and in risk attitudes.
decreasing in $\theta$.

The declining power of incentives in the election case should not be confused with the well-known fact that incentives will weaken in the labor market case if the unknown talent is fixed (see, for example, Gibbons and Murphy 1992). This happens because the posterior variance decreases with more observations if the unknown ability is fixed, so the weight the market assigns to output declines over time. On the other hand, the incentives in the election case decline on average even when the talent follows a random walk whose variance exactly offsets the effect of learning.

1.4 Service Activities and Term Limitation

The results of the previous section show that a legislator's incentives to work are powerfully influenced by how much she can affect the voter's beliefs about her ability. Incumbents work hard when they can have substantial influence on beliefs, and they shirk when they cannot. These results ignore one of the crucial decisions an incumbent makes, namely, how to allocate her effort among different tasks. Since one of the primary observations political scientists have made about Congressmen is that they devote too much time to casework and not enough to policymaking, this is a serious omission.

In this section I fill that gap by showing how the information structure interacts with the effects studied in the previous section to determine task allocation. The central result is that incumbents have incentives to bias their effort allocations toward tasks which have relatively little noise. This is intuitive; Bayesian updating puts more weight on a signal if its conditional variance is smaller, so the voter is more sensitive to easily monitored tasks. In this light, it makes perfect sense for incumbents to spend excessive time on casework. These tasks reveal more about the politician because they can be done independently, so there is no noise from the legislative bargaining process, and these tasks get results right away, while policy pays off in the future.

It is easiest to see why the allocation among tasks will be distorted due to differential observability by looking at a simple example. Assume that the incumbent takes two actions, $a_1$ and $a_2$. The voter gets signals $s_1$ and $s_2$, where $s_i = \theta + a_i + \epsilon_i$, and $\sigma_{\epsilon_1}^2 < \sigma_{\epsilon_2}^2$. If the voter expects $(a_1^*, a_2^*)$ and the incumbent takes action $(a_1, a_2)$, then the posterior mean has
a normal distribution with mean

$$\theta + \lambda_1 (a_1 - a_1^*) + \lambda_2 (a_2 - a_2^*)$$

where

$$\lambda_i = \frac{\sigma_d^2 + \sigma_{\epsilon_i}^2}{\sigma_d^2 + \sigma_{\epsilon_i}^2 + \sigma_{\epsilon_{-i}}^2}.$$  

Since the noise term for signal 1 has lower variance than the noise for signal 2, $\lambda_1 > \lambda_2$. This means that shifting effort from task 2 to task 1 will raise the voter’s estimate of the incumbent’s talent. Since effort incentives come from the incumbent’s desire to fool the voter about ability, the incentive to work on task 1 is stronger than the incentive to work on task 2.

This model fits well with the story told by Cain et al. (1987), who argue that constituency service accounts for a substantial part of the incumbency advantage. They find that voters who have benefited from casework are more likely to have a high opinion of the Congressman and are more likely to vote for him in the next election. They also find that incumbents who face poorer reelection prospects devote more time and staff resources to casework. Since the comparative statics results about effort continue to hold in the multitask case, this is just what we expect.

The above logic is familiar from the multitask moral hazard model of Holmström and Milgrom (1991), which was extended to the career concerns setting by Dewatripont, Jewitt, and Tirole (1999). Holmström and Milgrom point out that, in the multitask context, it can be optimal to provide no incentives at all. This will be the case if the agent will exert some effort without (external) incentives, and if the distortions in effort allocation due to different monitoring intensities are large. A similar result is true in our electoral context, and this fact can help us understand the demand for term limits.

To see how term limits might be valuable, consider a setting in which the incumbent must divide her time between casework and policymaking, as before. Now we assume that the incumbent’s preferences over effort allocations are exactly the same as the voter’s. This means there is no agency problem, and optimal effort can be elicited by providing no incentives at all. Unfortunately, the only way to implement this incentive scheme is to commit not to reelect the incumbent. A term limit is a mechanism for achieving low-
powered incentives. Because incumbents cannot commit to ignore their reelection chances, voters must tie their own hands. The common criticism that term limits are unnecessary, since voters can always vote out an incumbent, has no force here. Even if voters announced that they would never reelect an incumbent, they would not get undistorted effort choices. The incumbent would know the plan was not time consistent.

There is an important difference between the voter in my model and the principal in Holmström-Milgrom. In the explicit incentives case, the principal gets the first-best outcome when incentives are perfectly aligned. With career concerns, on the other hand, the principal cannot get the first-best, even with perfect alignment. Giving low powered incentives also means giving up the ability to select good types.

1.5 Committee Assignments

The implicit incentives from reelection concerns, studied above, are only a part of any politician's incentives. Both organized interest groups and the leadership of political parties try to influence the behavior of politicians. In this section I study the interaction of these incentives and the reelection incentives. For concreteness, I focus on the case of a legislature delegating the responsibility to acquire information about the effects of a proposed policy change. My results provide one reason that legislatures would choose more senior members for such information gathering roles. The key to this application is that constituency service and information gathering are substitutes in the incumbent's cost function. This is a quite reasonable assumption: if the incumbent is using her own time or her staff to do casework, then she can't use those resources to find out the effects of policies.

All of the results in this section are based on the following simple idea. The action for the voter (think of constituency service) competes with some other action (think of information gathering) for the incumbents time and resources. This competition means that an increase in the level of either activity raises the marginal cost of performing the other. Since we have seen above that veteran incumbents do less of the service activity than new members, they have a lower marginal cost of effort for the second activity. In turn, this makes the veteran a more attractive choice to be an agent for the party leadership or for the floor.

Gilligan and Krehbiel study a cheap-talk game in which a committee is the sender and the floor is the receiver. They find conditions under which information can be shared
efficiently under different rules for voting and amendments on the floor, and they analyze the incentives for the committee to invest in the skills needed to acquire information. Because I want to focus on the provision of incentives to acquire the information, rather than the efficiency with which it is transmitted, I analyze a simpler setting than Gilligan and Khrebiel do. In particular, I assume that the information is verifiable, so the subtleties of cheap-talk equilibrium are not needed.

Consider the median voter in a legislature who must decide to pass or reject a piece of legislation. The effect of the bill depends on an unknown state of the world, \( \omega \). A member of the legislature can be assigned the task of trying to determine \( \omega \). Let \( V \) denote the unconditional expected utility of the median voter when he votes optimally, and let \( V(\omega) \) be the expected utility of the median voter when he knows the state is \( \omega \) and votes optimally. Assume that \( \mathbb{E}(V(\omega)) > V \). A member can discover the true state of the world with probability \( p \) by exerting effort equal to \( c(a, p) \), where \( a \) is the effort she devotes to service activities for her constituents. Let \( B = \mathbb{E}(V(\omega)) - V \) be the benefit of information to the floor.

To formalize the intuition about substitute efforts, we let the date 1 incumbent take two actions: \( a \), which enters the voter’s utility as before, and \( p \in [0, 1] \), which is the probability that information is collected. The cost function \( c(a, p) \) is increasing and convex. I also assume that \( c_{ap} > 0 \), so a higher level of one action raises the marginal cost of the other action. The incumbent values money linearly, has limited liability, and has zero initial wealth.

If the floor offers a payment of \( \alpha \) conditional on receiving the information, then the incumbent maximizes

\[
R \left[ 1 - \Phi \left( \frac{-\bar{\theta} - \lambda(a - a^*)}{\sigma} \right) \right] + ap - c(a, p).
\]

Call the solution to this problem \( a^*(\alpha, R, \theta) \) and \( p^*(\alpha, R, \theta) \). These are the effort supply functions.

We can now derive comparative statics results for the incumbent’s problem, as \( R, \bar{\theta} \), and \( \alpha \) are varied. Since \( a \) and \( p \) interact only in the cost function, it is easy to see that the incumbent’s objective function is (strictly) supermodular in \((\alpha, p, \theta, -R, -a)\). This means that an increase in any of these variables raises the marginal benefit to increasing the others.
Consequently, $a^*$ is strictly increasing in $R$ and strictly decreasing in $\bar{\theta}$ and $\alpha$, while $p^*$ is strictly decreasing in $R$ and strictly increasing in $\bar{\theta}$ and $\alpha$.\footnote{The fact that the objective is differentiable implies that the results hold as strict inequalities. See Edlin and Shannon 1998.}

I assume that these effort supply functions are supermodular in $(\alpha, -\theta, R)$. This means that an increase in any of the exogenous parameters makes the incumbent more responsive to changes in the others. For example, a decrease in expected ability makes the incumbent more sensitive to higher incentive pay from the floor. This assumption guarantees that the second-order effects reinforce the first-order comparative statics results, and is a standard assumption in the literature on multi-task agency problems. (It is implied by the quadratic cost function often used, for example.)

Given these functions as incentive compatibility constraints, the floor will choose $\alpha$ to maximize

$$W(\alpha; B, R, \bar{\theta}) = (B - \alpha)p^*(\alpha, R, \theta),$$

where $B$ is the monetary value of the benefit. Let $\alpha^*(B, R, \theta)$ be the solution to this program.

**Proposition 4** The level of incentive pay ($\alpha$) is increasing in the value of the information and in the payoff to holding office, and it is decreasing in the expected ability.

**Proof** The derivative of the objective function with respect to $\alpha$ is

$$\frac{\partial W}{\partial \alpha} = -p^* + (B - \alpha)(\frac{\partial p^*}{\partial \alpha}).$$

The comparative statics are determined by the following three derivatives:

$$\frac{\partial^2 W}{\partial B \partial \alpha} = \frac{\partial p^*}{\partial \alpha},$$

$$\frac{\partial^2 W}{\partial R \partial \alpha} = -\frac{\partial p^*}{\partial R} + (B - \alpha)\frac{\partial^2 p^*}{\partial R \partial \alpha},$$

$$\frac{\partial^2 W}{\partial \theta \partial \alpha} = -\frac{\partial p^*}{\partial \theta} + (B - \alpha)\frac{\partial^2 p^*}{\partial \theta \partial \alpha}.$$

The comparative statics results for $p^*$, along with the assumed supermodularity of $p^*$, imply that the first two of these equations are positive and the last is negative. Now the claim
follows from standard results on monotone comparative statics. □

The most interesting of these results is the last; the floor has a lower "price" for incumbent effort the more favorable are the beliefs of the voter about the incumbent's ability. Notice that we get this result even though the talent of the incumbent does not influence her productivity in information gathering. The effect comes entirely from the fact that the incumbent is less concerned with impressing the voter, so she exerts less effort in that task. Because of the supermodularity of the cost function, this lowers her marginal cost of work for the principal. This means that the floor can offer less powerful explicit incentives to induce a fixed level of effort.

An immediate corollary of this result is that the floor gets a higher payoff from dealing with politicians with longer tenure. This higher payoff is the key to the value of seniority in my model. The floor delegates tasks to members who can take the necessary time away from their constituents without putting their seats at too much risk.

1.6 Conclusion

In this paper, I present a simple model of the dynamics of electoral careers with learning and moral hazard. There are several directions to go in extending this work.

First, in the multi-principal context of the last section, my results concern a single member of the legislature. The basic model should be embedded in an equilibrium model of an entire legislature, to see if the advantages of seniority persist. The model presented here is ideal for such an exercise, as it has a tractable closed-form solution.

Finally, the basic idea, that politicians get incentives for effort from career concerns, should also apply to parliamentary systems. Different electoral systems (i.e. first-past-the-post versus party list) will provide different incentives for politicians to supply effort to tasks. This will, in turn, give parties incentives to adopt different internal organizations. In future work, I plan to combine the model of this paper with insights from the economics literature on career concerns in organizations to study the covariation in electoral rules, party organization and effort allocations across legislatures in different countries.

29
1.7 Appendix

1.7.1 Proof of Proposition 1

We prove the proposition by constructing an equilibrium with the desired properties, working back from the end of the game. We saw in the text that a date 2 incumbent takes 0 effort and the voter elects the date 1 candidate with the best posterior.

Next we show that the date 1 action satisfies

$$\frac{\lambda R}{\sigma} \phi \left( \frac{\bar{\theta}}{\sigma} \right) = c'(a^*_1).$$  \hspace{1cm} (1.3)

The argument in the text shows that, if the voter expects effort described by this equation, then that effort satisfies the following first-order condition for:

$$-R \phi \left( \frac{\bar{\theta} - \lambda (a - a^*)}{\sigma} \right) \left( -\frac{\lambda}{\sigma} \right) - c'(a) = 0.$$  \hspace{1cm} (1.4)

The following lemma shows that the objective function is pseudoconcave, which implies that the first-order condition derived above is sufficient for a maximum. Let the voter expect a*^, given by (1.3). Then we have

**Lemma 2** If a < a*, then the derivative on the LHS of (1.4) is > 0. If a > a*, then the derivative on the LHS of (1.4) is < 0.

**Proof** First consider the case of a < a*. The derivative is

$$-\frac{\lambda R}{\sigma} \phi \left( \frac{\bar{\theta} - \lambda (a - a^*)}{\sigma} \right) - c'(a),$$

which (using the FOC) equals

$$- \int_a^{a^*} \left[ \frac{\lambda R}{\sigma^2} \alpha \phi(\alpha) - c'(\bar{\alpha}) \right] d\bar{\alpha},$$

where

$$\alpha = \frac{-\bar{\theta} - \lambda (\bar{\alpha} - a^*)}{\sigma}.$$  

I claim the function x → xΦ(x) is bounded between −(2πε)^{1/2} and (2πε)^{1/2}, so the assump-
tions on $R$ and $c''$ imply that the integrand is negative, so the derivative is positive.

The case of $a > a^*$ is handled similarly.

Proof of claim. Consider the function $x \mapsto x \phi(x)$. Clearly, it is zero at $x = 0$, and it approaches 0 as $x$ tends to either $\infty$ or $-\infty$, since the Gauss kernel tends to zero faster than any polynomial. The derivative is $\phi(x) - x^2 \phi(x)$, so the critical points are 1 and $-1$. At each of these points, the absolute value of the function is

$$\frac{1}{\sqrt{2\pi}}e^{-(1/2)}.$$

Now consider the problem of the voter in the date 0 election. It is clear that a date 0 incumbent whose date 1 posterior is less than zero is never reelected. This follows easily from the fact that effort, ability, and the option value for the date 1 election all decline as $\theta$ decreases from 0. Thus, we need to determine the voting rule only when the posterior is nonnegative.

I next show that, under the assumption that $R$ has the upper bound, the voter will use a cutoff rule at date 0. Recall that $\rho_0(\bar{\theta}) = 1$ if and only if $W(\bar{\theta}) \geq W(0)$, where

$$W(\theta) = \theta + a_1^*(\theta) + \mathbb{E}[\rho_1^*(\bar{\theta})|\bar{\theta} + a_2^*(\bar{\theta})] + (1 - \rho_1^*(\bar{\theta}))[a_2^*(0)].$$

The voter's expected value of a candidate with expected ability $\theta$ is the sum of three terms: the expected ability, the date 1 action, and an option value for the date 1 election. The difference in date 1 efforts is given by

$$(c')^{-1} \left[ \frac{\lambda R}{\sigma} \phi(\frac{\bar{\theta}}{\sigma}) \right] - (c')^{-1} \left[ \frac{\lambda R}{\sigma} \phi(0) \right],$$

which, by the Fundamental Theorem of calculus and the inverse function theorem, is

$$\int_0^\theta \frac{R}{\sigma^2} \alpha \phi'(\alpha) d\alpha / c''.$$

The numerator is between $-1$ and 1, while the denominator is greater than 1. Thus the integrand is greater than $-1$, so we have

$$a^*(\theta) - a^*(0) > -\theta,$$
which says the voter’s date 1 expected payoff is increasing in \( \theta \). Since the option value is clearly increasing in \( \theta \), the voter’s value of the incumbent is increasing in the posterior on the incumbent’s ability. This implies the voter will return the incumbent if and only the posterior is more than zero.

Finally, we show that there is an equilibrium for the date 0 action. Recall that \( V(\bar{\theta}) \) is the value of being a date 1 incumbent when beliefs have mean \( \bar{\theta} \). The date 0 incumbent solves

\[
\max_a \int_0^\infty V(\bar{\theta}) \phi \left( \frac{\bar{\theta} - \lambda(a - a^*)}{\sigma} \right) \, d\bar{\theta} - c(a).
\]

The first-order condition for this maximization is

\[
\frac{\lambda}{\sigma} \int_0^\infty V(\bar{\theta}) \phi \left( \frac{-\bar{\theta} - \lambda(a - a^*)}{\sigma} \right) \cdot \left( \frac{-\bar{\theta} - \lambda(a - a^*)}{\sigma} \right) \, d\bar{\theta} = c'(a).
\]

Notice that \( V(\bar{\theta}) \leq 2R \), since the maximum utility in each period is \( R \). With this observation, we can mimic the proof of lemma 2 to show that the first-order condition characterizes the optimum. Notice that, for any \( \alpha = \frac{\bar{\theta} - \lambda(a - a^*)}{\sigma} \),

\[
\frac{\lambda}{\sigma} \int_0^\infty -V(\bar{\theta}) \phi(\alpha) \, d\bar{\theta} \leq 2 \frac{\lambda R}{\sigma} \int_0^\infty -\alpha \phi(\alpha) \, d\bar{\theta} = 2 \frac{\lambda R}{\sigma} \phi \left( \frac{-\lambda(a - a^*)}{\sigma} \right),
\]

where we use the symmetry of the Gauss kernel to add a negative sign to the integrand. Now the argument of lemma 2 implies the objective is pseudoconcave. \( \square \)

### 1.7.2 Proof of Proposition 2

Recall that the date 1 action is maximal for an incumbent with \( \bar{\theta} = 0 \), who takes an action equal to

\[
(c')^{-1} \left[ \frac{\lambda R}{\sigma} \phi(0) \right],
\]

while the date 0 incumbent takes action equal to

\[
(c')^{-1} \left[ \frac{\lambda}{\sigma} \int_0^\infty -V(\bar{\theta}) \phi \left( \frac{\bar{\theta}}{\sigma} \right) \cdot \frac{\bar{\theta}}{\sigma} \, d\bar{\theta} \right].
\]
Since the inverse marginal cost function is increasing, we need only show that
\[
\frac{\lambda}{\sigma} \int_0^\infty -V(\bar{\theta}) \phi \left( \frac{\bar{\theta}}{\sigma} \right) \cdot \frac{\bar{\theta}}{\sigma} \, d\bar{\theta} > \frac{\lambda R}{\sigma} \phi(0).
\]

I claim that \( V(\bar{\theta}) > R \) for all \( \bar{\theta} \). To see that this will prove the proposition, notice that
\[
\frac{\lambda}{\sigma} \int_0^\infty -V(\bar{\theta}) \phi \left( \frac{\bar{\theta}}{\sigma} \right) \cdot \frac{\bar{\theta}}{\sigma} \, d\bar{\theta} > \frac{\lambda R}{\sigma} \int_0^\infty -\phi \left( \frac{\bar{\theta}}{\sigma} \right) \cdot \frac{\bar{\theta}}{\sigma} \, d\bar{\theta} = \frac{\lambda R}{\sigma} \phi(0).
\]

Finally, I prove the claim. Consider an incumbent at date 1. One feasible strategy is to set \( a = 0 \), for an immediate payoff of \( R \). Because the normal distribution gives every nondegenerate interval positive probability and there is always a sufficiently high signal which will lead to reelection, the do nothing strategy has a payoff strictly bigger than \( R \). The optimal strategy must be at least this. \( \square \)

### 1.7.3 Proof of Proposition 3

First, notice that, for a fixed candidate, the true abilities and the signals are joint normal random variables with a strictly positive covariance matrix, so they are affiliated.\(^{10}\) Furthermore, affiliation is preserved by integration, so all of the implied distributions over subsets of the variables are also affiliated. Finally, recall that affiliation implies the monotone likelihood ratio property for every pair of variables.

We use the following result.

**Theorem 1** (Milgrom, 1981) Let \( x, y \) be affiliated random variables, and let \( [a, b] \) and \( [c, d] \) be intervals with \( a \geq c \) and \( b \geq d \), at least one strict. Then the distribution of \( x \) conditional on \( y \in [a, b] \) MLR dominates the distribution \( x \) conditional \( y \in [c, d] \).

**proof of part i** An incumbent is reelected for a second term if and only if the signal is in \([0, \infty)\). Thus Theorem 1 implies that the distribution of ability conditional on reelection dominates the prior distribution of ability. \( \square \)

**proof of part ii** For any date 1 incumbent, there is a signal value \( \bar{u} \) such that the incumbent is reelected if and only if \( u \in [\bar{u}, \infty) \). Thus Theorem 1 implies the result. \( \square \)

\(^{10}\)See Milgrom and Weber [1982] for a treatment of affiliated random variables.
Chapter 2

An Equilibrium Model of Campaign Finance

2.1 Introduction

Our system of decentralized campaign finance is a major target of government reformers. They claim that the current system of decentralized finance leads to excessive special interest influence over policy and that incumbents are able to use the finance system to limit electoral competition.

The theoretical literature on elections contains some work which addresses these concerns. Baron (1994) (and others) have studied the influence of fundraising on policy outcomes. Unfortunately, most of this work avoids modeling voters at the level of preferences and beliefs, so the models cannot be used to directly address the reformers’ concern about voter welfare. Austen-Smith (1987) and Prat (2000a,b) have modeled fundraising and campaign spending with rational voters, but they assume symmetric candidates, so issues concerning incumbency are ignored.

I study a model with both rational voters and asymmetric candidates. Two candidates compete for an elective office. These candidates differ both in their ideological preferences and in their ability. The voters know the ideologies, but are uncertain about the candidates’ abilities. Voters prefer candidates who have high ability, everything else equal. The candidates thus have an incentive to reveal any information they may have which will enhance their reputations.
I assume that each candidate has some information about their own abilities. This is hard information, so it can be verifiably shown to the voters. However, the candidates cannot afford to buy the advertising space needed to reveal their information. They must turn to contributors to raise the needed funds. These contributors are motivated by the promise of favors contingent on their candidate winning the election. The value of such a promise depends on the probability that the candidate will win the election. This means that contributors are more willing to fund candidates who have good information to reveal or who have strong reputations to begin with.

This has two important implications for the ability of the model to reproduce the stylized facts concerning the role of campaign spending in congressional elections. First, a candidate is more likely to have funds to spend if she has good information to reveal, so candidates who spend are more likely to win, controlling for prior reputations. Second, incumbents who have good reputations are likely to not face strong challengers, because the potential challengers will not be able to get funds even with good information. This is because good news will not be enough to overcome the reputation of the incumbent.

A key feature of the model is that incumbents have an advantage in fundraising. This advantage arises endogenously, and is based on the incumbent’s superior ex-ante reputation. The voters believe that the incumbent has above average ability. Although not formally modeled, this belief can be shown to arise from rational learning through repeated elections (ch. 1 of this thesis, Zaller 1998). The basic mechanism in these papers is electoral selection. Prior to each election, the voters get information about the candidates, both from examining the record of the incumbent and from an informative campaign of the sort modeled here. The voters use this information to revise their beliefs about the candidates, and they vote for the one they believe to be best qualified. This voting rule implies that any candidate who has won an election is one about whom favorable information has been learned, and the more elections a candidate has won, the more good information has arrived. This means that conditioning on electoral success increases an observer’s assessment of a candidate’s ability.

These selection effects lead to an electoral advantage for incumbents: They have a higher ex-ante probability of winning the election. This advantage translates into an advantage in fundraising. To see why this is so, consider the incentives a potential contributor has when he is approached by a candidate. If he funds the campaign, he gets in exchange a promise
of favors contingent on the candidate winning the election. The key observation is that such a promise is worth more the higher is the probability that the candidate will win. Because of this, the level of favors a candidate must promise is decreasing in the prior probability that she is high quality. This creates an incumbency advantage in fundraising.

This incumbency advantage can in fact be so great that the challenger is completely foreclosed from mounting a campaign, even when she has quite good information. This arises when her interim probability of winning (the probability she wins conditional on her own information but not conditional on the incumbent's information) is low because of the prior strength of the incumbent.

If the voters perceive a large difference between "good" and "bad" candidates, and if the value of favors a candidate can promise to contributors is small, then this foreclosure outcome may be inefficient. The basic idea is as follows: if the challenger has good information, then it is likely she is high ability. This is not enough to ensure that she can advertise her information, however. That requires that she be able to promise sufficient rents to a contributor to overcome the ex-ante advantage of the incumbent. Unfortunately, this is the wrong comparison from a social welfare perspective. Instead of comparing the probability-weighted cost of advertising to the pledgeable rents, a social planner would compare the same weighted cost to the value of a high ability candidate times the probability that revealing the signal changes the outcome of the election. If this last quantity is large relative to the rents, then foreclosure is inefficient.

The welfare implications of the incumbent's spending rule are more straightforward. The value of advertising her marginal signal is exactly zero in equilibrium. Since the marginal ad is costly, she advertises too much in a foreclosure equilibrium.

In the equilibria in which both candidates advertise, the welfare analysis is similar to the case of the challenger in the incumbent only equilibrium. Instead of comparing the cost of ads to the social value of the information, the contributors compare the cost to the expected value of favors that candidates can promise. In general, this can lead to too much or too little advertising in equilibrium.

Several aspects of this story deserve comment. First, the information is hard information and cannot be falsified. While this is a strong assumption, there are many examples of hard information in campaigns. Prominent examples include endorsements, interest group ratings, and roll call votes on prominent bills. While candidates can lie about these things
in ads, doing so is risky. Opponents and journalists have strong incentives to uncover and publicize such lies. In fact, many news outlets (CNN, for example) have regular features examining the reliability of political advertising.

Voters will surely discount claims which they learn to be lies. Indeed, they may even draw adverse inferences about the candidate’s ability: if the candidate had good information to reveal, then why wouldn’t she reveal that rather than lie? If a voter reasons that way, he will lower his estimate of the candidate’s ability if she lies. Furthermore, being caught lying will diminish the candidate’s reputation for honesty. My assumption that information cannot be falsified corresponds to a situation in which these costs are so great that no candidate chooses to lie.

Second, I assume that the candidates can reveal information about themselves. This is not crucial for the analysis; since votes only care about the difference in quality between the two candidates, negative ads about an opponent are just as valuable as positive ads. This observation strengthens the case for assuming verifiable information, since it is often observed that negative ads contain more hard information than positive ads.

2.1.1 Related Literature

The work most closely related to this paper is Prat (2000a). He studies an election between two symmetric candidates who have unknown “valence”. Interest groups can observe the valence and condition their contributions on it. Voters then learn about valence by inverting the contribution schedule. He shows that an equilibrium exists with informative advertising, even thought the ads have no direct informational content. He also shows that for some parameter values the median voter wants to ban advertising, while for other values he does not. My paper expands on Prat’s welfare analysis by removing the symmetry assumption on the candidates, so there can be incumbency effects, and by considering a richer class of potential policy instruments.

The main difference between my model and Prat’s model is the treatment of the informational content of ads. I assume that ads contain hard information, while the informativeness of ads in Prat is a signaling phenomenon, which arises as part of the equilibrium. Indeed, Prat considers the effectiveness of advertising with no direct information one of the major stylized facts about elections. However, this conclusion is a bit premature. Ads (particularly negative ads) to contain verifiable information, such as votes on key bills, the state of the
economy, and endorsements. Furthermore, we should not assume that campaign promises provide no information just because they are not enforceable by courts. Banks (1990) shows that concern for reputation can lead candidates to choose promises which are informative in equilibrium. Harrington (1992) and Postlewaite and Aragones (2000) provide explicit dynamic models of this phenomenon.

Prat (2000b) differs from Prat (2000a) in two important way, both shared by my work: he treats incumbents and challengers asymmetrically, and the candidate, rather than interest groups, has all of the bargaining power. He shows that the results in his model are insensitive to the assignment of bargaining power. Since the focus of his paper is on the question of bargaining power, he simplifies the model by making the challenger a “dummy player”, who takes no actions. I use the opposite simplifying assumptions: I consider only a trivial bargaining game but the incumbent and challenger are symmetric with respect to their strategy sets. Allowing for strategic symmetry and reputational asymmetry at the same time is important for my results, since only in that case is it possible that a social planner would find it optimal to simultaneously reduce the expenditure of the incumbent and increase the expenditure of the challenger.

Snyder (1990) uses an arbitrage argument to derive a condition similar one of the equilibrium conditions I derive below, and tests it on data from House elections for open seats. He finds strong support for a specification which assumes the equation holds for close races while candidates who face token opposition decline to raise funds. My model provides an equilibrium justification for the decision of strong candidates to bypass fundraising, and it shows how to extend the arbitrage intuition to a broader range of competitive environments.

2.2 Model

Consider an election with two candidates, $L$ and $R$. Denote a generic candidate by $c \in \{L, R\}$. Each candidate has a fixed policy position $x^c \in \mathbb{R}$, with $x^L < 0 < x^R$. These positions are common knowledge. To keep the candidates as symmetric as possible, assume $x^L = -x^R$. Each candidate also has ability $\theta^c \in \{\theta, \overline{\theta}\}$, with $\overline{\theta} > \theta$.

There are a continuum of voters, indexed by their ideal points $x \in \mathbb{R}$. Ideal points are distributed symmetrically about their unique median, 0. Each voter has
preferences represented by the expectation of

$$\theta^W - |x^W - x|,$$

where $W \in \{L, R\}$ is the winner of the election and $x$ is the voter’s ideal policy.

With this specification, all voters agree that higher ability is better. (Political scientists call such issues “valence” issues, following Stokes 1963.) There are several interpretations of ability which imply such agreement. An obvious candidate is talent for handling international crisis, an important factor in elections for the Presidency. Another example, more relevant for congressional elections, is skill in providing pork for the district. Finally, ability could represent the absence of corruption.

Ability is unknown to the voters. (Whether or not candidates know $\theta$ is irrelevant.) They have prior beliefs that $\theta^e = \bar{\theta}$ with probability $q^e$. Assume, without loss of generality, that $L$ is the incumbent candidate. We represent this by assuming that $q^L \geq 1/2$, while $q^R = 1/2$.

The incumbent candidate has higher ability (on average) because of selection effects in repeated elections. A candidate is an incumbent because she won elections in the past. Presumably, this means that the voters observed favorable information about the candidate in past campaigns. Furthermore, voters get some information about an incumbent by observing outcomes while she is in office. If this information is not favorable, the candidate is likely to lose a subsequent election. Thus a candidate who has survived several elections is likely to be above average in ability.

Each candidate receives an informative signal $\sigma^c \in \Sigma = [\sigma, \bar{\sigma}]$. Higher signals are good news about ability. We represent this with the standard assumption that the conditional density $f(\sigma | \theta)$ satisfies the strict monotone likelihood ratio property (MLRP):

$$\sigma > \sigma' \Rightarrow \frac{f(\sigma | \bar{\theta})}{f(\sigma | \bar{\theta})} > \frac{f(\sigma' | \bar{\theta})}{f(\sigma' | \bar{\theta})}.$$

Let $f^c(\sigma)$ denote the unconditional likelihood that candidate $c$ receives signal $\sigma$.

If a voter is shown signal $\sigma^c = \sigma$, then his posterior belief about her ability $\theta^c$, $p^e(\sigma)$, is derived using Bayes’s rule. Our assumption that higher signals are good news about ability implies that $p^e(\sigma) > p^e(\sigma')$ whenever $\sigma > \sigma'$.

Candidates can buy advertising time on television to show their signal to the voters.
The cost of this advertising is $c > 0$. The signals are "hard" information, so the candidates cannot lie about their signals.

Candidates have no wealth and no access to credit markets. The winner of the election will have one divisible unit of favors she can provide while in office. Candidates can promise some of these favors to contributors, or they can hold them back, in anticipation of using them on their friends should they win the election.

Because we assume the winner of the election will provide a fixed amount of favors, regardless of how many she promised in the election, voters do not care about the level of favors promised during the campaign. Although this is not a realistic assumption, it allows us to focus on the role of the interim probability of winning in determining which candidates get funding, and on the divergence between the equilibrium and optimal levels of spending. In an extension to the basic model (discussed in section 6), we relax the assumption that voters do not care about the promised level of favors, and show that the insights of the basic model are robust to this extension.

Each candidate has the option of approaching two contributor groups and letting them compete for the right to make campaign contributions in exchange for promised favors. The competition takes the form of an auction: the groups simultaneously announce levels of favors such that they are willing to pay the cost of advertising in exchange for the rents contingent on the candidate winning. If both groups demand more than $R$, then the candidate has no access to funds. This Bertrand-style competition reflects the competitiveness of the market for favors.

The contributors have unlimited wealth and are risk-neutral.

Candidates care about their probability of winning the election and about the amount of favors they have to distribute ex post. Their preferences are lexicographic: if two strategies have different probabilities of winning, they prefer the strategy with the higher probability of winning, and if two strategies have the same probability of winning, they prefer the one in which they pledge fewer rents.

The assumption of lexicographic preferences implies that the candidates always want to trade favors for a higher probability of winning, but want to minimize the level of promised favors conditional on a fixed probability of winning. This is a natural pattern of behavior, and one which arises endogenously in the extension where voters think favors are costly and candidates are motivated purely by winning.
The timing is as follows:

1. Nature chooses $\theta^c$ and $\sigma^c$ for $c = L, R$.
2. Candidates decide to sell favors or not.
3. If they do, contributors compete over favors.
4. Candidates reveal information if they have funding.
5. Election.

We look for (trembling-hand) perfect equilibria. (See section 12.5 of Osborne and Rubinstein 1994.) We begin by pointing out some features which must hold in any perfect Bayesian equilibrium; since a perfect equilibrium induces an assessment which makes up a PBE, all of these properties must also hold in any PE. (Throughout the analysis, explicit notice will be given when the full requirements of perfection are used.)

2.3 Preliminary Results

We approach the solution of the model in two steps. First, we characterize the decisions to seek funds in terms of cutoff rules, so that any candidate with a signal above some threshold, $\tilde{\sigma}^c$, advertises. Along with showing that such rules are optimal, we provide a simple formula which expresses the quantity of favors promised as a function of the signal, the cost of ads, and beliefs about the other candidate's information. Our ability to write the equilibrium cutoffs in terms of this formula will be the key step in the welfare analysis of section 5.

Before turning to that welfare analysis, section 4 uses the results about how many favors a candidate must promise to complete the solution, actually determining the cutoffs. For incumbents with relatively weak reputations, the equilibrium features both candidates advertising high signals. For moderate reputations, the challenger is never able to raise funds, while the incumbent does advertise high signals. For incumbents with strong reputations, the equilibrium has no advertising.

Before we can characterize the cutoffs, we must determine the voters' optimal strategies. In the election, the voters will vote for whichever candidate offers the highest expected
utility. Let $\mu^c$ be the posterior probability that candidate $c$ is of ability $\overline{\theta}$. If candidate $c$ wins, a voter with ideal policy $x$ gets expected utility

$$\mu^c\overline{\theta} + (1 - \mu^c)\overline{\theta} - |x^c - x|,$$

which simplifies (up to a constant) to

$$\mu^c \Delta - |x^c - x|,$$

where $\Delta = \overline{\theta} - \bar{\theta} > 0$. Thus voter $x$ prefers $L$ to $R$ if and only if

$$(\mu^L - \mu^R) \Delta \geq |x^L - x| - |x^R - x|.$$

Because the voters’ utility functions are supermodular in $x$ and $x^c$, the Gans-Smart (1996) median voter theorem implies that the majority preference is exactly the preference of the voter with the median ideal policy. In our case, that is the voter with ideal point $x = 0$, so our symmetry assumption $x^L = x^R$ implies that $L$ wins exactly when $\mu^L \geq \mu^R$.

The voters’ beliefs are based on the information they receive during the campaign. The information about candidate $c$ is an element of $\Sigma \cup \{\emptyset\}$, where the symbol $\emptyset$ denotes no signal at all. To be part of a perfect Bayesian equilibrium, beliefs must be derived from Bayes’s rule and the equilibrium strategies. Define $\mathcal{A}^c$ to be the set of types of candidate $c$ who advertise in equilibrium. If $\sigma \in \mathcal{A}^c$, then beliefs must satisfy $\mu^c(\sigma) = p^c(\sigma)$. If a voter sees no signal from candidate $c$, he assumes her type is not in $\mathcal{A}^c$, and his posterior belief is

$$\mu^c(\emptyset) = \int_{\sigma \in \mathcal{A}^c} p^c(\sigma) f^c(\sigma) \, d\sigma.$$

To complete the specification of a perfect Bayesian equilibrium, we have to say what beliefs would be off the equilibrium path. Since we are only interested in PBE which are also perfect, this is straightforward. Beliefs must be the limit of beliefs derived by Bayes’s rule from some sequence of totally mixed strategies. Thus

$$\mu^c(\sigma) = \frac{q^c f(\sigma | \overline{\theta})}{f^c(\sigma)}$$

for all $\sigma$. Given this, the lowest type of each candidate will not advertise, since doing so
cannot increase the posterior belief about her and doing so is costly in terms of rents. Thus \( \mu^c(\emptyset) \) is always determined by the formula above.

With these preliminaries out of the way, we turn to the characterization of equilibrium in the fundraising stage. First, we show that the Bertrand-style competition over rents forces the contributor groups to exactly break even. Using this, we are then able to show that the equilibrium has a cutoff property—if any type advertises in equilibrium, then every higher type (of the same candidate) also advertises.

**Claim 1** Contributors break even in expectation: if candidate \( c \) advertises signal \( \sigma \), then the level of rents she promises to the contributor, \( r^* \), satisfies

\[
    r^* = \frac{c}{\Pr(\mu^{-c} \leq \mu^c(\sigma))}.
\]

**Proof** The payoff to a group which funds a campaign is

\[
    r^* \Pr(\mu^{-c} \leq p^c(\sigma)) - c.
\]

If this is greater than zero, then the other group could offer to fund the campaign for promised rents of \( r^* - c \); the candidate strictly prefers this offer and, for \( \epsilon \) small enough, the deviating group gets a positive payoff. Since that is better than the zero it would get when not funding the campaign, this is a profitable deviation.

If the groups payoff is less than zero, then it can do better demanding \( r > R \). Since this amount is not feasible for the candidate, the contributor gets 0, which is better. Thus the payoff must be zero. \( \square \)

This result has two important implications for the analysis. First, if

\[
    R < \frac{c}{\Pr(\mu^{-c} \leq p^c(\sigma))},
\]

then the candidate cannot get funding for advertising. In that case, we say the candidate is foreclosed from funds. The second important consequence of the claim is that, if

\[
    R > \frac{c}{\Pr(\mu^{-c} \leq p^c(\sigma))},
\]

then the candidate can get funding for a campaign if she wants it.
Our next result establishes the cutoff property of equilibrium strategies.

**Claim 2** If type $\sigma'$ of candidate $c$ advertises with positive probability, then any higher type $\sigma$ advertises with probability 1.

**Proof** Since rents are costly to the candidate, the fact that she advertises with signal $\sigma$ implies that

$$\Pr(\mu^{-c} \leq p^c(\sigma')) > \Pr(\mu^{-c} \leq \mu^c(\emptyset)).$$

This, along with MLRP, imply that if $\sigma > \sigma'$ then

$$\Pr(\mu^{-c} \leq p^c(\sigma)) > \Pr(\mu^{-c} \leq \mu^c(\emptyset)).$$

Furthermore, claim 1 implies that

$$r^*(c, \sigma) = \frac{c}{\Pr(\mu^{-c} \leq p^c(\sigma'))} > \frac{c}{\Pr(\mu^{-c} \leq p^c(\emptyset))},$$

so the candidate can afford to promise sufficient rents to get funding. □

### 2.4 Who Advertises?

The previous section characterized equilibrium behavior in the fundraising and electoral stages of the game, conditional on which types of each candidate desired campaign funds. To complete the description of the equilibrium, we need to determine who those types are. This section uses the incentive compatibility constraints that characterize equilibrium to identify which subsets of the parameter space lead to which types deciding to mount a costly campaign.

There are several types of equilibria to consider, depending on how different are beliefs about the incumbent and the challenger. When the two candidates are relatively evenly matched, the equilibrium features both candidates advertising for sufficiently good signals. On the other hand, when the incumbent candidate is sufficiently ahead, no one advertises in equilibrium. For intermediate levels of competition, the equilibrium features advertising only by the incumbent. This equilibrium has the feature that the incumbent wins exactly when she advertises.
First, consider a scenario in which neither candidate campaigns for any value of the signal. When can this be part of an equilibrium?

Given these strategies, the voters will have posterior beliefs equal to their priors. This means that the incumbent \( L \) will win the election. Given this, she has no incentive to deviate and raise funds—she is already getting the best possible payoff. The challenger does have an incentive to try to mount a campaign. If there is a value of the signal, \( \sigma \), such that the posterior probability of \( \hat{\theta} \) given \( \sigma \) is greater than the incumbent’s prior ability, then the challenger can win if the voters see \( \sigma \). Given this, interest groups would be happy to fund a campaign. Thus the only way no challenger type will deviate is

\[
\max_{\sigma \in \Sigma} \mathbb{E}(\theta^R|\sigma) \leq \mathbb{E}(\theta^L).
\]

By MLRP, this is true just when

\[
q^L > p^R(\overline{\sigma}).
\]

Thus a no advertising equilibrium exists only when \( p^R(\overline{\sigma}) \) is small relative to the expected ability of the incumbent. If the model is to be consistent with the fact that incumbent congressmen often win with large margins while spending very little money, it must be the case that this inequality is satisfied for a reasonably large range of incumbent reputations.

Next consider the case in which both candidates campaign for some values of the signal. From claim 2, we know that the equilibrium advertising levels follow cutoff rules. Write \( \overline{\sigma}^c \) for the cutoff describing candidate \( c \)'s strategy, so \( c \) shows any signal \( \sigma \geq \overline{\sigma}^c \). Given these cutoffs, the voters' beliefs about candidates who do not advertise are

\[
\mu^c(\emptyset) = \int_{\overline{\sigma}^c} q^c f(\sigma|\emptyset) \frac{f^c(\sigma)}{f^c(\sigma)} d\sigma.
\]

In this case, the competitive pressures of the campaign force both candidates to advertise all signals for which contributors are willing to give funding. The intuition for this is the familiar unraveling argument from Grossman (1981) and Milgrom (1981). The types who do not advertise are pooled together, and the voter assesses that their ability is the average of those types who do not advertise. The highest type in this pool can improve the voter’s perception of herself by deviating and showing her signal. Since she gets higher expected
payoff when beliefs about her are more favorable, this is a profitable deviation.

This suggests the following fact.

**Claim 3** If $(\bar{\sigma}^L, \bar{\sigma}^R)$ is a pair of cutoffs used in some equilibrium in which both candidates advertise, then $\bar{\sigma}^L$ is given by

$$\Pr(\mu^R \leq \mu^L(\bar{\sigma}^L)) = \frac{c}{R},$$

and similarly for $\bar{\sigma}^R$.

**Proof** By claim 1, we know that the probability must be at least ($c/R$). Furthermore, the argument above shows that a candidate will reveal signals all the way down to this limit if she can strictly increase her chance of winning by doing so. We just need to show that she can in fact strictly increase the probability in any perturbed game supporting a candidate perfect equilibrium. In fact, this is easy to see. Fix any two possible posteriors for $R$, $\mu^R$ and $\mu^{R'}$. With positive probability, $R$ will get a signal implying a posterior between these two, and with positive (conditional) probability, she will seek and get funding to reveal it. Thus $L$ can strictly increase her chance of winning by showing a signal which implies her probability of being high ability is between $\mu^R$ and $\mu^{R'}$. □

**Remark** Although the discussion leading up to the proposition seems to be enough to get the conclusion, more is in fact involved. That argument relies crucially on the increase in the probability of winning being strictly positive. However, partial pooling means the equilibrium cdf of a candidate's public beliefs is not everywhere strictly increasing, even though the cdf of the private posterior is. In fact, this is the first point where the restriction to trembling-hand perfect equilibria is crucial: allowing for trembles perturbs the cdf of the public beliefs so that it is strictly increasing. □

Finally, consider the case in which only one candidate advertises. (It must be the incumbent who advertises.) We make several observations: first, the challenger must win with positive probability when the incumbent does not advertise. If this were not true, then the incumbent could deviate by not advertising, keeping the rents, and still winning the election. This implies $\mu^L(\emptyset) \leq 1/2$. Given this, the cutoff rule the incumbent follows must
be the $\tilde{\sigma}$ such that

$$\mu^L(\tilde{\sigma}) = \frac{1}{2}.$$ 

If the cutoff were higher, the incumbent could advertise the signals between the cutoff and $\tilde{\sigma}$ and change her fortune from losing to winning. On the other hand, no lower cutoff is possible because no contributor will fund a candidate with a lower signal.

For this to be an equilibrium, candidate $R$ must be not deviate and try to finance a campaign. By MLRP, it suffices to show that the $R$ candidate cannot deviate when she has signal $\overline{\sigma}$. For this to be the case, we must have

$$\Pr(\mu^L \leq p^R(\overline{\sigma})) < \frac{c}{R^*}.$$ 

We can rewrite this condition in terms of the primitives of the game. The voters will believe that the incumbent’s probability of being high ability is less than $p^R(\overline{\sigma})$ exactly when the incumbent gets a signal lower than $\tilde{\sigma}$, where $\tilde{\sigma}$ satisfies

$$p^L(\tilde{\sigma}) = p^R(\overline{\sigma}).$$

Thus the necessary condition for an incumbent-only equilibrium is

$$F^L(\overline{\sigma}) < \frac{c}{R^*}.$$ 

If this inequality is not satisfied, then the candidate could raise funds if she wanted to, by claim 1 above. Furthermore, she would mount a campaign, because it would strictly increase her chance of winning the election.

Collecting these facts, we can determine which type of equilibrium we will see as a function of the prior belief that the incumbent is high ability. Our characterization will use the following two numbers: let

$$\overline{q} = \frac{f(\overline{\sigma} | \theta)}{f(\overline{\sigma} | \overline{\theta}) + f(\overline{\sigma} | \theta)},$$

48
and let \( q \) solve

\[
q F(\sigma(q) \mid \theta) + (1 - q) F(\sigma(q) \mid \theta) = \frac{c}{R}.
\] (2.1)

These numbers have intuitive interpretations. \( \bar{q} \) is the least favorable prior for the incumbent such that no signal for the challenger can lead to a posterior better than the incumbent’s prior. \( q \) is the most favorable prior for the incumbent such that some contributor can break even funding a campaign for the challenger.

Before proceeding, we must make sure \( q \) is well defined.

**Claim 4** Equation (2.1) has exactly one solution.

**Proof** We must show that a solution exists and that it is unique. For uniqueness, notice that if \( q = 1 \), then the LHS is 0, while if \( q = 0 \) the LHS is 1. Since \( (c/R) \in (0, 1) \) and the LHS is continuous, the intermediate value theorem implies that there is a solution.

Next, we show that the LHS is strictly decreasing in \( q \), so the solution is unique. Consider two values of \( q, q > q' \). Since \( q \) indexes \( F^L \) in the sense of first-order stochastic dominance, \( \sigma(q) \) is decreasing in \( q \), so \( \sigma(q') > \sigma(q) \). Thus we have

\[
F^L(\sigma(q') \mid q') > F^L(\sigma(q') \mid q)
\]

\[
> F^L(\sigma(q) \mid q),
\]

where the first inequality follows from \( F(\cdot \mid \theta) \geq F_{OSD}(\cdot \mid \theta) \) and the second follows from the fact that \( F \) is strictly increasing. \( \square \)

Using these numbers, we can collect the results above as the following proposition.

**Proposition 5** If the incumbent’s reputation is weak \( (q^L \in [1/2, \min(q, \bar{q})] \) \), then in any equilibrium both candidates advertise.

If the incumbent’s reputation is strong \( (q^L \in [\max(q, \bar{q}), 1]) \), then there is an equilibrium in which no candidate advertises.

For an incumbent with a moderate reputation, there are two possibilities: if \( \underline{q} < q^L < \bar{q} \), then in any equilibrium only candidate L advertises; and if \( \bar{q} < q^L < q \), then there is an equilibrium in which both candidates advertise and one in which no one advertises.
Intuitively, as the incumbent's reputation improves, contributors become more skeptical that challengers will end up in a position to make good on their promises of favors. Consequently, they demand more (contingent) favors, eventually demanding more than the challenger can promise.

In general, \( \bar{q} > q \) and \( \bar{q} < q \) are both possible.

**Example** Consider the case where \( \Sigma = [0,1] \), \( R = 1 \), \( f(\sigma | \bar{\theta}) = 1 \), and \( f(\sigma | \theta) = 2\sigma \). Direct calculation shows that \( \bar{q} = 2/3 \). Before we find \( q \), we need to know \( \bar{\sigma}(q) \), which, recall, is the value of \( \sigma^L \) which leaves a voter indifferent between an incumbent with prior ability \( q \) and a challenger who has the best possible signal. Here, \( \bar{\sigma} \) satisfies

\[
\frac{q2\sigma}{q2\sigma + (1 - q)} = \frac{2}{3},
\]

so

\[
\bar{\sigma}(q) = \frac{1 - q}{q}.
\]

Thus \( q \) solves

\[
q \left( \frac{1 - q}{q} \right)^2 + (1 - q) \left( \frac{1 - q}{q} \right) = c.
\]

Simplify this to get a quadratic in \( q \) whose roots are

\[
q = \frac{(4 + c) \pm \sqrt{(4 + c)^2 - 16}}{4}.
\]

Now it's easy to see that \( q \) can be greater than or less than \( \bar{q} \). If \( c = 1 \), then the unique root is \( 1 > 2/3 = \bar{q} \). If \( c = 0 \), then the root in \([0,1]\) is \( 1/2 < 2/3 = \bar{q} \).

Before turning to the welfare analysis, it is useful to look at some simple comparative statics. As the prior on the incumbent's ability increases from \( 1/2 \) to \( 1 \), the equilibrium probability that the incumbent advertises first increases, and then falls to \( 0 \). Her chance of winning increases the whole way. This means that a cross-sectional statistical analysis would find a relatively weak relationship between spending and the probability the incumbent wins. This is important because one of the robust findings of the empirical literature is exactly that weak relationship.
2.5 Social Welfare

In this section we consider the relationship between the equilibria found above and the decision rule that would be used by a surplus maximizing social planner. There is an important difference between the case with both advertising and only the incumbent advertising. In the first case, the cutoff levels of the signal are sensitive to the cost of advertising. In the case of only incumbent advertising, on the other hand, the cutoff is independent of the cost. This means that the welfare effects of advertising will be different in the two cases.

Total surplus is equal to sum of the policy payoff, the value of office to the winner, the value of the favors and the winner’s quality minus the cost of any advertising. Since the value of office must be received by one of the candidates, and the favors will be received by someone (either contributors or friends of the winner), we can ignore those terms in the analysis.

Furthermore, we will ignore the effects of different campaign regulations on the expected policy payoff. Since the distribution of ideal points and the candidate locations are both symmetric about the median voter, the ex-post policy surplus is the same from either candidate. We are left with comparing the benefits of better sorting in the election to the cost of information revelation.

We are interested in constrained optimality. Clearly, we must restrict the planner to use the same signals and information channels as the private actors. However, these considerations do no exhaust the relevant constraints. We also insist that policies respect the decentralization of the campaign: decisions about revealing $L$’s signal should not depend on the value of $R$’s signal. (They will depend on the prior distribution of the signal.)

To see why this restriction is important, consider the following mechanism: both candidates report their signals to the planner. He computes the optimal choice for the median voter and sends a single message—an instruction of who to vote for. (This can be improved still further—send a message only when $R$ is the optimal choice.) Such a mechanism can be designed to dominate any decentralized mechanism. It is not, however, a sensible description of the optimal policy. This mechanism would be subject to corruption and is very undemocratic. Furthermore, it bears no resemblance to any feasible reform plan. In particular, any policy of taxing or subsidizing contributions, or limiting fundraising will be decentralized in this sense.
To avoid mechanisms like that, we focus on the question of the interim optimality of a candidate’s advertising plan, holding fixed both the plan of the other candidate and the voters’ strategies.\(^1\) The candidates’ plans will (generically) not be interim optimal, even given these constraints.

Consider first the case where only the incumbent advertises. We know from the description of equilibrium above that the incumbent’s cutoff \(\tilde{\sigma}\) satisfies \(\mu^L(\tilde{\sigma}) = 1/2\). We also know that the incumbent wins exactly when she advertises. If she does not advertise, then the challenger is elected and the voters get a high ability office-holder with probability 1/2. The key point is that these two probabilities of high ability are the same; there is no social benefit from revealing \(\tilde{\sigma}\). Thus the social value of the revelation is \(-c < 0\). This inefficiency arises because the equilibrium decision to advertise or not is independent of the cost of advertising in this regime.

Now consider the challenger. Is it interim constrained efficient for her to not show the signal \(\tilde{\sigma}\)? Let \(\tilde{\sigma}\) solve

\[\mu^L(\tilde{\sigma}) = \mu^R(\tilde{\sigma}).\]

Advertising \(\sigma^R = \tilde{\sigma}\) changes the outcome of the election only if \(\sigma^L \in [\tilde{\sigma}, \tilde{\sigma}]\). The social benefit of advertising is the probability of this event, \(\Pr(\sigma^L \in [\tilde{\sigma}, \tilde{\sigma}])\), times the expected improvement in the probability that the winner is high ability conditional on this event, \(\mathbb{E}(\mu^R(\tilde{\sigma}) - \mu^L \mid \sigma^L \in [\tilde{\sigma}, \tilde{\sigma}])\), times the incremental benefit to high ability, \(\Delta\). This benefit must be traded off against the cost, \(c\). Putting all of this together, the social value of advertising the highest signal for the challenger is

\[\Pr(\sigma^L \in [\tilde{\sigma}, \tilde{\sigma}]) \cdot \mathbb{E}(\mu^R(\tilde{\sigma}) - \mu^L \mid \sigma^L \in [\tilde{\sigma}, \tilde{\sigma}]) \cdot \Delta - c.\]

It is instructive to write the constrained optimality condition a little differently. Notice that the challenger wins whether she advertises or not if the incumbent’s signal is less than \(\tilde{\sigma}\). Thus the expected value of the change in the winner’s expected ability given advertising is zero for that range of signals. Writing \(\mu^W\) for the posterior probability that the winner

\(^1\)This is related to the idea of a Nash Social Welfare Optimum from the literature on general equilibrium with incomplete markets. See Grossman 1977.
is high ability, we can rewrite the efficiency condition as

$$\Pr(\mu^L < \mu^R(\bar{\sigma})) \cdot \mathbb{E}(\mu^W - \mu^L \mid \mu^L < \mu^R(\bar{\sigma})) \cdot \Delta \geq c.$$  

Comparing this to the equilibrium condition we derived, we see that the exclusion of R from campaigning when L has reputation \(\bar{\sigma}\) is inefficient if the expected change in the ability of the winner, given that \(\bar{\sigma}^R\) is pivotal for the election,

$$\mathbb{E}(\mu^W - \mu^L \mid \mu^L < \mu^R(\bar{\sigma})) \cdot \Delta,$$

is greater than the value of pledgable favors, \(R\).

The conditional expectation in the marginal benefit term depends only on the distributions of the signals, the incumbent’s reputation, and the ratio of advertising cost to pledgable favors. It is independent of \(\Delta\). This implies that we can easily construct examples in which the planner would not want the challenger to advertise; just let \(\Delta\) be close to 0. On the other hand, by taking \(\Delta\) large enough, we can create examples where the planner would like to subsidize advertising for the challenger. It is striking that we can construct examples in which the planner wants to simultaneously decrease the advertising of incumbents and increase the advertising of challengers.

Next consider an equilibrium in which both candidates advertise. The social value of advertising \(\bar{\sigma}^c\) is given by an expression like that for the challenger in the case above. For example, the condition for \(L\) is

$$\Pr(\mu^R \leq \mu^L(\bar{\sigma}^L))\mathbb{E}(\mu^W - \mu^L \mid \mu^R \leq \mu^L(\bar{\sigma}^L))\Delta - c.$$  

Like the case above, we can vary \(\Delta\) to get examples where there is too much advertising and where there is too little, but now it is possible for the condition for both candidates to go either way.

In this case, the equilibrium decision does take into account the cost of the ad, but the cost and benefits are not given the socially correct weights. If the pledgable rents are large relative to \(\bar{\theta} - \theta\), then there will be too much advertising. Otherwise, there will be too little.

We collect these observations in the following proposition.

**Proposition 6** In an equilibrium in which both candidates advertise, there can be too much
or too little advertising. In an equilibrium in which only the incumbent advertises, the incumbent always advertises too often. It may be socially efficient to increase the challenger's advertising even while decreasing the incumbent's.

2.6 Conclusion

It is tempting to draw policy conclusions from the model in this paper. Zaller (1998) reports that the median challenger of a one-term incumbent in 1990 spent only $36,000, about one tenth the amount spent by the median loser in open-seat elections. Sophomore incumbents, on the other hand, spent almost as much as the winners of open seat elections. Although far from conclusive evidence, this certainly suggests that the typical case is that sophomore incumbents feel the need to run expensive campaigns, while their challengers are unable to do so. This is the case where the model has an unambiguous welfare conclusion: the marginal incumbent is spending too much.

Drawing such conclusions is probably premature. There are several extensions which should be made to the model before it is brought to bear on policy. First, a commonly expressed concern about campaign finance is that fundraising forces politicians to expand their provision of favors, so the fundraising is costly for the voters. Our model abstracts from this. However, the basic insights of the model are robust to this extension. Since voters will evaluate candidates on the basis of both ability and expected favors, candidates will only mount campaigns when they have information good enough to overcome the penalty due to promised favors. Since the challenger will still have to promise more favors than the incumbent, the incumbency advantages in fundraising studied above will persist in the richer model.

Furthermore, the equilibrium in which only the incumbent advertises is suggestive of an entry deterrence story. The challenger is unable to raise funds, but if it were to become common knowledge that the incumbent would not advertise regardless of her signal, the challenger could raise funds. However, this story needs to be made dynamic. Giving the incumbent the chance to raise funds prior to the simultaneous fundraising game will allow the incumbent to build a war chest in an attempt to deter entry.

On the institutional side, the model could be enriched to include "soft money", which is money given to parties rather than individual candidates. This institution may be able to
overcome the inefficiency studied here. Soft money allows the funding decision to be partly separated from the interim probability that the candidate will win, because the party will be able to pay off contributors even if the specific candidate loses. Of course, it is far from obvious that parties will actually use their soft money in an informationally efficient way. They will also be sensitive to probabilities of winning, and they presumably have objectives not limited to providing high ability office holders. It is an interesting question how these factors trade off against one another to determine which institutional form is best at providing information to citizens, especially in light of the recent interest in soft money bans among policy-makers.
Chapter 3

Rational Learning and the Incumbency Advantage in Congressional Elections

Joint with Aaron Hantman

3.1 Introduction

In recent elections, Democratic incumbent congressmen have received 75% of the two-party vote, on average. At the same time, Democrats running in open-seat elections received an average of about 55% of the two-party vote. Numbers for Republican incumbents are similar. Numbers such as these raise understandable concerns about the competitiveness of elections. However, before we conclude that policy needs to be changed to encourage more competition, we need to identify what factors drive the success of incumbents.

There are at least two reasons incumbents might do so well in elections. First, elections might be efficiently matching representatives to districts. This matching involves both ideological fit between the representative and the district and the incumbent having high ability to perform tasks for her constituents. Incumbents are likely to have both good fit to the district and high ability; after all, voters have had a chance to learn about their representatives, and vote out those who disappointed them. The second reason incumbents might win so often is that elections are slanted in their favor. This is referred to as the
incumbency advantage in the literature. These two explanations have different implications for policy. The first suggests that high reelection margins are a good sign, since they show that voters have high quality representatives. The second, on the other hand, suggests that voters might get stuck with low quality representatives because high quality challengers cannot overcome their competitive disadvantage.

The early attempts to estimate the incumbency advantage were based on intuitive comparisons: the "sophomore surge", which is the average increase in vote share between the first win and first reelection bid, or the "retirement slump"\(^1\), which is the average decline in a party's support in a district following a retirement. Later papers have combined these measures, looking at a simple average called the surge. Although quite intuitive, these measures were not given any formal justification.

In an important paper, Gelman and King (1990) examined the foundations of these estimates. They assume that incumbency adds a constant share of the vote, irrespective of any other characteristics. Under this assumption, they showed that all three measures are biased, because of the way they select elections. Gelman and King also provide a simple regression-based estimator of the incumbency advantage. Their estimate is just the coefficient on an indicator variable for incumbency from a regression of the Democrat share of the vote on incumbency and lagged Democrat share.

Although widely recognized as a major advance, Gelman and King's procedure has not escaped criticism. Ansolabehere, Snyder, and Stewart (1999), for example, criticize Gelman and King's procedure because the use of lagged vote to capture the normal vote is not well motivated theoretically, and because the coefficient on lagged vote is substantially less than one, contrary to the empirical justification offered by Gelman and King.

We address both of those concerns. We present a simple model of an election driven by voter learning about candidate quality. After taking a linear approximation to the model, basic properties of Bayesian updating yield precisely the equation estimated by Gelman and King. We are also able to address the second point: the (linearized) model predicts that the coefficient on lagged vote should be 1, but this prediction is only valid for a range of races where the linear approximation is valid.

We estimate the model on a subsample of races which are expected to be "close". We select the sample by taking districts in which the last election was close. These races should

---

\(^1\)See, for example, Alford and Brady 1988.
be close because, in the context of our model, these are districts in which the incumbent has not been successful in building a strong reputation. The key assumption justifying this selection rule is that there is no predictable trend in voters’ beliefs about an incumbent. This is an implication of Bayesian learning, but it is also consistent with a variety of boundedly rational learning rules.

The results are supportive of the model. With the exception of three elections, the restrictions implied by the learning model are not rejected by the data. The exceptions are years that are well known to be bad for incumbents. Our estimates of the treatment effect of incumbency are uniformly lower than Gelman and King’s. This indicates that the restriction to races where our model implies their procedure is valid is substantively important.

Although our main purpose in introducing the model of rational learning is to better understand estimates of the incumbency advantage, our test of the learning model is of independent interest. Several recent papers assume that voters are learning about candidate ability over time and study the interaction of the learning and electoral incentives. Examples are Lohmann (1998) on the bias toward special interests, Persson and Tabellini on corruption with different electoral rules (2000, ch. 9), and chapter 1 of this thesis on the dynamics of effort allocation over Congressional careers. Our evidence supporting the rational learning assumption lends important support to the analysis of these papers.

3.2 A Simple Model

A candidates for office are characterized by a level of quality, \( \theta \in \mathbb{R} \). All candidates are ex-ante identical with respect to quality. Uncertainty about quality is completely symmetric throughout the game. Each period the candidate holds office, the public gets an informative signal, \( \sigma \), of her unknown quality. Signals are independent conditional on \( \theta \).

The signal could take several forms. The voters could learn about their incumbent from the media, or the campaign could be informative. Even if voters are not motivated to follow politics closely enough to get information from these sources, they could learn about the incumbent just by observing their own utility. For example, a high quality candidate might be better able to secure local public goods for the district. This will directly affect the voters’ utility.
If a candidate has been in office for \( t \) periods, then \( h^t = (\sigma^1, \ldots, \sigma^t) \) is her history of signals, where \( \sigma^n \) is the signal from her \( n \)th term in office. A candidate in an open seat election or a challenger has an empty history, denoted \( h^0 = \emptyset \).

Voters are myopic, and vote to maximize the expectation of

\[
\theta - (p - p^i)^2,
\]

where \( \theta \) is the quality of the election’s winner and \( p \) is the policy she is committed to implement. Voter \( i \)'s ideal point is \( p^i \). Ideal points in the population are distributed according to some logconcave distribution \( F \). This distribution is symmetric about its median, \( p^m \).

Expected policy positions for the candidates are fixed at \( p^D \) for the Democrat and \( p^R \) for the Republican, with \( p^D < p^R \).

Assume WLOG that the incumbent is a Democrat. Voter \( i \) will vote for the incumbent if and only if

\[
\mathbb{E}(\theta|h^t) - (p^D - p^i)^2 \geq -(p^R - p^i)^2.
\]

This implies a cutpoint election rule where all voters with ideal points to left of \( C_{DR} \) vote for the incumbent. The cutpoint satisfies

\[
C_{DR} = \frac{1}{2} \frac{\mathbb{E}(\theta|h^t)}{(p^R - p^D)} + \frac{(p^R + p^D)}{2}.
\]

Since the median voter is decisive, the incumbent wins if \( p^m \leq C_{DR} \). The share of the vote going to the incumbent is just the fraction of the voters with ideal points to the left of \( C_{DR} \), \( F(C_{DR}) \).

The fact that voters are myopic is important in establishing this voting rule. If voters are forward looking, they will take into account the option value of the candidates, not just their current beliefs about mean abilities. If two candidates are equal in terms of expected ability, but the variance of ability is higher for one, the voters will prefer the one with higher variance. We assume that voters ignore this factor when making decisions. If voters are not myopic, we would expect that, of two candidates with the same expected ability, the one with shorter tenure will be preferred by the voters. This is because beliefs will be more
precise for the candidate who has been around longer.

Consider first the case where ideal points are uniformly distributed on \([-a,a]\). Then \(F(C_{DR})\) is equal to

\[
\frac{C_{DR} + a}{2a}.
\]

Substitute the equation determining \(C_{DR}\) to get

\[
\alpha \mathbb{E}(\theta | h^t) + \gamma,
\]

where \(\alpha\) and \(\gamma\) are constants depending on \(p^R, p^D,\) and \(a\).

This version of the model has strong implications for a regression of vote share on lagged vote share. Write \(SHARE_t\) for the Democrat’s share of the vote at date \(t\). Then we have (using the martingale property of beliefs)

\[
SHARE_t = \alpha \mathbb{E}(\theta | h^t) + K
\]

\[
= \alpha (\mathbb{E}(\theta | h^{t-1}) + \epsilon) + k
\]

\[
= \alpha \mathbb{E}(\theta | h^{t-1}) + K + \alpha \epsilon
\]

\[
= SHARE_{t-1} + \alpha \epsilon,
\]

where \(K\) is a constant depending on the district’s ideological character and \(\epsilon\) is a prediction error. Because \(\epsilon\) is the error in the optimal prediction of \(\theta\), \(SHARE_{t-1}\) and \(\alpha \epsilon\) are orthogonal. Note that this orthogonality does not depend on any sort of conditional independence assumptions if learning follows Bayes’s rule. Thus, if we estimate

\[
SHARE_t = \beta_0 + \beta_1 * SHARE_{t-1},
\]

the theory implies that \(\beta_1 = 1\) and \(\beta_0 = 0\).

This result clearly relies on the assumed uniform distribution. Indeed, for general \(F\),
the above argument becomes

\[
SHARE_t = F(\alpha \mathbb{E}(\theta \mid h^t) + \gamma)
\]
\[
= F(\alpha \mathbb{E}(\theta \mid h^{t-1}) + \alpha \epsilon + \gamma)
\]
\[
= F(F^{-1}(SHARE_{t-1}) + \alpha \epsilon).
\]

Now the expected value of \(SHARE_t\) is

\[
\mathbb{E}(F(F^{-1}(SHARE_{t-1}) + \alpha \epsilon)),
\]

which is not in general equal to \(SHARE_{t-1}\).

If, for example, \(F\) is concave, then

\[
\mathbb{E}_{t-1}(SHARE_t) = \mathbb{E}_{t-1}(F(F^{-1}(SHARE_{t-1}) + \epsilon_t))
\]
\[
< F(\mathbb{E}_{t-1}(F^{-1}(SHARE_{t-1}) + \epsilon_t))
\]
\[
= F(F^{-1}(SHARE_{t-1}))
\]
\[
= SHARE_{t-1},
\]

where the inequality is Jensen's inequality.

Our strategy is to take a linear approximation to the cdf of voter ideal points around the median, justifying the analysis based on the uniform distribution. This approximation is valid for close races if the distribution of voter ideal points is approximately linear around the mean ideal point. This is true, for example, for the normal distribution. Of course, we can't select based on closeness in the current election. However, we can select for close last elections without introducing selection bias. The absence of bias is due to the martingale property of beliefs generated by Bayesian updating.

We have seen that, in a setting with no incumbency advantage, the lagged vote is a sufficient statistic for the incumbent's ability conditional on date \(t-1\) information and for the characteristics of the district. When we allow for an incumbency advantage, this is no longer true. We can only infer the incumbent's ability from the lagged vote if we know whether that vote was achieved with or without the advantage. Thus we must condition on lagged vote and lagged incumbency. These variables are jointly sufficient for ability
and district characteristics. This suggests that the coefficient on lagged incumbency should equal minus the coefficient on incumbency in the previous election.

3.3 Data

Our data come from the ICPSR. We observe the results of House elections in each Congressional district in every election from 1944 until 1998. For each observation (election) we have the names and ICPSR codes of the candidate for each party (Democrat, Republican, and other) and the number of votes each received.\(^2\)

We have chosen to work with only the post-World War II period for several reasons. This is the period commonly studied in the literature, so for sake of comparison, it is most useful to focus here. Also, it is reasonable to believe that patterns in the data may vary over time. For example, we find that the advantage of incumbents in elections appears to increase significantly over the six decades we study. For our purposes, looking at a longer period of time might confuse issues.

We have eliminated observations based on three other criteria. We only study regular elections, so we eliminate elections from odd-numbered years. We control for year fixed effects, so the inclusion of special elections in odd years would create problems. Our central results are separate regressions for each election year. The number of elections in off years would be far too small for study. Also, note that special elections, by definition, do not have incumbents, thus eliminating much of the interest for our purposes. The number of these elections (136) is small enough that it does not raise concern.

We also eliminated elections in which a third-party candidate is the incumbent. As explained below, our regressand (the Democratic share of the vote) is based only the Democratic and Republican votes. As a result, third-party candidates don’t fit our model. This only accounts for 19 out of nearly ten thousand observations in the data, so it not of substantial concern.

Finally, we also eliminate the elections following Census years (1952, 1962, 1972, 1982, and 1992). These years are when redistricting occurs, which would raise significant problems for our analysis. We rely on data from the same district in the previous election in our model. Not only are these lagged variables central to our model, but they control for

\(^2\)We thank Jim Snyder for providing us with this data.
district-specific effects, such as social conditions and political tendencies. Since a particular district, as indicated by state and number, may not really be the same area in 1950 and 1952, our results would be biased. Even the indicated "incumbent" may not be correct, as two districts may have merged to a large degree. So, for example, it would be possible that both candidates had served in the House in the previous term.

Since the key variables in our regressions are constructed from the data, a brief explanation of their computation is in order. The Democratic share of the vote (DS) is simply the number of votes the Democratic candidate received divided by that number plus the number of votes the Republican received. We leave third-party votes out of "total votes" because not doing so would consistently bias the variable downward. Third-party votes would, in essence, be counted as Republican votes. The model, as explained in Sections 2 and 4, depends on a two party system.

The incumbency variable (INC) takes one of three values. It has a value of 1 if the Democrat is an incumbent, -1 if the Republican is an incumbent, and 0 if there is no incumbent. So, for example, take a particular observation where there is a Democratic incumbent. The coefficient on INC gives the additional fraction of the vote that she received as a result of incumbency. This formulation restricts the incumbency advantage to be equal for both parties. This restriction is standard in the literature.

3.4 Econometric model

Our goal is to test the theoretical predictions of the model presented in Section 2. The regression model directly reflects the theoretical result in equation (3.2). We can test whether the coefficient on lagged share is, in fact, equal to one. We can also study the effect of incumbency within a framework supported by a theoretical understanding of election behavior.

In order to implement the analysis, we must change the model slightly. The model, as expressed in equation (3.2), assumes there is an incumbent and predicts the share of votes she will receive. The SHARE variable is defined as the vote share of the incumbent. The incumbency variable in Section 2 is a binary variable, and its coefficient gives the advantage the incumbent has. Of course, to estimate this coefficient, we also must look at elections without an incumbent. This would create a serious problem if it were not for the two party
system. In an open seat election, we would have no incumbent for whom to define SHARE.

We simply replace the SHARE variable with the Democratic share (DS) and redefined INC as explained in the Data section. In other words, we have three cases: INC = 1 and DS = SHARE; INC = -1 and DS = -1×SHARE; and the open seat case which allows us to estimate the effect of incumbency.

With that modification, our regression model follows directly from the theoretical model. For one election year, it is:

\[ DS_{sd} = \beta_0 + \beta_1 LAGDS_{sd} + \beta_2 INC_{sd} + \beta_3 LAGINC_{sd} + \epsilon_{sd} \]

where s indicates state and d indicates district. The theory in Section 2 predicted that we should get \( \beta_1 = 1, \beta_0 = 0, \) and \( \beta_2 = -\beta_3. \) One can break down the predictions of this model into cases. An incumbent who was in incumbent in the previous election should expect to get the same share of the votes as last time. An incumbent who was not an incumbent last time (INC = 1, LAGINC = 0) should expect a greater vote share (the sophomore surge). A candidate of the same party as a retired incumbent (INC = 0, LAGINC = 1) should expect a lower vote share than her predecessor received (the retirement slump).

This model has a couple of nice properties. First, running separate regressions for each year (or including year dummy variables in a pooled regression) controls for national level variables that might influence the results. For example, an economic slowdown or recession might have an adverse effect on the electoral prospects of all incumbents. Note, however, that events that particularly favor or disfavor one party will not be controlled by this. We see what appear to be such effects in three cases.

Second, the inclusion of the lagged Democratic share allows us to ignore all district-specific effects. Any variable is constant for each district will affect LAGDS as much as DS and thus be controlled for when we include LAGDS in the regression.

We considered the possibility that the effect of lagged vote might behave differently for incumbents and non-incumbents. We ran our regressions with an additional variable that was equal to LAGDS×INC. The results didn’t show any significant effect, and so we chose to leave the variable out.

It is also possible to run regression used multiple years of data. In such a case, we include year dummy variables to account for year fixed effects. We often use this procedure
when we want to find average coefficients for the entire data range or a particular decade.

The final issue of concern is the linearity condition discussed in Section 2. As explained there, unless the distribution of voter preferences is roughly uniform, the results will be biased. Empirically, we want to see the effects of limiting the data to close races before we go on with our analysis on that basis. Of course, we can only do this on the basis of the lagged share because selecting on the regressand would introduce bias.

3.5 Results

As it will determine the set of data we will use for the rest of our analysis, we first look at the effect of limiting the data to close races, based on specific ranges of LAGDS. Since LAGDS has a range of 0 to 1 and 0.5 would be a tie, these ranges are defined as values in a range $[0.5 - x, 0.5 + x]$, where $x$ is a constant. We found that results for different values of $x$ were largely consistent with each other but that they were noticeably different from the unrestricted regression (i.e., $x = 0.5$).

Table 1 shows coefficients on LAGDS for three ranges ($x = 0.1, 0.2, 0.3$) and for the unrestricted regression. The columns indicate different time periods, while the rows are show each restriction possibility. The coefficients in the first three rows are statistically indistinguishable from each other. They are, however, higher and almost always statistically different from the unrestricted regression coefficients in the bottom row.

Figure 1 (and the data for it displayed in Table 2) shows a year-by-year comparison between the coefficients on LAGDS for the unrestricted regression and those for the regression limited to the range 0.3 to 0.7. It shows the coefficient to be higher for the restricted regression in all but four years. The averages of these two series are 0.791 and 0.621, respectively (shown on Table 1).

Another, more direct test of the linearity assumption is to run the regressions with higher order terms for LAGDS (i.e., squared and cubed terms). Such regressions over the entire range gave highly jointly significant coefficients on the higher order terms. This implies that the linearity assumption is invalid. By comparison, regressions over the limited ranges of data, as described above, give small, insignificant coefficients on higher order terms. This leads to the conclusion that linearity is a good approximation for these ranges.

These results fits our hypothesis that a strictly logconcave (non-uniform) distribution
of voter preferences would bias this coefficient downward. It is reasonable to believe that voter preferences have such a distribution, with a substantial mass of “moderate” voters and fewer at the extremes. For the purposes of the rest of our analysis, we will restrict the data to where LAGDS is in the range 0.3 to 0.7. We do this rather than estimating a non-linear model because our theoretical model makes strong predictions for the linear case, so the linear case is convenient for testing.

Tables 3 and 4 show regression results for the basic model in each year. Recall that our hypothesis was that the coefficient on LAGDS would be one. We see that the results are typically lower than one (in all but three cases). However, on a year-by-year basis, the difference is typically statistically insignificant. Figure 3 shows these results with 95% confidence intervals.

The average value appears to be about 0.8. We have two possible explanations for this. One is that the voter myopia hypothesis is incorrect. As discussed in Section 2, this is testable by looking at tenure effects. We looked at this possibility by running regressions with tenure as an additional regressor, but found that there did not seem to be a tenure effect beyond the simple incumbency effect. That is, having some tenure helps but having more tenure does not seem to help more.

The other explanation is that the estimated effect is being attenuated by unobserved changes in the distribution of voter preferences in at least some districts. We believe this may cause a form of measurement error that explains a number of our results. We are relying on LAGDS to account for district characteristics or, more precisely, the distribution of voter preferences (and location of the median voter) in the district. This distribution may not be stable in all districts at all times. If this is the case, there will be “measurement error” in the representation by LAGDS of these preferences, since LAGDS is subject to small shocks in each election. It is a well-known result that measurement error causes the estimated coefficient to be biased downward (toward zero) and also biases the other estimated coefficients in the regression.

The theory predicted that the regression constant ($\beta_0$) would be equal to zero. Tables 3 and 4 show the estimates. While for many years the constant does appear to be zero, there are many other years in which it is positive and significant. On average, it is 0.114. This solves an interesting issue raised by the results on LAGDS. That coefficient being less that one on average would result in erosion of the Democratic vote. Take, for example, a
race that was very close in one year with a Democratic share of the vote approximately 0.5. If the constant term were zero, we would predict that the Democratic share in the next election would be 0.4 (assuming $\beta_1 = 0.8$). With $\beta_0 = 0.1$, however, this problem is solved and the predicted vote share is about 0.5.

Moreover, there appears to be a pattern to when the constant is positive and significant and when it is not. Precisely in those years where the coefficient on LAGDS drops below one, the constant appears to grow and offset it. A bivariate regression of one of these sets of coefficients on the other shows a highly statistically significant, negative relationship with only 22 observations. As will be discussed, this provides strong evidence for the measurement error hypothesis.

Our results for the incumbency effect appear to be consistent with the literature (see Gelman and King 1990, for example) in two ways. First is that there does appear to be an advantage to incumbents that is significant both practically and statistically. In particular, we find, like Gelman and King, that the advantage existed as far back as we have data. The second is that this advantage seems to have grown noticeably over the last six decades (see Figure 2). We have reason, however, to differ with some of the conclusions.

Tables 3 and 4 contain the year-by-year estimates for the incumbency effect. Note that these coefficients are statistically significant at the 95% level for every year but 1946. There appears to be a rise in size of this effect of the years observed. The effect seems to be about 2-4% in the 1940s and 1950s but rises to 7-10% in the later decades. There are a couple of years that don’t fit this pattern, but overall there seems to be a shift. Figure 2 shows these results (the solid line).

Figure 2 also includes the same estimates using the full range of data (i.e., not that restricted to LAGDS values of 0.3 to 0.7). The values plotted on this figure are shown in Table 2. This is essentially the equation estimated by Gelman and King. The restriction to close races reduces the estimated advantage to incumbents. We believe this is a more accurate estimate of the effective incumbency advantage in elections. Since we saw above that there are significant nonlinearities in the full sample, a regression which ignores the higher-order terms will suffer from omitted variable bias.

We estimate that the average incumbency advantage in the years 1946-1960 is about 2.7%. The average for the later period is about 8.5%. These estimates lie between the traditional estimates and those of Gelman and King. Given the appearance of an increase
over time, the “average” is only partly informative. However, our estimation of a smaller effect with the limited range of data is consistent over time.

We believe, however, that these estimates may still be biased upward by the measurement error due to shifts in voter tastes. The theory predicted $\beta_1 = 1$, $\beta_0 = 0$, and $\beta_2 = -\beta_3$. Measurement error in LAGDS would bias $\beta_1$ downward. However, it still must be true that the predicted vote share of the average candidate is equal to the average actual vote share. For this to hold true, some linear combination of $\beta_0$, $\beta_2$, and $\beta_3$ would be biased upward. This explains the above result about $\beta_0$. The results as a whole support a hypothesis that (a) the estimated incumbency advantage is being biased upwards in our results by measurement error and (b) the observed increase in the size of this effect is being at least partly driven by an increase in the measurement error.

Examination of Tables 3 and 4 show a pattern that fits this theory. There is a tendency for the estimates to go wrong (relative to our predictions) all at the same time and in the directions predicted by the measurement error hypothesis. On average, $\beta_1$ is too low and the other estimates are too high. The results for the 1940s and 1950s are close to the predictions of the model, but in the later decades, the other coefficients tend to rise and to do so particularly when $\beta_1$ is lower. This could be driven by increasing instability in the political preferences of districts. There are well-known shifts in the party preferences of certain districts, particularly the movement of large parts of the South towards the Republican party and shifts toward the Democrats in Northern states (McCarty, Poole, and Rosenthal 1997). The exact effects of measurement error bias on the estimated coefficients of other variables is hard to predict. In fact, the effects may not be entirely consistent, thus explaining some of the unstable results we observe.

A test of this hypothesis would need to correct for the measurement error in LAGDS using instrumental variables techniques. Using an instrument for LAGDS, we could correct the bias and thus obtain a consistent estimate. At this time, we do not know of a proper instrument for this test.

3.6 Conclusion

We provide a theoretical justification for the estimator of the incumbency advantage introduced by Gelman and King. Our model provided guidance about when that procedure is
valid; in particular, we show it is only expected to work for "close" races. Turning to the
data, we find empirical support both for the model and for the importance of restricting
the sample to avoid bias from nonlinearities.

Our estimates of the advantage lie in between the traditional estimates based on the
"sophomore surge" or the "slurge" and the estimates in Gelman and King. Like all of
these papers, we find that the advantage has increased dramatically over time. The rate
of increase that we find, however, is less than that found by Gelman and King. We lend
support to Gelman and King's claim that a significant advantage existed well before the
eye 1960's, contrary to the earlier literature.

We are still concerned about additional bias caused by measurement error in the lagged
Democratic share. This error may arise because of shifts in the distribution of voter pref-
erences over time. There is substantial evidence from other sources that such shifts have
occurred at an important level between 1960 and today. The pattern of our results is largely
consistent with what one would expect if there is measurement error. Further evidence that
measurement error might be a problem is provide by the correlation in timing of the "creep-
ing realignment" and rise in our estimates of the incumbency advantage. This suggests that
at least part of the rise in the measured incumbency advantage may be due to this bias.
Dealing with this measurement error problem is an important topic for future research on
the incumbency advantage.
Table 3.1: Coefficients on LAGDS for different ranges of LAGDS included in regression

<table>
<thead>
<tr>
<th>Range</th>
<th>all years</th>
<th>1940s</th>
<th>1950s</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
<th>1990s</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4-0.6</td>
<td>0.783</td>
<td>0.931</td>
<td>0.871</td>
<td>0.657</td>
<td>0.638</td>
<td>0.837</td>
<td>0.805</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.124)</td>
<td>(0.092)</td>
<td>(0.119)</td>
<td>(0.152)</td>
<td>(0.192)</td>
<td>(0.141)</td>
</tr>
<tr>
<td>0.3-0.7</td>
<td>0.791</td>
<td>0.904</td>
<td>0.874</td>
<td>0.747</td>
<td>0.630</td>
<td>0.809</td>
<td>0.819</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.056)</td>
<td>(0.043)</td>
<td>(0.056)</td>
<td>(0.066)</td>
<td>(0.071)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>0.2-0.8</td>
<td>0.805</td>
<td>0.906</td>
<td>0.882</td>
<td>0.883</td>
<td>0.605</td>
<td>0.899</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.043)</td>
<td>(0.034)</td>
<td>(0.039)</td>
<td>(0.041)</td>
<td>(0.044)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>0.0-1.0</td>
<td>0.621</td>
<td>0.792</td>
<td>0.697</td>
<td>0.562</td>
<td>0.538</td>
<td>0.504</td>
<td>0.637</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.019)</td>
<td>(0.016)</td>
<td>(0.021)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td>(0.026)</td>
</tr>
</tbody>
</table>

Standard errors are in parentheses.
<table>
<thead>
<tr>
<th>Year</th>
<th>LAGDS (0.3-0.7)</th>
<th>LAGDS (full)</th>
<th>INC (0.3-0.7)</th>
<th>INC (full)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1946</td>
<td>0.966</td>
<td>0.931</td>
<td>0.009</td>
<td>0.008</td>
</tr>
<tr>
<td>1948</td>
<td>0.899</td>
<td>0.691</td>
<td>0.028</td>
<td>0.048</td>
</tr>
<tr>
<td>1950</td>
<td>0.833</td>
<td>0.770</td>
<td>0.022</td>
<td>0.030</td>
</tr>
<tr>
<td>1954</td>
<td>1.006</td>
<td>0.666</td>
<td>0.027</td>
<td>0.054</td>
</tr>
<tr>
<td>1956</td>
<td>0.880</td>
<td>0.790</td>
<td>0.024</td>
<td>0.026</td>
</tr>
<tr>
<td>1958</td>
<td>0.799</td>
<td>0.556</td>
<td>0.051</td>
<td>0.073</td>
</tr>
<tr>
<td>1960</td>
<td>0.782</td>
<td>0.820</td>
<td>0.029</td>
<td>0.030</td>
</tr>
<tr>
<td>1964</td>
<td>0.756</td>
<td>0.462</td>
<td>0.038</td>
<td>0.081</td>
</tr>
<tr>
<td>1966</td>
<td>0.475</td>
<td>0.516</td>
<td>0.077</td>
<td>0.087</td>
</tr>
<tr>
<td>1968</td>
<td>0.852</td>
<td>0.592</td>
<td>0.066</td>
<td>0.089</td>
</tr>
<tr>
<td>1970</td>
<td>0.930</td>
<td>0.683</td>
<td>0.068</td>
<td>0.091</td>
</tr>
<tr>
<td>1974</td>
<td>0.456</td>
<td>0.458</td>
<td>0.095</td>
<td>0.113</td>
</tr>
<tr>
<td>1976</td>
<td>0.817</td>
<td>0.567</td>
<td>0.073</td>
<td>0.092</td>
</tr>
<tr>
<td>1978</td>
<td>0.571</td>
<td>0.525</td>
<td>0.107</td>
<td>0.125</td>
</tr>
<tr>
<td>1980</td>
<td>0.737</td>
<td>0.601</td>
<td>0.093</td>
<td>0.110</td>
</tr>
<tr>
<td>1984</td>
<td>0.856</td>
<td>0.592</td>
<td>0.087</td>
<td>0.109</td>
</tr>
<tr>
<td>1986</td>
<td>0.781</td>
<td>0.494</td>
<td>0.110</td>
<td>0.134</td>
</tr>
<tr>
<td>1988</td>
<td>1.059</td>
<td>0.463</td>
<td>0.083</td>
<td>0.147</td>
</tr>
<tr>
<td>1990</td>
<td>0.547</td>
<td>0.466</td>
<td>0.100</td>
<td>0.120</td>
</tr>
<tr>
<td>1994</td>
<td>0.593</td>
<td>0.653</td>
<td>0.095</td>
<td>0.099</td>
</tr>
<tr>
<td>1996</td>
<td>0.769</td>
<td>0.529</td>
<td>0.050</td>
<td>0.073</td>
</tr>
<tr>
<td>1998</td>
<td>1.115</td>
<td>0.748</td>
<td>0.096</td>
<td>0.124</td>
</tr>
</tbody>
</table>
Table 3.3: Regression Results by Year, Part I

<table>
<thead>
<tr>
<th>Year</th>
<th>LAGDS</th>
<th>INC</th>
<th>LAGINC</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1946</td>
<td>0.966*</td>
<td>0.010</td>
<td>-0.005</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>(0.090)</td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>1948</td>
<td>0.899*</td>
<td>0.028</td>
<td>0.004</td>
<td>0.130</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.010)</td>
<td>(0.009)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>1950</td>
<td>0.833*</td>
<td>0.025</td>
<td>0.036</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>(0.101)</td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>1954</td>
<td>1.006*</td>
<td>0.028</td>
<td>-0.011</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.008)</td>
<td>(0.011)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>1956</td>
<td>0.880*</td>
<td>0.025</td>
<td>-0.004</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>(0.090)</td>
<td>(0.008)</td>
<td>(0.012)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>1958</td>
<td>0.799</td>
<td>0.041</td>
<td>0.026</td>
<td>0.190</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.009)</td>
<td>(0.012)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>1960</td>
<td>0.782</td>
<td>0.029</td>
<td>-0.001</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>(0.094)</td>
<td>(0.008)</td>
<td>(0.011)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>1964</td>
<td>0.756</td>
<td>0.041</td>
<td>-0.013</td>
<td>0.181</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.011)</td>
<td>(0.010)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>1966</td>
<td>0.475</td>
<td>0.076</td>
<td>0.024</td>
<td>0.232</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.075)</td>
</tr>
<tr>
<td>1968</td>
<td>0.852*</td>
<td>0.063</td>
<td>-0.027</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td>(0.124)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.056)</td>
</tr>
<tr>
<td>1970</td>
<td>0.930*</td>
<td>0.084</td>
<td>-0.042</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td>(0.012)</td>
<td>(0.014)</td>
<td>(0.051)</td>
</tr>
</tbody>
</table>

Standard errors are in parentheses.
Data limited to values of 0.3 to 0.7 for LAGDS.
* indicates not different than 1 at the 95% confidence level.
Table 3.4: Regression Results by Year, Part 2

<table>
<thead>
<tr>
<th>Year</th>
<th>LAGDS</th>
<th>INC</th>
<th>LAGINC</th>
<th>Constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974</td>
<td>0.456</td>
<td>0.096</td>
<td>-0.024</td>
<td>0.360</td>
</tr>
<tr>
<td></td>
<td>(0.120)</td>
<td>(0.013)</td>
<td>(0.012)</td>
<td>(0.059)</td>
</tr>
<tr>
<td>1976</td>
<td>0.817*</td>
<td>0.073</td>
<td>-0.009</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>(0.134)</td>
<td>(0.012)</td>
<td>(0.011)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>1978</td>
<td>0.571</td>
<td>0.105</td>
<td>0.006</td>
<td>0.198</td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td>(0.015)</td>
<td>(0.012)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>1980</td>
<td>0.737*</td>
<td>0.093</td>
<td>-0.044</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.014)</td>
<td>(0.012)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>1984</td>
<td>0.856*</td>
<td>0.090</td>
<td>-0.024</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(0.124)</td>
<td>(0.014)</td>
<td>(0.011)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>1986</td>
<td>0.781*</td>
<td>0.114</td>
<td>-0.029</td>
<td>0.152</td>
</tr>
<tr>
<td></td>
<td>(0.146)</td>
<td>(0.015)</td>
<td>(0.013)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>1988</td>
<td>1.059*</td>
<td>0.093</td>
<td>-0.046</td>
<td>-0.026</td>
</tr>
<tr>
<td></td>
<td>(0.153)</td>
<td>(0.017)</td>
<td>(0.015)</td>
<td>(0.076)</td>
</tr>
<tr>
<td>1990</td>
<td>0.547</td>
<td>0.106</td>
<td>-0.018</td>
<td>0.231</td>
</tr>
<tr>
<td></td>
<td>(0.146)</td>
<td>(0.018)</td>
<td>(0.016)</td>
<td>(0.085)</td>
</tr>
<tr>
<td>1994</td>
<td>0.593</td>
<td>0.094</td>
<td>-0.017</td>
<td>0.133</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>1996</td>
<td>0.769</td>
<td>0.049</td>
<td>-0.021</td>
<td>0.148</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.041)</td>
</tr>
<tr>
<td>1998</td>
<td>1.115*</td>
<td>0.101</td>
<td>-0.031</td>
<td>-0.058</td>
</tr>
<tr>
<td></td>
<td>(0.108)</td>
<td>(0.012)</td>
<td>(0.011)</td>
<td>(0.067)</td>
</tr>
</tbody>
</table>

Standard errors are in parentheses.
Data limited to values of 0.3 to 0.7 for LAGDS.
* indicates not different than 1 at the 95% confidence level.
Figure 1: Effect of Lagged Democratic Vote Share over Time
Figure 2: Incumbency advantage by Year
Figure 3: 95% Confidence Intervals for LAGDS coefficient estimate, by year
Appendix A

References


