ESSAYS ON THE QUALITY OF GOVERNMENT

RAINER ALBRECHT HEINZ SCHWABE

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
ECONOMICS

ADVISOR: JOHN B. LONDREGAN

SEPTEMBER 2010

ABSTRACT

Essays on the Quality of Government

Rainer Schwabe

I examine how voters use elections to select the candidate most suited to a given office.

In chapter 1, I propose a model of representative democracy in which the value of holding office and candidate identity are endogenous. There is a trade-off between limiting the rents from holding office and being able to attract society's brightest to public service. The quality of government is jointly determined by equilibrium levels of candidate ability and allocation of resources to the public good. Comparative statics suggest that while increasing the power of a political post may attract higher ability candidates, it may also have a negative effect on the quality of government. The model also provides insights into the motivation of weak challengers.

In chapter 2, I study a model of infinitely repeated elections in which voters attempt simultaneously to select competent politicians and to provide them with

incentives to exert costly effort. Voters are unable to incentivize effort if they base their reelection decisions only on incumbent reputation. However, equilibria in which voters use reputation-dependent performance cutoffs (RDC) to make reelections decisions exist and support positive effort. In these equilibria, politicians' effort is decreasing in reputation, and expected performance is decreasing in tenure. Like the equilibria in Ferejohn 1986, RDC equilibria rely on voters being indifferent between reelecting incumbents and electing challengers. I show that this voter-indifference condition is closely related to weak renegotiation-proofness (Farrell and Maskin 1989).

In chapter 3, I present a model of campaign finance in primary elections in which campaigns supply hard information about candidates' electability. Focusing on a class of equilibria in which informed voters vote according to the signal they observe, I show that bandwagons can arise in equilibrium when a third party is financing campaigns, and should be accompanied by a cessation of funding for the trailing candidate.

To address the controversy surrounding the timing of presidential primaries in the United States, I examine the welfare effects of making changes to the electoral calendar. For relatively low campaign costs, a calendar with a block of voters voting simultaneously early in the process, followed by the remaining voters voting consecutively, is optimal for voters and the party. This result provides a rationale for "Super Tuesdays" in U.S. presidential primaries. For higher campaign costs, a sequential calendar is optimal. Donors always prefer a sequential electoral calendar.

Acknowledgements

I would like to thank John B. Londregan for his tireless support and encouragement, for our many stimulating conversations about the topics discussed in this dissertation as well as many others which await further development, and for his unwavering belief in my capacity. I also owe much to Scott Ashworth, Stephen Morris, and Adam Meirowitz for their guidance and support. I benefited from conversations with many faculty and students including Dilip Abreu, Marco Battaglini, Roland Bénabou, Faruk Gul, Karen Kaiser, Navin Kartik, and Wolfgang Pesendorfer.

The administrative staff at Princeton's Department of Economics were wonderfully helpful throughout my time there. Special thanks are due to Kathleen DeGennaro.

I am grateful to my parents, Heinz and Paulina Schwabe, and my wife, Cathleen Mitchell, for their love and support throughout.

Contents

CHAPTER 1

Recruiting, Elections, and the Quality of Government

1.1. Introduction

> The object of every political constitution is, or ought to be, first to obtain for rulers men who possess most wisdom to discern, and most virtue to pursue, the common good of the society; and in the next place to take the most effectual precautions for keeping them virtuous whilst they continue to hold their public trust.
>
> Publius (James Madison), The Federalist Papers 57 (p. 343)

The presumption that a representative democracy will be able to attract the best of its citizens to public service was an important part of the Federalists' argument in favor of this form of government, and it continues to be relevant in today's assessment of our governmental institutions.

As Madison points out, it is not enough to grant positive incentives to politicians so that their position is attractive. We must also be sure that electoral competition is effective at keeping our politicians focused on the public good.

The tension between these two necessities is the central theme of this paper. Limiting the diversion of resources has a positive effect on the quality of government, *given the ability of those in office.* However, it can have a negative effect on the ability of entrants into politics as rent-extraction is part of the attraction of holding office. The net effect of these forces depends on the relative size of fixed rewards from holding office, such as salaries, and the amount of resources over which a politician has control. It also depends on the strength of the incumbent, if there is one, and the opportunity costs of going into politics.

This paper presents a simple model which takes these factors into consideration, characterizes its equilibria, and attempts to describe how the quality of government depends on the parameters of the model. In doing this, I call into question the commonly held view that higher ability candidates provide better quality government. Section 4 provides a counterexample in which society is able to attract more able politicians by increasing the resources available to them, but these provide a lower quality of government than was previously received.

1.1.1. Related Literature

As discussed in a survey article by Timothy Besley (2005), formal political theory has generally abstracted from questions of politicians' ability and political selection. Research that has emphasized the role of ability have tended to assume that candidates are randomly drawn from a fixed pool of potential candidates (e.g. Rogoff and Sibert 1988).

Besley and Coate and Osbourne and Slivinsky's models of a representative democracy focused attention on the entry decisions of potential politicians but, rather than emphasize questions of competence, they stress ideological motivations for running for office. Later work by LeBorgne and Lockwood (2002) and Casselli and Morelli (2004) used the citizen-candidate framework to explore the determinants of the competence of politicians. The central difference between their approach and the one taken in this paper is that while LeBorgne and Lockwood and Casselli and Morelli assume that more skilled politicians will provide more of the public good, I allow for the possibility that politicians will divert resources for their private gain inasmuch as electoral competition allows. Therefore, while other models of competence present elections as screening mechanisms, this paper emphasizes the disciplinary role of elections.

In highlighting the role of electoral competition in limiting rent-seeking, I follow Polo (1998)[1] who uses a probabilistic voting framework to model the trade-offs between vote share and rent-taking. Thus, while extending the theory of political selection to include rent-seeking behavior, this paper can also be seen as extending the theory of rent-seeking in competitive elections to include the effects of political selection.

This paper is also related to recent work on political careers by Matozzi and Merlo (2007) who focus on the incentive effects of lucrative post-politics careers in the private sector. Besley, Pande and Rao (2006) provide empirical evidence for the importance of politician identity to the quality of government. The model below also contributes a fuller theoretical account of the motivation of weak challengers

[1]Also discussed in Persson and Tabellini (2000, ch.4).

in congressional elections than had previously been given by, for example, Banks and Kiewiet (1989) and Canon (1993).

The rest of the paper is organized as follows: Section 1.2 describes the model, Section 1.3.1 describes the equilibria of the model when ability and private sector income are perfectly correlated, and Section 1.3.2 does the same in the case where the correlation is imperfect. Section 1.4 discusses comparative statics. Section 1.5 concludes.

1.2. Model

There is a community (polity) consisting of a continuum of agents characterized by an income distribution F(y). An agent's income in the private sector y_i is a perfect indicator of that agent's private sector ability. Public sector ability (γ_i) is correlated with y_i:

$$\gamma_i = (\mu + \theta y_i)\varepsilon_i$$

where $E(\varepsilon_i)=1$. $\ln\varepsilon_i$'s are distributed i.i.d. with cdf H(). I use $\hat{\gamma}_i = E[\gamma_i] = \mu + \theta y_i$ to denote the expected competence of a given agent i. Agents without experience in the public sector do not know their own γ but, as in Londregan and Romer (1993), it is publicly revealed through the campaign process before elections take place.

There is one political post which needs to be filled via a simple majority election. This post commands exogenously fixed resources R. A politician's public sector skill level scales this resource pool so that effective resources available when i is in power are $\gamma_i R$. These resources can be used either to provide the public good P or for the politician's private benefit r_i (I will also refer to r_i as rents) so that

$$\gamma_i R = P + r_i.$$

I denote the proportion of effective resources used to provide the public good

$$q_i \equiv \frac{P}{\gamma_i R}.$$

Thus, the product γq is a measure of the effectiveness with which government resources are being used. I call γq the *quality of government.*

Political office provides a salary S consisting of monetary compensation and ego rents.

Each agent (politicians included) i has a utility function

$$u_i = C_i + \alpha P$$

where P is a public good and C_i is private consumption, be it from private sector earnings (y) or benefits extracted from public office (S+r). Throughout, I assume $\alpha < 1$ so that private consumption is more important to our citizens than the government-provided public good[2]. Note here that the question of politician motivation is moot since they are taken to be ordinary citizens. The specified preferences are over policy and private consumption, and the fixed rewards of office imply an interest in winning, making this model consistent with Wittman (1983).

[2]The units of public good and private income are necessarily comparable since politicians can choose to use government resources for one or the other. If α were greater than one, politicians would never have an incentive to offer less than the highest amount of public good possible.

S and r together are a politician's private consumption. Thus, a politician's utility when in office is

$$u_i(\text{i in office}) = (1 - q_i)\gamma_i R + S + \alpha q_i \gamma_i R = (1 - q_i(1 - \alpha))\gamma_i R + S.$$

If an agent runs for office and loses, she enjoys the public good provided by her competitor but is deprived of private income so that u_i(i runs for office and loses) $= \alpha q_j \gamma_j R$ where j refers to the opposing candidate.

There are two political parties: A and B. The parties are permanent institutions of the polity and have duopoly power over candidate selection. Before the election, each political party recruits the candidate with highest expected ability from those in the population willing to run, or if there is an incumbent only the party out of office recruits a candidate. Throughout, I will use a superscript A (B) to identify the parameters of party A's (B's) equilibrium choice of candidate.

Parties are important in this model mainly because they keep the number of candidates to two, thus keeping the platform selection stage tractable. One may think of several party objective functions which would induce the selection of the highest expected ability candidates[3][4]. This would be the case if utility were derived directly from the quality of candidates selected, which can be taken as shorthand

[3]Carrillo and Mariotti (2001) develop a model of parties predicting similar behavior assuming that parties maximize the probability of winning. In this context, we cannot rule out high ability candidates sacrificing some probability of a win in favor of greater rents in the case of a win.

[4]Although I do not explicitly model primary elections, one may consider the assumption that the highest expected ability candidate is able to win a primary. Alternatively, in a polity where the primary is the main hurdle to gaining office, the model is applicable to primaries with two competing candidates.

for unmodelled party reputation concerns. Parties receiving a proportion of the rents extracted from office would also do. However, because I view this behavior as intuitive, but a full theory of political parties as beyond the scope of this paper, I opt for modelling the parties as mechanically selecting for quality.

Once recruited, candidates select platforms $\{q^A, q^B\}$. I make two assumptions about candidate platforms. The first is that these can be modelled as binding commitments, as they classically are in one-period models of electoral competition[5]. The second is that these commitments are made at an early stage of the campaigning process, before information about ability is revealed.

Given that politicians cannot affect private sector incomes, all citizens will prefer the candidate who offers a larger quantity of the public good.

To summarize, the timing of events is as follows:

(1) Citizens simultaneously decide whether or not they will run if asked to.

(2) Parties select their candidates simultaneously from among those willing to run.

(3) Candidates simultaneously make resource allocation commitments q^A and q^B.

(4) Candidates' ability is revealed through the campaign process.

(5) Voters cast their ballots.

(6) The winning candidate implements the policy promised at stage 3.

[5]See Persson and Tabellini (2000) ch. 3.

Throughout, I consider two versions of the model. In the first, which I call the *incumbent game*, there is an incumbent of known ability representing party A in the election. In the second, the *open-seat game*, both parties must recruit candidates.

1.3. Equilibrium

As is standard, I solve for subgame perfect equilibria of the game by analyzing its stages in reverse order. All formal proofs are relegated to the Technical Appendix.

I begin by considering a perfect information version of the model where private and public ability are perfectly correlated ($\varepsilon_i \equiv 1$). This will help to highlight the importance of uncertain ability in this model as well as lead to some interesting insights about candidate motivation. Furthermore, it will help the reader understand the structure of the model in a transparent way.

In Section 1.3.2, I turn my attention to the game with uncertainty over political ability.

1.3.1. Known Political Ability

Because only one candidate is selected by each party, there are many equilibria in which the entry decisions of citizens who are not selected vary. Primarily, I focus on equilibria where a citizen runs for office purely for the private benefits, that is, she expects that if she does not run for office, another candidate of the equilibrium ability will run in her place. I call these equilibria *regular*. I also describe equilibria in which candidates believe that if they do not run, nobody else will. I call these equilibria *arm-twisting* since one can think of parties twisting reluctant candidates' arms by making it clear to them that they are the polity's only hope for a competitive election. Arm-twisting equilibria involve private provision of a public good, or dragon-slaying in the sense of Bliss and Nalebuff (1984)[6]. As will become clear below, these equilibria identify a possible motivation for weak challengers in congressional elections: they may be willing to run to force the strong candidate to make more campaign promises. In the rest of this Section I call a candidate *weak* if she has no chance of winning the election. Bliss and Nalebuff show how, within a war of attrition framework, the highest ability "knight" will step forward and provide the public good. In this model, I argue that because parties play a part in equilibrium selection, they will recruit the highest ability candidates possible in an arm-twisting equilibrium.

I call an equilibrium *symmetric* if $\gamma^A = \gamma^B$ and $q^A = q^B$. I assume that candidates win with probability $\frac{1}{2}$ if voters are indifferent between them.

When there is no uncertainty about abilities, the highest ability candidate will always win the election. To do so, she must offer at least as much of the public good as her competitor is able to do since, otherwise, the other candidate would have incentives to up her offer and win the election. This logic is summarized in

[6]Palfrey and Rosenthal (1984) also discuss this type of equilibria.

the following Lemma which describes the Nash equilibrium of the platform setting subgame.

Lemma 1. ***(Electoral Equilibrium)*** *With perfect information, the unique Nash equilibrium of the platform selection subgame is as follows:*

- *If $\gamma^A > \gamma^B$, candidate A wins the election by setting $q^A = \frac{\gamma^B}{\gamma^A}$ and candidate B loses setting $q^B = 1$. The reverse is true for $\gamma^B > \gamma^A$.*
- *If $\gamma^A = \gamma^B$, candidates both set q=1 and each wins with probability 1/2.*

Having solved the platform selection subgame, the payoffs to entry are well-defined. Action sets are {enter with either A or B, enter only with B, enter only with A, do not enter with either A or B} for each citizen.

I make the following technical assumptions to ensure that if an agent is successful enough in the private sector, her best option is to stay in the private sector, and that there are always citizens who are too successful for politics. This rules out corner solutions in which society's very best become politicians. One can think of the converse of A1 as a sufficient condition for attracting society's best to politics, as is discussed in Proposition 5. In the game with uncertainty, a generalization of the following conditions serves the same purpose.

A1. $\theta R < 1$.

A2. F is strictly increasing over $[0, \frac{2(\mu R+S)}{1-\theta R}]$.

A1 guarantees that the net benefit of running for office is downward sloping in γ. A2 ensures that individuals talented enough to find politics unappealing are part of our polity. Together, they guarantee that the candidates selected will be indifferent between running for office and staying in the private sector. The main implications of A1 are summarized in the following lemma.

Lemma 2. ***(Entry Conditions)*** *Assume A1 and A2 hold in the game without uncertainty. Then, in a regular equilibrium, entering candidates will be characterized by the indifference condition* $y^B = max\{0, (\gamma^B - \gamma^A)R + S\}$ *. In any arm-twisting equilibrium* $y^B \leq max\{0, (\gamma^B - \gamma^A)R + S\} + \alpha R \min\{\gamma^B, \gamma^A\}$ *.*

As is discussed in the following propositions, multiple equilibria are possible. I focus on pure strategy equilibria involving the highest ability candidates.

When there is an incumbent, his ability along with the fixed rewards of holding office (S) determine the quality of the challenger. If the incumbent's ability is low enough and S is large enough, a challenger strong enough to defeat the incumbent will find running attractive. Conversely, if the incumbent is of high ability, or if the fixed rewards of holding office are low, the opposition party will be forced to recruit a weak challenger. This relation is made specific in the following proposition.

Proposition 3. ***(Equilibrium with an Incumbent)*** *Given A1 and A2, the equilibria of the incumbent game with perfect information is as follows:*

- *If* $\gamma^A < \theta S + \mu$ *there exists a regular equilibrium in which the challenger wins with probability one,* $\gamma^B = \frac{\theta(S - \gamma^A R) + \mu}{1 - \theta R}$*, and* $q^B = \frac{\gamma^A}{\gamma^B}$*.*
- *If* $\gamma^A > \theta S + \mu$ *there exists a regular equilibrium in which the incumbent defeats a challenger* $\gamma^B = \mu$ *and sets* $q^A = \frac{\mu}{\gamma^A}$*.*

- *If* $\gamma^A > \frac{\mu}{1-\alpha\theta R}$*, there exists an arm-twisting equilibrium where the incumbent defeats a challenger* $\gamma^B = \frac{\mu}{1-\alpha\theta R}$ *and sets* $q^A = \frac{\gamma^B}{\gamma^A}$.
- *If* $\gamma^A < \frac{\theta S+\mu}{1-\alpha\theta R}$*, there exists an arm-twisting equilibrium where the challenger* $\gamma^B = \mu + \theta\frac{R(\mu+\gamma^A(\alpha-1))+S}{1-\theta R}$ *wins and sets* $q^B = \frac{\gamma^A}{\gamma^B}$.

When there is no incumbent, expectations of a party's (or perhaps and individual's) success play a key role in the recruiting process. If one expects party A to win because they will be more successful recruiters, the expectations become self-fulfilling. On the other hand, it is generally not rational to expect a tie to occur in equilibrium. If candidates of a certain ability are willing to run for half the fixed rewards of holding office (in expected terms), then a slightly better qualified candidate would surely find it profitable to enter politics and take all the spoils. The following proposition describes and qualifies this asymmetry.

Proposition 4. ***(Equilibrium with an Open Seat)*** *Given A1 and A2:*

- *For any S>0, there exists a regular equilibrium where party A (B) recruits a candidate* $\gamma^A = \mu + \frac{\theta S}{1-\theta R}$ *who wins the election by setting* $q^A = \frac{\gamma^B}{\gamma^A}$*. Party B (A) recruits a weak candidate* $\gamma^B = \mu$ *and loses setting* $q^B = 1$.
- *For any S>0, there exists an arm-twisting equilibrium where party A (B) recruits a candidate* $\gamma^A = \frac{\mu}{1-\alpha\theta R} + \frac{\theta S}{1-\theta R}$ *who wins the election by setting* $q^A = \frac{\gamma^B}{\gamma^A}$*. Party B (A) recruits a weak candidate* $\gamma^B = \frac{\mu}{1-\alpha\theta R}$ *loses setting* $q^B = 1$.

- *If* $S < \frac{\alpha\mu R}{\frac{1}{2}-\alpha\theta R}$ *there are symmetric arm-twisting equilibria in which* $\gamma^A = \gamma^B = \gamma$ *and* $q^A = q^B = 1$. *In the best of these equilibria,* $\gamma = \frac{\frac{1}{2}\theta S+\mu}{1-\alpha\theta R}$.

The converse of A1 is:

CA1. $\theta R > 1$.

This makes the portion of the entrants' incentive constraint corresponding to potential winners upward sloping in γ. The following Proposition makes precise the sense in which this is a sufficient condition for government to attract society's highest ability citizens. One may think of this rule of the skilled as the rule of the Natural Aristocracy (Jefferson 1998, pgs. 579-80). Denote these citizen's private sector income by $\bar{y} = max\{y|f(y) > 0\}$. Their corresponding level of ability is $\bar{\gamma} = \mu + \theta\bar{y}$.

Proposition 5. ***(Natural Aristocracy)****If CA1 holds, then*

- *In any equilibrium of the incumbent game, if* $\gamma^A \leq S + \mu R + \bar{y}(\theta R - 1)$ *then a challenger* $\bar{\gamma}$ *enters and wins the election.*
- *In any regular equilibrium of the open-seat game, the winning candidate will be society's best* $\bar{\gamma}$. *If* $\frac{\frac{1}{2}\theta S+\mu}{1-\alpha\theta R} \leq S + \mu R + \bar{y}(\theta R - 1)$ *, the same is true for all arm-twisting equilibria.*

Note that comparative statics of the equilibria (on S, R, incumbent ability, etc.) described above are uninteresting over wide ranges of these variables. For example, in a regular equilibrium in which the challenger defeats the incumbent

(Proposition 3, first bullet point), the amount of the public good provided and the quality of government is constant in S as long as S is large enough for the relevant inequalities to hold. Thus, I leave comparative statics exercises to the case of unknown political ability, discussed in the next section, where uncertainty generates smooth comparative statics.

1.3.2. Unknown Political Ability

I now turn to a (more realistic) world in which candidate ability is unknown until after the campaign is finished[7]. This uncertainty will affect the entry decisions of citizens as well as the policy choices of candidates.

When there is uncertainty regarding the public sector ability of candidates, the ex-ante probability of A winning the election is a function of the distributions of ε^A and ε^B. Candidate A wins if voters get more public good from A than from B, that is if $q^B R\hat{\gamma}^B\varepsilon^B < q^A R\hat{\gamma}^A\varepsilon^A$ or $\frac{\varepsilon^B}{\varepsilon^A} < \frac{q^A\hat{\gamma}^A}{q^B\hat{\gamma}^B}$ or $\ln\frac{\varepsilon^B}{\varepsilon^A} < \ln q^A\hat{\gamma}^A - \ln q^B\hat{\gamma}^B \equiv \Delta$. Thus, the probability of A winning is G(Δ) where G is the cdf of $\ln\frac{\varepsilon^B}{\varepsilon^A}$. In the incumbent case, ε^A is known so that G is the cumulative distribution function of $\ln\varepsilon^B$ (H). Let g(x)=$\frac{\partial G(x)}{\partial x}$ be the density function associated with G($\cdot$). Thus, as in Londregan and Romer (1993), uncertainty over candidate's ability generates a platform selection decision analogous to that in probabilistic voting (Lindbeck and

[7]One may also think of uncertainty being reduced by the same amount for each candidate, without reaching full revelation.

Weibull 1987). Most of the analysis in this section applies to both open seat and incumbent games.

The role of uncertainty being clear, one can write down the candidates' objective functions in the platform selection subgame and solve for the Cournot-like equilibria of the subgame. The following Lemma formalizes this.

Lemma 6. ***(Electoral Equilibrium)*** *Given $\hat{\gamma}^A$ and $\hat{\gamma}^B$, the platform selection subgame has a Nash equilibrium which is characterized by the first-order conditions:*

$$G(\Delta)(R\hat{\gamma}^A \mathrm{E}(\varepsilon^A|Awins)(\alpha-1) + \tfrac{\partial \mathrm{E}(\varepsilon^A|Awins)}{\partial q^A}(1-q^A(1-\alpha))R\hat{\gamma}^A)$$
$$+\tfrac{1}{q^A}g(\Delta)((1-q^A(1-\alpha))R\hat{\gamma}^A \mathrm{E}(\varepsilon^A|Awins) + S - \alpha q^B R\hat{\gamma}^B \mathrm{E}(\varepsilon^B|Bwins))$$
$$+\tfrac{\partial \mathrm{E}(\varepsilon^B|Bwins)}{\partial q^A}(1-G(\Delta))(\alpha q^B R\hat{\gamma}^B) + (\nu-\lambda) = 0 \qquad (1)$$
$$(1-G(\Delta))[(\alpha-1)\hat{\gamma}^B R\mathrm{E}(\varepsilon^B|Bwins) + \tfrac{\partial \mathrm{E}(\varepsilon^B|Bwins)}{\partial q^B}(1-q^B(1-\alpha))R\hat{\gamma}^B]$$
$$+\tfrac{1}{q^B}g(\Delta)((1-q^B(1-\alpha))R\hat{\gamma}^B \mathrm{E}(\varepsilon^B|Bwins) + S - \alpha q^A R\hat{\gamma}^A \mathrm{E}(\varepsilon^A|Awins))$$
$$+\tfrac{\partial \mathrm{E}(\varepsilon^A|Awins)}{\partial q^B}G(\Delta)\alpha q^A R\hat{\gamma}^A + (\nu-\lambda) = 0 \qquad (2)$$

Where υ and λ are Lagrange multipliers which are zero in any interior solution.

In general, this equilibrium will involve positive rents ($q^j < 1$) even if candidates are of identical expected ability, a point made by Polo (1998) but derived from the general principle that the uncertainty in elections permits candidates to propose non-optimal (from the voter's perspective) platforms without discretely hurting their chances of winning.

Note the presence of platform effects on the conditional expectation of ability: offering a lower q lowers the probability of winning, but also means that a win will only take place if the candidate is of relatively higher ability so that private and public benefits are at high levels. Conversely, it makes it easier for the opponent to win, so that our expectation of her ability conditional on victory is lower.

The first order conditions above can be solved for q^A and q^B and thus for the expected value to i of running for office as A (or B's) candidate. I analyze only regular equilibria here, so that the expected value of staying out of politics is:

$\hat{u}_i = y_i + \alpha[G(\Delta)q^A\hat{\gamma}^A E(\varepsilon^A|Awins)R + (1 - G(\Delta))q^B\hat{\gamma}^B E(\varepsilon^B|Bwins)R]$

Thus, agent i assumes that if she does not run for office under party A's banner, someone else of equilibrium competence will. The choice of whether to work in the private sector or take a chance in the political arena boils down to a comparison of expected private benefits. That is, the net expected benefit of running for office for a citizen of expected ability $\hat{\gamma}$ is:

- $G(\Delta)((1-q^A)\hat{\gamma}E(\varepsilon|Awins)R+S)-\frac{1}{\theta}(\hat{\gamma}-\mu)$ when running as A's candidate.
- $(1 - G(\Delta))((1 - q^B)\hat{\gamma}E(\varepsilon|Bwins)R + S) - \frac{1}{\theta}(\hat{\gamma} - \mu)$ when running as B's candidate.

Let q^A and q^B represent their equilibrium values given $\hat{\gamma}^A$ and $\hat{\gamma}^B$. Given $\hat{\gamma}^B$, define $\tilde{\gamma}^A \equiv \inf\{\hat{\gamma}^A|G(\Delta)((1 - q^A)\hat{\gamma}^A E(\varepsilon^A|Awins)R + S) - \frac{1}{\theta}(\hat{\gamma}^A - \mu) < 0\}$ and $\tilde{\gamma}^B$ symmetrically.

We now generalize A1 and A2:

A1'. $\sup\limits_{\hat{\gamma}^B} \tilde{\gamma}^A$ and $\sup\limits_{\hat{\gamma}^A} \tilde{\gamma}^B$ are finite and, for each $\hat{\gamma}^B$ ($\hat{\gamma}^A$), the net value of running for office as A's (B's) candidate is negative for all $\hat{\gamma} > \tilde{\gamma}^A$ ($\tilde{\gamma}^B$).

A2'. F places strictly positive probability on the interval [0,2max{ $\tilde{\gamma}^A$, $\tilde{\gamma}^B$}].

A1' again relies on θR being sufficiently small. The presence of the conditional expectation in the expression makes it difficult to provide globally sufficient conditions for A1' to be true. However, in most examples θR<1 suffices. Figure 1.1 shows two possible net expected benefit curves in which A1' is violated, as well as a typical case in which A1' holds.

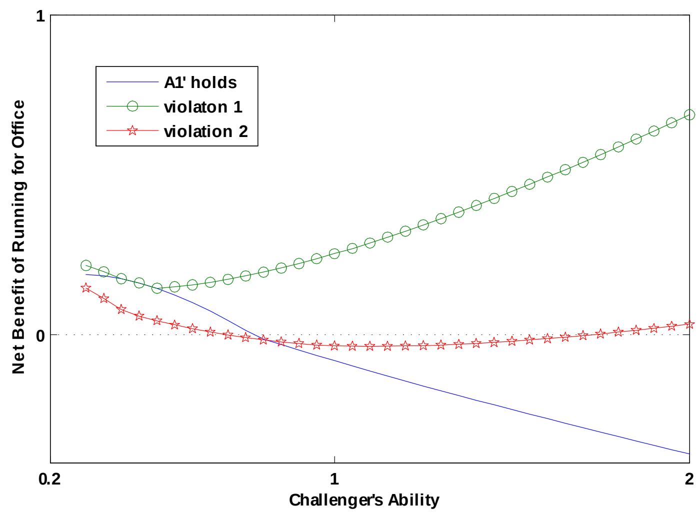


Figure 1.1: The role of A1'.

Lemma 7. ***(Entry Conditions)*** *Given that parties choose the most competent agents available, and given A1' and A2', candidate selection will be characterized by the indifference conditions*[8]*:*

$$G(\Delta)((1-q^A)\hat{\gamma}^A E(\varepsilon^A|Awins)R+S)-\frac{1}{\theta}(\hat{\gamma}^A-\mu)=0 \quad (3)$$

$$(1-G(\Delta))((1-q^B)\hat{\gamma}^B E(\varepsilon^B|Bwins)R+S)-\frac{1}{\theta}(\hat{\gamma}^B-\mu)=0 \quad (4)$$

[8]Recall that $\hat{\gamma}_i=\mu+\theta y_i$ so $y_i=\frac{1}{\theta}(\hat{\gamma}_i-\mu)$.

Lemmas 5 and 6 characterize the regular equilibria of the game with uncertainty. The following proposition gathers the results.

Proposition 8. *Under the assumptions above, an equilibrium always exists and is characterized by equations 1-4.*

In general, many equilibria of the open seat game may exist as one party may be more successful than the other in recruiting candidates. However, in contrast with the forced asymmetry of the perfect information equilibria, one can find symmetric equilibria when there is uncertainty as long as a weak version of A1 holds.

Remark 9. ***(Symmetric Equilibria)*** *If* $\theta R < \frac{2}{E(\varepsilon^i|\varepsilon^i > \varepsilon^j)}$ *, the open seat game always has a symmetric equilibrium in which* $q^A = q^B$ *and* $\hat{\gamma}^A = \hat{\gamma}^B$*.*

To see this, one must simply observe that because the $\ln\varepsilon_i$'s are i.i.d., the distribution of their difference (G) must be symmetric around a zero mean. Therefore, equilibrium conditions (1) and (2) are identical with the party labels switched, as are (3) and (4), so that it is sufficient to solve one pair, be it (1) and (3) or (2) and (4) to find a symmetric equilibrium of the game. A detailed proof is given in the Technical Appendix.

1.4. Comparative Statics

Because in the tails of common distributions for G, changes in expected ability conditional on winning and second order effects regarding the distribution can become very important, it is very hard to find assumptions that ensure that derivatives are globally of a certain sign. Assuming an interior solution (so that $\lambda = \nu = 0$ in both A and B's first order conditions) equations (1), (2), (3) and (4) define a system of implicit functions so that it is possible to do comparative statics without explicitly solving for the equilibrium values of q^A, q^B, y^A, and y^B. However, these local results are of limited significance here. For similar reasons, global monotonicity of objective functions cannot easily be established so that monotone comparative statics are not applicable here. Instead, I simulate changes in important parameters using common families of distributions such (exponential, gamma, and lognormal) for ε. I illustrate only the case where ε is exponential as other cases are qualitatively similar.

The lack of general comparative statics results does not mean that nothing can be learned from these exercises. Figures 1.2 and 1.3 illustrate one of the main results of this paper: that equilibria with higher ability candidates need not be desirable because they may involve lower levels of the quality of government. Note that the quality of government and the ability of the challenger have opposite slopes here. The simulated comparative statics in the figure constitute a proof by counterexample.

Figures 1.2-1.3 show how as government resources are increased, better candidates are attracted to politics but equilibria are less honest as the temptation of diverting resources increases. Figures 1.4-1.7 illustrate the effect on the quality of government of varying the fixed rewards from office (S) or the ability of the incumbent (γ^A). Both of these relations are positive. It is not surprising that increasing the fixed rewards of holding office increases the equilibrium quality of government, as this relation has been observed in a related setting by Ferejohn (1986). All figures illustrate the incumbent game.

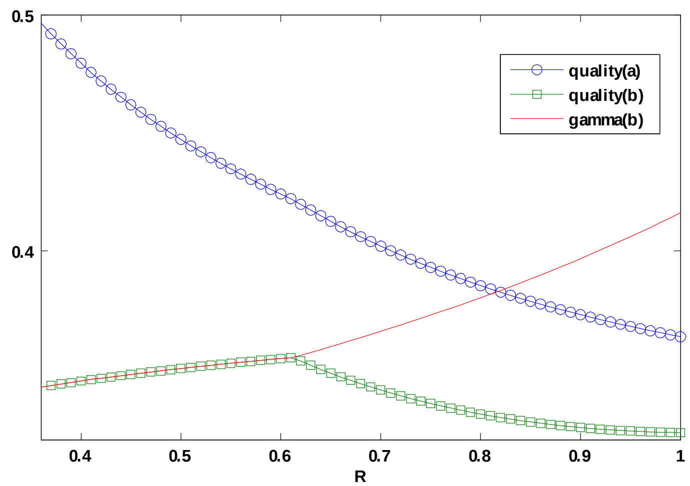


Figure 1.2: Quality of Government vs. Government Resources-R

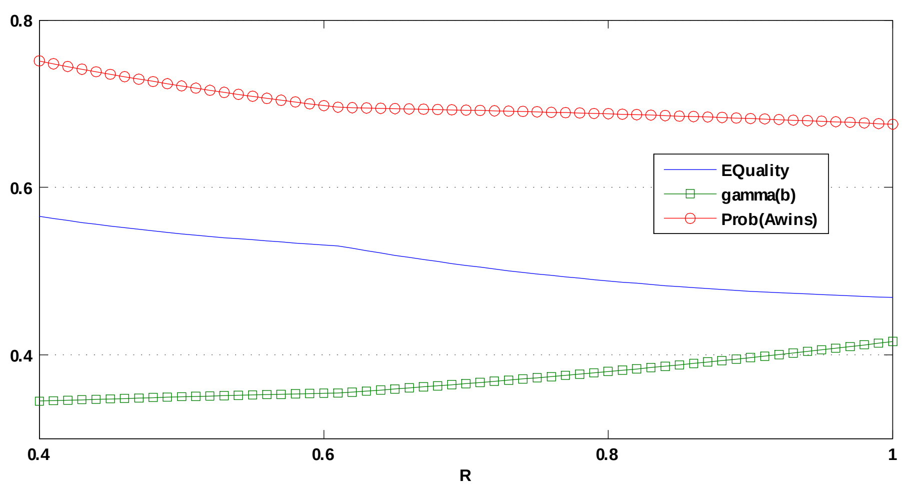


Figure 1.3: Expected Quality of Government vs. Government Resources-R

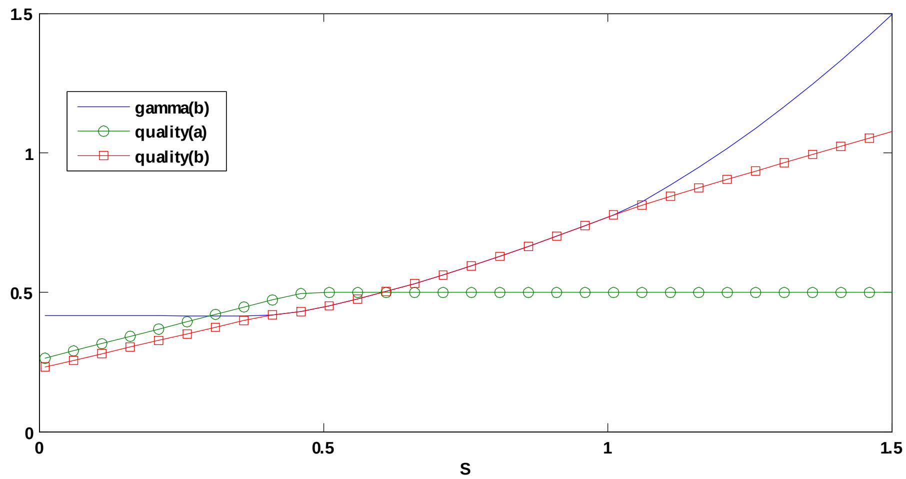


Figure 1.4: Quality of Goverment vs. Fixed Rewards from Office-S

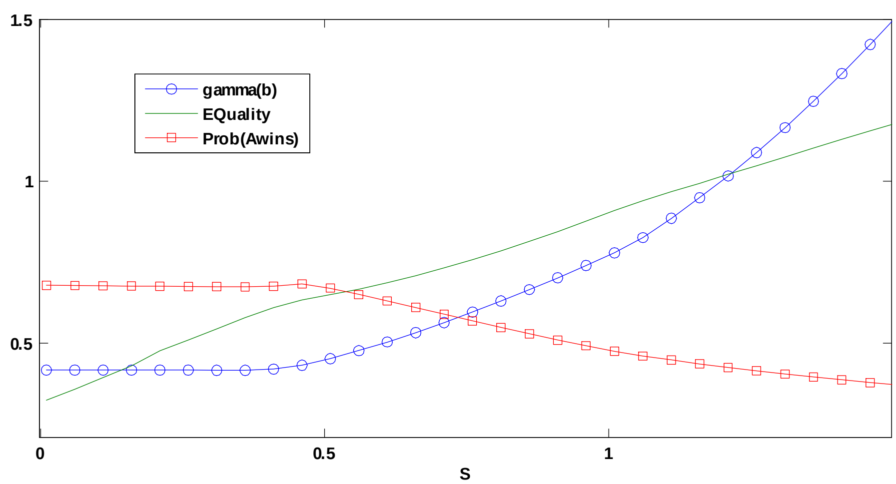


Figure 1.5: Expected Quality of Government and Ability vs. S

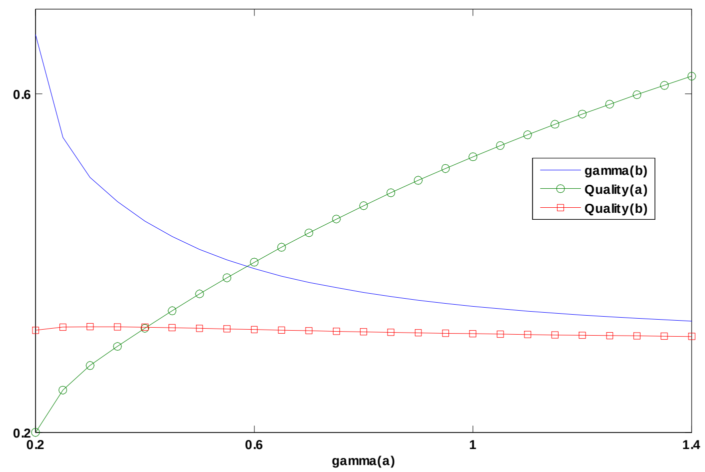


Figure 1.6: Quality of Government vs. Incumbent Ability

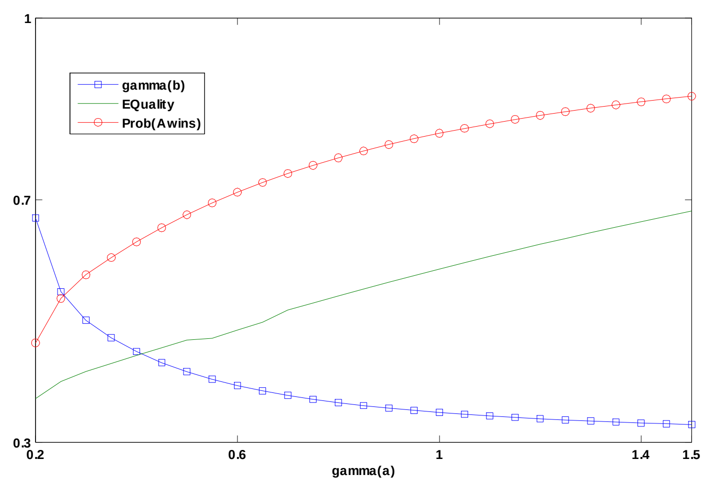


Figure 1.7: Expected Quality of Government vs. Incumbent Ability

1.5. Concluding Remarks

Citizens find careers in politics appealing for different reasons: wages, ego rents and rent extraction among them. This paper has highlighted the importance of the relative and absolute size of these motivating factors for determining the quality of government. Politicians motivated primarily by the fixed rewards of office will commit to limit their pursuit of rents. Those motivated primarily by the rents themselves, however, cannot be expected to do the same. By the same logic, the analysis suggests that if a society is to endeavor to attract better qualified people to public office, it should be wary of doing so by increasing the power and resources of this post.

1.6. Technical Appendix

Proof of Lemma 1.

If $\gamma^A > \gamma^B$ candidate A's best response to any q^B is $q^A = \frac{q^B \gamma^B}{\gamma^A}$ which ensures electoral victory while maximizing private benefits. On the other hand, any q^B is a best response for B(A) to $q^A \geq \frac{\gamma^B}{\gamma^A}$ since she stands no chance of winning. B's best response to any $q^A < \frac{\gamma^B}{\gamma^A}$ is $q^B = \frac{q^A \gamma^A}{\gamma^B}$ ensuring victory for herself. Thus, $q^B < 1$ is not an equilibrium since it is not a best response to any q^A which is itself a best response to this q^B.

Candidates of equal ability will compete each others private rents away since a small increase in the public good offered leads to a discrete increase in payoffs. Thus, the voters will be indifferent between the candidates and each candidate will win with probability 1/2 by assumption.

Proof of Lemma 2.

In a regular equilibrium, potential candidates believe that the public good will be provided at the equilibrium level whether or not they decide to run for office. Thus, given γ^A,the incentive constraint for a party B candidate of equilibrium competence is $max\{0, (\gamma - \gamma^A)R + S\} - y = max\{0, (\gamma - \gamma^A)R + S\} - \frac{1}{\theta}(\gamma - \mu) \geq 0$. For a winning candidate, a unit increase in γ has an opportunity cost of $\frac{1}{\theta}$ and a benefit of R. A1 guarantees that $\frac{1}{\theta}$>R so that the incentive constraint is weakly decreasing in γ, and crosses zero at most once for $\gamma > \mu$. Equilibrium rewards from running for office are bounded above by the rewards of running unopposed:

$(\theta y+\mu)R+S = \gamma R+S$ so that the indifferent citizen will have income $y = \frac{\mu R+S}{1-\theta R} < \frac{2(\mu R+S)}{1-\theta R}$. Thus, A2 ensures that the indifferent citizen is part of our polity.

In an arm-twisting equilibrium, the incentive constraint becomes $max\{0, (\gamma- \gamma^A)R+S\} + \alpha R \min\{\gamma, \gamma^A\} - y \geq 0$. A1 and A2 serve the same purpose as in the regular equilibrium.

Proof of Proposition 3.

A winning challenger (necessarily $\gamma^B > \gamma^A$) gets a private payoff of $(\gamma^B- \gamma^A)R + S$ since she must offer infinitesimally more of the public good than the incumbent is able to do. Given A1, γ^B is determined by the indifference condition $\gamma^B = \theta[(\gamma^B- \gamma^A)R + S] + \mu$ so that $\gamma^B = \frac{\theta(S-\gamma^A R)+\mu}{1-\theta R}$ which is greater than γ^A if and only if $\gamma^A < \theta S + \mu$.

An unopposed incumbent will set q^A=0. Thus, if the incumbent is above this ability threshold, party B will turn to lower skilled citizens for whom it makes sense to sacrifice their private income in the interest of forcing the incumbent to offer some public good, i.e. weak candidates. The best of these is determined by the indifference condition y=$\alpha(\mu + \theta y)R$ which implies y=$\frac{\alpha\mu R}{1-\alpha\theta R}$. Given this critical value of y, I observe that if $\gamma^A \geq \mu + \theta\frac{\alpha\mu R}{1-\alpha\theta R} = \frac{\mu}{1-\alpha\theta R}$ the hopeless candidate will indeed lose the election. Finally, there must be only one citizen willing to challenge since otherwise incentives to run for office disappear as each potential candidate is sure that someone else will run for office and force the incumbent to provide some public good.

Citizens that would be willing to challenge the incumbent only to force him to provide some public good may also be able to win the election. The income of the best citizen party B could recruit in an arm-twisting equilibrium in which party B wins is determined by the indifference condition y=$(\theta y + \mu - \gamma^A)R + S + \alpha\gamma^A R$ ($q^A = 1$ by lemma 1) since the candidate will win office and she (correctly) believes that no one will run if she does not. Thus, $\gamma^B = \mu + \theta\frac{R(\mu+\gamma^A(\alpha-1))+S}{1-\theta R}$ which is greater than γ^A if and only if $\gamma^A < \frac{\theta S+\mu}{1-\alpha\theta R}$.

Proof of Proposition 4.

To see that this is a Nash equilibrium consider the strategies above. If A is recruiting optimistically and citizens believe A will win the election, the best available candidate is determined by the indifference condition $\gamma^A = \theta[(\gamma^A - \gamma^B)R + S] + \mu = \theta[(\gamma^A - \mu)R + S] + \mu$ so that $\gamma^A = \mu + \frac{\theta S}{1-\theta R}$. Since potential candidates believe (correctly) that B has no chance of winning, the best candidate the party can recruit in a regular equilibrium is a weak challenger μ.

In an arm-twisting equilibrium where only one citizen agrees to run for A the indifference condition specifying the wealthiest citizen willing to run for A is $\gamma^A = \theta[(\gamma^A - \frac{\mu}{1-\alpha\theta R})R + S + \frac{\alpha R\mu}{1-\alpha\theta R}] + \mu$ which implies $\gamma^A = \frac{\mu}{1-\alpha\theta R} + \frac{\theta S}{1-\theta R}$.

In an arm-twisting equilibrium in which candidates tie, their participation constraint is $\frac{1}{2}S + \alpha\gamma R \geq \frac{1}{\theta}(\gamma - \mu)$. This is satisfied with equality at $\gamma = \frac{\frac{1}{2}\theta S+\mu}{1-\alpha\theta R}$. If a candidate slightly better than those running is to stay out of the race, we must have $S < \frac{1}{\theta}(\gamma - \mu)$ which simplifies to $S < \frac{\alpha\mu R}{\frac{1}{2}-\alpha\theta R}$. Once again, A1 and A2 guarantee that this is all we need to check.

Proof of Proposition 5.

In the incumbent game, because the incentive constraint is upward sloping for winners, we need only check it for society's best. Thus, we need $(\bar{\gamma}-\gamma^{A})R+S \geq \bar{y}$ which can be rewritten as $\gamma^{A} \leq S+\mu R+\bar{y}(\theta R-1)$.

In the open seat game, in any regular equilibrium in which one party loses for sure, the losing party can only attract candidates μ. Thus, running for office is profitable for a candidate with zero income if she expects to win, and because the net expected benefit of running is upward sloping in γ, it is profitable for any citizen to run with the winning party. In an arm-twisting equilibrium, the losing party can at best recruit a weak candidate $\gamma^{B}=\frac{\mu}{1-\alpha\theta R}$ so that, following the proof for the incumbent game, we get the necessary condition $\frac{\mu}{1-\alpha\theta R} \leq S+\mu R+\bar{y}(\theta R-1)$. To rule out ties we need either S$>\frac{\alpha\mu R}{\frac{1}{2}-\alpha\theta R}$ so that ties are ruled out as in Proposition 4, or $\frac{\frac{1}{2}\theta S+\mu}{1-\alpha\theta R} \leq S+\mu R+\bar{y}(\theta R-1)$ so that the incentive constraint for $\bar{\gamma}$ is satisfied as above.

Proof of Lemma 6.

By continuity of the objective functions, a Nash Equilibrium exists (Glicksberg 1952). Candidate A chooses a policy platform q^{A} by maximizing her expected utility taking her opponent's platform as given:

$\max_{q^{A}} G(\Delta)((1-q^{A}(1-\alpha))RE(\gamma^{A}|Awins)+S)+(1-G(\Delta))(\alpha q^{B}RE(\gamma^{B}|Bwins)).$

$s.t.\ q^{A} \in [0,1]$

The corresponding Lagrangian is:

$\Gamma(q^{A},\lambda,\upsilon)=G(\Delta)((1-q^{A}(1-\alpha))RE(\gamma^{A}|Awins)+S)$

$+(1-G(\Delta))(\alpha q^B RE(\gamma^B|Bwins)) + \lambda(1-q^A) + \upsilon q^A$

which has the following first order (necessary) conditions:

$\partial q^A: \quad G(\Delta)(R\hat{\gamma}^A E(\varepsilon^A|Awins)(\alpha-1) + \frac{\partial E(\varepsilon^A|Awins)}{\partial q^A}(1-q^A(1-\alpha))R\hat{\gamma}^A)$

$+\frac{1}{q^A} g(\Delta)((1-q^A(1-\alpha))R\hat{\gamma}^A E(\varepsilon^A|Awins) + S - \alpha q^B R\hat{\gamma}^B E(\varepsilon^B|Bwins))$

$+\frac{\partial E(\varepsilon^B|Bwins)}{\partial q^A}(1-G(\Delta))(\alpha q^B R\hat{\gamma}^B) + (\nu - \lambda) = 0$

$\partial\lambda: \quad 1-q^A = 0$ and $\lambda > 0$ or $\lambda = 0$

$\partial\nu: \quad q^A = 0$ and $\nu > 0$ or $\nu = 0$

Symmetrically, candidate B solves:

$\max_{q^B}(1-G(\Delta)((1-q^B(1-\alpha))RE(\gamma^B|Bwins)+S)+G(\Delta)(\alpha q^A RE(\gamma^A|Awins))$

$s.t.\ q^B \in [0,1]$

Which leads to first order conditions:

$\partial q^B: \quad (1-G(\Delta))[(\alpha-1)\hat{\gamma}^B RE(\varepsilon^B|Bwins) + \frac{\partial E(\varepsilon^B|Bwins)}{\partial q^B}(1-q^B(1-\alpha))R\hat{\gamma}^B]$

$+\frac{1}{q^B} g(\Delta)((1-q^B(1-\alpha))R\hat{\gamma}^B E(\varepsilon^B|Bwins) + S - \alpha q^A R\hat{\gamma}^A E(\varepsilon^A|Awins))$

$+\frac{\partial E(\varepsilon^A|Awins)}{\partial q^B} G(\Delta)\alpha q^A R\hat{\gamma}^A + (\nu - \lambda) = 0$

$\partial\lambda: \quad 1-q^B = 0$ and $\lambda > 0$ or $\lambda = 0$

$\partial\nu: \quad q^B = 0$ and $\nu > 0$ or $\nu = 0$.

Proof of Lemma 7.

A1' ensures the net value of running for office is negative for high y individuals. Note that an individual with zero private sector income always finds running for office desirable since they have a small chance of winning S>0. This fact, together with A2' and the continuity of the net expected value of running for office function, ensures that there is indeed a citizen in our polity which is made indifferent between running for office or not. It is easy to see that a player cannot benefit from deviating from this strategy profile; i.e. cannot gain from volunteering to run when her net value of running is negative, or choosing not to run when her net value of running is positive.

Proof of Proposition 8.

Lemmas 4 and 5 together provide a system of equations (1)-(4) in four unknowns ((1),(2) and (4), with three unknowns in the incumbent game). Existence in the incumbent game is easy to see given the previous lemma. In the open seat game, existence is guaranteed through A1' and A2' since A1' lets us concentrate only on citizens of ability in [0,max{ $\bar{\gamma}^A$, $\bar{\gamma}^B$}], and thus a solution to 3 and 4 constrained to 1 and 2 holding exists by the Glicksberg fixed point theorem.

Proof of Remark 9.

If $\theta R < \frac{2}{E(\varepsilon^i | j wins, q^i\gamma^i = q^j\gamma^j)} = \frac{2}{E(\varepsilon^i|\varepsilon^i > \varepsilon^j)}$, the open seat game always has a symmetric equilibrium in which $q^A = q^B$ (q's may be mixed) and $\hat{\gamma}^A = \hat{\gamma}^B$.

Proof:

Step 1: In the open seat game, G is symmetric around zero.

Recall that G is the cdf of $(ln\varepsilon^B - ln\varepsilon^A)$ where $ln\varepsilon^B$ and $ln\varepsilon^A$ are i.i.d. random variables.

Clearly $E(ln\varepsilon^B - ln\varepsilon^A) = E(ln\varepsilon^B) - E(ln\varepsilon^A) = 0$.

Also, $G(\Delta) = 1 - G(-\Delta)$. To see this, recall h is the pdf of lnε.

Let $\hat{G}$ be the cdf of $-(ln\varepsilon^B - ln\varepsilon^A)$. Note that

$G(\hat{x}) = \int_{-\infty}^{\hat{x}} \int_{-\infty}^{\infty} h(\ln \varepsilon^A) h(x + \ln \varepsilon^A) d \ln \varepsilon^A dx$

$= \int_{-\infty}^{\hat{x}} \int_{-\infty}^{\infty} h(\ln \varepsilon^B) h(x + \ln \varepsilon^B) d \ln \varepsilon^B dx = \hat{G}(\hat{x})$

Then, note that $1 - G(\Delta) = \hat{G}(-\Delta) = G(-\Delta)$ which proves that $G(\Delta) = 1 - G(-\Delta)$.

Step 2: If $\hat{\gamma}^A = \hat{\gamma}^B = \gamma$ then the utility functions of A and B (the objective functions in the platform selection subgame) are symmetric. Thus, A and B's best response functions are identical.

Utility functions are:

$G(\Delta)((1 - q^A(1-\alpha))RE(\gamma^A|Awins) + S) + (1 - G(\Delta))(\alpha q^B RE(\gamma^B|Bwins))$

for A,

$(1 - G(\Delta)((1 - q^B(1-\alpha))RE(\gamma^B|Bwins) + S) + G(\Delta)(\alpha q^A RE(\gamma^A|Awins))$

$= G(-\Delta)((1-q^B(1-\alpha))RE(\gamma^B|Bwins)+S)+(1-G(-\Delta))(\alpha q^A RE(\gamma^A|Awins))$

for B.

That is, B's utility function is the same as A's except with B variables playing the part of B variables.

Say q* maximizes utility for A in q^A given q^B=q̂, then q* maximizes B's utility in q^B when q^A=q̂. Thus, A and B's best response functions are identical.

Step 3: FACT: Any 2 player game with symmetric and continuous payoffs and compact and convex strategy sets has a symmetric equilibrium.

Follow the standard proof of the existence of Nash Equilibrium, generalized to work for infinite but compact and convex strategy sets.

To use Glicksberg's (1952) generalization of the Kakutani fixed point theorem, best response correspondences ($B : \Sigma \rightarrow \Sigma$) must satisfy:

i) B(x) is non-empty for all x.

True since utility functions are continuous and action spaces are compact so that the theorem of the maximum applies.

ii) B(x) is convex for all x.

Consider two points in B(x). Then the two points yield the same level of utility, and so does any mixture between them. Thus B(x) is convex-valued.

iii) B(x) has a closed graph (is upper hemi-continuous).

The standard argument relies only on continuity of the utility function.

suppose $(x(n), y(n)) \rightarrow (x, y)$ with $y(n) \in B(x(n))$ but $y \notin B(x)$. Then there is $\epsilon > 0$ and y' such that $u(x, y\prime) > u(x, y) + 3\epsilon$. By continuity of u and convergence of $(x(n), y(n))$, for n sufficiently large we have

$u(x(n), y\prime) > u(x, y\prime) - \epsilon > u(x, y) + 2\epsilon > u(x(n), y(n)) + \epsilon$

Thus y' does strictly better than y(n) against x(n), which is a contradiction.

By applying the fixed point theorem to player A's best response function find an x such that x=B(x). But this means x is also a best response for player B when player A plays x, so that x is a symmetric equilibrium.

Step 4: If $q^A = q^B$ then there is an entry equilibrium with $\hat{\gamma}^A = \hat{\gamma}^B$.

By similar arguments to those in step 1, entry conditions for A and B are symmetric:

$G(\Delta)((1-q)\hat{\gamma}^A E(\varepsilon^A|Awins)R + S) - \frac{1}{\theta}(\hat{\gamma}^A - \mu) = 0$ for A

$(1 - G(\Delta)((1-q)\hat{\gamma}^B E(\varepsilon^B|Bwins)R + S) - \frac{1}{\theta}(\hat{\gamma}^B - \mu)$

$= G(-\Delta)((1-q)\hat{\gamma}^B E(\varepsilon^B|Bwins)R + S) - \frac{1}{\theta}(\hat{\gamma}^B - \mu) = 0$ for B

Furthermore, if we look for a solution where $\hat{\gamma}^A = \hat{\gamma}^B$ the entry conditions become:

$\frac{1}{2}((1-q)\hat{\gamma}E(\varepsilon^i|jwins)R + S) - \frac{1}{\theta}(\hat{\gamma} - \mu) = 0$

so $\hat{\gamma} = \frac{\frac{\mu}{\theta} + \frac{S}{2}}{\frac{1}{\theta} - \frac{1}{2}(1-q)E(\varepsilon^i|jwins, q^i\gamma^i = q^j\gamma^j)R}$ which is positive by assumption.

To sum up, the symmetric equilibrium in the platform selection subgame supports the symmetric entry equilibrium. A1' guarantees that more able challengers for either A or B will not find it worthwhile to enter.

A fully symmetric always exists, although it may involve mixing in the platform selection subgame. Quasiconcavity of utility functions would be sufficient to guarantee symmetric pure strategy equilibrium. I have not been able to prove this in general.

CHAPTER 2

Reputation and Accountability in Repeated Elections

2.1. Introduction

In a representative democracy, voters have the power to choose which citizens will occupy government posts. Even if they cannot directly observe politicians' actions, voters may harness this power to induce incumbent politicians to work in their interest by conditioning reelection on performance. This understanding of the relationship between voter and politician has been studied by Key (1966), Barro (1973), Ferejohn (1986), and others, and is the driving force behind all models of political agency.

If, as seems likely, politicians differ in their ability or preferences, an additional consideration must be taken into account by voters when making reelection decisions. There is a trade-off between having a reelection rule which effectively aligns the interests of the incumbent with the voters', and one which focuses on reelecting the type of politicians who are most able or willing to work in the voters' interest. The first of these is commonly referred to as sanctioning, while the second is called selection.

At an intuitive level, the two goals need not be entirely at odds. If good performance is the primary means by which voters can identify 'good' politicians, then focusing on selection means rewarding good performance with reelection. This should motivate all politicians to work in the voters' interest as 'bad' politicians try to appear 'good' in order to secure another term in office. Thus, selection and sanctioning are at least partly complementary.

In spite of this apparent complementarity, the view that elections are best understood in terms of selection only has gained considerable traction among political scientists[1]. To cite a representative and influential example, Fearon (1999, p. 77) writes: "when it comes time to vote it makes sense for the electorate to focus completely on the question of type: which candidate is more likely to be principled and share the public's preferences?" He argues that, while voters might like to use a retrospective voting rule which incentivizes incumbents optimally, they cannot commit to doing so because politicians who are more likely to be 'good' are also more likely to perform well in the future. Thus, if voters are rational, they will focus exclusively on keeping 'good' politicians in office.

While this argument is unqualifiedly true in the model studied by Fearon, it is important to keep in mind the assumptions on which it relies. As I show in the literature review below, these assumptions are common in many related works. First, Fearon studies a two period model, so that electoral incentives cannot be provided during an incumbent's second term. This assumption makes sense in many contexts, but it is not apt for studying settings where there are no term

[1]See the literature review below and Ashworth, Bueno de Mesquita, and Friedenberg (2009) for further discussion.

limits, such as the U.S. Congress. Second, he assumes that differences among politicians are such that some will perform better than others even in the absence of electoral incentives. This seems natural when studying differences in integrity or preferences, but it is not clear that it is true when politicians differ only in their competence or ability.

If one modifies these assumptions, it may no longer be true that ‘good’ politicians are always more likely to perform well in the future. Rather, future performance will depend on expectations of future behavior. Therefore, one can no longer conclude that voters must focus entirely on selection without looking closely at how voters and politicians expect their relationship to proceed. For example, if voters always reelect incumbents with high enough reputation, once a politician has developed a strong reputation for being ‘good’ he will have little reason to worry about his job security, and will thus have little motivation to exert costly effort. Therefore, voters may prefer to take their chances with an inexperienced politician who will work hard in order to make a name for herself rather than reelect a venerable incumbent who is not motivated to perform.

In this paper I study a simple, infinite-horizon model of repeated elections with no term limits in which politicians differ in their competence. I find that, in this setting, there is no equilibrium in which competent politicians exert positive effort while voters condition reelection only on reputation. Furthermore, I find a class of equilibria in which voters use performance cutoffs to induce incumbents to work in their interest. These equilibria predict that politicians will work less as their reputation improves.

There are two types of politician in my model: H (high) and L (low). H-types are competent: by exerting costly effort they can improve the expected utility of voters. L-types, on the other hand, are incompetent: they do not have the ability to improve outcomes, or it is prohibitively costly for them to do so. If H-types are believed to exert some effort, the voters' beliefs about the likelihood of an incumbent being an H-type will evolve along with his observed performance. I refer to these beliefs as a politician's reputation.

Because of the repeated nature of the elections, the set of equilibria of this model is large and complex. In fact, any pure reelection strategy may be supported as part of a sequential equilibrium (see Proposition 11). This does not mean that any level of effort can be sustained in equilibrium. Nevertheless, the fact that arbitrary behavior on the part of voters can be derived as a prediction of this model highlights the importance of equilibrium selection. The incumbent's reputation is a payoff-relevant state variable in this model, so the Markov perfection refinement has some intuitive appeal. Furthermore, in a Markov perfect equilibrium, the voters' reelection decision depends only on the incumbent's reputation so that these equilibria can be interpreted as those in which voters focus only on selection. In the first of my main results (Proposition 15), I establish that the set of Markovian strategies is not rich enough to allow the voter to incentivize politicians to provide effort.

A slightly more permissive refinement, which has the added benefit of having a clear interpretation in a political context, is Weak Renegotiation-Proofness (WRP, Farrell and Maskin 1989). The logic behind it is as follows: if the relationship between an incumbent of a given reputation and voters can proceed in two different ways (e.g., reelect or not reelect the incumbent), it cannot be that the payoffs associated with one are strictly higher than the payoffs associated with the other for both the incumbent and the voter. If they were, incumbent and voter could come to an agreement, i.e. renegotiate, to proceed in the mutually beneficial way. In particular, WRP rules out equilibria in which voters throw incumbents with good reputations out of office even if, under different circumstances, they would perform better in office than a challenger. Typically (e.g. Banks and Sundaram 1993), these equilibria rely on a belief that, if the incumbent is reelected, he will shirk, expecting to be thrown out of office regardless of his productivity.

To develop some intuition about what WRP might mean in a political context, consider the following example. Keep in mind that this example is meant to clarify the concept of WRP and is not meant to provide an explanation of actual events. Suppose voters in Washington D.C. had certain standards for behavior and outcomes which, if violated by an incumbent, would lead them to elect a relatively unknown challenger in the next election. Furthermore, assume that these standards of behavior included zero-tolerance for illegal acts. From 1979 to 1990, Marion Barry served as mayor and met these performance standards while developing a reputation for being a capable politician. However, in 1990 he was accused and convicted of drug use and possession as well as tax evasion, behavior which clearly violated our assumed standards of behavior for the voters. Nevertheless, in 1995, after working through his legal troubles, Barry was once again elected mayor of D.C.

One interpretation of this type of voter behavior is that Barry was able to convince voters that, if elected, he would behave as if he expected to be held to the same performance standard that he would have been held to had he not broken the law. Because he had a reputation for political ability, this meant that voters could expect a better performance from him, on average, than from an inexperienced challenger. In this sense, Barry and D.C. voters were able to renegotiate their implicit (and hypothetical) electoral contract. Clearly, if this type of renegotiation is feasible, any commitment by voters to expel high reputation politicians after they have violated performance standards is not credible if high reputation politicians are normally believed to outperform challengers.

WRP addresses a commitment problem quite similar to that highlighted by Fearon (1999) and others and which I discuss above. If voters believe that there is a feasible and mutually beneficial way for their relationship with the incumbent to proceed, then it is not credible for the voters to commit to throw such an incumbent out of office. If it were the case that the best achievable future performance by an incumbent were increasing in reputation, rational voters would focus only on selection.

In this model, WRP is qualifiedly equivalent to the condition that the voters' expected payoffs be constant across incumbent reputations (see Claim 16 and Proposition 17 for details). If this is the case, voters face no commitment problem when making reelection decisions because they will be indifferent between having the incumbent or an inexperienced challenger in office. Note the similarity with the equilibria in a seminal work on political agency, Ferejohn 1986, in which voters commit to a reelection rule based on a fixed performance standard. In Ferejohn 1986 this voter indifference condition arises automatically from the assumption that politicians are identical. Thus, if one considers this assumption to be too strong, one may worry about the robustness of the proposed equilibria. However, I find that voter indifference has an important theoretical justification in a model with heterogeneous politicians. Thus, my results provide fresh perspective on, and microfoundations for, the equilibria of Ferejohn 1986.

My second main result (Theorem 19) establishes existence of a class of WRP equilibria in which H-types are incentivized to exert positive effort. In these equilibria, voters condition their reelection strategy only on reputation and current performance. Incumbents are reelected only if their observed performance exceeds a cutoff which varies with the incumbent's reputation at the beginning of his term. Crucially, these performance cutoffs vary in such a way as to make it incentive compatible for politicians to exert just enough effort to leave voters indifferent between reelecting the incumbent and electing an inexperienced challenger, thus making the voters' value function constant across reputations. I refer to this class of equilibria as equilibria in reputation-dependent performance cutoffs (RDC). I view RDC equilibria as a natural generalization of the strategies in Ferejohn 1986 because they rely on the same basic insights. First, performance cutoffs are an intuitive and effective way to provide incentives. Second, voters can credibly commit to using these strategies if they are indifferent between reelecting an incumbent and electing an inexperienced challenger.

An implication of voter indifference is that politicians are able to appropriate the benefits of increases in their reputation by exerting lower effort. This highlights a tension between the selection and incentivizing roles of elections. Voters could do better by committing to a reelection rule which optimally incentivized incumbents. However, such a commitment is not WRP and, thus, not credible. RDC strategies reconcile this tension in a way that is as simple as possible, while passing a stringent test of their credibility and providing politicians with incentives to exert costly effort.

Because, in a model with heterogeneous politicians, selection will play some role in explaining voter behavior, one may reasonably expect that a veteran politician who has developed a reputation for being of a certain type will be treated differently by voters than a first-termer with no record. This, in turn, suggests that a model of electoral control which simultaneously contemplates selection and sanctioning will help us understand the dynamics of political careers. That is, there is likely to be an interplay between an incumbent's reputation, tenure, and behavior, and the standards to which he is held by voters. In RDC equilibria, politicians of high reputation exert lower effort. Also, in expectation, reputation is positively related to tenure so that, for a given politician, tenure is negatively related to performance (see Claim 21).

The paper proceeds as follows. In the following Subsection, I discuss related work and its relationship to this paper. In Section 2.2, I describe the model and its assumptions. Section 2.3 addresses the problem of multiplicity of equilibria, and uses some simple equilibria of the model to motivate equilibrium selection criteria. In Section 2.4, I define RDC equilibria and prove their existence. In Section 2.4.1, I look at what RDC equilibria can tell us about the career dynamics of politicians. Section 2.5 concludes.

2.1.1. Related Literature

There is a growing number of papers which study the selection and incentivizing roles of elections in a unified framework. Much of this work builds on work by Holmström (1999) on career concerns, with the relationship most directly apparent in Persson and Tabellini (2000, ch. 4.5). Notable contributions include Reed (1994), Banks and Sundaram (1998), Fearon (1999), Berganza (2000), Ashworth (2005), and Besley (2006, ch. 3.3). Each of these works studies a model in which voters consider both the selection and incentivizing roles of elections and politicians face term limits. Additionally, several papers have applied similar models to the study of subjects such as transfers to special interest groups (Coate and Morris 1995 and Lohmann 1998), the incumbency advantage (Ashworth and Bueno de

Mesquita 2006), constituency service (Ashworth and Bueno de Mesquita 2008), and CEO activism (Dominguez-Martinez, Swank, and Visser 2008) to name a few.

There are two important differences between the models cited in the previous paragraph and this paper. First, imposing term limits means that last period behavior is easily solved for, and reelection rules are derived by backward induction. In this paper there are no term limits, so voters and politicians face a dynamic problem at every stage. The second is the type of politician heterogeneity studied. In the papers above, voters are assumed to benefit from having a high type in office even if the prospect of reelection is not available to the voter as an incentivizing tool. In this paper, high types differ from low types only in their ability to induce outcomes preferred by the voters. However, improving outcomes is costly to high types so, in the absence of electoral incentives, average performance is equal for high and low types. I feel that this is a more natural way of modeling differences in ability or competence, while the alternative approach is best suited to modeling differences in honesty or alignment of preferences.

Banks and Sundaram (1993) study the selection and disciplining roles of elections in a fully dynamic framework with no term limits. However, they focus on stationary strategies in which voters hold all incumbents, regardless of reputation, to a single performance standard. Therefore, there is no place in their analysis for career dynamics: for a given politician, effort and the probability of reelection are constant in tenure and reputation and independent of the history of play. Additionally, because it is supported by trigger-strategy punishments, the equilibria they propose are not weakly renegotiation-proof; they require that players follow continuation strategies whose associated payoffs are strictly Pareto dominated by other continuation strategies played in equilibrium. Snyder and Ting (2008) use a similar model to study how voter oversight can limit the influence of special interest groups. Gallego and Pitchik (2004) use a related model to explain the timing of the overthrow of dictators.

Duggan (2000) and Banks and Duggan (2006) study a model of repeated elections in which politicians differ in their spatial policy preferences. Voters use the incentive of reelection to induce politicians to temper their policy choices while in office. However, because there is no uncertainty in the execution of policy and strategies are stationary, there is no evolution of beliefs about the incumbent's preferences beyond their first period in office.

Meirowitz (2007) proposes a model of repeated elections in which two long-lived parties, differing in their policy preferences and valence, compete in elections each period. Voters are uncertain about the set of feasible policies rather than about the parties' characteristics or the policy choices made. He shows that, while electoral control is impossible if voters are constrained to using pure strategies, perfect control is possible in mixed strategies. If mixed strategies are to be used, each party must provide the same expected utility to the voter when in office. This leads to a voter indifference condition analogous to the one emphasized in this paper.

Smart and Sturm (2006) present a model of repeated elections in which incumbents' actions are publicly observable, but the underlying state of the world which determines which action is good for the voters is observed only by the incumbent. In this context, they prove that the best Markov perfect equilibrium in the absence of term limits involves all politicians taking the same action regardless of the state of the world. They go on to argue that imposing term limits may help voters by decreasing the incentives for politicians to conform. Their result on the limits of Markov perfect equilibria are in the spirit of the first main result of this paper. The existence of RDC equilibria in my model suggests that allowing voters to condition on more information than Markov perfection allows is an alternative way to increase their expected payoffs which may dominate term limits.

This paper is related to the literature on dynamic principal-agent interactions outside of the political sphere. The approach taken here differs from that taken in much that literature in two main dimensions. First, this paper focuses on the use of a retention rule rather than a compensation contract as an incentivizing mechanism. Second, in most of the literature on principal-agent relationships the principal is assumed to be a Stackelberg first-mover, leaving the agent only his reservation utility. In this paper, I look at Nash equilibria which admit the possibility that the gains from interaction may be shared. Indeed, in the RDC equilibria which I focus on, the agent reaps all of the benefits from increases in his reputation and enjoys utility strictly greater than his reservation value.

Mailath and Samuelson 2001 and Hörner 2002 study related models of reputation formation in which firms attempt to convince consumers that they are competent. Consumers are willing to pay more for a competent firm's products only if they expect the firm to exert high effort. In a result reminiscent my Proposition 2, Mailath and Samuelson show that, with persistent types, Markov perfect equilibria cannot support high effort. They share my skepticism of trigger strategy equilibria but, rather than using a renegotiation-proofness argument to make this point, they argue instead for a restriction to Markovian strategies. They show that effort can be incentivized if a firm's type changes with some positive probability every period so that reputation cannot become 'too good'. Hörner studies a similar model in which many firms compete with each other for marketshare while developing a reputation for competence. He shows that, even with persistent types, effort can be incentivized if firms believe that customers will leave them for a competitor after a bad outcome. Intriguingly, his equilibria involve customers being left indifferent among firms of varying reputations as high reputation firms charge higher prices. However, this is assumed as an equilibrium condition and supported by appropriate beliefs off the equilibrium path. While the equilibria in Hörner's model are Markov perfect, this relies on having only two possible outcomes so that histories of play can be inferred from reputations. Having continuous outcomes would make this partition impossible and it becomes clear that his equilibria have a similar strategic complexity to the RDC equilibria studied in this paper.

2.2. The Model

I study a discrete-time, infinite horizon model of a democratic society. In order to focus on the problems of selecting competent politicians and providing them with incentives to perform well, I abstract from ideological differences in the electorate. Instead, I model citizens as a single, infinitely-lived representative voter.

2.2.1. Preferences, Timing, and Information in the Stage Game

Each period (indexed by $t \in \{1, 2, ...\}$), the voter must select a politician to carry out a task. There is an infinite set P of potential politicians from which the voter may choose. Each politician is infinitely-lived and may serve for as many periods, or terms, in office as the voter asks him to. Once replaced by a challenger, however, a politician cannot return to office.

After the voter elects a politician, the politician exerts effort $a \in \mathbb{R}_+ = [0, \infty)$. This effort impacts, but does not perfectly determine, results $r \in \mathbb{R}$ which I interpret as the voter's stage-game utility.

In order to consider differences in competence, I assume that politicians are one of two types: H or L. For H-types, effort is related to results via a conditional distribution function $F(r|a)$ with density $f(r|a)$. For ease of notation, I normalize units of effort so that effort exerted equals expected results: $E(r|a) = \int_{-\infty}^{\infty} rf(r|a)dr = a$.

H-types receive per-period utility $u(a)$ when in office, and 0 otherwise. The utility function $u(a)$ is twice continuously differentiable and strictly concave. Effort is costly so that u is weakly decreasing in a with $u'(0) = 0$ and $u'(a) < 0 \; \forall a > 0$. I also assume that $u(a) > 0$ for all $a \in [0, \bar{a})$ and some $\bar{a} > 0$.

L-type politicians, on the other hand, are unable[2] to affect the distribution of r so that it is always $F(r|0)$ when an L-type is in office. They receive a payoff $u_L > 0$ when in office and 0 otherwise, so that they are always willing to serve if elected. Because L-types are always willing to hold office but cannot make choices which influence payoffs in this game, I will focus on the behavior of H-types.

As is standard in games with imperfect monitoring (Abreu, Pearce, and Stachetti 1990), I assume that the distribution of results has full support: $f(r|a) > 0$ for all r and a. This guarantees that effort levels can never be perfectly inferred by observing results. I also make the following assumptions for analytical convenience. First, that $f(r|a)$ is twice continuously differentiable in both arguments. Second, that changing a does not change the shape of the distribution: $f(r|a) = f(r + k|a + k)$ for any $k \in \mathbb{R}$. This also implies that outcomes can be written as the sum of the effort choice and a zero-mean stochastic component (ε), a common modeling choice: $r = a + \varepsilon$. Finally, I assume that $f(r|a)$ is symmetric around its mean.

Because the evolution of beliefs about incumbent types is central in this paper, it is useful to make assumptions ensuring that good results are more likely when

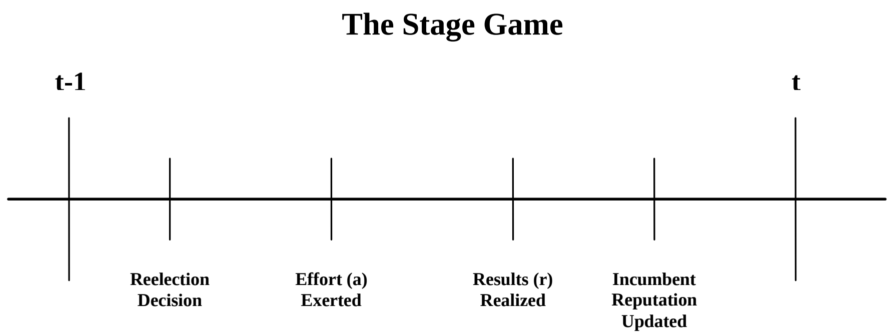
[2]Alternatively, effort is too costly for L-types for it to be worthwhile exerting. $-u_L'(0) > \frac{\delta}{1-\delta} u_L(0) f(0|0)$ is sufficient for this if I make the same assumptions on u_L as on u.

effort is high. Thus, I assume that $f(r|a)$ satisfies the monotone likelihood ratio property (Milgrom 1981): $\frac{f(x|a)}{f(x|a')} > \frac{f(y|a)}{f(y|a')}$ whenever $x > y$ and $a > a\prime$.

In order to guarantee that the politician's objective function is concave so that I may work with first order conditions, I make the following joint assumption on $u(a)$ and $f(r|a)$: $-u''(a) > \max_Q \int f_{aa}(r|a)Q(r)dr$ where Q is any function $Q : \mathbb{R} \to [u(0), \frac{u(0)}{1-\delta}]$ and $f_{aa}(r|a)$ is the second derivative of $f(r|a)$ with respect to a. If $f(r|a)$ is the density of the normal distribution with mean a and variance 1, $-u''(a) > 0.4839 \left(\frac{\delta}{1-\delta}u(0)\right)$ for all a is a sufficient condition.

A politician's type is the private information of the politician. The voter assigns a probability μ_j of being an H-type to politician j. I call μ_j politician j's *reputation*. For ease of notation, when referring to the incumbent's reputation I drop the subscript j. Note that the expected stage-game payoff to the voter when an H-type incumbent exerts effort a is μa, so that reputation is payoff relevant.

The proportion of H-types among the set of potential politicians P is μ_0. Because new politicians are selected randomly from this set, μ_0 will also be the reputation of any politician at the beginning of his first term.



2.2.2. Histories, Strategies, and the Repeated Game

At time t, the voter and all politicians will have information about who has been in office and what rewards the voter has received in all previous periods, $1, 2, ..., t-1$. I call this information a t-history and label it h_t. Let H denote the set of all possible t-histories.

A *reelection strategy* is a measurable function $\sigma : H \to [0,1]$ denoting the probability with which the voter will reelect the incumbent, conditional on all currently available information.

Similarly, an *effort strategy* is a measurable function $a : H \to \mathbb{R}_+$ denoting the effort which a given politician will exert conditional on being in office and on all currently available information.

A belief function $\hat{\mu} : H \to [0,1]^\infty$ is a measurable function specifying the probability with which the voter believes each politician in P to be an H-type. For any politician i who has never previously held office, $\hat{\mu}_i = \mu_0$ regardless of the t-history h_t. In equilibrium, beliefs about a politician's type evolve according to Bayes' rule:

$$\hat{\mu}_j(h_{t+1}) = \frac{\hat{\mu}_j(h_t) f(r_{t+1}|a(h_t))}{\hat{\mu}_j(h_t) f(r_{t+1}|a(h_t)) + (1 - \hat{\mu}_j(h_t)) f(r_{t+1}|0)}$$

Because the distribution f satisfies the monotone likelihood ratio property, $\hat{\mu}_j(h_t)$ is strictly increasing in r_t. For ease of notation, in what follows I drop the subscript when referring to beliefs about the incumbent so that $\hat{\mu}(h_t)$ denotes the probability that the incumbent at time t is an H-type.

It is important to note that different histories can lead to the same incumbent reputation. I can group these together to define a coarser partition of the set of all histories as follows: if $\hat{\mu}(h^1) = \hat{\mu}(h^2)$ for $h^1, h^2 \in H$ then $h^1, h^2 \subset \hat{h} \in \hat{H}$. Note that h^1 and h^2 need not be of the same length. I will refer to this as a Markovian partition of histories and I will use this definition of $\hat{H}$ in the following section to define the Markov perfection refinement.

Given a strategy profile (σ, a) and beliefs $\hat{\mu}$, the voter can compute his expected future payoffs at h_t. Keeping in mind that σ, a and $\hat{\mu}$ denote functions while $\sigma(h_t)$, $a(h_t)$ and $\hat{\mu}(h_t)$ are particular values, I write $V(\sigma, a, \hat{\mu}; h_t)$ for the voter's value function. Letting $h_{t+1}(r)$ denote the t+1-history reached from h_t after a result r is observed, it may be defined recursively:

$$V(\sigma, a, \hat{\mu}; h_t) = \hat{\mu}(h_t) a(h_t) + \delta \int_{-\infty}^{\infty} V(\sigma, a, \hat{\mu}; h_{t+1}(r)) f(r|a(h_t)) dr$$

Where $\delta \in (0, 1)$ is a discount factor common to the voter and all politicians. Note that I do not explicitly write the reelection probability $\sigma(h_{t+1}(r))$ here. Instead, $h_{t+1}(r)$ captures whether the incumbent is reelected or an inexperienced politician of reputation μ_0 is elected.

Similarly, I denote the value function of an incumbent H-type politician $Q(\sigma, a, \hat{\mu}; h_t)$. It may be defined recursively as:

$$Q(\sigma, a, \hat{\mu}; h_t) = u(a(h_t)) + \delta \int_{-\infty}^{\infty} \sigma(h_{t+1}(r))Q(\sigma, a, \hat{\mu}; h_{t+1}(r))f(r|a(h_t))dr$$

Note that I explicitly write the reelection probability $\sigma(h_{t+1}(r))$ in the politician's value function to highlight that the reelection decision determines whether an incumbent will receive $Q(\sigma, a, \hat{\mu}; h_{t+1}(r))$ the following period, or 0 if he is not reelected.

Definition 10. *A sequential equilibrium (Kreps and Wilson 1982) is a strategy profile* (σ^*, a^*) *and a belief function* $\hat{\mu}$ *such that:*

(1) $V(\sigma^*, a^*, \hat{\mu}; h_t) \geq V(\sigma', a^*, \hat{\mu}; h_t)$ for all σ' and h_t.

(2) $Q(\sigma^*, a^*, \hat{\mu}; h_t) \geq Q(\sigma^*, a', \hat{\mu}; h_t)$ for all a' and h_t.

(3) $\hat{\mu}$ evolves according to Bayes' rule[3] using the strategies (σ^*, a^*).

[3]The full support assumption ensures that Bayes' rule is always applicable since all histories are reached with positive density.

r	Voter's stage-game utility.
a	Politician's effort.
$u(a)$	Politician's stage game utility.
$f(r\|a)$	pdf of r given a.
V	Voter's value function.
Q	Politician's value function.
σ	Voter's reelection strategy.
μ	Incumbent's reputation.

Summary of Important Notation

2.3. Equilibrium Selection

As in any infinitely repeated game, I expect there to be a large set of sequential equilibria. In this section, I discuss the problem of the multiplicity of equilibria and some possibilities for narrowing my focus to those equilibria which are most appealing. I begin with the following result which starkly outlines the problem of multiplicity. Then, I proceed by describing several classes of equilibria of this model and using them to motivate several equilibrium refinements.

Proposition 11. *Any pure reelection strategy σ can be supported as part of a sequential equilibrium.*

To see that this is true, I first identify the equilibrium with the lowest payoffs for all players in the equilibrium set, which I call an equilibrium in *grim strategies.* Suppose H-type politicians always choose $a = 0$. Then, the voter is left indifferent among all politicians and may choose any reelection rule. In particular, it is a best response for him never to reelect a politician, regardless of his performance. This reelection strategy makes $a = 0$ a best response.

Next, I note that this equilibrium may be used as part of other sequential equilibria as a credible punishment to the voter for not following a prescribed reelection strategy. Because the voter's expected payoff can never be worse than 0, the following is an equilibrium for any pure reelection strategy σ: the voter plays σ on the equilibrium path while politicians play a best response to σ. If the voter ever deviates from σ, equilibrium play switches to grim strategies.

One may object to the equilibria above by arguing that it is implausible that all politicians in P will coordinate on playing grim strategies in the continuation game. Since the physical environment is identical each time a politician is elected to his first term, it seems natural to focus on equilibria in which strategies are the same every time the voter begins a fresh relationship with a politician. This, of course, implies that the value of the outside option for the politician is constant through all histories. In a sense, this is a stationarity condition which I will call *challenger-stationarity.* Because it is sufficient for my purposes and a weaker condition, I define challenger-stationarity in terms of the value of electing an inexperienced politician rather than the continuation strategies played.

Definition 12. *An equilibrium satisfies challenger-stationarity if the value of electing an inexperienced politician is history-independent.*

In a closely related paper, Banks and Sundaram (1993) describe an equilibrium of the repeated elections game which satisfies challenger-stationarity (following Banks and Sundaram, I call these *simple equilibria*). All politicians are held to a single performance standard. When this performance standard is not met, the politician is not reelected. This is the case for politicians of any reputation, even though the expected rewards to the voter are increasing in the incumbent's reputation. This is enforced through the following trigger strategy: after a politician has missed his performance target once, he never expects to be reelected again and will therefore never again exert effort.

A serious criticism of simple equilibria, in my view, is that after a politician with high reputation misses a performance target, both the voter and the politician would benefit from agreeing to keep the politician in office and continue play as if the incumbent had not violated the voter's performance standard. Therefore, the punishment prescribed by the equilibrium is not credible. More formally, the equilibria are not *Weakly Renegotiation-Proof* (Farrell and Maskin 1989)[4]. There is a continuation equilibrium with associated payoffs which strictly Pareto dominate those specified as following the history in question.

[4]Weak renegotiation-proofness is a condition of internal consistency in that it makes comparisons between the continuation payoffs of a given equilibrium strategy profile. Competing notions of renegotiation-proofness, such as that advocated by Pearce (1987), call for external consistency so that comparisons are made across equilibria. In particular, Pearce argues that comparisons should be made among the the lowest continuation payoffs of equilibria. Because not reelecting politicians (giving them continuation payoff of zero) is the voter's only effective tool for providing incentives, this approach is unlikely to narrow the set of equilibria in this game.

Farrell and Maskin's definition of WRP equilibrium is as follows: an equilibrium strategy σ is WRP if there do not exist continuation equilibria σ^1 and σ^2 of σ such that σ^1 strictly Pareto dominates σ^2 (i.e. payoffs under σ^1 are strictly greater for both players than under σ^2). To adapt the definition of WRP to the current game, I must take into account that the politician's reputation is payoff relevant, so that continuation payoffs when the politician's reputation is μ may not be feasible when his reputation is $\mu' \neq \mu$. The following definition formalizes this notion.

Definition 13. *A sequential equilibrium is Weakly Renegotiation-Proof (WRP) if, for any two histories $h^1, h^2 \in H$ leading to a reputation μ, i.e. $h^1, h^2 \subset \hat{h} \in \hat{H}$, $V(h^1) > V(h^2)$ implies $Q(h^1) \leq Q(h^2)$ (and therefore $Q(h^1) > Q(h^2)$ implies $V(h^1) \leq V(h^2)$).*

As is well known, any Markov perfect equilibrium is WRP. Furthermore, given that reputation is the only payoff relevant state variable, it is natural to look for Markov perfect equilibria in which strategies depend only on reputation. In addition to the standard arguments for Markov perfect equilibrium (Maskin and Tirole 2001) which focus on the simplicity of Markovian strategies, it is also important to address this possibility because previous work on related models has tended to predict that voters will use a simple reputation cutoff as a reelection rule[5]. Additionally, related work on repeated elections by Meirowitz (2007), Duggan (2000), and Banks and Duggan (2006) has focused on Markov perfect equilibria. Banks and Sundaram 1993 (p. 310) end their article by asking whether 'interesting' equilibria which are stationary in reputation exist. Proposition 15 below answers in the negative, at least for this slightly simpler setting.

Definition 14. *A sequential equilibrium is Markov perfect if strategies* (σ, a) *are measurable with respect to the Markovian partition* $\hat{H}$*:* $(\sigma, a) : \hat{H} \to [0,1] \times \mathbb{R}_+$.

Markov perfection takes the idea that history can matter only through the state variable even further than WRP[6]. Once again, existence is easy to check as equilibrium in grim strategies provides a trivial example of a Markov perfect equilibrium. However, the following result makes clear that the Markovian criterion is too strict to allow for the voter to effectively incentivize H-type politicians.

Proposition 15. *There is no Markov perfect equilibrium with positive value for the voter* ($V > 0$)*.*

A full proof is provided in the Appendix (Section 2.6.2). To develop some of the intuition behind the proof, suppose that politicians of all reputations provide effort

[5]See Reed (1994), Banks and Sundaram (1998), Fearon (1999), Berganza (2000), and Ashworth (2005). In these models with term limits, it is assumed that high types perform better than low types in the last term in which no incentives for effort can be provided. Therefore, incumbents are reelected if their expected type is higher than that of a replacement.

[6]See Farrell and Maskin 1989 for further discussion of the relation between WRP and Markov Equilibria.

of at least $\alpha > 0$ in equilibrium. If reelection strategies depend only on reputation, the politician's ex-ante value of acquiring a reputation μ is $\sigma(\mu)Q(\mu) = \hat{Q}(\mu)$. Because posterior reputation is increasing in performance, in order to provide incentives for effort the function $\hat{Q}(\mu)$ must be increasing in reputation. As a politician's reputation nears 1, the change in his reputation for a fixed but wide set of outcomes (r) approaches 0. Therefore, the politician's value function $\hat{Q}$ must increase at least a fixed amount (itself dependent on α) in each of an infinity of ever smaller intervals. However, I know that $\hat{Q}$ is bounded above by the value of holding office forever while exerting zero effort: $\frac{u(0)}{1-\delta}$. Therefore, providing incentives for effort at least α for all reputations is infeasible. Conversely, if politicians of reputation at least μ do not provide effort, it is not worthwhile for the voter to reelect them. This in turn, means that politicians should avoid ending up with a reputation higher than μ, and they can only do this by providing lower effort, leading to an unraveling of incentives for incumbents of all reputations.

The result, and its proof, echo Proposition 2 in Mailath and Samuelson 2001 (henceforth M-S). There are, however, important differences. For instance, the relation between prices and reputation assumed in M-S gives the agent built-in incentives to improve his reputation which are absent in the current setting. Interestingly, the presence of continuous noise (and, thus, continuous outcomes) in my model rules out the type of partition of reputation space which makes mixed strategy equilibria with positive effort possible in M-S.

If one persists in looking for equilibria in which positive effort is exerted while insisting that strategies depend on history in the simplest way possible, the natural next step is to allow for strategies to depend on both the politician's current reputation and his reputation the previous period. Such strategies are Markovian if we take (μ_{t-1}, μ_t) rather than μ_t as the state variable. This condition is equivalent to constraining strategies to depend only on reputation and performance: (r_t, μ_t) or (μ_{t-1}, r_t). In Section 2.4, I define a class of equilibria which satisfy this condition, discuss their relation to previous work, and a prove their existence. These equilibria are also WRP.

2.4. Equilibria in Reputation-Dependent Performance Cutoffs (RDC)

Because the strategies I will consider in this section depend only on reputation at the beginning of the term and current performance, I drop the notation emphasizing the dependence of the voter's and the politicians' value functions V and Q on the entire history of play and strategy profiles. Instead, I emphasize their dependence on incumbent reputation by writing $V(\mu)$ and $Q(\mu)$.

In order to find equilibria in which the voter provides incentives for H-types to provide positive effort but that are WRP and depend on history in the simplest way possible, I look to the structure of the equilibria in the baseline models of political agency. In my view, this has the added virtue of providing some continuity in the modeling and understanding of electoral incentives. The seminal work of Ferejohn 1986 makes two important observations:

- Performance cutoffs are effective means of providing incentives to politicians.
- Voter indifference over incumbents and replacements can be exploited to sustain equilibria with performance cutoffs.

In this model, politicians differ only in their perceived probability of being an H-type - their reputation. An incumbent's reputation will evolve as his record of performance grows and, once he has served at least one term, it will never (with probability zero) be exactly the same as that of a challenger (μ_0). Therefore, for the voter to be kept indifferent between reelecting an incumbent and electing a challenger, it must be that politicians of different reputations provide the same expected utility to the voter: $V(\mu) = V(\mu_0)$, at least for $\mu > \mu_0$. Therefore, I speak of voter indifference and a constant voter value function interchangeably.

Intuitively, this voter indifference condition can be seen as a formalization of the often-voiced sentiment: "One politician is as bad as another." This does not mean that there are no differences in competence among politicians, but that they all exploit the system in their favor to the point where expected performance is constant across politicians.

In addition to the connection to earlier models of political accountability, the voter indifference condition is connected in the current model to the concept weak renegotiation-proofness (WRP, Definition 13). Clearly, voter indifference implies WRP since continuation payoffs are the same for the voter after any history of play, ruling out Pareto improvements.

Claim 16. *Any equilibrium in which* $V(\mu) = V(\mu_0)$ *for all* μ *which are reelected with positive probability is weakly renegotiation-proof (WRP).*

The following Proposition goes some way toward establishing the reverse implication; i.e. that WRP implies voter indifference. Specifically, the indifference condition will hold for a set of reputations of positive measure, and strategies outside of this set will be "uninteresting". In order to do so, I assume that the effort strategies of newly elected politicians do not depend on prior history (i.e. equilibria are challenger-stationary, see Definition 12). This seems natural in the current context where each time a politician is elected for the first time, the continuation game looks identical to the start of the game at time 0.

Proposition 17. *In any equilibrium satisfying weak renegotiation-proofness (WRP) and challenger-stationarity the following conditions hold:*

- *There is a subset of reputation space of strictly positive measure* $S \subset [0, 1]$ *such that, for any* $\mu \in S$*, if* $\hat{\mu}(h) = \mu$ *then* $V(h) = V(\mu_0)$*.*
- *For any* $\mu \in S^C = [0, 1] \backslash S$*, if* $h^1, h^2 \subset \hat{h}(\mu)$ *then, either* $\sigma(h^1) = \sigma(h^2) = 1$ *or* $\sigma(h^1) = \sigma(h^2) = 0$*. That is, strategies in the complement of* S *are Markovian and degenerate.*

Proof. If an equilibrium does not provide positive value for the voter, then the voter's value function is constant at 0 and the conditions above are trivially satisfied. Thus, in what follows, I look at equilibria in which $V(\mu_0) > 0$.

Consider any reputation μ such that one can find histories h^1 and h^2 satisfying $\hat{\mu}(h^1) = \hat{\mu}(h^2) = \mu$, $\sigma(h^1) = 1$ and $\sigma(h^2) = 0$ (or strategies are mixed but may lead to reelection after h^1 and dismissal after h^2). Then WRP implies that, because $Q(h^1) > Q(h^2)$, $V(h^1) \leq V(h^2)$. Also, because it is a best response to reelect after h^1, $V(h^1) \geq V(\mu_0)$. Because it is a best response not to reelect after h^2, $V(h^2) = V(\mu_0)$. From this I conclude that $V(h^1) = V(h^2) = V(\mu_0)$.

This leaves reputation levels at which incumbents are always reelected or always thrown out of office. However, any reelection strategy leading to this sort of behavior over almost all reputations is an essentially Markovian reelection strategy. By the generalization of Proposition 15 in the appendix, this contradicts the premise that the equilibrium in question provides positive value for the voter. □

As it relates to the model of Ferejohn 1986, the relationship between WRP and voter indifference solidifies the microfoundations of equilibria in performance cutoffs. Even if one allows for heterogeneity among politicians, there is an intuitively appealing equilibrium refinement (WRP) which leads back to voter indifference. Thus, its use as a commitment device is both credible and focal.

If we are to preserve voter indifference, we must use performance cutoffs which adjust to the incumbent's reputation. Otherwise, expected results will be increasing in reputation as in Banks and Sundaram 1993's simple equilibria.

In order to keep the voter indifferent between incumbents and replacements ($V(\mu) = V(\mu_0)$), it must be that

$V(\mu) = \mu a(\mu) + \delta \int_{-\infty}^{\infty} [\sigma(\hat{\mu}(r,\mu))V(\hat{\mu}(r,\mu)) + (1-\sigma(\hat{\mu}(r,\mu)))V(\mu_0)]\, f(r|a(\mu))dr$

Solving for $a(\mu)$ and substituting $V(\mu) = V(\mu_0) = V$, we find that $V = \mu a(\mu) + \delta V$. Solving for the incumbent's effort strategy: $a(\mu) = \frac{V(1-\delta)}{\mu}$. Denoting $v = V(1-\delta)$ I write the identity for effort levels which keep the voter indifferent among politicians as:

$$a(\mu) = \frac{v}{\mu}$$

I refer to v as the *value to the voter* of an effort profile $a(\mu)$. Note that $a'(\mu) = -\frac{v}{\mu^2}$ so that effort is decreasing in reputation. Clearly, any equilibrium with positive value to the voter ($v > 0$) will involve a lowest reputation politician which will ever be elected, since $a(\mu) \to \infty$ as $\mu \to 0$. I denote this lowest reelectable reputation $\mu_{\min}$.

Because effort strategies $a(\mu)$ keep the voter indifferent among reelection strategies, if there exists a performance cutoff function $r(\mu) : [0,1] \to \mathbb{R}$ which makes $a(\mu)$ a best response, this will be a sequential equilibrium.

Definition 18. *An equilibrium in reputation-dependent performance cutoffs (RDC) with value v is a sequential equilibrium in which:*

- *Politicians follow an effort strategy $a(\mu_t) = \frac{v}{\mu_t}$.*
- *The voter follows a reelection strategy* $\sigma(r_t, \mu_t) = \begin{cases} 1 \text{ if } r_t \geq r(\mu_{t-1}) \\ 0 \text{ otherwise} \end{cases}$

The following Theorem states the existence of equilibria in reputation-dependent cutoff strategies.

Theorem 19. *There exists a class of equilibria in reputation-dependent performance cutoffs (RDC) in which the voters use a reputation-dependent performance cutoff as their reelection strategy, are indifferent among politicians of all reputations above some threshold* $\mu_{\min}$*, and receive strictly positive expected utility.*

The proof (in Section 2.6.1) proceeds as follows: let $Q(\mu)$ be any bounded and well-behaved candidate for the politician's value function. If I have chosen v carefully, it will be obtainable under $Q(\mu)$ in an RDC equilibrium since $Q(\mu)$ is bounded below by $u(0)$. I then define an operator $T(Q)(\mu) = u(a(\mu)) + \delta \int_{r_Q(\mu)}^{\infty} Q(\hat{\mu}(r,\mu)) f(r|a(\mu))dr$ where $r_Q(\mu)$ is a reputation-dependent performance cutoff implementing v. A fixed point of T will be a value function Q with associated cutoff function $r_Q(\mu)$ implementing an effort strategy $a(\mu) = \frac{v}{\mu}$. Because this effort strategy leaves the voter indifferent between reelecting the incumbent or not, the cutoff function describes a reelection strategy which is a best response. Therefore, once I check sufficient conditions for a fixed point of T, I have found an RDC equilibrium.

The equilibria constructed in the proof of Theorem 19 use performance cutoffs which are above the expected performance of high types (i.e. $r(\mu) - a(\mu) > 0$), and therefore politicians are always reelected with probability strictly less than

$\frac{1}{2}$. Because of the relatively low reelection probability, the politician's value function is lower than it would be in an equilibrium with higher reelection rates, and therefore the highest level of implementable effort would likely be higher in this alternative scenario. Generally, I would expect similar equilibria using cutoffs below expected performance to exist and guarantee reelection rates strictly higher than $\frac{1}{2}$. However, moving performance cutoffs below expected performance allows for the possibility that you may be reelected when your reputation has decreased, and thus that a politician will be reelected even if it is infeasible for him to be incentivized to provide the required effort to keep indifference. Whether this takes place will depend on the slope of $r(\mu)$, which in turn depends on the shape of $Q(\mu)$, which is an endogenous object. Therefore, whether RDC equilibria with performance cutoffs below expected performance exist remains an open question.

2.4.1. Career Dynamics and Comparative Statics

In this Section, I describe some properties of RDC equilibria with positive value for the voter. A straight-forward implication of RDC equilibria is that effort decreases with reputation. Additionally, in any RDC equilibrium expected reputation increases with tenure. This is easy to see by the following argument. Because, in equilibrium, the voter correctly anticipates the incumbent's behavior as a function of his type, the expected reputation of the incumbent after serving a term is the same as his reputation at the beginning of the term. However, those incumbents who end the term with the lowest reputation will be thrown out of office, leading us to conclude that expected reputation will increase every time an incumbent is reelected. This implies a negative relationship between expected performance and tenure for a given politician (though not across politicians). Because incompetent incumbents have no reason to exert effort, or to vary their effort exertion with their reputation, effort is decreasing in tenure across politicians, on average.

Claim 20. *The expected level of reputation conditional on tenure is strictly increasing in tenure. High types will fully reveal their type if they can stay in office forever,* $\lim_{tenure\to\infty} E\left(\mu|tenure, H\right) = 1$. *However, any incumbent will be thrown out of office in finite time,* $\lim_{tenure\to\infty} \Pr(i \text{ in office}|tenure) = 0$.

Proof. The first part of the claim is argued in the text above. That high types fully reveal their type if they can stay in office forever, consider that in an RDC equilibrium with value $\frac{a}{1-\delta}$, any competent incumbent exerts at least effort a. As is well known, the sample average of outcomes will be greater than a with probability which converges to one as the sample size increases without bound.

To see that any incumbent is thrown out of office in finite time with probability one, observe first that competent incumbents always survive with a higher probability than incompetent incumbents. This is the case because competent incumbents exert effort so that the distribution of outcomes when they are in office first order stochastically dominates that of incompetent incumbents. In order to incentivize positive effort, there must be a subset of outcomes of positive mass on which the incumbent is not reelected for almost every μ since, otherwise, we would have an absorbing state in which no effort is provided. Because the set of reputations is compact, there is a maximum probability of reelection $p < 1$. Therefore, the probability that any incumbent survives at least n periods in office is $\Pr(i \text{ in office} | tenure = n) \leq p^{n-1}$. Clearly, $\lim_{n \to \infty} p^{n-1} = 0$. □

Claim 21. *For a given politician, expected effort and performance are negatively related to tenure. Across politicians, expected effort is negatively related to tenure.*

This is a prediction which has been emphasized by others, including Banks and Sundaram (1998) and Ashworth (2005), though their derivation relies on last-period effects. As Ashworth (2005) points out, the prediction fits well with the negative correlation between tenure and personal constituent services in the U.S. House of Representatives examined in Cain, Ferejohn and Fiorina (1990). In a study of the U.S. senate, Levitt (1996) finds some evidence of a positive correlation between ideological shirking and tenure.

Recent work by Galasso, et al. (2009) finds a negative relationship between tenure and attendance in the Italian legislature. Attendance may be interpreted as an observation of performance in this context if I reinterpret the model to fit Italy's parliamentary system. In this case, politicians are directly accountable to their party rather than to the voters. One might imagine that parties face a similar retention problem to that faced by voters in democracies with direct representation, and thus may use RDC strategies.

Because of the importance of the voter's outside option in RDC equilibria, it should not be surprising that the best payoff achievable for voters in an RDC equilibrium is increasing in the average reputation of new politicians (μ_0). Indeed, given an RDC equilibrium under μ_0, the same strategies may be used when new politicians are more likely to be H-types ($\mu_0' > \mu_0$). This is because RDC strategies do not depend on the initial reputation of politicians. This implies that the highest achievable voter utility is weakly increasing in μ_0. However, we know that as $\mu_0 \to 0$, so do the feasible payoff levels for voters since $\frac{a}{\mu_0} \to \infty$. Therefore, for any RDC equilibrium with positive value to the voter, there is a μ_0 at which this value is not feasible and, therefore, there is a μ_0 at which the highest expected payoff to the voter is strictly increasing. Similar logic applies to the comparison among equilibria when we vary incumbent payoffs by a constant.

Claim 22. *The highest expected payoff to the voter in an RDC equilibrium is weakly increasing in the proportion of high types in P (μ_0), and is strictly increasing for some μ_0.*

2.4.2. An Example with Perfect Monitoring: Reelection Rates and Efficiency

In this Subsection, I present a simple example in which the issues discussed above and the dynamics implied by RDC equilibria are clarified. In order to derive explicit expressions for the voters' and incumbent's payoffs, I propose a model in which effort and outcomes are perfectly correlated. Predictably, this makes the moral hazard problem manageable. Nevertheless, voter-optimal reelection strategies are performance cutoffs supported by trigger strategies, RDC strategies are optimal among renegotiation-proof strategies, and the example exhibits the key elements of the dynamics described above as well as an intuitive incumbency advantage effect.

The model is a special case of the general model described in Section 2.2. There is an infinite set of potential politicians N from which voters draw when they wish to replace an incumbent. A proportion $1-\mu$ are incompetent and will provide utility $r=0$ to voters. A proportion μ are competent and can improve voter utility by exerting costly effort $a=r$. Competent incumbents receive utility $U-a^{\rho}$ each period they are in office.

Let 1 denote the state in which the incumbent is known to be competent, and 0 the state in which the incumbent is believed to be competent with probability μ. Once the incumbent is known to be competent, voters can elicit effort $a=r^{*}$ by playing the following reelection strategy:

$$\sigma = \begin{matrix} 1 \text{ if } r \geq r^{*} \\ 0 \text{ otherwise} \end{matrix}$$

If the following incentive constraint is satisfied:

$$Q_1 = U - a^{\rho} + \delta Q_1 \geq U$$

For any level of supportable effort a, the incumbent's state 1 value function is $Q_1 = \frac{U-a^\rho}{1-\delta}$. The maximum level of effort which can be supported is $\bar{a} = (\delta U)^{\frac{1}{\rho}}$. Intuitively, this increasing in δ and U and decreasing in ρ but, interestingly, unresponsive to μ.

In state 0, before the incumbent's type has been revealed, the voter can implement effort a using the same type of reelection strategies as above as long as the following incentive constraint is satisfied:

$$Q_0 = U - a^\rho + \delta Q_1 \geq U$$

Clearly, if the voters elicit the same level of effort in both states, $Q_0 = Q_1$. Thus, the maximum level of effort that can be incentivized from a competent incumbent when his type is unknown is $\bar{a}$.

However, as is argued in the Section 2.3, these strategies are not weakly renegotiation-proof if the expected value to the voter of being in state 1 is greater than his expected value of being in state 0, which is clearly the case when the same level of effort is elicited in either case:

$V_1 = \frac{a}{1-\delta} > \mu a + \delta \left(\mu \frac{a}{1-\delta} + (1-\mu) V_0\right) = V_0$

$V_0 = \frac{\mu a}{(1-\delta)(1-\delta(1-\mu))}$

To see this, consider the off-equilibrium outcome in which a competent incumbent (whether we are in state 0 or 1 does not matter here) exerts effort $a' < a$. He is now revealed to be competent and could provide the voter with continuation utility $\frac{a}{1-\delta}$ if the voter and the incumbent agreed to continue play as if expectations of performance had been met. Because $\frac{a}{1-\delta} > \frac{\mu a}{(1-\delta)(1-\delta(1-\mu))}$ and $Q_1 > 0$, this agreement would be strictly beneficial to both parties.

In order to make the threat of electing a challenger credible, more effort should be elicited of unknown incumbents than of known competent incumbents. Specifically, we can derive the following relation:

$$V = \frac{a_1}{1-\delta} = \mu a_0 + \delta V \Rightarrow a_0 = \frac{a_1}{\mu}$$

The state 0 incentive constraint is now:

$$Q_0 = U - a^\rho + \delta Q_1 = U - \left(\frac{a}{\mu}\right)^\rho + \frac{\delta}{1-\delta}\left(U - a^\rho\right) \geq U$$

So the highest level of effort which can be supported of incumbents who are known to be competent is $a^{RDC} = \frac{(\delta U)^{\frac{1}{\rho}}\mu}{(1-\delta+\delta\mu^\rho)^{\frac{1}{\rho}}}$. This increasing in δ, μ, and U and decreasing in ρ.

Note that there is a stark incumbency advantage exhibited in both the equilibria discussed above. The ex-ante probability of reelection in state 1, that is for incumbents who have survived one term, is 1. The ex-ante probability of reelection following an incumbent's first term is μ. Although there is no difference in reelection rates, RDC equilibria predict that incumbents will work less once their type is revealed, whereas trigger strategy equilibria predict no variation in effort or performance levels over a given politician's career.

To measure the loss voters suffer because of their inability to commit to trigger strategy equilibria, we can look at the ratio of the voters' value functions at the best trigger strategy and RDC equilibria respectively:

$$\frac{\bar{V}}{V^{RDC}} = \frac{\frac{\bar{a}\mu}{(1-\delta)(1-\delta(1-\mu))}}{\frac{a^{RDC}}{1-\delta}} = \frac{\bar{a}}{a^{RDC}} \frac{\mu}{(1-\delta(1-\mu))}$$

$$= \frac{(\delta U)^{\frac{1}{\rho}}(1-\delta+\delta\mu^{\rho})^{\frac{1}{\rho}}}{(\delta U)^{\frac{1}{\rho}}\mu} \frac{\mu}{(1-\delta(1-\mu))} = \frac{(1-\delta(1-\mu^{\rho}))^{\frac{1}{\rho}}}{(1-\delta(1-\mu))}$$

The following Claim points out two implications of the expression above.

Claim 23. *The efficiency loss due to the voters' commitment problem vanishes as voters and politicians become arbitrarily patient, or as the adverse selection problem becomes arbitrarily small:*

- $\lim_{\delta\to 1} \frac{\bar{V}}{V^{RDC}} = 1$ *and* $\lim_{\delta\to 0} \frac{\bar{V}}{V^{RDC}} = 1$
- $\lim_{\mu\to 1} \frac{\bar{V}}{V^{RDC}} = 1$

The efficiency loss is monotonically decreasing in μ*, but hump-shaped in* δ*. That is, there exists a* $\delta^* \in (0,1)$ *such that* $\frac{\partial \frac{\bar{V}}{V^{RDC}}}{\partial\delta} > 0$ *for any* $\delta < \delta^*$ *and* $\frac{\partial \frac{\bar{V}}{V^{RDC}}}{\partial\delta} < 0$ *for any* $\delta > \delta^*$*.*

Proof. The limit results can be derived from inspection of the expression for $\frac{\bar{V}}{V^{RDC}}$:

$$\lim_{\delta\to 1} \frac{\bar{V}}{V^{RDC}} = \lim_{\delta\to 1} \frac{(1-\delta(1-\mu^{\rho}))^{\frac{1}{\rho}}}{(1-\delta(1-\mu))} = \frac{\mu^{\frac{\rho}{\rho}}}{\mu} = 1$$

$$\lim_{\delta\to 0} \frac{\bar{V}}{V^{RDC}} = \lim_{\delta\to 0} \frac{(1-\delta(1-\mu^{\rho}))^{\frac{1}{\rho}}}{(1-\delta(1-\mu))} = \frac{1^{\frac{1}{\rho}}}{1} = 1$$

$$\lim_{\mu\to 1} \frac{\bar{V}}{V^{RDC}} = \lim_{\mu\to 1} \frac{(1-\delta(1-\mu^{\rho}))^{\frac{1}{\rho}}}{(1-\delta(1-\mu))} = \frac{1^{\frac{1}{\rho}}}{1} = 1$$

To establish the second result, we must take derivatives of $\frac{\bar{V}}{V^{RDC}}$:

$$\frac{\partial \frac{\bar{V}}{V^{RDC}}}{\partial\mu} = \frac{\frac{1}{\rho}(1-\delta(1-\mu^{\rho}))^{\frac{1-\rho}{\rho}}\delta\rho\mu^{\rho-1}(1-\delta(1-\mu)) - \delta(1-\delta(1-\mu^{\rho}))^{\frac{1}{\rho}}}{(1-\delta(1-\mu))^2}$$

$= \frac{\delta\left((1-\delta(1-\mu^\rho))^{\frac{1-\rho}{\rho}}\mu^{\rho-1}(1-\delta(1-\mu))-(1-\delta(1-\mu^\rho))^{\frac{1}{\rho}}\right)}{(1-\delta(1-\mu))^2}$

Which is negative if $(1-\delta(1-\mu^\rho))^{\frac{1-\rho}{\rho}}\mu^{\rho-1}(1-\delta(1-\mu)) < (1-\delta(1-\mu^\rho))^{\frac{1}{\rho}}$

Or: $(1-\delta(1-\mu^\rho))^{\frac{-\rho}{\rho}}\mu^{\rho-1}(1-\delta(1-\mu))$

$= (1-\delta(1-\mu^\rho))^{-1}\mu^{\rho-1}(1-\delta(1-\mu)) < 1$

That is: $\mu^{\rho-1}(1-\delta(1-\mu)) < (1-\delta(1-\mu^\rho))$

$\mu^{\rho-1}(1-\delta) + \delta\mu^\rho < (1-\delta) + \delta\mu^\rho$

Which always holds if $\mu^{\rho-1} < 1$, or equivalently $\mu < 1$.

Which confirms that $\frac{\bar{V}}{V^{RDC}}$ is decreasing in μ.

$\frac{\partial\frac{\bar{V}}{V^{RDC}}}{\partial\delta} = \frac{(1-\mu^\rho)\frac{1}{\rho}(1-\delta(1-\mu^\rho))^{\frac{1-\rho}{\rho}}(1-\delta(1-\mu))-(1-\mu)(1-\delta(1-\mu^\rho))^{\frac{1}{\rho}}}{(1-\delta(1-\mu))^2}$

Which is positive if:

$(1-\mu^\rho)\frac{1}{\rho}(1-\delta(1-\mu^\rho))^{\frac{1-\rho}{\rho}}(1-\delta(1-\mu)) > (1-\mu)(1-\delta(1-\mu^\rho))^{\frac{1}{\rho}}$

Or, $(1-\mu^\rho)(1-\delta(1-\mu)) > \rho(1-\mu)(1-\delta(1-\mu^\rho))$

$1-\mu^\rho-\rho(1-\mu) > \frac{\delta}{1-\delta}\left[\rho\mu^\rho(1-\mu)-\mu(1-\mu^\rho)\right]$

Note that $\rho\mu^\rho(1-\mu)-\mu(1-\mu^\rho) < 0$

and $1-\mu^\rho-\rho(1-\mu) < 0$.

Therefore, $\frac{\partial\frac{\bar{V}}{V^{RDC}}}{\partial\delta}$ is positive whenever

$\frac{\delta}{1-\delta} > \frac{1-\mu^\rho-\rho(1-\mu)}{\rho\mu^\rho(1-\mu)-\mu(1-\mu^\rho)}$

Because $\frac{\delta}{1-\delta}$ approaches zero as $\delta \to 0$ and grows without bound as $\delta \to \infty$, $\frac{\partial\frac{\bar{V}}{V^{RDC}}}{\partial\delta}$ is positive for small values of δ and negative for high enough δ. □

The following figures illustrate the result above.

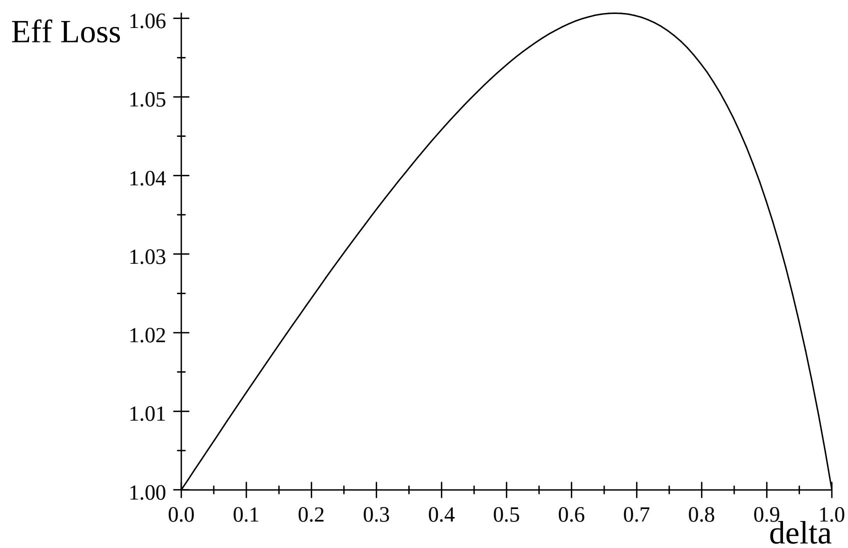


Figure 2.1: Efficiency of RDC equilibria as δ varies ($\mu = 0.5$, $\rho = 2$).

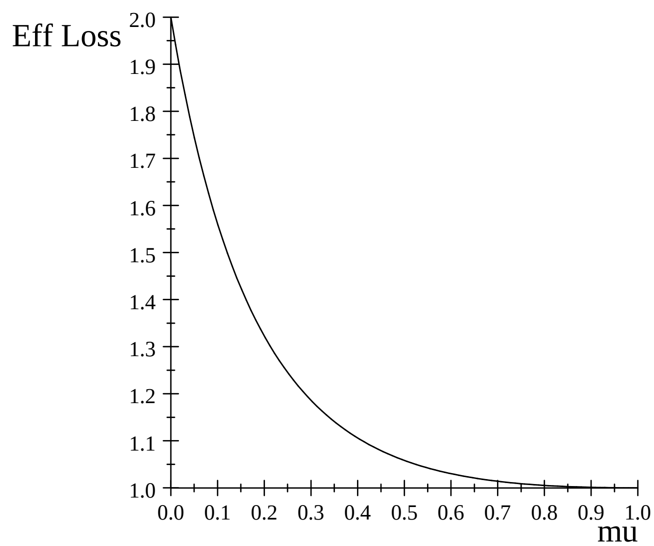


Figure 2.2: Efficiency of RDC equilibria as μ varies ($\delta = .75$, $\rho = 2$).

2.5. Conclusions

The aim of this paper has been to improve the general understanding of the dual role of elections: selecting competent politicians and incentivizing them to exert costly effort to the benefit of the electorate. In particular, I have focused on the potential interaction between a politician's reputation, the voter's willingness to replace him with a less experienced candidate, and the politician's performance. I have done so in the context of a simple model of repeated elections without term limits which does not assume that competence is desirable to the voter even in the absence of incentivizing mechanisms.

As in many infinitely repeated games, the problem of equilibrium selection takes center stage. However, attention paid to this issue has been rewarded in unexpected ways. I have shed light on the question of whether voters can effectively incentivize politicians by simply conditioning reelection on reputation. The answer is no (Proposition 15), at least in the simple model I study. I have uncovered an interesting relationship between weak renegotiation-proofness and the condition that the voter be left indifferent among politicians of different reputations and, therefore, between reelecting an incumbent and electing an inexperienced challenger (Claim 16 and Proposition 17). This has given us fresh perspective on a seminal work in political agency (Ferejohn 1986) and increased confidence in its underlying logic. Finally, I have considered some of the virtues and limitations of the large set of equilibria in trigger strategies.

My exploration of the equilibrium set and its refinements led me to generalize the equilibria of Ferejohn 1986 to a model with non-homogeneous politicians (RDC equilibria, Section 2.4). The use of voter indifference to support performance cutoffs which, in turn, allow the voter to incentivize effort from politicians is consistent with several intuitively appealing equilibrium refinements. Additionally, after establishing existence (Theorem 19), I go on to explore the predictions of the model for political careers. The results presented in Section 2.4.1 replicate those derived in similar models with term limits and in which incumbent type directly affects voter utility. That they continue to hold when there are no term limits and politicians differ only in competence should encourage researchers to look for evidence of these career dynamics in contexts such as the U.S. Congress and understand them as a consequence of political agency.

I conclude by pointing to two promising avenues for future research. First, one might expect related models to yield rich predictions about the variation in reelection rates across politicians of differing tenure and reputation. Because of technical difficulties which I discuss at the end on Section 2.4, I have not been able to derive such implications from this paper's model. Second, a model of repeated elections which allows for differences in politicians' integrity instead of, or in addition to, differences in competence will make different predictions about voter behavior and political careers. In particular, I conjecture that, in stark contrast to the results of this paper, stationary or Markovian strategies would serve the voter well if 'good' politicians perform as well as any incumbent can. Once a politician's reputation is high enough, his expected performance will necessarily be better than that of a challenger, and voters would like to keep such an incumbent in office as long as possible.

2.6. Technical Appendix

For easy reference in the proofs that follow, I rewrite and label the assumptions on the density function $f(r|a)$ discussed in Section 2.2.1.

(A1.) Full support: $f(r|a) > 0$ for all r and a.

(A2.) $f(r|a)$ is twice continuously differentiable in both arguments.

(A3.)

Monotone Likelihood Ratio Property: $\frac{f(x|a)}{f(x|a')} > \frac{f(y|a)}{f(y|a')}$ $\forall$ $x > y$ and $a > a\prime$.

(A4.) Immutability: $f(r|a) = f(r + k|a + k)$ for any $k \in \mathbb{R}$.

(A5.) Symmetry: $f(r|a)$ is symmetric around its mean.

(A6.)

$$\text{Strict Concavity: } -u''(a) > \max_Q \int f_{aa}(r|a)Q(r)dr \text{ for } Q : \mathbb{R} \to [u(0), \frac{u(0)}{1-\delta}].$$

It is useful to note that A3. and A4. imply that $f(r|a)$ is log-concave in r (see Bagnoli and Bergstrom 2005 for some implications). I use this fact in the proof of Lemma 28 below.

Lemma 24. *If $f(r|a)$ is twice continuously differentiable and it satisfies the monotone likelihood ratio property and immutability, it is log-concave in r.*

Proof. Let $x\prime > x$ and $y\prime > y$.

A density function satisfies the monotone likelihood ratio property if:

$\frac{f(x',y')}{f(x',y)} > \frac{f(x,y')}{f(x,y)}$

Taking logs on both sides:

$\ln f(x'|y') - \ln f(x'|y) > \ln f(x|y') - \ln f(x|y)$

If f is twice continuously differentiable, $\ln f(x'|y') - \ln f(x'|y) \approx \frac{\partial \ln f(x|y)}{\partial y}(y' - y)$ when $(y' - y)$ is small. Thus, the inequality above implies:

$\frac{\partial \ln f(x'|y)}{\partial y} > \frac{\partial \ln f(x|y)}{\partial y}$

Because this most hold for all $x\prime > x$, this is equivalent to $\frac{\partial^2 \ln f(x|y)}{\partial y \partial x} > 0$.

Immutability states that $f(x|y) = f(x - y|0)$.

Therefore, it must be that $\frac{\partial^2 \ln f(x|y)}{\partial y \partial x} = -\frac{\partial^2 \ln f(x-y|0)}{\partial x^2} > 0$. Since this holds for all x and y, I conclude that f is log concave. □

In what follows, $f_a(r|a) = \frac{\partial f(r|a)}{\partial a}$, $f_{aa}(r|a) = \frac{\partial^2 f(r|a)}{\partial a^2}$ and $\hat{\mu}_2(r,\mu) = \frac{\partial \hat{\mu}(r,\mu)}{\partial \mu}$.

2.6.1. Existence of RDC equilibria - proof of Theorem 19

I proceed by determining reputation-dependent performance cutoffs which implement effort levels which make the voter's expected utility constant across reputations. Once I have done this, I define an operator which, for any well-behaved candidate value function for the incumbent, determines performance cutoffs and a new candidate value function. A fixed point of this operator gives us an incumbent value function and associated performance cutoffs. Because, at every reputation point, the voter is indifferent between reelecting the incumbent and electing a challenger, using these performance cutoffs as a reelection strategy is sequentially rational for the voter. Thus, the following four elements describe a sequential equilibrium: value functions for the politician and the voter, effort strategies which keep the voter's value function constant, and reelection strategies which use the derived performance cutoffs to make reelection decisions. In order to guarantee the existence of a fixed point, I must check that the conditions for Schauder's fixed point theorem hold. I do so in a series of Lemmas.

When facing a reputation-dependent performance cutoff, an H-type politician with reputation μ solves the problem:

$$\max_a \left\{ u(a) + \delta \int_{r(\mu)}^{\infty} Q(\hat{\mu}(r,\mu)) f(r|a) dr \right\}$$

To implement performance v (or effort strategy $a(\mu) = \frac{v}{\mu}$) with a reputation-dependent cutoff $r(\mu)$ the politician's first order condition (FOC) must be satisfied at $a(\mu)$:

$$u'(a(\mu)) + \delta \int_{r(\mu)}^{\infty} Q(\hat{\mu}(r,\mu)) f_a(r|a(\mu))dr = 0$$

The FOC uniquely determines the incumbent's action since, by assumption A6.,

$$u''(a(\mu)) + \delta \int_{r(\mu)}^{\infty} Q(\hat{\mu}(r,\mu)) f_{aa}(r|a(\mu))dr < 0$$

The FOC must hold at every reputation point μ so that the derivative of the F.O.C. with respect to μ must be 0:

$u''(a(\mu))a'(\mu) - \delta r'(\mu) Q(\hat{\mu}(r(\mu),\mu)) f_a(r(\mu)|a(\mu))$

$+\delta \int_{r(\mu)}^{\infty} Q'(\hat{\mu}(r,\mu))\hat{\mu}_2(r,\mu) f_a(r|a(\mu)) + Q(\hat{\mu}(r,\mu)) f_{aa}(r|a(\mu))a'(\mu)dr = 0$

Solving for $r'(\mu)$:

$$r'(\mu) = \frac{\frac{1}{\delta}u''(a(\mu))a'(\mu) + \int_{r(\mu)}^{\infty} Q'(\hat{\mu}(r,\mu))\hat{\mu}_2(r,\mu) f_a(r|a(\mu))}{f_a(r(\mu)|a(\mu))Q(\hat{\mu}(r(\mu),\mu))} \tag{2.1}$$

$$+\frac{Q(\hat{\mu}(r,\mu)) f_{aa}(r|a(\mu))a'(\mu)dr}{f_a(r(\mu)|a(\mu))Q(\hat{\mu}(r(\mu),\mu))} \tag{2.2}$$

The Fundamental Theorem of Differential Equations guarantees the existence of a function $r(\mu)$ satisfying the equation above as long as the first order condition is feasible and I can bound $r(\mu)$ away from the point where $f_a(r(\mu)|a(\mu)) = 0$ (for symmetric distributions, this point is $a(\mu)$), since the RHS of the expression above is continuous and the domain of $r(\mu)$ is compact.

Before presenting a proof of existence of these equilibria, I select a feasible value for the voter: $v > 0$. For analytical convenience, I focus on cutoffs where $r(\mu) - a(\mu) > 0$ and $f_a(r(\mu)|a(\mu)) > 0$.

A lower bound for the value of holding office is $\bar{Q} = u(0)$. To emphasize its dependence on v, I write $a(\mu, v) = \frac{v}{\mu}$ for the incumbent's effort strategy. Using this lower bound as a hypothetical constant value function and invoking the immutability assumption A4.:

$$-u'(a(\mu, v)) = \delta \int_{r(\mu)}^{\infty} \bar{Q} f_a(r|a(\mu, v))dr = \delta \bar{Q} f(r(\mu)|a(\mu, v))$$

Clearly, this equality cannot hold for v large enough. However, as $v \to 0$, $a(\mu, v) \to 0$ and therefore $u'(a(\mu, v)) \to 0$. However, $\bar{Q} > 0$, so that the equation must hold for appropriate $r(\mu)$ for v low enough (but still strictly positive). Indeed, I can guarantee that a strictly positive v may be sustained as above even if I restrict attention to cutoffs satisfying $r(\mu) - a(\mu) > L$ for any given lower bound L. This will be useful when proving Lemma 28.

I now present the fixed point problem, referring to the derivations above as they become useful.

Definition 25. *Let $C([0,1])$ be the space of bounded, continuous functions $f : [0,1] \to \mathbb{R}$.*

Let $\hat{C} \subset C([0,1])$ be the restriction of this space to functions with K-bounded first derivative and codomain $[u(0), \frac{u(0)}{(1-\delta)}]$.

It is clear that $\hat{C}$ is non-empty, bounded, closed, and convex.

Definition 26. *The operator $T : \hat{C} \to \hat{C}$ is:*

$T(Q)(\mu) = u(a(\mu)) + \delta \int_{r_Q(\mu)}^{\infty} Q(\hat{\mu}(r,\mu)) f(r|a(\mu)) dr$

A fixed point of this operator will define a value function for the politician in a reputation-dependent cutoff equilibrium. To prove the existence of a fixed point, I will use Schauder's fixed point theorem. Schauder's theorem is a generalization of Brouwer's fixed point theorem to infinite-dimensional spaces. For a proof, see Lusternik and Sobolev (1974).

Theorem 27 (Schauder's Fixed Point Theorem)**.** *Let X be a bounded subset of $\mathbb{R}^m$, and let $C(X)$ be the space of bounded continuous functions on X, with the sup norm. Let F be non-empty, closed, bounded and convex. If the mapping $T : F \to F$ is continuous and the family $T(F)$ is equicontinuous, then T has a fixed point in F.*

I must first verify that T maps $\hat{C}$ to $\hat{C}$.

That $T(Q)$ is continuously differentiable in μ is immediate from the differentiability of f, Q, $a(\mu)$, and $r_Q(\mu)$.

That $T(Q)$ has a K-bounded derivative is verified in the following Lemma.

It will be useful in proving the Lemma to note that the first derivative with respect to μ of the Bayesian updating function is:

$$\frac{\partial \hat{\mu}(r,\mu)}{\partial \mu} = \hat{\mu}_2(r,\mu) = \frac{f(r|a(\mu))f(r|0) + \mu(1-\mu)a'(\mu)f_a(r|a(\mu))f(r|0)}{(\mu f(r|a(\mu)) + (1-\mu)f(r|0))^2}$$

It is useful to note that $\hat{\mu}_2(r,\mu) \to 1$ as $a(\mu) \to 0$. The second term in the numerator converges to zero since $f_a(r|a(\mu))f(r|0)$ is uniformly bounded above.

Lemma 28. *For any continuously differentiable function Q with absolutely K-bounded first derivative, $\left|\frac{\partial T(Q)}{\partial \mu}\right| < K$ for any $\mu \in [\mu_0, 1]$ and small enough $v > 0$.*

Proof. $\frac{\partial T(Q)}{\partial \mu} = \frac{\partial \left[u(a(\mu)) + \delta \int_{r(\mu)}^{\infty} Q(\hat{\mu}(r,\mu)) f(r|a(\mu)) dr\right]}{\partial \mu} =$

$u'(a(\mu))a'(\mu) + \delta \int_{r(\mu)}^{\infty} a'(\mu) Q(\hat{\mu}(r,\mu)) f_a(r|a(\mu)) dr$

$+\delta \int_{r(\mu)}^{\infty} Q'(\hat{\mu}(r,\mu))\hat{\mu}_2(r,\mu) f(r|a(\mu)) dr - \delta r'(\mu) Q(\hat{\mu}(r(\mu),\mu)) f(r(\mu)|a(\mu))$

The first two terms add up to zero by the politician's F.O.C. Substituting equation 2.1 into the fourth term:

$\delta \int_{r(\mu)}^{\infty} Q'(\hat{\mu}(r,\mu))\hat{\mu}_2(r,\mu) f(r|a(\mu)) dr$

$-\frac{f(r(\mu)|a(\mu))}{f_a(r(\mu)|a(\mu))}(u''(a(\mu))a'(\mu) + \delta \int_{r(\mu)}^{\infty} Q'(\hat{\mu}(r,\mu))\hat{\mu}_2(r,\mu) f_a(r|a(\mu))$

$+Q(\hat{\mu}(r,\mu)) f_{aa}(r|a(\mu))a'(\mu) dr)$

I first consider the terms which include Q'. Combining them gives:

$\left|\int_{r(\mu)}^{\infty} \hat{\mu}_2(r,\mu) Q'(\hat{\mu}(r,\mu)) \left(f(r|a(\mu)) - \frac{f(r(\mu)|a(\mu))}{f_a(r(\mu)|a(\mu))} f_a(r|a(\mu))\right) dr\right|$

$< K \left|\int_{r(\mu)}^{\infty} \hat{\mu}_2(r,\mu) \left(f(r|a(\mu)) - \frac{f(r(\mu)|a(\mu))}{f_a(r(\mu)|a(\mu))} f_a(r|a(\mu))\right) dr\right|$

$< K \left| \frac{f(r(\mu)|a(\mu))}{f_a(r(\mu)|a(\mu))} \int_{r(\mu)}^{\infty} \hat{\mu}_2(r,\mu) f_a(r|a(\mu))dr \right|$

Where I use Lemma 24 to derive both inequalities as it guarantees that the terms involving $f(r|a(\mu))$ and $f_a(r|a(\mu))$ will not change sign.

Because $\lim_{a\to 0} \hat{\mu}_2(r,\mu) = 1$ and using Lemma 24 again I can say that the last expression is finite for any $r(\mu) > a(\mu)$ and low enough $a(\mu)$. Therefore, $\lim_{r(\mu)\to\infty} \frac{f(r(\mu)|a(\mu))}{f_a(r(\mu)|a(\mu))} \int_{r(\mu)}^{\infty} \hat{\mu}_2(r,\mu) f_a(r|a(\mu))dr = 0$.

As argued in the text preceding the Lemma, since Q is bounded below, I may choose $r(\mu)$ as large as I like while still supporting positive effort. In particular, I may choose $r(\mu)$, and thus $a(\mu)$, so that the following inequality holds for all $\mu > \mu_0$:

$\delta \int_{r(\mu)}^{\infty} \hat{\mu}_2(r,\mu) f(r|a(\mu))dr - \frac{f(r(\mu)|a(\mu))}{f_a(r(\mu)|a(\mu))} \delta \int_{r(\mu)}^{\infty} \hat{\mu}_2(r,\mu) f_a(r|a(\mu))dr < .9$

Therefore,

$\delta \int_{r(\mu)}^{\infty} \left[Q'(\hat{\mu}(r,\mu)) f(r|a(\mu)) - Q'(\hat{\mu}(r,\mu)) \frac{f(r(\mu)|a(\mu))}{f_a(r(\mu)|a(\mu))} f_a(r|a(\mu)) \right] \hat{\mu}_2(r,\mu)dr < .9K.$

The maximum value Q can take is the value of exerting minimum effort (0) and holding office forever: $\frac{u(0)}{1-\delta}$. Thus, because

$\left| \int_{r(\mu)}^{\infty} f_{aa}(r|a(\mu))dr \right| < B$ for some $B > 0$ I can conclude that

$\left| \int_{r(\mu)}^{\infty} a'(\mu) Q(\hat{\mu}(r,\mu)) f_{aa}(r|a(\mu))dr \right| < B \frac{v}{\mu_0^2} \frac{u(0)}{1-\delta}$ where v is determined by the choice of $r(\mu)$ made above.

Similarly, by assumption $|u''(a(\mu))| < \infty$. I may focus on a closed interval $a \in [0, \frac{v}{\mu_0}]$ so that the second derivative is uniformly bounded above:

$|u''(a(\mu))| < U$ for some $U > 0$.

Using these bounds, I have that the absolute value of the derivative above is bounded by:

$0.9K + \frac{f(r(\mu)|a(\mu))}{f_a(r(\mu)|a(\mu))} \frac{v}{\mu_0^2} \left(U + B \frac{u(0)}{1-\delta} \right) < K$

The first term is strictly less than K. The second term does not depend on K, so that choosing K high enough makes it strictly less than $0.1K$. □

Having a bounded derivative also ensures that the class of functions $T(\hat{C})$ is equicontinuous. A class of functions is equicontinuous if, given $\varepsilon > 0$, there is a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$ for any x in the domain of f and any $f \in T(\hat{C})$.

Lemma 29. *Let $T(\hat{C})$ be a class of bounded, continuous and differentiable functions with a uniformly bounded derivative. Then $T(\hat{C})$ is equicontinuous.*

Proof. For any $f \in T(\hat{C})$, $\frac{|f(x)-f(y)|}{|x-y|} \approx f'(x)$. Because $|f'(x)| < B$, $|f(x) - f(y)| < B\,|x - y|$.

Because the bound B on the derivative is the same for all $f \in T(\hat{C})$, if we choose x and y such that $|x - y| < \frac{\varepsilon}{B}$, $|f(x) - f(y)| < \varepsilon$ for any $f \in T(\hat{C})$. Therefore $T(\hat{C})$ is equicontinuous. □

Next I verify that the operator T is continuous.

Lemma 30. *The operator T is continuous.*

Proof. Let $\{Q_i\}_{i \in \mathbb{N}} \subset \hat{C}$ be a sequence of functions converging to Q in the sup norm.

Then, for any $\beta > 0$ $\exists j \in \mathbb{N}$ such that $\forall i > j$, $\|Q_i - Q\| < \beta$.

$(T(Q_i) - T(Q))(\mu) = \delta \int_{r_Q(\mu)}^{\infty} [Q_i(\hat{\mu}(r,\mu)) - Q(\hat{\mu}(r,\mu))] f(r|a(\mu))dr$

$+\delta \int_{r_{Q_i}(\mu)}^{r_Q(\mu)} Q_i(\hat{\mu}(r,\mu)) f(r|a(\mu))dr$

if $r_{Q_i}(\mu) > r_Q(\mu)$. For the reverse case, an identical argument may be used.

The first term converges to zero by definition of Q_i.

The second term converges to zero because $r_{Q_i}(\mu) \to r_Q(\mu)$. To see this, consider the following equality derived from the politician's F.O.C.:

$\int_{r_Q(\mu)}^{\infty} [Q_i(\hat{\mu}(r,\mu)) - Q(\hat{\mu}(r,\mu))] f_a(r|a(\mu))dr = \int_{r_{Q_i}(\mu)}^{r_Q(\mu)} Q_i(\hat{\mu}(r,\mu)) f_a(r|a(\mu))dr$

Again, the term on the LHS converges to zero by convergence of Q_i. Hence the RHS must also converge to zero. However, because $r_{Q_i}(\mu) > a(\mu)$ and $Q_i(\hat{\mu}(r,\mu)) \geq u(0)$, the terms inside the integral are bounded away from zero. Therefore, it must be that $r_{Q_i}(\mu) \to r_Q(\mu)$.

I have now established that $\|T(Q_i) - T(Q)\| \to 0$ so that T is a continuous operator.

I may now apply Schauder's FPT to find a value function and a reputation-dependent cutoff function $r(\mu)$ implementing effort strategy $a(\mu, v)$.

This completes the proof of existence. $\square$

2.6.2. Impossibility of Markov perfect equilibria with positive effort - proof of Proposition 15

In this section I present a proof of a slightly more general version of Proposition 15. Specifically, I generalize the statement to include strategies which are Markovian with probability 1.

Definition 31. *An equilibrium is essentially Markov perfect if strategies* (σ, a) *are measurable with respect to the Markovian partition for a set of reputations* $M \subset [0,1]$ *of Lebesgue measure* 1*.*

Note that any Markovian strategy is also essentially Markovian. Although the distinction is not of interest in and of itself, I make it here as it is useful in establishing Proposition 17 in Section 2.4. The extension does not significantly complicate the proof since it requires only that we note that non-Markovian strategies which are played with probability 0 do not affect the strategic calculus of players involved.

Proposition 32. *There is no essentially Markov perfect equilibrium with positive value for the voter.*

In what follows, for ease of exposition I write $\hat{Q}(\hat{\mu}(r,\mu))$ for $\sigma(\hat{\mu}(r,\mu))Q(\hat{\mu}(r,\mu))$.

The proof proceeds as follows. First, I consider the case in which effort is bounded below for some interval $[m, 1)$ of reputations and $\hat{Q}$ is weakly monotonic. This leads me to conclude that $\hat{Q}$ is unbounded, a contradiction.

Then, I generalize the result in several ways. First, if $\hat{Q}$ is not weakly monotonic, I show that one may look at a moving average of $\hat{Q}$ and that repeated application of the moving average operator leads to a function which is monotonic or approximately constant over an interval $[z, 1)$, and thus to the same contradiction as above.

Once this is done, I am left with the possibility that effort is not bounded below. However, I show that, if positive effort is ever incentivized, politicians with high reputation must be reelected with positive probability and, if that is the case, there must be politicians of arbitrarily high reputation who exert effort above some fixed lower bound. Therefore, I am able to complete the argument by showing that these minimum conditions are enough to lead to the conclusion that $\hat{Q}$ is unbounded. Thus, there can be no Markov perfect equilibrium supporting positive effort if the politician's payoffs are bounded.

Consider first the case in which there is a lower bound $b > 0$ on the effort exerted by politicians with reputation in $[x, 1)$. Using the politician's FOC, I know that his value function must satisfy

$$\delta \int_{-\infty}^{\infty} \hat{Q}(\hat{\mu}(r, \mu)) f_a(r|a(\mu)) dr \geq -u'(b) = B > 0$$

By A4. I can rewrite $\hat{Q}(\hat{\mu}(r, \mu)) f_a(r|a(\mu))$ as $\hat{Q}(\hat{\mu}(r + a(\mu), \mu)) f_a(r|0))$.

Because $\int_0^{\infty} f_a(r|0) dr < \infty$, I can find a value $r^* \in \mathbb{R}_+$ such that

$\delta \int_{-r^*}^{r^*} \hat{Q}(\hat{\mu}(r + a(\mu), \mu)) f_a(r|0) dr \geq \delta \int_{-\infty}^{\infty} \hat{Q}(\hat{\mu}(r + a(\mu), \mu)) f_a(r|0) dr - \varepsilon$

for some fixed $\varepsilon \in (0, \frac{B}{2})$.

Suppose $\hat{Q}$ is weakly monotonic. If $\hat{Q}$ is weakly decreasing, the integrals above will be weakly negative since, by A5., $\delta \int_{-\infty}^{\infty} \hat{Q}(\hat{\mu}(r + a(\mu), \mu)) f_a(r|0) dr = \delta \int_0^{\infty} \left[\hat{Q}(\hat{\mu}(r + a(\mu), \mu)) - \hat{Q}(\hat{\mu}(-r + a(\mu), \mu)) \right] f_a(r|0) dr < 0$. Thus, the F.O.C. will not be satisfied. Suppose $\hat{Q}$ is weakly increasing. By the monotone likelihood ratio property (A3.), I know that there is a unique point at which $f_a(r|0) = 0$ with the derivative being negative to the left and positive to the right of that point. Because $f(r|0)$ is symmetric (A5.), this point is 0. Then,

$\delta \int_{-r^*}^{r^*} \hat{Q}(\hat{\mu}(r + a(\mu), \mu)) f_a(r|0) dr$

$\leq \delta \int_0^{r^*} \hat{Q}(\hat{\mu}(r^* + a(\mu), \mu)) f_a(r|0) dr + \delta \int_{-r^*}^{0} \hat{Q}(\hat{\mu}(-r^* + a(\mu), \mu)) f_a(r|0) dr$

$\leq \delta \left[\hat{Q}(\hat{\mu}(r^* + a(\mu), \mu)) - \hat{Q}(\hat{\mu}(-r^* + a(\mu), \mu) \right] k$

where $k = \int_0^{r^*} f_a(r|0) dr$.

Therefore, $\hat{Q}(\hat{\mu}(r^* + a(\mu), \mu)) - \hat{Q}(\hat{\mu}(-r^* + a(\mu), \mu) \geq \frac{B}{2\delta k} > 0$ for all μ.

Given μ and r^*, there is a μ' such that $\mu = \hat{\mu}(-r^* + a(\mu'), \mu')$. Therefore, $\hat{Q}$ must increase by at least $\frac{B}{2\delta k}$ over $[\hat{\mu}(-r^* + a(\mu'), \mu'), \hat{\mu}(r^* + a(\mu'), \mu')]$. Because this process can be repeated indefinitely, this implies that $\hat{Q}$ grows without bound, which is a contradiction. Therefore, there can be no Markov reelection strategy leading to a weakly monotonic $\hat{Q}$ over any interval $[x, 1]$ while effort is bounded below by $b > 0$.

I am left with the possibility of a $\hat{Q}$ which is non-monotonic over every interval of the form $[x, 1]$. Suppose I have found such a $\hat{Q}$. Then,

$\delta \int_{-r^*}^{r^*} \hat{Q}(\hat{\mu}(r + a(\mu), \mu)) f_a(r|0) dr \geq \frac{B}{2}$ for all μ.

Define $\hat{Q}(x) = \hat{Q}(\hat{\mu}(r + a(x), x))$ for $x \in [m, 1]$. Then,

$\delta \int_{-r^*}^{r^*} \hat{Q}(x) f_a(r|0) dr \geq \frac{B}{2}$

Therefore, $\delta \int_{-r^*}^{r^*} \frac{1}{\hat{\mu}(r^*,\mu)-\mu} \int_{\mu}^{\hat{\mu}(r^*,\mu)} \hat{Q}(x) dx f_a(r|0) dr \geq \frac{B}{2}$

$\frac{1}{\hat{\mu}(r^*,\mu)-\mu} \int_{\mu}^{\hat{\mu}(r^*,\mu)} \hat{Q}(x) dx$ is a moving average of $\hat{Q}$. We may apply this operator repeatedly defining $\hat{Q}_0 = \hat{Q}$ and $\hat{Q}_i(\mu) = \frac{1}{\hat{\mu}(r^*,\mu)-\mu} \int_{\mu}^{\hat{\mu}(r^*,\mu)} \hat{Q}_{i-1}(x) dx$. The following Lemma establishes a basic but useful fact about the moving average operator.

Lemma 33. *Given a function $\hat{Q}$, there exists an interval of positive length $[z, 1)$ such that $\hat{Q}_2$ is either weakly monotonic or approximately constant on $[z, 1)$.*

Proof. After the moving average operator has been applied once, $\hat{Q}_1$ is continuous and differentiable with derivative

$\hat{Q}_1'(\mu) = \frac{\partial\left(\frac{1}{\hat{\mu}(r^*,\mu)-\mu}\right)}{\partial\mu} \int_{\mu}^{\hat{\mu}(r^*,\mu)} \hat{Q}_0(x) dx + \frac{1}{\hat{\mu}(r^*,\mu)-\mu} \left(\hat{\mu}_2(r^*,\mu)\hat{Q}_0(\hat{\mu}(r^*,\mu)) - \hat{Q}_0(\mu)\right)$.

Therefore $\hat{Q}_2$ is continuously differentiable and

$\hat{Q}_2' = \frac{\partial\left(\frac{1}{\hat{\mu}(r^*,\mu)-\mu}\right)}{\partial\mu} \int_{\mu}^{\hat{\mu}(r^*,\mu)} \hat{Q}_1(x) dx + \frac{1}{\hat{\mu}(r^*,\mu)-\mu} \left(\hat{\mu}_2(r^*,\mu)\hat{Q}_1(\hat{\mu}(r^*,\mu)) - \hat{Q}_1(\mu)\right)$.

$\hat{Q}_2'' = \frac{\partial^2\left(\frac{1}{\hat{\mu}(r^*,\mu)-\mu}\right)}{\partial\mu^2} \int_{\mu}^{\hat{\mu}(r^*,\mu)} \hat{Q}_1(x) dx + 2\frac{\partial\left(\frac{1}{\hat{\mu}(r^*,\mu)-\mu}\right)}{\partial\mu} \left(\hat{\mu}_2(r^*,\mu)\hat{Q}_1(\hat{\mu}(r^*,\mu)) - \hat{Q}_1(\mu)\right)$
$+ \frac{1}{\hat{\mu}(r^*,\mu)-\mu} \left(\hat{\mu}_{22}(r^*,\mu)\hat{Q}_1(\hat{\mu}(r^*,\mu)) + \hat{\mu}_2(r^*,\mu)\hat{Q}_1'(\hat{\mu}(r^*,\mu)) - \hat{Q}_1'(\mu)\right)$.

Because $\hat{Q}$ is bounded, $\hat{Q}_2'$ and $\hat{Q}_2''$ are bounded. Let $B > 0$ denote the bound on $\hat{Q}_2''$.

Given an $\varepsilon > 0$, there is a z such that if $\left|\hat{Q}_2'(\mu)\right| > \varepsilon$ for some $\mu \in [z, 1)$ then $\hat{Q}_2$ is strictly monotonic over $[z, 1)$. This is because the most $\hat{Q}_2'$ can change in a distance less than $1 - z$ is $B(1 - z) < \varepsilon$ for z close enough to 1. If there is no $\mu \in [z, 1)$ such that $\left|\hat{Q}_2'(\mu)\right| > \varepsilon$, then $\left\|\hat{Q}_2 - C\right\|_\infty < \eta$ for some constant function

C and a η which becomes arbitrarily small as $\varepsilon \to 0$. Thus, $\hat{Q}_2$ is approximately constant. $\square$

If $\hat{Q}_2$ is weakly monotonic over $[z, 1)$, I may now repeat the arguments for weakly monotonic functions on $\hat{Q}_2$ starting at the point z. Since a bounded $\hat{Q}$ should imply a bounded $\hat{Q}_2$, I am once again left with a contradiction. If $\hat{Q}_2$ is merely approximately constant, I note that $\delta \int_{-r^*}^{r^*} C f_a(r|0) dr = 0$ by symmetry of $f(r|0)$ (A5.) and, for μ such that $\hat{\mu}(-r^* + a(\mu), \mu)) > z$,
$\left| \int_{-r^*}^{r^*} \hat{Q}_2(\hat{\mu}(r + a(\mu), \mu)) f_a(r|0) dr - \int_{-r^*}^{r^*} C f_a(r|0) dr \right|$
$= \left| \int_{-r^*}^{r^*} \hat{Q}_2(\hat{\mu}(r + a(\mu), \mu)) f_a(r|0) dr \right| < \frac{B}{2}$ (if η is chosen small enough) which contradicts the derived properties of $\hat{Q}$.

Now, I consider the case where there is no lower bound on effort exerted. The following Lemmas provide constraints on what can happen in such a hypothetical equilibrium.

Lemma 34. *In any Markov equilibrium with positive value $V(\mu_0) > 0$, every interval of the form $[\mu, 1]$ must contain reputation points at which politicians are reelected with strictly positive probability.*

Proof. Suppose not. Let $\hat{r}(a)$ denote the outcome which would keep the politician's reputation constant:

$$\hat{r}(a) = \{r | \hat{\mu}(r, \mu) = \mu\}$$

Note that, using assumption A3., $\hat{r}(a) < a$ (if r is normally distributed $\hat{r}(a) = \frac{a}{2}$).

Then consider the first order condition of a politician with the *highest reputation which is reelected with positive probability* μ:

$$u'(a(\mu)) + \delta \int_{-\infty}^{\hat{r}(a)} Q(\hat{\mu}(r,\mu)) f_a(r|a(\mu))dr < 0 \text{ for any } a(\mu).$$

Because $f_a(r|a(\mu))$ is negative for all values below $a(\mu)$. Therefore, $a(\mu) = 0$ and μ is an absorbing state. Since I assumed $V(\mu_0) > 0$, it is not a best response for the voter to reelect a politician with reputation μ, contradicting the definition of μ. □

Lemma 35. *Consider a Markov perfect equilibrium with positive value for the voter* $V(\mu_0) > 0$*. In every reputation interval of the form* $[\mu, 1]$ *there must be a subset of positive measure in which politicians exert effort above some fixed lower bound* $b > 0$*.*

Proof. Suppose not. Then, choose a lower bound $b < \frac{1}{2}V(\mu_0)$ and let $[\mu, 1]$ be an interval over which effort is bounded above by b almost everywhere. V is bounded above by the constant function $\bar{V} = \frac{\bar{a}}{1-\delta}$ where $u(\bar{a}) = 0$. Let k satisfy $\sum_{i=0}^{k} \delta^i b + \sum_{i=k+1}^{\infty} \delta^i \bar{V} < V(\mu_0)$. By Lemma 34, there must be reputations arbitrarily close to 1 which are reelected with positive probability. Because effort is bounded, I may choose a reputation (call it $\hat{\mu}$) which is reelected with positive probability and from which the probability of transitioning out of $[\mu, 1]$ in k periods or fewer (call it p) is arbitrarily small. In particular, if I choose $p < \frac{V(\mu_0)}{\delta \bar{V}}$, an upper bound on the value to the voter of having a politician with reputation $\hat{\mu}$ in office ($V(\hat{\mu})$) is:

$$V(\hat{\mu}) < (1-p)\left(\sum_{i=0}^{k} \delta^i b + \sum_{i=k+1}^{\infty} \delta^i \bar{V}\right) + p\delta\bar{V} < V(\mu_0)$$

If the politician is reelected in each of his first k terms. Note that the probability of transitioning to a point in $[\mu, 1]$ at which effort higher than b is exerted is zero because this may happens only on a subset of measure 0, and therefore this possibility does not affect the calculation of expected rewards.

If he does not survive k terms, then $V(\hat{\mu})$ is less than:

$$V(\hat{\mu}) < b + \delta V(\mu_0) < V(\mu_0)$$

Therefore, it is not a best response to reelect a politician when his reputation is $\hat{\mu}$, which contradicts the definition of $\hat{\mu}$. □

Given Lemma 35, if I have a weakly monotonic value function I need only to modify the arguments above as follows. Instead of moving to a reputation satisfying $\mu = \hat{\mu}(-r^* + a(\mu'), \mu')$ I move to one satisfying $a(\mu') > b$ and $\mu < \hat{\mu}(-r^* + a(\mu'), \mu')$. Once again, I conclude that $\hat{Q}$ must increase by at least a fixed amount $\frac{B}{2\delta k}$ infinitely many times, contradicting its boundedness.

To deal with non-monotonic candidate value functions $\hat{Q}$ I note that, given Lemma 35, repeated application of the moving average operation ensures that the value of all integrals $\delta \int_{-r^*}^{r^*} \hat{Q}_i(\hat{\mu}(r,\mu)) f_a(r|a(\mu))dr$ will be positive. Because these are defined on a closed set $[\mu', 1]$, there exists a minimum value of these integrals. Now, I may apply the same arguments as above: $\hat{Q}_2$ includes a weakly monotonic segment $[z, 1)$, and this contradicts of the boundedness of $\hat{Q}$.

Finally, note that in all the arguments above, having a function Z which differs from $\hat{Q}$ only on a set of Lebesgue measure 0 will not change any of the results, because the integrals will yield the same values under both functions. Therefore, it is immediate that the result extends to rule out essentially Markov perfect equilibria with positive value for the voter.

CHAPTER 3

Super Tuesday: Campaign Finance and the Dynamics of Primary Elections

3.1. Introduction

> "People don't lose campaigns. They run out of money and can't get their planes in the air. That's the reality."
>
> -Robert Farmer, fundraiser for Michael Dukakis and Bill Clinton (quoted in Brown et al. 1995)

Presidential nomination campaigns in the United States are lengthy processes. In 2008, the most recent, the first votes were cast in the Iowa caucus on January 3rd while the last votes were cast five full months later in Montana and South Dakota on June 3rd. During the course of the primary season candidates entered and left the race, and their popularity with voters and donors fluctuated as the process unfolded. Many of the phenomena of interest to both the popular media and academics are inherently dynamic in nature: candidates' momentum, the potential effects of changing the electoral calendar, what voters learn from early results and how information affects their voting decisions, etc.

Money plays an important role throughout the nomination process. Running a competitive campaign is very costly and candidates depend on donors to keep their bids alive. During the 2008 primaries, candidates for the Democratic nomination raised a staggering $787 million, while Republicans raised $477 million[1]. Donors learn about candidates as the primary season progresses and donations fluctuate through time as candidates' performance in early states informs future donation decisions. As the opening quote highlights, contenders typically know they have lost the election when they can no longer raise enough funds to continue campaigning competitively. Clearly, donors are major players in presidential primaries, and their behavior has a first-order impact on the dynamic phenomena mentioned above.

This paper presents a game-theoretic, microfounded model of primary elections which examines the role of campaign finance in determining the unfolding of presidential nomination campaigns. I take the view that campaigns are a means of providing information to the public (as in Coate 2004 and Ashworth 2006), and that the election itself is an information aggregation mechanism through which information dispersed in the population is elicited in order to make the best possible choice of nominee (as in Feddersen and Pesendorfer 1996). Policy differences within a party are taken to be negligible and the information that is aggregated by the elections and revealed through campaigns is about the candidates' electability:

[1]Numbers from The Campaign Finance Institute. See http://www.cfinst.org/pr/prRelease.aspx?ReleaseID=205.

the qualities which determine how likely a candidate is to win the general election for the party.

Within this framework, donating to a political campaign is a way of increasing the amount of information available to voters. Thus, it is a way of helping the party select a better candidate and increase its chances of winning the general election. Given mixed evidence (Ansolabehere, de Figueiredo and Snyder 2003), I am agnostic as to the motivation of donors. In the main body of the paper I speak of a single special interest group (SIG) which sees donations as investments which will yield future benefits in the form of access, policy favors, agenda setting, or other services if the receiving candidate wins the general election. The SIG has an interest in helping the party select the most electable candidate because that is the group's only chance of benefiting from favorable policies. In Section 5, I present extensions of the model, with many special interest groups (Proposition 46) and with altruistic donors (Proposition 47), which lead to equivalent donor behavior. The model of altruistic donations is of special interest as it reconciles the small average size of individual donations (individuals are currently subject to a $2400 donation limit) with donor behavior which is responsive to circumstances in ways suggestive of the expectation that a particular donation will affect outcomes and/or elicit future favors. Thus, it contributes to the theoretical understanding of campaign finance.

The predictions of the model are in line with stylized facts observed in U.S. presidential primaries:

- Donors give gradually to candidates (McCarty and Rothenberg 2000).
- Money follows electoral success (Aldrich 1980 [2][3], Hinckley and Green 1996, Mayer 1996, Damore 1997).
- Candidates drop out under financial duress (Mayer 1996 ch. 2, Norrander 2000, Haynes et al. 2004).

Another point of interest is the effect of making changes in the electoral calendar. In the 2008 primaries, much controversy was sparked by Florida and Michigan's decision to hold their primaries in January, ignoring the parties' order that they be held no sooner than February 5th. Much speculation surrounded the largest Super Tuesday ever, held on February 5th, in which 22 states voted and over half of all delegates were pledged. Indeed, the trend toward frontloading, holding more events sooner in the primary season, has been a subject of debate since it began in 1988 (Busch and Mayer 2004). Proposals for reform of the electoral calendar abound and include a national primary, voting in regional blocks, a scheduling lottery, and others (Smith and Springer 2009).

Because the electoral calendar determines what information donors will have when deciding whether to fund a campaign, the model studied in this paper allows one to examine the welfare implications of adopting different electoral calendars. Donors prefer to have strictly sequential primaries so that the decision of whether to fund each campaign can be made individually, thus minimizing the expected cost of the process (Proposition 40). However, stakeholders who do not bear the cost of the campaign, such as voters and parties, prefer to have as many campaigns funded as possible. These stakeholders may be best served by electoral calendars which are 'lumpy' and force donors to choose whether to fund campaigns in groups. Under the right cost conditions, these electoral calendars will maximize the expected amount of donations made and, thus, the expected amount of information revealed before a nominee is selected (Theorem 44). These blocks of voters, voting simultaneously early in an election, are reminiscent of Super Tuesdays in U.S. presidential primaries. One of the main contributions of this paper is to provide a game theoretic rationale for the existence of Super Tuesdays. My conclusion is that a frontloaded or Super Tuesday calendar may be preferable to a sequential one if the cost of campaigning is low enough for competitive challengers to raise adequate funds for early primaries. Otherwise, a sequential election will be more effective at helping voters select the most competitive nominee.

3.1.1. Related Literature

Sequential elections were first studied in a game-theoretic setting by Dekel and Piccione (2000). Their main result is that equilibria of a simultaneous election game are also equilibria of all sequential versions of the game. Because voters condition their vote on being pivotal, it does not matter whether some information is revealed before a voter casts his ballot. This result left scholars to wonder under what conditions the dynamic phenomena mentioned in the introduction, especially momentum, might arise. Battaglini (2005) shows that, if voting is costly, voters will abstain once a candidate takes a sufficiently large lead. Callander (2007) shows that bandwagons can arise when voters prefer to vote for the eventual winner. Ali and Kartik (2008) show that voting according to posterior beliefs is an equilibrium and can lead to herding. Gershkov and Szentes (2009) present a model where voters must decide whether to acquire costly information prior to voting. They characterize voting mechanisms which maximize the quality of the decisions taken in equilibrium.

These papers have established a canonical model of sequential elections in which there are two candidates and two states of the world. Voters receive private signals about the true state of the world and their utility depends on whether the election selects the 'right' candidate. In this paper I adhere to this canonical framework as far as possible.

Two of the most influential works on the dynamics of primary elections in the political science literature are Bartels (1988) and Aldrich (1980). Both present empirical and anecdotal evidence of momentum and other dynamic phenomena. Bartels emphasizes the role of the media in influencing voter preferences while mostly ignoring the role of money[2]. Aldrich focuses more on campaign finance and, in [3], he models momentum as explicitly arising from a feedback mechanism where electoral success increases donations which, in turn, make electoral success more likely. However, he stops short of explicitly modelling the decisions of voters

[2]According to Mayer 1996, "Bartels mentions campaign finance on exactly three pages, and then only in passing."

and donors that are behind this feedback mechanism. Both Bartels' and Aldrich's work serve as a starting point for the modelling done in this paper, and one of my goals is to reconcile some of their arguments and evidence with the abstract literature on sequential voting.

While the effect of campaign spending on voting behavior (e.g. Haynes, Gurian and Nichols 1997) and the importance of accumulating campaign funds early in a contest (e.g. Goff 2004) has been widely studied, little attention has been paid to the timing of donations and the effect of campaign finance on the dynamics of primaries. A notable exception is McCarty and Rothenberg (2000) who propose a model of the timing of donations and provide empirical support for their conclusions. Their focus, however, is on the bargaining between candidates and PACs rather than on the effect of donations on the dynamics of the election itself. Klumpp and Polborn (2006) study a game-theoretic model of campaign spending and its effects on the dynamics of primary elections. They point out that fewer resources will be spent when the electoral calendar is sequential rather than simultaneous. However, they do not explicitly model donors, assuming instead that campaign funds are available but costly to candidates.

Early models of campaign finance took the relation between spending and votes as given[3]. More recent work has taken the position that campaign spending plays

[3]See Morton and Cameron 1992, Stratmann 2005, and Ashworth 2008 for excellent surveys.

an informational role. One strand of the literature (e.g. Pratt 2002, Roumanias 2005) has argued that campaign spending is indirectly informative, revealing private information held by donors. A second strand posits that campaigns are directly informative, revealing information about the candidates which cannot be falsified (e.g. Coate 2004, Ashworth 2006). This is the position I take in this paper.

3.2. Model

A political party must nominate a candidate to represent it in a general election. It does so by means of a primary election, decided by majority rule.

3.2.1. Candidates

There are two candidates running for the party's nomination: A and B. Candidates care only about winning the primary election. They differ in their electability $e \in \{h, l\}$ with $1 \geq h > l \geq 0$. Electability is a summary variable capturing charisma, political ability, and other characteristics which help a candidate win elections. It is further interpreted as the probability with which the candidate will win the general election if nominated. I will use the terms highly electable, high type, and h-type interchangeably.

There are two states of the world: A and B. In state A, candidate A has electability h while candidate B has electability l. In state B, the reverse is true.

While candidates known their own electability[4], voters and donors do not. Rather, I model them as Bayesian learners with prior beliefs over the state of the world $Pr(A) = p = \frac{1}{2}$.

The limitation to two candidates may seem severe, especially in the context of U.S. presidential primaries where several serious candidates typically seek the nomination. I stick to this narrow focus primarily to keep the model tractable and for continuity with previous theoretical research on sequential elections (see Section 3.1.1). Nevertheless, there are two ways in which the model may be interpreted that make the assumption seem less stringent. First, one may consider the model as pitting the front-runner versus the field. Second, some researchers (e.g. Kessel 1992) divide the nomination process into stages. During the first, non-competitive candidates are winnowed out. During the second, the contest begins in earnest. This model may be interpreted as studying only the second phase of the primary.

3.2.2. Voters

There are 5 voters. When thinking of presidential primaries, I may take voter i to be a representative voter from state i, so that I am modeling a primary with 5 states. Let V be the set of voters with typical element $v \in V$.

All voters have identical preferences:

[4]One could argue that candidates do not know any more than the public about their own appeal. I explore the consequences of making this alternative assumption in Section 3.5.3.

$$u = \begin{cases} 1 & \text{if the h-type wins the primary} \\ 0 & \text{otherwise} \end{cases}$$

This utility function can be interpreted as an expected utility function where the value to the voters of having their party win the general election is $\frac{1}{h-l}$: $u(e) = \frac{e-l}{h-l}$, where e takes on the value (h or l) of the primary winner's electability. Thus, if the h-type wins the primary: $u(h) = \frac{h-l}{h-l} = 1$. If the l-type wins the primary: $u(l) = \frac{l-l}{h-l} = 0$.

3.2.3. Campaigns and Donors

Candidates may send an informative signal to voter j by running a campaign at cost c. For simplicity, I assume throughout that the cost of campaigning is constant across states. If a campaign is run, the voter receives a signal $s \in \{a, b\}$. For notational convenience, I say that a voter who does not receive a signal receives $s = \varnothing$. It is common in the literature (e.g. Feddersen and Pesendorfer '96) to call voters who receive $s \in \{a, b\}$ informed, and voters with $s = \varnothing$ uninformed.

If both candidates campaign actively, the signal has accuracy q, or more precisely, $Pr(a|A) = Pr(b|B) = q > \frac{1}{2}$.

If only one candidate campaigns, he is able to manipulate the information so that a signal favorable to him is sent. Thus, no information is revealed. As is shown in the proof of Theorem 1, a campaign by a single candidate will only be run on the equilibrium path if one of the candidates cannot afford to campaign.

A signal is the private information of the voter to which it is directed. This is meant to capture the effect of face-to-face impressions achieved through town hall meetings, rallies, TV commercials on local channels, etc.

There is one special interest group (SIG) who may choose to provide campaign funds to the candidates. A total donation to candidate i of d_i provides economic benefits of $b(d_i)$[5] if i wins the primary election *and* goes on to win the general election. If the party's candidate does not win the general election, the SIG gets a payoff of 0. Because h (l) represents the probability with which an h-type (l-type) will win the general election, the SIG's expected payoffs conditional on the primary winner's political talent, total donations d to that candidate, and donations $\hat{d}$ to the primary loser, are $hb(d) - d - \hat{d}$ ($lb(d) - d - \hat{d}$).

I assume that the SIG maximizes its expected payoffs. Thus, the SIG, like the voters, has an interest in having a highly electable nominee. However, while voters would prefer signals be sent to all states, the SIG faces a costly information acquisition problem and may decide to stop funding campaigns if it considers that the benefits of additional information are outweighed by its costs.

The assumption of a single SIG is a strong one, particularly in the context of American presidential primaries. In Section 3.5, I develop two alternative models of multiple donors which lead to equivalent equilibrium behavior. The first identifies a cooperative equilibrium of the game with an arbitrary number K of identical

[5]Some models of campaign finance explicitly model the bargaining between candidate and SIG which leads to the function b. I leave such an extension to future work.

SIGs. If all contribute c/K when collectively optimal, and the payoff functions are scaled down to $b(d)/K$, the strategic decision of a particular SIG is identical to that of a single SIG. The second takes seriously the role of individual small contributors by postulating a behavioral model of morally motivated campaign donations.

Let the net benefit of making donations d to an l-type nominee $lb(d) - d$ be a strictly concave function with an interior maximum $d^l > 5c$. This inequality guarantees that the SIG is willing, ex-ante, to fund all feasible campaigns for the winner. Given this condition, when the SIG pays out $2c$ to elicit an informative signal, only c is considered an informational cost, since the remaining c is an investment in political influence that would be made in any case. In other words, the SIG will make donations d^e to the party's nominee, but it will also have made donations $\hat{d}$ to the primary's loser. These donations to the loser pay off by enabling a competitive campaign which provided information to the voters. The assumption is also consistent with the observation in McCarty and Rothenberg (2000) that most campaign donations are made after the primary season is over.

I assume that the electability of the primary winner is revealed to the SIG before it makes its final contribution decision at time 6. I make this assumption to keep the analysis as simple as possible. Without it, the SIG's payoffs would depend on the final vote count which determines the posterior probability that the nominee is an h-type, and therefore the size of the SIG's donation to the candidate.

Let d^h solve $\max hb(d) - d$. Then the SIG donates d^h if the primary winner is type h, and d^l otherwise. Therefore, the benefit to the SIG of having the h-type candidate win the primary is $\Delta b = \left(hb(d^h) - d^h\right) - \left(lb(d^l) - d^l\right) > 0$.

To summarize, the SIG's total payoffs, given that the primary winner's type is $e \in \{h, l\}$, and that it made donations $\hat{d}$ to the losing candidate, will be:

$$eb(d^e) - d^e - \hat{d}$$

3.2.4. Timing

There are 6 periods or dates, indexed by $t \in \{1, 2, 3, 4, 5, 6\}$. The primary election may take place in periods 1 through 5. The general election takes place at $t = 6$.

An electoral calendar $\Gamma : V \rightarrow \{1, 2, 3, 4, 5\}$ is a function specifying the time at which each voter will cast his ballot. The electoral calendar is set prior to the start of the game and is generally taken as exogenous, except in Section 3.4 where I consider the design of optimal electoral calendars. The electoral calendars most commonly studied in the literature are simultaneous (all voters vote at the same date) and sequential (one voter per date). In this paper, I consider all possible calendars.

Without loss of generality, I assume that states vote in order: $\Gamma(1) \leq \Gamma(2) \leq ... \leq \Gamma(5)$. Also, when some voters vote simultaneously so that the image of Γ contains fewer than five elements $|\Gamma(V)| < 5$, the voting takes place at $t = 1, ..., |\Gamma(V)|$, so dates without voting come at the end of the process.

Each period except 6, the timing of events is as follows:

(1) The SIG decides how many campaigns to fund that period.

(2) Candidates decide whether to spend the campaign funds. If they do, campaigns are run and signals sent to voters.

(3) Voters cast their ballots.

As is discussed above, the following takes place at time 6:

(1) The primary winner's type is revealed.

(2) The SIG makes its final donations.

(3) Nature decides the outcome of the general election.

(4) The SIG's payoffs are realized.

To sum up, an *election* is an extensive form game with 5+3=8 players (5 voters, the SIG, and two candidates), where the order of moves and the information structure is as described above.

3.3. Equilibrium

Let H_t denote the set of all public date-t histories h_t. Each h_t includes information on all prior votes, as well as on which voters received a signal, what donations each candidate received at each date, and whether and where those donations were spent by the candidates. Let $H = \cup_{t=1}^{6} H_t$.

A voting strategy for voter i is a function $\sigma_i : H_{\Gamma(i)} \times \{a, b, \varnothing\} \to \{A, B\}$. Thus, a voting strategy specifies who a voter will vote for given everything that has happened in the election and the voter's signal. Note that the amount of information observed by each voter is determined by the electoral calendar. For instance, in a sequential election voter 4 has access to h_4, but in a simultaneous election he can condition his vote only on his signal and h_1.

I define *simple voting* to be the following voting strategy:

$\sigma(a) = A$

$\sigma(b) = B$

$\sigma(\varnothing) = A$ if no uninformed voter has voted in the past.

$\sigma(\varnothing) = B$ if the previous uninformed voter voted A.

$\sigma(\varnothing) = A$ if the previous uninformed voter voted for B.

Simple voting has several attractive characteristics. First, it makes minimal requirements of voters' rationality and computational ability. Indeed, it is consistent with a wide range of theories of voter behavior, including expressive voting (Brennan and Lomasky 1993). Second, it is maximally informative, ensuring that the information contained in the campaigns is revealed to other voters and to the SIG and has a maximal effect on the outcome of the election. This last point is important in this model because the donation decision of the SIG depends on how much of a difference a campaign will make to the probability of a highly electable candidate being selected.

That it is in the interest of uninformed voters to vote in such a way as to let the votes of informed voters determine the outcome of the election has been pointed out by Feddersen and Pesendorfer '06. Their behavior here is analogous to abstention in the Feddersen-Pesendorfer model.

A funding strategy for the SIG is a function $d : H \to \mathbb{R}^2_+$ specifying how much the SIG will donate to each candidate as a function of past play. At date t, the SIG plays the date t strategy $d_t : H_t \to \mathbb{R}^2_+$.

I look at a class of perfect Bayesian equilibria involving simple voting. I call these *simple equilibria*:

Definition 36. *A simple equilibrium is a perfect bayesian Nash equilibrium in which:*

(1) *Voters use simple voting strategies.*

(2) *Candidates always campaign when financially feasible.*

(3) *The SIG maximizes its expected payoffs at each stage. That is, d_t solves* $\max E\left[eb(d^e) - d^e - \hat{d}|h_t\right]$.

Candidates will campaign whenever they have the funds to do so. This is enforced by an off-equilibrium belief, held by voters and the SIG, that failure to campaign is a sure sign of a low electability candidate. Because an l-type would have more to gain from impeding the flow of information, this restriction on beliefs is in the spirit of the Divinity condition on off-equilibrium beliefs (Banks and Sobel 1987).

If voters are voting simply and candidates campaign whenever feasible, the investment decision of the SIG is strategically equivalent to that of a statistician who must decide how many costly experiments to perform before making a binding decision. The following Theorem confirms that simple equilibria always exist.

Theorem 37. *Every election has a simple equilibrium.*

Proof. In the appendix. □

The following Proposition uses the simple equilibria of Theorem 1 to confirm the intuition developed by Aldrich (1980) that momentum may arise because donors are more willing to fund candidates who have been successful in early contests. In this model, when the electoral calendar is sequential, when one of the candidates develops a large enough lead he will receive all subsequent votes. This happens because the SIG will stop funding the opposition, making further updating impossible, while the frontrunner continues to receive funds and campaign. A key difference from other game-theoretic models in which bandwagons arise such as Ali and Kartik (2008) and Callander (2007), however, is that once a bandwagon begins only one candidate continues to receive donations and spend money on campaigns, leading to uninformative signals. This also means that, in contrast to previously held conventional wisdom, one should be able to empirically observe the start of a bandwagon as it will have consequences for the flow of funds to candidates. Finally, I point out that the bandwagon does not arise from learning by voters, but rather from learning by the donor. Nevertheless, the voting behavior I refer to as a bandwagon is the same as in these previous models.

While previous research has referred to bandwagons as situations in which voters are ignoring informative signals, in this context in which signals are costly and are not automatically sent, I propose a definition of a bandwagon which is based on the probability of receiving votes.

Definition 38. *A bandwagon has formed if the leading candidate will receive all subsequent votes with probability 1.*

Proposition 39. *Suppose the electoral calendar is sequential. Then there exist campaign costs* $0 < c_{\text{seq}} < \bar{c}$ *such that, if* $c \in (c_{\text{seq}}, \bar{c})$*, a bandwagon forms with positive probability.*

Proof. In the appendix. □

3.4. Optimal Electoral Calendars

Another point of interest is the design of an optimal primary schedule. It is important to keep in mind that we do not have uniqueness in terms of Nash equilibria. However, as I discuss in the preceding Section, I find the simple equilibria to be particularly convincing and, in what follows, I evaluate electoral calendars assuming that simple equilibria will be played in each case.

It is useful to introduce some easy to understand notation for particular electoral calendars. I use brackets { } to denote a calendar. The first number in brackets is the number of voters voting at date 1. The second number, separated from the first by a dash, denotes the number of voters voting at date 2. I repeat this process until all voters are accounted for. Thus, the sequential calendar is {1-1-1-1-1}, while a simultaneous election corresponds to the calendar {5}.

As a first step, I establish the optimal calendar for the SIG. That the SIG will prefer a sequential calendar is a result established in a different setting by DeGroot (1970). Intuitively, the sequential calendar lets the SIG condition its funding decision on the latest poll results, therefore giving the SIG a larger strategy set. The following Proposition describes the optimal primary schedule for the SIG.

Proposition 40. *(DeGroot '70) The donor-optimal primary schedule is sequential {1-1-1-1-1}.*

Although the SIG-optimal primary schedule is of a simple form, it does not coincide with presidential primary calendars in the United States. This may be because, in the United States, the primary schedule is determined by the National Committees of the parties[6]. The members of these committees have more in common with voters than with the SIG. That is, they do not bear the cost of the campaign. Thus, it is interesting to consider the primary schedule which maximizes the probability of selecting the highly electable candidate.

Definition 41. *Given c, b, and q, an electoral calendar is voter-optimal if it maximizes the ex-ante probability with which the h-type candidate wins the election.*

Definition 42. *An electoral calendar is said to dominate another if the ex-ante probability of the h-type candidate winning the election is weakly higher for all c, b, and q.*

[6]In practice, the National Committee sets certain ground rules and state Committees individually decide when to hold their primary election. However, setting the rules goes a long way toward determining the final outcome.

A calendar strictly dominates another if it dominates it and the ex-ante probability of the h-type winning the election is strictly higher for some triple c, b, q.

If the relation holds only for certain values of c, b, and q, I say that a calendar (strictly) dominates another over the relevant ranges of c, b, q.

There are many possible electoral calendars in a world with five voters (16 in fact, they are listed in the appendix). The following Lemma helps to narrow the field to some particularly interesting candidates: a pure sequential election ({1-1-1-1-1}) and a mixed calendar in which there is a block of three voters voting simultaneously at date 1 followed by the two remaining voters voting sequentially ({3-1-1})[7]. I call this second calendar a Super Tuesday calendar because of its structural similarity to presidential primary calendars in which large blocks of states vote on the same "Super Tuesday" early in the primary season.

Lemma 43. *For any c, b, and q, one of the following electoral calendars is voter-optimal among all possible calendars: sequential {1-1-1-1-1} or Super Tuesday {3-1-1}.*

Proof. In the appendix. □

[7]The calendars {3-1-1} and {1-2-1-1} are strategically equivalent so I could refer to either as a Super Tuesday calendar. Perhaps the second is more reminiscent of Super Tuesday since it allows for a single early vote, like Iowa and New Hampshire might be in the U.S. presidential primary, to happen before the block of voters are scheduled.

The proof proceeds by listing all possible calendars in a five voter election. Then, I derive equivalence relationships which allow me to focus on a subset of calendars. For example, because of the symmetry of the election, which candidate takes a 1-0 lead after voter 1's informed vote does not matter for the SIG's funding strategy. That is, conditioning on the outcome of the first vote will have no effect on the SIG's funding strategy at that moment. Therefore, any calendar in which the first voter votes before all others is strategically equivalent to the calendar identical to it but in which voter 1 votes at the same time as voter 2. I then establish dominance relationships among the remaining calendars until I am left only with the candidates listed above. For example, a Super Tuesday calendar will always lead to the election of the high type candidate with at least as high a probability as a {4-1} calendar and, for some values of c, it will do so with strictly higher ex-ante probability.

The preceding Lemma sets the stage for the following result describing the voter-optimal electoral calendars. Given that I know that either a sequential or a Super Tuesday calendar is optimal, it is much easier to make comparisons among them and arrive at a result describing when one dominates another. In some ways, the following Theorem is the central result of this paper. It is the first result in this literature in which a hybrid calendar (not strictly sequential or simultaneous) plays a major role. It also provides a game-theoretic, effectiveness-based explanation for the existence of Super Tuesdays.

Theorem 44. *For given b and q, there exist values $0 < c_{\text{seq}} < c_{st} < \hat{c}_{\text{seq}}$ such that:*

- *The Super Tuesday calendar dominates all other calendars, and strictly dominates the sequential calendar, when $c \in (c_{\text{seq}}, c_{st})$.*
- *The sequential calendar dominates all other calendars, and strictly dominates the Super Tuesday calendar, when $c \in (c_{st}, \hat{c}_{\text{seq}})$.*

Proof. In the appendix. □

The simultaneous calendar, which has been widely studied and is usually used as a point of comparison to the sequential calendar, is dominated by the Super Tuesday calendar in this model. That is not to say, however, that it is never optimal. It is optimal whenever the costs of campaigning are low enough for the SIG to fund all five campaigns in a simultaneous election, or high enough so that at most one campaign is funded under any electoral calendar. The following Proposition clarifies.

Proposition 45. *For given b and q, there exist campaign costs $c_{sim} \in (0, c_{\text{seq}})$ and $\bar{c} > \hat{c}_{\text{seq}}$ such that the simultaneous electoral calendar is voter-optimal whenever $c < c_{sim}$ or $c > \bar{c}$.*

Proof. In the appendix. □

3.5. Extensions and Alternative Modelling Approaches

Although it is common in the micro-founded literature on campaign finance to assume a small number of donors (Ashworth 2006, Coate 2004 to name just two examples), the assumption is rather restrictive. During the 1999-2000 U.S. election cycle 21 million individuals donated to the candidates' campaigns (Ansolabehere, de Figueiredo and Snyder 2003).

I argued above that the single SIG could be interpreted as an aggregate of the objective functions of many special interest groups or even of many voters for whom political donations are a particular type of consumption. In the following sections I substantiate these claims.

3.5.1. Multiple SIGs

Consider an arbitrary number K of identical special interest groups. SIG i receives benefits $\$\frac{b(d_i)}{K}$ from making total donations d_i to a candidate who goes on to win both the primary and the general election.

Proposition 46. *There exists an equilibrium of the game with K identical special interest groups in which the collective behavior of the SIGs is identical to that of a single large SIG.*

Proof. Consider the funding decision of SIG i when the other K-1 SIGs are following the funding strategy of the single SIG described in Theorem 1. SIG i will fund a campaign if $\frac{c/K}{\Delta b/K} = \frac{c}{\Delta b}$ is smaller than the increase in the probability of the h-type winning the nomination. That is, i's problem is identical to that of a single SIG. □

3.5.2. Donations as Altruistic Behavior

Ansolabehere, de Figueiredo and Snyder 2003 argue that small average donation sizes and the large number of donors to political campaigns make any theory of campaign finance in which donations are seen as investments which are expected to produce returns in the form of altered results or influence on policy-making implausible. They survey 40 articles which attempt to find a link between donations and voting records and find little evidence of a link. However, research looking at the behavior of donors (e.g. Brown et al. 1995 and Gordon et al. 2007) finds that at least some groups of donors behave *as if* their donation were an investment with policy implications or could change the outcome of the election.

One way to reconcile these pieces of evidence is to propose a model of campaign donations as an altruistic act, motivated by its moral implications. One popular explanation of altruistic behavior holds that agents often use a simple version of Kant's categorical imperative to evaluate an action's moral salience (Harsanyi 1980, Brekke et al. 2003). In particular, an action is morally salient if, when adopted by everyone, it maximizes a social welfare function. This type of motivation, known as rule-utilitarian, has been used to explain altruistic behavior in recycling, community service, voting (Feddersen and Sandroni 2006, Coate and

Conlin 2004), information acquisition by voters (Feddersen and Sandroni 2006), and other prosocial behavior.

Consider a set M of identical moral agents of Lebesgue measure 1. Moral agents are a minority in each state, as they are outnumbered by selfish agents in a set S of measure 2. Each agent must decide whether to donate to a candidate's campaign. Even though a single donation cannot influence the outcome of the election, the agents receive satisfaction from performing morally salient actions and receive ex-post utility $1_{(h)} + \sum_{t=1}^{5} \left(m1_{(M,t)} - d_t\right)$, where $1_{(h)}$ is an indicator function which takes the value 1 if the highly electable candidate wins the primary and 0 otherwise, $1_{(M,t)}$ is an indicator function which takes on the value 1 if the agent performed the morally salient action concerning the i'th campaign and 0 otherwise, and d_t denotes the donations made at date t.

An donation d_t is morally salient if it satisfies:

$$d_t = \arg\max_{d_t} E\left[\int_{M\cup S}(1_{(h)} - d1_{(j\in M)})dj|h_t\right]$$
$$= \arg\max_{d_t} \left\{3P(h|h_t, d_t) - E\left[\sum_{i=1}^{5} d_i|h_t, d_t\right]\right\}$$

Where $1_{(h)}$ is the utility of voters and $d = \sum_{i=1}^{5} d_i$ is the total donated by morally motivated agents. It is clear that the expectation of $1_{(h)}$ is $P_t(h|d)$, the probability that the high electability candidate will win the nomination given donations d_t at time t.

In my model, where total donations of $2c$ are necessary for candidates to continue informative campaigning, $d = 2c$ is the only relevant level since there is no loss from funding future state campaigns at a later date, making higher donations redundant and lower donations are merely wasteful[8]. Furthermore, because there is a continuum of morally motivated agents, an individual's decision not to donate will not affect outcomes. Thus, I need not consider strategic deviations sacrificing m now in order to make future morally salient actions cheaper.

A candidate's campaign will continue to be funded as long as the satisfaction of undertaking the morally salient action, $m > 0$, is higher than the cost, and as long as it is socially optimal for the campaign to be funded:

$m > 2c$

$3\Delta P_{t,t+1}(h) > E\left[\sum_{i=t}^{5} d_i | h_t, d_t\right]$

where $\Delta P_{t,t+1}(h)$ denotes the increase in the probability of nominating candidate h given that at least one more campaign is funded. In contrast, a SIG will fund an additional campaign if $\Delta b \Delta P_{t,t+1}(h) > E\left[\sum_{i=t}^{5} \hat{d}_i | h_t, \hat{d}_t\right]$. Note that the SIG only considers donations to the eventual loser as informational costs, while morally motivated donors see all donations as informational expenses. Therefore, keeping a campaign going implies informational costs of c for the SIG and 2c for morally motivated donors. Nevertheless, the problem they are solving is isomorphic.

Proposition 47. *If $m > 2c$, and there are morally motivated voters, then donations to political campaigns will be made as if a single SIG with $\Delta b = \frac{3}{2}$ were funding the campaigns using the strategy of Theorem 1.*

[8]Mixed strategies in which agents give c in expectation are also solutions to the maximization above, but can be ruled out by making the utility cost of money convex.

3.5.3. Symmetric Information about Types

It is plausible to think that candidates are uncertain about their own electoral appeal and learn about their electability through the primary process along with the voters and the SIG.

If candidates are solely office motivated, however, a leading candidate will never agree to campaign even if he is well-funded. Consider a candidate with reputation better than a half. If he campaigns he will lose the nomination with positive probability even if he is, in fact, the h-type. If he refuses to campaign, however, no new information is revealed and the primary season will end with the same beliefs, so that the only rational way for voters to vote is for the leading candidate. Thus, we are left with no campaigning or donations if the prior $p \neq \frac{1}{2}$, and campaigns only in the first period if $p = \frac{1}{2}$.

3.5.3.1. Policy-Motivated Candidates. If candidates also care about whether a candidate from their party wins the general election, they may willingly choose to campaign in order to increase the probability that the most electable candidate from their party is selected. Define a candidate's utility as:

$g(P, W) = \gamma P + (1 - \gamma)W$

Where P is an indicator function for a candidate from his party winning the general election, W is an indicator function for whether the candidate in question has won the primary, and $\gamma \in [0, 1]$. $\gamma = 0$ corresponds to the case discussed above in which the candidate only cares about winning the primary. $\gamma = 1$ describes a candidate motivated only by policy.

Proposition 48. *Given any election, there exists $\bar{\gamma} > 0$ such that, if $\gamma > \bar{\gamma}$, then a candidate will always campaign when feasible.*

Proof. Suppose that a simple equilibrium is played. For this to be incentive compatible for the candidates it must be that at each possible t-history in which A has reputation $R > \frac{1}{2}$,

$$(1-\gamma) + \gamma[R + (1-R)l] < (1-\gamma)P_{t,t+1}(Aw) + \gamma[P_{t,t+1}(h) + (1 - P_{t,t+1}(h))l]$$

or $\gamma \Delta P_{t,t+1}(h)(1-l) > (1-\gamma)(1 - P_{t,t+1}(Aw))$

The left-hand-side is always positive. Because I are dealing with a finite game, there is an absolute maximum value which the right hand side can take. Thus, a $\bar{\gamma}_A > 0$ can be found for which the inequality above is satisfied at all t-histories in which A is leading as long as $\gamma > \bar{\gamma}$. A symmetric argument finds a $\bar{\gamma}_B$ which serves the same purpose for t-histories in which B is in the lead. Taking $\bar{\gamma} = \max\{\bar{\gamma}_A, \bar{\gamma}_B\}$ completes the proof. □

3.6. Concluding Remarks

I have set out to illustrate how campaign finance can be the main driving force behind the dynamics of primary elections. I have done so by presenting a microfounded, game-theoretic model of the interaction between voters, candidates, and donors. Underpinning my analysis is the conception of a primary election as a means by which like-minded members and adherents of a party acquire credible information about the electability of the candidates. The results in Section 3.3 show that it is possible for the process to unfold in such a way that all parties involved are willing to participate, information is revealed, and it is used effectively. I also show that bandwagons can arise as a consequence of learning by donors. This provides an alternative to previous theories of bandwagons based on learning by voters.

In Section 3.4, I present results characterizing the optimal electoral calendar for both donors and voters. Donors prefer to have a sequential primary so that funding decisions can be made gradually and the costs of funding the eventual loser to elicit information minimized. Voters coincide with this preference for sequential elections when campaign costs are relatively high. However, when campaign costs are in a given range, voters prefer electoral calendars in which a group of voters vote simultaneously early in the electoral calendar (at date 1 or 2). This type of calendar is reminiscent of those in recent U.S. presidential primaries which include "Super Tuesdays" in which several states hold their elections on the same day. These results are especially interesting in light of the ongoing controversy surrounding the scheduling of presidential primary elections in the United States. My analysis concludes that a frontloaded calendar, including a Super Tuesday, is optimal for voters and parties as long as competitive candidates are able to fully fund their campaigns in these early stages.

Given the simplicity of the model studied here, several generalizations seem likely to add richness to my conclusions. First, my focus on a contest between two candidates seems inadequate given the large number of candidates who generally contest presidential primaries in the United States, at least in the early stages of the process. Furthermore, campaign spending is, in reality, not a discrete variable. Different amounts of spending can lead to different results. Similarly, something may be learned from the margin of victory in a given district beyond what is revealed by a win or a loss. I am currently exploring these and other extensions.

3.7. Technical Appendix

Proof of Theorem 1.

Proof of existence of simple equilibria with 5 voters.

I proceed by considering each voter's decision problem, from last to first. In evaluating the utility effects of deviations, I assume that the voter knows how earlier voters have voted. If this is not the case, and there are voters of a lower number voting at the same time, the expected benefit of a deviation will be a weighted average of those considered. Thus, if deviating is never worthwhile when the voter is able to condition on previous voters' votes, it is not profitable when the voter cannot condition on this information.

I begin by establishing some basic facts about the SIG's donation strategy which will help simplify the arguments in the proof of Theorem 1 other results. The following Claim states that whenever the SIG funds at least one, or at least two campaigns, it is optimal for it to fund voter 1 and 2's.

Claim 49. *Whenever it is optimal for the SIG to fund any campaigns, it is optimal for the SIG to fund voter 1's campaign. Furthermore, whenever it is optimal for the SIG to fund at least two campaigns, it is optimal for the SIG to fund voter 2's.*

Proof. If the SIG is going to fund at least one campaign, and because voters are identical except for the order they vote in, it loses nothing by having it be voter

1's. On the other hand, its choice set, conditional on whether voter 1's campaign was funded, is weakly larger.

Similarly, because the information revealed by voter 1's vote is not relevant to the SIG's decision to fund additional campaigns, the SIG gains nothing by waiting until this information is revealed. If the SIG funds voter 2's campaign, it is left with more alternatives when considering the campaigns of voters 3-5. □

Definition 50. *A donation strategy d_t is relevant if there is an electoral calendar Γ and a history h_t such that d_t is a best response to simple voting for some triple c, b, q.*

Lemma 51. *In any relevant donation strategy, the maximum (over histories) total number of campaigns funded by the SIG is either odd or zero. Thus, under simple voting, uninformed votes will never decide the election, and it is a best response for them to vote simply given that other voters are also voting simply.*

Proof. Suppose that voters are voting simply. Consider first the SIG's decision whether to fund one or two campaigns. Funding one campaign leads to selecting an h-type with probability q. Funding two campaigns leads to an h-type nominee with probability $q\left(q + 2\frac{1}{2}(1-q)\right) = q$. Intuitively, conditional on the first vote, adding an additional campaign can only tie the informative vote count or increase the frontrunners lead. At worst, the frontrunner's posterior will be $\frac{1}{2}$ for each candidate and does not change the optimal choice of candidate. The same logic applies to the difference between funding three and four campaigns. In that case, funding three campaigns selects the h-type with probability $q^2(q+3(1-q)) = q^2(3-2q)$. Funding four leads to a probability of success of $q^2\left(q^2+4q(1-q)+6\frac{1}{2}(1-q)^2\right) = q^2(3-2q)$.

Therefore, the SIG will fund zero campaigns, one campaign, or fund until one candidate receives two out of three or three out of five informed votes. That is, an election will end with an even number of campaigns funded only if one candidate has a 2-0, 3-1, or 4-0 lead in informed votes. When there is an odd number of campaigns funded, whenever a voter does not receive an informative signal (i.e. at least one candidate does not have sufficient funds to campaign for that voter) there will be another voter in the same situation. Thus, according to simple voting, the first voter in question will vote A and the second B, allowing informed voters to determine the outcome of the election. Because informed voters vote their signal, this means that the candidate who finishes the election with the highest posterior probability of being an h-type will win. Therefore, voting simply is a best response for uninformed voters. After a 2-0 lead in informed votes, if no additional campaigns are run, simple voting specifies that the final vote count will be 4-1 or 3-2 with the frontrunner winning, so that simple voting is indeed a best response for uninformed voters. After a 3-1 or 4-0 informed vote lead, the election is decided and simple voting is also a best response for the remaining uninformed voters. □

In what follows, I build on the preceding results and consider feasible deviations for each voter, given an any relevant continuation funding strategy for the SIG.

Fifth voter:

The fifth voter is either pivotal or irrelevant. Whenever a voter is pivotal, it is a strict best response for him to vote his signal. The probability of getting it right is $q > 1 - q$.

Fourth voter:

It is also true that the fourth voter will only be relevant if he is pivotal, i.e. if the vote is 1-2 or 2-1. Otherwise, the election is already decided.

Suppose the informative vote count is 2-1.

If voter 4 receives a signal a, voting his signal leads to a win by candidate A having received 3 positive signals. The worst outcome for A in terms of signals from this point on is 3-2, where A is the state of the world with prob. q>1-q.

If voter 4 receives a signal b, voting his signal will lead to the correct candidate being chosen with prob. q. Voting for A leads to A winning the election which will be the correct choice with prob. $\frac{1}{2} < q$.

If the vote is 1-2, the arguments are symmetric.

If the informative vote count is 1-1, 2-0, or 0-2 (i.e. one of the first three voters did not receive an informative signal), voter 4 will receive an informative signal only if it is the last one of the election by Lemma 51. Therefore, voter 4 is either pivotal (after 1-1) or irrelevant (after 2-0).

Third Voter:

Voter 3 can receive an informative signal when the informative voting has been 1-1, 2-0, or 0-2.

If the vote is 2-0 and the voter receives a signal a, state A will have received three positive signals and is the correct choice with probability at least q. Similarly, if the vote is 0-2 and voter 3 receives a b signal, it is a strict best response for him to vote for B and end the election.

That leaves scenarios in which the signal count, including 3's signal, is 2-1 or 1-2.

Suppose the signal count is 2-1.

The probability of a correct outcome if 3 votes his signal and all future campaigns are financed is:

$q(1-(1-q)^2)+(1-q)q^2 = -q^2(2q-3)$

The probability of a correct outcome if 3 votes B and all future campaigns are financed is:

$q^3+(1-q)(1-(1-q)^2) = q(2q^2-3q+2)$

Clearly, $q(1-(1-q)^2)+(1-q)q^2 > q^3+(1-q)(1-(1-q)^2)$ since

$-q(2q-3)-(2q^2-3q+2) = -4q^2+6q-2$

The first derivative of this difference is: $6-8q$

$-4q^2+6q-2$ has roots at 1 and .5, so the two expressions are equal at .5 and 1, while the derivative of the expression is positive at .5 and negative after 3/4, meaning that the expression is positive for all $q \in [.5, 1]$.

If the campaign ends with 3's vote, he is pivotal and voting his signal is a strict best response by the arguments made above.

Second Voter:

Voter 2 necessarily inherits either a 1-0 or 0-1 vote count. Therefore, if he receives an informative signal, Claim 49 confirms that his signal count is either 2-0 or 1-1.

If the SIG will finance the third campaign regardless of 2's vote, a deviation by 2 is equivalent to a deviation by 3, which we have seen above is never profitable.

If the SIG will stop funding the trailer if the vote goes to 2-0, then one must verify directly that a deviation is not profitable.

Let the signal count be 2-0. If 2 votes his signal, A is elected which is the correct choice with probability $\frac{q^2}{q^2+(1-q)^2}$.

If 2 deviates making the vote count 1-1, the correct choice will be made with prob. at most:

$\frac{q^2}{q^2+(1-q)^2}\left(q^2\left(1+2(1-q)\right)\right)+\frac{(1-q)^2}{q^2+(1-q)^2}\left(q^2\left(1+2(1-q)\right)\right)=\left(q^2\left(1+2(1-q)\right)\right)=-q^2\left(2q-3\right)$

$\frac{q^2}{q^2+(1-q)^2}+q^2\left(2q-3\right)=2q^2\left(2q-1\right)\frac{(q-1)^2}{2q^2-2q+1}>0$

Which is clearly less than $\frac{q^2}{q^2+(1-q)^2}$ since $2q<3$.

Now, let the signal count be 1-1. If 2 deviates, he will end the election with a choice that is correct with prob. .5. If he votes his signal, the campaign continues and, because other voters are voting informatively and the SIG is willing to fund campaigns so that a correct decision can be made with probability at least q, the correct outcome will be chosen with prob. at least q>.5.

First Voter:

If voter 1's vote will decide the election, it is a strict best response for him to vote his signal, as argued above.

If this is not the case, the campaign will continue regardless of his vote. Thus, a deviation by 1 will have the same effects as if voter 2 voted first (informatively) and then voter 1 deviated. We have seen in the step above that deviations by the second voter are not profitable.

For electoral calendars which are not strictly sequential, the calculus of deviations is very similar.

Lemma 52. *If no deviation by voters is profitable in a sequential election for any relevant funding strategy, no deviation is profitable for voters under any electoral calendar.*

Proof. When voting at the same date as other voters, the voter must consider a weighted average of the effect of his deviation conditional on the vote of those voters who are voting at the same time. Because, as I have shown, a deviation is never profitable, this is a weighted average of negative numbers and, thus, itself negative. Therefore, simple voting is always an equilibrium. □

Finally, I must make clear that it candidates will always campaign when they have the funds to do so. This is easily enforced by off-equilibrium beliefs on the part of the voters that only an $l-$type would avoid campaigning. That makes not campaigning equivalent to losing the election if it is done when the election is still in play. These beliefs are not arbitrary. One may argue that it is more likely that an $l-$type candidate has more to gain (or hide) by not campaigning than an $h-$type, since he will go on to win the election with lower probability. Therefore, when one sees such a deviation from equilibrium play, it may be considered infinitely more likely to have come from an $l-$type. This is an application of the logic behind the Divinity refinement used in signalling games (Banks and Sobel 1987). If the election has been won, the winning candidate may continue to spend campaign funds as he does not value other uses for these funds.

Proof of Proposition 39

I begin by explaining the mechanics of a bandwagon in this model. Under simple equilibria, when the SIG stops making informational donations, it continues to make service-motivated donations to the frontrunner. The frontrunner, in turn, will campaign for the remaining voters. This means that all remaining voters will receive positive signals about the frontrunner. They will vote for him even though they know that the signals are not informative.

This particular series of events, specified by simple equilibrium, is not necessary for bandwagons to form. Rather, I focus on it because it allows a particularly simple specification of equilibrium strategies. If voters are aware that no more informative signals will be sent, it is a best response for them to vote for the frontrunner in order to assure his victory, as the frontrunner will finish the election with the highest posterior probability of being the h-type.

For the SIG to fund the first campaign, it must be that $c < \bar{c} = q - \frac{1}{2}$.

If the SIG stops funding after one informed vote has been cast, the frontrunner will win all remaining votes and thus a bandwagon will be trivially observed.

Suppose that one of the candidates has a 2-0 lead in informed votes in a sequential election. The SIG's posterior belief about the probability that the leading candidate is the correct choice is $\frac{q^2}{q^2+(1-q)^2}$. That is also the probability of the correct candidate winning the election if the SIG funds no further elections. If the SIG does continue to fund campaigns, it must be willing to do so until a candidate reaches 3 votes. This increases the probability of electing the correct candidate to:

$$\frac{q^2}{q^2+(1-q)^2}\left(q+(1-q)\,q+(1-q)^2\,q\right)+\frac{(1-q)^2}{q^2+(1-q)^2}q^3=\frac{q^3}{2q^2-2q+1}\left(2q^2-5q+4\right)$$

So that the increase in the probability of electing the correct candidate is:

$$\frac{(1-q)^2}{q^2+(1-q)^2}q^3-\frac{q^2}{q^2+(1-q)^2}\left(1-q\right)^3=q^2\left(2q-1\right)\frac{(q-1)^2}{2q^2-2q+1}$$

This strategy brings with it an additional cost of:

$$c\left(\frac{q^2}{q^2+(1-q)^2}q+\frac{(1-q)^2}{q^2+(1-q)^2}\left(1-q\right)\right)+2c\left(\frac{q^2}{q^2+(1-q)^2}\left(1-q\right)q+\frac{(1-q)^2}{q^2+(1-q)^2}q\left(1-q\right)\right)+$$
$$3c\left(\frac{q^2}{q^2+(1-q)^2}\left(1-q\right)^2+\frac{(1-q)^2}{q^2+(1-q)^2}q^2\right)=\frac{c}{2q^2-2q+1}\left(2q^4-4q^3+3q^2-q+1\right)$$

Therefore, the SIG will fund voter 3 after a 2-0 start if:

$$c<B\frac{q^2(2q-1)\frac{(q-1)^2}{q^2+(1-q)^2}}{\frac{1}{q^2+(1-q)^2}(2q^4-4q^3+3q^2-q+1)}=Bq^2\left(2q-1\right)\frac{(q-1)^2}{2q^4-4q^3+3q^2-q+1}=c_{\text{seq}}$$

Therefore, the frontrunner will win all remaining votes in the primary after a 2-0 or 1-0 start whenever $c \in (c_{\text{seq}}, \bar{c})$.

Proof of Lemma 43.

I begin by listing all possible electoral calendars in a five voter election. I then prove through a series of Claims that we may focus on only three. Given Theorem 37, I assume throughout that voters vote simply. This allows me to focus on the SIG's funding decisions.

In what follows, I use the following notation: brackets signal that I am referring to a calendar, I separate the dates at which voting takes place with a dash and write the number of voters who vote at each date. Thus, the sequential electoral calendar is {1-1-1-1-1} while the simultaneous calendar is {5}. A question mark stands in for all possible variations of the missing values. For instance, {3-?}=[{3-1-1},{3-2}]. Numbers separated by a dash, but not in brackets denote a partial vote count. For example 2-1 means that one candidate has a 2 vote to 1 vote lead over his competitor.

(1) Simultaneous {5}.

(2) Sequential {1-1-1-1-1}.

(3) Super Tuesday {3-1-1}.

(4) {2-3}

(5) {3-2}

(6) {4-1}

(7) {1-4}

(8) {2-2-1}

(9) {2-1-2}

(10) {1-2-2}

(11) {1-1-3}

(12) {1-3-1}

(13) {1-1-1-2}

(14) {1-1-2-1}

(15) {1-2-1-1}

(16) {2-1-1-1}

Lemma 53. *One of the following electoral calendars is voter-optimal among all possible calendars: sequential, simultaneous, or Super Tuesday.*

In order to prove this Lemma, I establish a series of facts which will, together, make the result clear.

Claim 54. *Any calendar in which only voter 1 votes at date 1 is strategically equivalent to a calendar identical except that voter 1 votes at time 2 with the following block of voters.*

Proof. It is clear that the result of the first vote reveals no new information to the donor, that is, the donor knows that the election will either be 1-0 or 0-1 after one informative vote. Because of symmetry, these two situations are strategically equivalent. Therefore, the funding decision of the second voter's campaign will be the same regardless of whether the donor can condition on the outcome of the first vote. □

This allows us to ignore calendars 7, 10, 11, 12, 13, 14, 15, and 16 and look only at 1, 5, 4, 6, 9, 8, 3, and 2 respectively (or vice-versa).

Claim 55. *Any calendar in which voters 1 through 3 vote at different dates than voters 4 and 5, who vote simultaneously (calendars 5, 9, 10, and 13), is weakly dominated by a calendar identical to it but in which voters 4 and 5 vote sequentially (calendars 3, 16, 15, and 2).*

Proof. If voters 1 through 3 are funded, the election will either be 2-1 or 3-0. Only the first case is relevant since the election is over if it is 3-0. If the last two voters vote simultaneously, the SIG will either fund both or neither since only two votes against the front-runner can change the result. Whenever c is such that the SIG funds both of the last two voters, it will also fund voter 4 in an election ending in {-1-1}, and fund voter 5 if needed. This is because the expected benefit of both strategies is the same, but the expected cost is strictly lower in the sequential case.

As observed above, if the SIG will fund any campaigns at all, it will fund voter 1's. Furthermore, if the SIG intends to fund more than one campaign, it should fund voter 2's since voter 1's vote does not provide new information useful for future funding decisions (the election will be 1-0 either way). If the SIG does not fund the third campaign, it may choose to fund one additional campaign only if the election is tied 1-1 (a 2-0 lead cannot be overcome by two votes, and 2-1 lead cannot be overcome by one vote).

Therefore, the only remaining question is whether voter 3's campaign funding could be adversely affected by having voters 4 and 5 vote sequentially rather than simultaneously. Suppose first that the election is tied after the first two voters have gone to the polls. The election will be 2-1 after an informative 3rd vote. If the SIG does not fund voter 3's campaign, it will fund at most one more campaign, but in this case it may as well fund voter 3's. If the election is 2-0 after two votes, it is only optimal to fund further campaigns if the SIG is prepared to fund campaigns until a candidate reaches a 3 vote majority. This can be accomplished more cheaply if voters 4 and 5 vote sequentially. If voter 3 votes at the same time as voter 2, the calculus involves an odds weighted average of these two scenarios, so the conclusions continue to hold. Thus, the SIG's funding decisions are more likely to lead to a correct decision if the final two voters vote sequentially. □

Claim 56. *The calendar {4-1} (no. 6) (and thus {1-3-1} (no. 12)) is dominated by Super Tuesday {3-1-1}.*

Proof. In all cases, the 5th voter campaign, when considered independently, will only be funded if the voter is strongly pivotal (i.e. if the vote total is tied).

If all four voters in the first block of {4-1} are funded, it must be that the donor would fund the fourth voter conditional on the election being 2-1 since it will either be 2-1 or 3-0, in which case the election is over. Therefore, if the fourth voter is funded in a {4-1}, it is also funded in a {3-1-1} when the election is still in play.

In a {4-1}, the SIG will never fund only 3 date-1 campaigns. Funding three voters in the first block of a {4-1} means that voter 5 will never be funded because voter 5 is funded only if the election is tied, which is impossible when an odd number of informative votes have been cast thus far. Moreover, if the SIG funds 2 date-1 campaigns, he can make the funding decision for the third (voter 5) after conditioning on the outcome of the first two (i.e. fund it only if the informative vote count is 1-1 and not 2-0). Therefore, it is strictly better for the SIG to fund 2 campaigns on date 1 and then fund voter 5 if the informative vote count is tied, thus giving the same probability of success at a strictly lower expected cost.

If only two campaigns are funded in the first block of {4-1}, at least two will be funded in a {3-1-1}. In both cases, only one additional campaign may be funded: if the election is 2-0 after the first block, the lead cannot be overcome, if it is 1-1 one additional vote will make it 2-1 and the last vote cannot overcome that lead. Therefore, if it is optimal to fund the two voters in the {4-1} it is also optimal to do so in {3-1-1}. □

The following Claim shows that the sequential calendar dominates any calendar beginning {1-1-?} or {2-?}.

Claim 57. *Any calendar in which voters 1 and 2 vote at different dates than voters 3, 4 and 5 is dominated by a calendar identical except that voters 3, 4 and 5 vote sequentially.*

Proof. By Claim 49, voter 1's campaign will be funded whenever the comparison of these calendars is in question. Because the SIG can condition its choice on the outcome of 2's vote after it has taken place, if the SIG is going to fund more than one campaign, it is optimal for it to fund voter 2's. Therefore, I compare calendars conditional on two informative votes having been cast.

Suppose a candidate has a 2-0 lead after voters 1 and 2 have voted. Then, the SIG will only fund further campaigns if it is willing to fund campaigns until one candidate has received three favorable informed votes. This may be done at a lower expected cost when the final three voters vote sequentially because the SIG can choose to stop funding as soon as one candidate reaches 3 votes. Therefore, having the final three voters vote sequentially dominates all other arrangements of the last three voters conditional on the first two voters voting informatively for the same candidate.

Now suppose the election is tied after voters 1 and 2 have gone to the polls. The SIG will fund voter 3's campaign since the cost of previous campaigns is sunk and it was willing to fund the campaign of voter 1. One candidate will have a 2-1 lead after voter 3's vote. Because of the symmetry of the game, it does not matter which candidate it is for the SIG's funding decision and therefore a calendar in which voters 3 and 4 vote simultaneously is strategically identical to one in which they vote sequentially. By Claim 55, if the first three campaigns have been funded, the calendar with voters 4 and 5 voting sequentially dominates the one in which they vote simultaneously. □

Application of these Claims leaves us with three contenders for the voter-optimal electoral calendar: sequential, simultaneous, and Super Tuesday. The simultaneous calendar is dominated by the Super Tuesday calendar. However, because of its special role in the literature, I will examine it more closely than other dominated calendars. I include the proof of this dominance relation in the following proof.

Proof of Theorem 44 and Proposition 45

Theorem 58. *There exist values $c_{sim} < c_{\text{seq}} < c_{st} < \hat{c}_{\text{seq}} < \bar{c}$ such that:*

- *The Super Tuesday calendar dominates all other calendars, and strictly dominates the sequential calendar, when $c \in (c_{\text{seq}}, c_{st})$.*
- *The sequential calendar dominates all other calendars, and strictly dominates the Super Tuesday calendar, when $c \in (c_{st}, \hat{c}_{\text{seq}})$.*
- *The simultaneous calendar weakly dominates all other calendars when $c < c_{sim}$ or $c > \bar{c}$.*

Proof. Funding one (or two) campaigns results in the h-type winning the nomination with probability q. The first voter will be funded under any electoral calendar if $c < \Delta b \left(q - \frac{1}{2}\right)$.

In a simultaneous election, funding three (or four) campaigns leads to selecting the correct candidate with probability:

$q^2 \left(q + 3\left(1 - q\right)\right) = q^2 \left(3 - 2q\right)$

The increase in the probability of selecting the h-type resulting from funding three campaigns rather than one is:

$q^2(3-2q) - q = -q(2q^2 - 3q + 1) = -2q^3 + 3q^2 - q > 0$

The additional cost, given that one campaign is being funded, is $2c$. Therefore, the SIG will fund three campaigns if:

$c < \Delta b \frac{1}{2}(-2q^3 + 3q^2 - q)$

Funding five campaigns leads to selecting the correct candidate with probability:

$q^3\left(q^2 + 5q(1-q) + 10(1-q)^2\right) = q^3(10 - 15q + 6q^2)$

Subtracting the first expression from the second, we get the increase in probability of success from funding voters 4 and 5:

$q^3(10 - 15q + 6q^2) - q^2(3-2q) = 3q^2(2q-1)(q-1)^2 = 6q^5 - 15q^4 + 12q^3 - 3q^2$

The difference in cost between the two funding strategies is $2c$, so the SIG will fund all five campaigns if:

$c < \frac{1}{2}\Delta b(6q^5 - 15q^4 + 12q^3 - 3q^2) = c_{sim}$

In a Super Tuesday election, funding all three date-1 campaigns leads to selecting the right candidate with probability:

$q^3\left[1 + 3(1-q) + 6(1-q)^2\right] = 10q^3 - 15q^4 + 6q^5$

at an expected cost of:

$3c + c\left(1 - q^3 - (1-q)^3\right) + c\left(6q^2(1-q)^2\right) = 3c(1 + q + q^2 - 4q^3 + 2q^4)$

Funding only two campaigns in the first block leads to selecting the right candidate with probability:

$q^2(1+2(1-q)) = q^2(3-2q)$

at an expected cost of:

$2c + c\left(1-q^2-(1-q)^2\right) = 2c(1+q(1-q)) < 3c$

Note that the SIG will be willing to fund this strategy for higher c than it is to fund three campaigns in a simultaneous election because the difference in cost from funding only one campaign to following this strategy is $c+2cq(1-q) < 2c$, while the benefits of the change are the same.

The increase in the probability of nominating an h-type from funding voter 3's campaign is:

$6q^5 - 15q^4 + 12q^3 - 3q^2$

Subtracting the expected cost of the fund 2 strategy from that of the fund 3 I find the difference in expected cost:

$3c(1+q+q^2-4q^3+2q^4) - 2c(1+q(1-q)) = c(6q^4-12q^3+5q^2+q+1) = c(1+q(1-q)(1+6q(1-q)))$

Therefore, the SIG will fund all 3 date-1 campaigns in a Super Tuesday election if:

$c < \Delta b \frac{6q^5-15q^4+12q^3-3q^2}{1+q(1-q)(1+6q(1-q))} = c_{st}$

Because $q(1-q)$ reaches a maximum for $q \in (0,1)$ at $\frac{1}{4}$,

$q(1-q)(1+6q(1-q)) \leq \frac{5}{8}$ and therefore $c_{st} > c_{sim}$.

Funding only one campaign leads to selecting the h-type with probability q. The increase in the probability of success from this strategy to funding two date-1 voters in a Super Tuesday election is:

$q^2(3-2q)-q=-q(2q^2-3q+1)=-2q^3+3q^2-q>0$

The increase in cost from funding one campaign to funding two date-1 campaigns in a Super Tuesday election is:

$c+2cq(1-q)<2c$

Therefore, at least two date-1 campaigns will be funded in a Super Tuesday election if:

$c<\Delta b\left(\frac{-2q^3+3q^2-q}{c+2cq(1-q)}\right)=\bar{c}$

Suppose that one of the candidates has a 2-0 lead in informed votes in a sequential election. The SIG's posterior belief about the probability that the leading candidate is the correct choice is $\frac{q^2}{q^2+(1-q)^2}$. That is also the probability of the correct candidate winning the election if the SIG funds no further elections. If the SIG does continue to fund campaigns, it must be willing to do so until a candidate reaches 3 votes. This increases the probability of electing the correct candidate to:

$\frac{q^2}{q^2+(1-q)^2}\left(q+(1-q)q+(1-q)^2q\right)+\frac{(1-q)^2}{q^2+(1-q)^2}q^3=\frac{q^3}{2q^2-2q+1}(2q^2-5q+4)$

So that the increase in the probability of electing the correct candidate is:

$\frac{q^3}{2q^2-2q+1}(2q^2-5q+4)-\frac{q^2}{q^2+(1-q)^2}=q^2(2q-1)\frac{(q-1)^2}{2q^2-2q+1}$

or,

$\frac{(1-q)^2}{q^2+(1-q)^2}q^3-\frac{q^2}{q^2+(1-q)^2}(1-q)^3=q^2(2q-1)\frac{(q-1)^2}{2q^2-2q+1}$

This strategy brings with it an additional cost of:

$c+c\left(\frac{q^2}{q^2+(1-q)^2}(1-q)+\frac{(1-q)^2}{q^2+(1-q)^2}q\right)+c\left(\frac{q^2}{q^2+(1-q)^2}(1-q)^2+\frac{(1-q)^2}{q^2+(1-q)^2}q^2\right)$

$=\frac{c}{q^2+(1-q)^2}(2q^4-4q^3+3q^2-q+1)$

or,

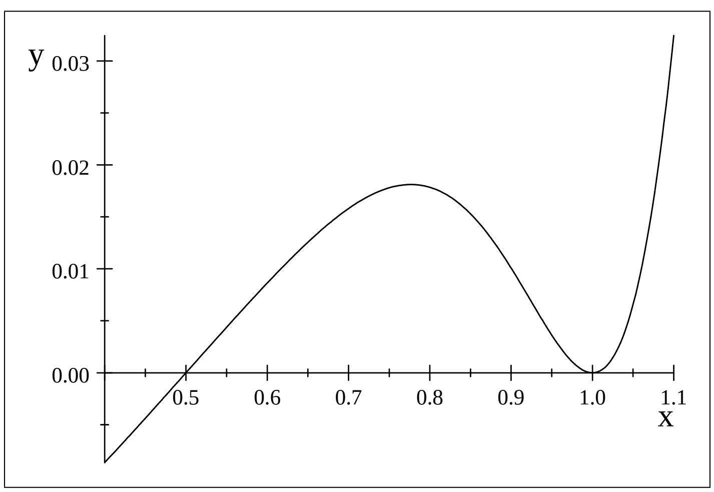
$$c\left(\frac{q^2}{q^2+(1-q)^2}q+\frac{(1-q)^2}{q^2+(1-q)^2}(1-q)\right)+2c\left(\frac{q^2}{q^2+(1-q)^2}(1-q)\,q+\frac{(1-q)^2}{q^2+(1-q)^2}q\,(1-q)\right)$$
$$+3c\left(\frac{q^2}{q^2+(1-q)^2}(1-q)^2+\frac{(1-q)^2}{q^2+(1-q)^2}q^2\right)=\frac{c}{2q^2-2q+1}\left(2q^4-4q^3+3q^2-q+1\right)$$

Therefore, the SIG will fund voter 3 after a 2-0 start if:

$$c<\Delta b\frac{q^2(2q-1)\frac{(q-1)^2}{q^2+(1-q)^2}}{\frac{1}{q^2+(1-q)^2}(2q^4-4q^3+3q^2-q+1)}=\Delta bq^2\left(2q-1\right)\frac{(q-1)^2}{2q^4-4q^3+3q^2-q+1}=c_{\text{seq}}$$

If $c_{\text{seq}} < c_{st}$, then there will be circumstances under which a Super Tuesday calendar outperforms a sequential calendar. This is because, when $c \in (c_{\text{seq}}, c_{st})$, the Super Tuesday calendar will continue to fund campaigns when they start 2-0 and go to 2-1, while with the sequential calendar funding would stop at 2-0. The Super Tuesday calendar takes advantage of the uncertainty about whether the election will start 2-0 or 1-1. Because the two calendars are identical after voter 3, this advantage is the only difference in this range. It is difficult to verify algebraically that $c_{\text{seq}} < c_{st}$, but straight forward to do so numerically as we need only check that the inequality holds for values of q in $\left(\frac{1}{2}, 1\right)$:

$$c_{st}-c_{\text{seq}}=\frac{6q^5-15q^4+12q^3-3q^2}{1+q(1-q)(1+6q(1-q))}-q^2\left(2q-1\right)\frac{(q-1)^2}{2q^4-4q^3+3q^2-q+1}>0$$



In a sequential election, if three campaigns have been financed leading to a 2-1 vote lead by one of the candidates, continuing to fund campaigns makes sense for the SIG only if it is willing to fund until one candidate has three votes. This leads to electing the correct candidate with probability $q^2(3-2q)$, while stopping funding now means the frontrunner will win the election, which is the correct choice with probability q. The increase in the probability of selecting the correct candidate is therefore:

$$q^2(3-2q) - q = -q(2q^2 - 3q + 1)$$

This strategy leads to additional expected costs of

$$c + c(2q(1-q)) = c(-2q^2 + 2q + 1).$$

Or if we derive it differently:

$$c\left(q^2 + (1-q)^2\right) + 4cq(1-q) = c(-2q^2 + 2q + 1).$$

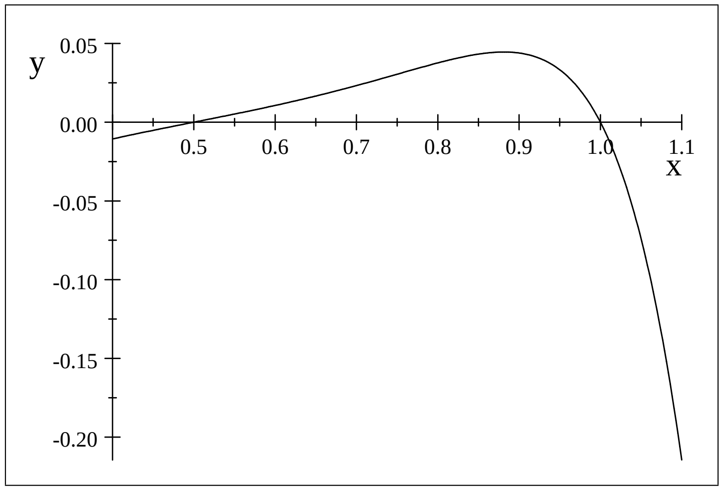
Therefore, the SIG will continue this funding if:

$$c < \Delta b \tfrac{-q\left(2q^2-3q+1\right)}{-2q^2+2q+1} = \hat{c}_{\text{seq}}$$

On the other hand, if the SIG funds only two date-1 campaigns in a Super Tuesday election it will fund a third if the election is tied 1-1 in informed votes after the first block has voted, but will never fund more than that because a 2-1 lead which would ensue could never be overcome by a single informed vote. Therefore, there may be a range of costs, $c \in (c_{st}, \hat{c}_{\text{seq}})$, over which the sequential calendar strictly overperforms the Super Tuesday calendar in expected terms.

$$\hat{c}_{\text{seq}} - c_{st} = \tfrac{-q\left(2q^2-3q+1\right)}{-2q^2+2q+1} - \tfrac{6q^5-15q^4+12q^3-3q^2}{1+q(1-q)(1+6q(1-q))} > 0$$

Again, I verify this inequality numerically.


y
0.05
0.00
-0.05
-0.10
-0.15
-0.20
0.5
0.6
0.7
0.8
0.9
1.0
1.1
x

□

References

[1] Abreu, D., D. Pearce, and E. Stachetti. 1990. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica* 58(5): 1041-63.

[2] Aldrich, J. H. 1980. *Before the Convention: Presidential Nomination Campaigns.* University of Chicago Press.

[3] Aldrich, J. H. 1980. "A Dynamic Model of Presidential Nomination Campaigns." *The American Political Science Review* 74(3): 651-669.

[4] Ali, S.N. and N. Kartik. 2008. "Social Learning in Elections." Unpublished manuscript.

[5] Ansolabehere, S., J.M. de Figueiredo, and J.M. Snyder. 2003. "Why Is There So Little Money in U.S. Politics?" *The Journal of Economic Perspectives* 17(1): 105-130.

[6] Ashworth, S. 2005. "Reputational Dynamics and Political Careers." *Journal of Law, Economics, & Organization* 21(2): 441-466.

[7] Ashworth, S. 2006. "Campaign Finance and Voter Welfare with Entrenched Incumbents." *The American Political Science Review* 100(1): 55-68.

[8] Ashworth, S. 2008. "The Economics of Campaign Finance." in S.N. Durlauf and L.E. Blume, editors, *The New Palgrave Dictionary of Economics.* Palgrave Macmillan, 2nd edition.

[9] Ashworth, S. and E. Bueno de Mesquita. 2006. "Delivering the Goods: Legislative Particularism in Different Electoral and Institutional Settings" *The Journal of Politics* 68: 168-179.

[10] Ashworth, S. and E. Bueno de Mesquita. 2008. "Electoral Selection, Strategic Challenger Entry, and the Incumbency Advantage." *The Journal of Politics* 70: 1006-1025.

[11] Ashworth, S., E. Bueno de Mesquita and A. Friedenberg. 2009. "Sanctioning and Learning in Elections." Unpublished manuscript.

[12] Bagnoli, M. and T. Bergstrom. 2005. "Log-concave Probability and its Applications." *Economic Theory* 26: 445-469.

[13] Banks, J. and J. Duggan. 2006. "A Dynamic Model of Democratic Elections in Multidimensional Policy Spaces." Unpublished manuscript.

[14] Banks, J. S. and D. R. Kiewiet. 1989. "Explaining patterns of candidate competition in congressional elections." *American Journal of Political Science* 33: 997-1015.

[15] Banks, J.S. and J. Sobel 1987. "Equilibrium Selection in Signalling Games." *Econometrica* 55(3): 647-661.

[16] Banks, J. and R. Sundaram. 1993. "Adverse Selection and Moral Hazard in a Repeated Elections Model." In *Political Economy: Institutions, Information, Competition, and Representation*, eds. W. Barnett, M.J. Hinich and N. Schofield. New York, Cambridge University Press.

[17] Banks, J. and R. Sundaram 1998. "Optimal Retention in Agency Problems." *Journal of Economic Theory* 82: 293-323.

[18] Barro, R. 1973. "The Control of Politicians: An Economic Model." *Public Choice* 14: 19-42.

[19] Bartels, L.M. 1988. *Presidential Primaries and the Dynamics of Public Choice.* Princeton University Press.

[20] Battaglini, M. 2005. "Sequential Voting with Abstention." *Games and Economic Behavior* 51(2): 445-463.

[21] Berganza, J. C. 2000. "Two Roles for Elections: Disciplining the Incumbent and Selecting a Competent Candidate." *Public Choice* 105: 165-193.

[22] Besley, T. 2006. *Principled Agents? The Political Economy of Good Government.* New York: Oxford University Press.

[23] Besley, T. 2005. "Political Selection." *Journal of Economic Perspectives* 19: 43-60.

[24] Besley, T. and S. Coate. 1997. "An Economic Model of Representative Democracy." *Quarterly Journal of Economics* 112: 85-114.

[25] Besley, T., R. Pande and V. Rao. 2006. "Political Selection and the Quality of Government: Evidence from South India." mimeo, London School of Economics.

[26] Bliss, C. and B. Nalebuff. 1984. "Dragon-slaying and Ballroom Dancing: The Private Supply of a Public Good." *Journal of Public Economics* 25: 1-12.

[27] Bobonis, G. and L.R. Cámara-Fuerte 2009. "Does Exposing Corrupt Politicians Reduce Corruption?" Unpublished manuscript.

[28] Brekke, K.A., S. Kverndokk, and K. Nyborg. "An Economic Model of Moral Motivation." *Journal of Public Economics* 87(9-10): 1967-1983.

[29] Brennan, G. and L. Lomasky. 1993. *Democracy and Decision: The Pure Theory of Electoral Preference.* Cambridge University Press.

[30] Brown, C.W., L.W. Powell, and C. Wilcox. 1995. *Serious Money: Fundraising and Contributing in Presidential Nomination Campaigns.* Cambridge University Press.

[31] Cain, B., J. Ferejohn, and M. Fiorina 1990. *The Personal Vote.* Cambridge MA, Harvard University Press.

[32] Callander, S. 2007. "Bandwagons and Momentum in Sequential Voting." *Review of Economic Studies* 74(3): 653-684.

[33] Carrillo, J.D. and T. Mariotti. 2001. "Electoral Competition and Political Turnover." *European Economic Review* 45: 1-25.

[34] Casselli, F. and M. Morelli. 2004. "Bad Politicians." *Journal of Public Economics* 88: 759-782.

[35] Canon, D. T. 1993. "Sacrificial Lambs or Strategic Politicians? Political Amateurs in U.S. House Elections." *American Journal of Political Science* 37: 1119-1141.

[36] Coate, S. 2004. "Pareto Improving Campaign Finance Policy." *The American Economic Review* 94(3): 628-684.

[37] Coate, S. and M. Conlin. 2004. "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence." *The American Economic Review* 94(5): 1476-1504.

[38] Coate, S. and S. Morris 1995. "On the Form of Transfers to Special Interests" *The Journal of Political Economy* 103(6): 1210-1235.

[39] Damore, D.F. 1997. "A Dynamic Model of Candidate Fundraising: the Case of Presidential Nomination Campaigns." *Political Research Quarterly* 50(2): 343-364.

[40] Degroot, M. 2004. *Optimal Statistical Decisions.* Wiley-Interscience.

[41] Dekel, E. and M. Piccione. 2000. "Sequential Voting Procedures in Symmetric Binary Elections." *The Journal of Political Economy* 108(1): 34-55.

[42] Dominguez-Martinez, S., O. Swank, and B. Visser. 2008. "In Defense of Boards" *Journal of Economics & Management Strategy* 17(3): 667-682.

[43] Duggan, J. 2000. "Repeated Elections with Asymmetric Information." *Economics and Politics* 12(2): 109-135.

[44] Farrell, J. and E. Maskin 1989. "Renegotiation in Repeated Games." *Games and Economic Behavior* 1: 327-360.

[45] Fearon, J. 1999. "Electoral Accountability and the Control of Politicians: Selecting Good Candidates versus Sanctioning Poor Performance." In *Democracy, Accountability, and Representation*, eds. A. Przeworski, S.C. Stokes and B. Manin. New York, Cambridge University Press.

[46] Feddersen, T. and A. Sandroni. 2006. "Ethical Voters and Costly Information Acquisition." *Quarterly Journal of Political Science* 1(3): 287-311.

[47] Feddersen, T. and A. Sandroni. 2006. "A Theory of Participation in Elections." *The American Economic Review* 96(4): 1271-1282.

[48] Feddersen, T. and W. Pesendorfer. 1996. "The Swing Voter's Curse." *The American Economic Review* 86(3): 408-424.

[49] Ferejohn, J. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50: 5-25.

[50] Galasso, V., M. Landi, A. Merlo and A. Matozzi 2009. "The Labor Market of Italian Politicians." Unpublished manuscript.

[51] Gallego M. and C. Pitchik (2004) "An Economic Theory of Leadership Turnover." *Journal of Public Economics* 88 (12), 2361–2382.

[52] Gershkov, A. and B. Szentes. 2009. "Optimal Voting Schemes with Costly Information Acquisition." *Journal of Economic Theory* 144(1): 36-68.

[53] Glicksberg, I.L. 1952. "A Further Generalization of the Kakutani Fixed Point Theorem, with Application to Nash Equilibrium Points." *Proceedings of the American Mathematical Society* 3: 170-174.

[54] Goff, M.J. 2004. *The Money Primary: The New Politics of the Early Presidential Nomination Process.* Rowman & Littlefield Publishers.

[55] Gordon, S.C., C. Hafer, and D. Landa. 2007. "Consumption or Investment? On Motivations for Political Giving." *The Journal of Politics* 69(4): 1057-1072.

[56] Hamilton, A., J. Jay and J. Madison, 1987, *The Federalist Papers.* (Penguin Books, New York).

[57] Harsanyi, J.C. 1980. "Rule Utilitarianism, Rights, Obligations and the Theory of Rational Behavior." *Theory and Decisions* 12(2): 115-133.

[58] Haynes, A.A., P.H. Gurian, and S.M. Nichols. 1997. "The Role of Candidate Spending in Presidential Nomination Campaigns." *The Journal of Politics* 59(1): 213-225.

[59] Hinckley, K.A. and J.C. Green. 1996. "Fund-raising in Presidential Nomination Campaigns: The Primary Lessons of 1988." *Political Research Quarterly* 49(4): 693-718.

[60] Holmström, B. 1999. "Managerial Incentive Problems: A Dynamic Perspective." *Review of Economic Studies* 66: 169-182.

[61] Hörner, J. 2002. "Reputation and Competition." *Journal of Political Economy* 92: 644-663.

[62] Jefferson, T., 1998, *The life and selected writings of Thomas Jefferson.* Koch, A. and W. Peden eds. (Random House,. New York).

[63] Kessel, J.H. 1992. *Presidential Campaign Politics.* Brooks-Cole, 4th edition.

[64] Key, V. O., Jr. 1966. *The Responsible Electorate.* Cambridge MA, Harvard University Press.

[65] Klumpp, T. and M.K. Polborn. 2006. "Primaries and the New Hampshire Effect." *Journal of Public Economics* 90(6-7): 1073-1114.

[66] Kreps, D. and R. Wilson 1982. "Sequential Equilibria" *Econometrica* 50: 863-894.

[67] Le Borgne, E. and B. Lockwood. 2002. "Candidate Entry, Screening, and the Political Budget Cycle." IMF Working Paper No. 02/48.

[68] Levitt, S.D. 1996. "How Do Senators Vote? Disentangling the Role of Voter Preferences, Party Affiliation, and Senator Ideology" *American Economic Review* 86(3): 425-41.

[69] Lindbeck, A. and J. Weibull. 1987. "Balanced-budget Redistribution as the Outcome of Political Competition." *Public Choice* 52: 273–297.

[70] Lohmann, S. 1998. "An Information Rationale for the Power of Special Interests" *American Political Science Review* 92(4): 809-827.

[71] Londregan, J. and T. Romer, 1993, "Polarization, incumbency, and the personal vote." in W. Barnett, M.J. Hinich, and N.J. Schofield eds., *Political*

Economy: Institutions, Competition and Representation. Cambridge University Press.

[72] Lusternik, L.A. and V.I. Sobolev 1974. *Elements of Functional Analysis.* New York, John Wiley and Sons.

[73] Mailath, G.J. and L. Samuelson 2001. "Who Wants a Good Reputation?" *Review of Economic Studies* 68: 415-441.

[74] Maskin, E. and J. Tirole 2001. "Markov Perfect Equilibrium: I. Observable Actions." *Journal of Economic Theory* 100: 191-219.

[75] Matozzi, A. and A. Merlo. 2005. "Political Careers or Career Politicians." mimeo, University of Pennsylvania.

[76] Mayer, W.G. 1996. "Comment: Of Money and Momentum." *Political Research Quarterly* 49(4): 719-726.

[77] Mayer, W.G. 1996. *The Divided Democrats: Ideological Unity, Party Reform, and Presidential Elections.* Westview Press.

[78] Mayer, W.G. and A.E. Busch. 2004. *The Frontloading Problem in Presidential Nominations.* Brookings Institution Press.

[79] McCarty, N. and L.S. Rothenberg. 2000. "The Time to Give: PAC Motivations and Electoral Timing." *Political Analysis* 8(3): 239-259.

[80] Meirowitz, A. 2007. "Probabilistic Voting and Accountability in Elections with Uncertain Policy Constraints." *Journal of Public Economic Theory* 9(1): 41-68.

[81] Milgrom, P.R. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* 12(2): 380-391.

[82] Morton, R. and C. Cameron. 1992. "Elections and the Theory of Campaign Contributions: A Survey and Critical Analysis." *Economics & Politics* 4(1): 79-108.

[83] Norrander, Barbara. 2000. "The End Game in Post-Reform Presidential Nominations." *The Journal of Politics* 62(4): 999-1013.

[84] Pearce, D. 1987. "Renegotiation-Proof Equilibria: Collective Rationality and Intertemporal Cooperation." Cowles Foundation Discussion Paper No. 855.

[85] Pratt, A. 2002. "Campaign Advertising and Voter Welfare." *Review of Economic Studies* 69(4): 999-1017.

[86] Osbourne, M. J. and A. Slivinski. 1996. "A Model of Political Competition with Citizen Candidates." *Quarterly Journal of Economics* 111: 65-96.

[87] Palfrey, T. and H. Rosenthal. 1984. "Participation and Provision of Discrete Public Goods: a Strategic Analysis." *Journal of Public Economics* 24: 171-193.

[88] Persson, T. and G. Tabellini 2000. *Political Economics: Explaining Economic Policy.* Cambridge MA, MIT Press.

[89] Polo, M. 1998. "Electoral Competition and Political Rents." mimeo, Bocconi University.

[90] Reed, R.W. 1994. "A Retrospective Voting Model With Heterogeneous Politicians." *Economics and Politics* 6(1): 39-58.

[91] Rogoff, K. and A. Sibert. 1988. "Elections and Macroeconomic Policy Cycles." *Review of Economic Studies* 55: 1-16.

[92] Roumanias, C. 2005. "Signalling Through Political Campaigns: Elections as a Revelation Mechanism." *Economics and Politics* 17(3): 367-392.

[93] Smart, M. and D.M. Sturm (2006) "Term Limits and Electoral Accountability." CEPR Discussion Paper no. 4272.

[94] Smith, S.S. and M.J. Springer. 2009. *Reforming the Presidential Nomination Process.* Brookings Institution Press, 1st edition.

[95] Snyder, J.M. and M.M. Ting. 2008. "Interest Groups and the Electoral Control of Politicians." *Journal of Public Economics* 92: 482-500.

[96] Stratmann, T. 2005. "Some Talk: Money in Politics - A (Partial) Review of the Literature." *Public Choice* 124(1): 135-156.

[97] Wittman, D. A. 1983. "Candidate Motivation: A Synthesis of Alternative Theories." *American Political Science Review* 77: 142-157.