Electoral Control with Behavioral Voters

Daniel Diermeier* and Christopher Li†

September 9, 2015

Abstract

We present a model of electoral control with behavioral voters. The model captures two main regularities of voting behavior found in empirical studies: voters are forgetful and are influenced by extraneous events beyond the control of public officials (e.g. rainfall). Specifically, we assume the voters’ propensities to reelect the incumbent is governed by a stochastic reinforcement process instead of strategic reasoning. We study electoral control (i.e. public officials’ incentive exercise effort) in such an environment. We show that even in the context of low-rationality, electoral control of public officials can work well. However, the extent of control depends on the properties of the election and the electorate. Extraneous events that decrease a voter’s propensity to reelect the incumbent benefit electoral control, as the incumbent must exert greater effort to ensure reelection. Increasing the benefits of holding office also has a positive effect, while the degree of voter forgetfulness helps electoral control if and only if elections are held frequently.

JEL classification: D03, D72, D83

Keywords: electoral control, adaptive behavioral voting, electoral accountability

*Dean and Emmott Dedman Professor, Harris School of Public Policy, The University of Chicago, 1155 60th Street, Chicago, IL 60637, ddiermeier@uchicago.edu and Canadian Institute for Advanced Research, CIFAR Program on Institutions, Organizations, and Growth, CIFAR, Toronto, ON M5G 1Z8.
†Department of Economics, Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208, christopher.li@u.northwestern.edu.
1 Introduction

The question of electoral control of public officials is of central concern in political science. Some political theorists, e.g., Riker (1982; 9) following Madison (Federalist 39), have even argued that democracy consists of the control of public officials and little else. This has led to the question of how well electoral accountability actually works. The formal political science literature\(^1\) has focused on the effect of asymmetric information on electoral control.\(^2\) Some papers, like Barro (1973) and Ferejohn (1986), focus on moral hazard, where the actions ("effort") of the incumbent are unobservable and the electorate chooses a retrospective voting rule to induce high effort from the incumbent. Other approaches, like Rogoff (1990) and Ashworth (2005), consider the case of adverse selection, where the ability of the incumbent is unknown and the electorate needs to screen out low ability incumbents. In a typical model, the electorate and the incumbent interact in a dynamic game, and the prediction is based on Nash Equilibrium. One of the main insights of this literature is that some level of electoral accountability can be maintained under informational asymmetry, albeit at a cost to the voters.\(^3\)

In such models voters are assumed to be rational. They process information via Bayesian updating and act strategically in a game form that is assumed to be common knowledge. This approach has been heavily criticized by political scientists who empirically study voting behavior and public opinion. Concerns go back to some of the early studies of voters conducted by the Columbia and Michigan School (e.g. Berelson, Lazarsfeld and McPhee 1954, Campbell, Converse, Miller, and Stokes 1960).

In his summary of post-war public opinion research, Stimson states (Stimson 2004; p. 13)

> What those studies found was that ordinary Americans knew almost nothing about public affairs and appeared to care about issues as much as they knew: almost not at all. Their beliefs were a scattering of unrelated ideas, often mutually contradictory. Structure was nowhere to be found.

---

\(^1\)See Ashworth (2012) for a recent overview.

\(^2\)Under perfect information, the voters can trivially achieve complete electoral control by conditioning reelection on the implementation of (voter) welfare maximizing policy.

\(^3\)In Rogoff (1990), for example, the voters can distinguish high ability incumbent from low ability incumbent in a separating equilibrium, although the incumbent will implement inflationary fiscal policy, which decreases welfare.
Recent research has further argued that not only do voters lack basic information or coherent policy positions, their reasoning processes are heavily biased and bear no resemblance to Bayesian rationality (Achen and Bartels 2004a). First, in evaluating the incumbent’s performance, voters tend to be forgetful, i.e. they rely predominantly on the more recent events instead of the overall record (e.g. Achen and Bartels 2004a, Bartels 2008, Bartels and Zaller 2001, Erickson 1989). Second, voters are affected by irrelevant factors and events; they are swayed by rhetoric, framing, and advertising and hold incumbents accountable for events that are clearly beyond the office holder’s control.⁴ One such factor is facial features (Todorov et al. 2005, Bellew and Todorov 2007), including facial similarities between candidates and voters (Bailenson et al. 2009). Another example are events that are clearly unrelated to the efforts of the office holders but nevertheless influence voters’ attitudes. Studies have shown that the voters’ decisions are affected by shark attacks (Achen and Bartels 2004a), rainfall (Cole, Healy and Werker 2012, Gasper and Reeves 2011), the global oil price (Wolfers 2007) and the success or failure of local college football teams (Healy, Mo and Malhotra 2010).⁵

The aforementioned issues are particularly relevant in the domain of economic voting. On the one hand, there is a large literature that documents the correlation between favorable economic conditions and the reelection rates of incumbents (e.g. Kramer 1971, Fair 1978, Lewis-Beck 1988, Erickson 1989, Erickson 1990, Duch and Stevenson 2008). But establishing this correlation by itself is not enough. For the standard principal-agent model to hold, voters must be able to reward incumbents for good actions, but filter out external events ("luck") that are clearly beyond the incumbent’s control.

Wolfers (2007) provides evidence that this is not the case. Wolfers investigates the impact of national economic conditions and the global oil price on reelection rates for U.S. state governors. Consistent with previous studies of economics voting, Wolfers (2007) shows that state economic performance impacts gubernatorial reelection rates, but, crucially, he also shows that a rise in oil prices has a positive effect in oil producing states (such as Alaska, Wyoming, and Texas), but a negative effect on rust belt states which are net consumers of

---

⁴These concerns not only apply to models with fully rational voters, but also accounts of "reasoning voters", where voters uses cues, endorsements by trusted parties, media coverage, debate performance etc. to make, while not fully rational, at least competent decisions (e.g. Popkin 1991, Lupia and McCubbins 1998, Lau and Redlawsk 2006).

⁵Some external events do allow rational voters to deduce incumbent’s performance. For example, voter response will be influenced by disaster relief efforts (Cole, Healy, and Werker 2011, Healy and Malhotra 2010, Gasper and Reeves 2011). The point is not that the incumbent’s performance (on the local economy or other matters of public importance, e.g. disaster management) has no effect on reelection rates, but that rational voters should be able to ignore irrelevant factors, but the evidence suggests that they aren’t.
oil (e.g. Michigan and Indiana). In other words, while governors can do little to affect global oil prices, they are still held accountable by the voters.

The failure of rational information filtering is not the only problem faced by rational accounts of economic voting. As Hibbs (2006; p.570) has pointed out, one implication of the principal-agent approach to electoral accountability is that "the electorate should evaluate performance over the incumbent’s entire term of office, with little or no backward time discounting of performance outcomes". In practice, however, much of the empirical economic voting literature has used periods close to the election dates or overweighted recent periods (e.g. Kramer 1971, Tufte 1978, Erickson 1989; Hibbs 2000; Bartels and Zaller 2001). Achen and Bartels (2004a) argue that such restrictions are not an accident, but a reflection of a fundamental feature of an electorate who systematically ignores, discounts, or simply forgets relevant information that occurred earlier in the term of an elected official.6

Voter sophistication has potentially important normative consequences for the assessment of democratic governance structures. Plato famously argued in the Republic that an ill-informed and irrational public renders democracy unsuitable as a form of government.7 Many contemporary political scientists believe that the public control of officials does not function in an environment with an ignorant and uninterested public (e.g. Achen and Bartels 2004b). Nonetheless, instead of rigorously evaluating the implications of these empirical findings, much of the existing debate is centered around the interpretation of empirical findings. On the one hand, behaviorally oriented researchers take the findings to be proof of voter irrationality. On the other hand, scholars from the rational choice tradition have either tried to undermine the validity of the evidence or argued that rational choice models of electoral control do a satisfactory job of accounting for the evidence (e.g. Ashworth 2012, Ashworth and Bueno de Mesquita 2013a).8

---

6 A separate line of argument, originally pointed out by Fearon (1999), has identified a tension between forward looking selection of good office-holders and backward looking strategies that maximize incentives for office-holders to engage in costly effort (Alt, de Mesquita and Rose 2011, Ashworth and de Mesquita 2008, Ashworth, de Mesquita and Friedenberg 2012). The election rule that selects good types may be different from the election rule that maximized the office-holder’s incentives to take costly action. Therefore, electoral control can increase if voters disregard some information about incumbents (Ashworth and de Mesquita 2013a). Ashworth and de Mesquita (2013b) show that in the presence of both adverse selection and moral hazard, a rational retrospective voting rule does not necessarily maximize incentives for effort nor ex ante welfare. This leaves open the possibility that a different voting rule (with limited rationality) may induce higher effort welfare. The results in our model are not driven by direct strategic interaction between voters and politicians since voters are non-strategic by definition. Moreover, we set aside issues of incumbent quality and focus exclusively on electoral control as in the original Ferejohn (1986) model.

7 For a discussion of this and related views see Dahl (1989).

8 In a recent paper Healy and Lenz (2014) provide evidence that the myopia exhibited by voters is based on
In this paper, we take a different approach. We shall develop a model that captures behavioral features identified in the empirical behavioralist literature and assess its empirical and normative consequences for electoral control of public officials. We will, for the sake of the argument, set aside the rational voting model and assume that voters indeed behave according to the behavioral tradition. Specifically, our model will capture three main features of voting behavior identified in the empirical studies.

1. Voting partially depends on the actions ("effort") of the incumbent, as highlighted in the economic voting literature.

2. Voting partially depends on extraneous events that are beyond the control of the incumbent.

3. Voters are forgetful. They overweigh their experiences of the recent past in forming their attitudes toward the incumbent.

Our model provides new insights on the relationship between voter sophistication and electoral control. Our result in Section 4.1 shows that extraneous events (e.g. shark attacks, global oil prices) can be differentiated in terms of their effects on electoral control. This observation comes from recognizing that such events differ in many aspects. For example, events such as natural disasters and shark attacks are found to depress a voter’s propensity to reelect the incumbent; such events are thus negative for the incumbent. Other events, such as global oil prices, may either increase or decrease voters’ propensities. We find that positive and negative events have asymmetric effects on electoral control. In particular, electoral control is increasing in the salience of negative events, and decreasing in the salience of positive events. In other words, electoral control is enhanced when voters “pay greater attention” to negative events. Extraneous events also differ in terms of frequency of occurrence and the degree of correlation among voters. For example, a change in the global oil price has broad economic consequences, and therefore it can have an effect on many voters’ sentiments. Results in Section 5.1 suggest that events that occur frequently and affect a

---

9 For recent experimental results that support these findings see Huber, Hill, and Lentz (2012).

10 In the case of global oil prices, Wolters (2007) finds that a rise in prices hurts the incumbent and a drop in prices helps the incumbent.
large proportion of the electorate have the most pernicious effect on electoral control, while rare, approximately independent events can enhance electoral control.

By using a behavioral approach we can also investigate the impact of cognitive characteristics of voters, such as the degree of forgetfulness, on electoral control (see Section 4.2 and 5.2). The empirical literature on electoral behavior has argued that incumbents’ performances close to the date of the election have a disproportionately large effect on voters’ decisions. One explanation is that voters have limited memory. We are able to parametrize voter forgetfulness within our model. We find that the effect of voter forgetfulness depends crucially on the frequency of elections. When elections are frequent, higher level of forgetfulness can be beneficial to electoral control. Indeed, when election is held in every period, electoral control is maximized if voters are maximally forgetful, or "satisfice" as in Simon (1955) models of bounded rationality. However, when elections are infrequent, satisficing voters harm electoral control. Generally speaking, with forgetful voters, increasing election frequency improves electoral control.

Our modeling approach is along the lines of adaptive models of voting (Bendor, Diermeier, and Ting 2003, Bendor, Diermeier, Siegel and Ting 2011, Bendor, Kumar, and Siegel 2010, Andonie and Diermeier 2012). In this literature, voter behavior is directly described by stochastic reinforcement processes rather than expected utility maximization. The reinforcement process only relies on basic feedback mechanisms that are well-founded in the psychological learning literature. In contrast, politicians are assumed to be rational, maximizing long-run discounted utility. We contribute to the adaptive voting literature by being the first to examine interactions between strategic agents (i.e. the politician) and behavioral agents (i.e. the voters) in the context of electoral accountability. Moreover, unlike many existing adaptive voting models where solutions are obtained numerically, our model can be solved analytically. This is mainly driven by two features: first, we dispense with endogenous aspiration levels, and second, we consider a continuum of voters.

The rational electoral control model closest to ours is Ferejohn’s (1986) moral hazard model. As in the Ferejohn (1986), we model politicians as rational forward looking agents whose utility depends on the value of office and the level of effort exercised. Unlike in Ferejohn (1986), voters in our model are not strategic; the voters do not take into account the effect

---

11 There are a few models that explore other aspects of politicians in the context of behavioral voting. Bisin, Lizzeri, and Yariv (2011) and Lizzeri and Yariv (2012) consider time-inconsistent voters. Achen and Bartels (2002) consider a model of two-candidate competition with uninformed voters. Their modeling approaches are substantively different than ours.
of their behavior on the politician’s incentive nor do they infer about politicians effort in a Bayesian rational manner. In spite of this, we confirm some of the comparative statics that are known in the electoral control literature and have been supported by empirical studies. For example, higher values for holding office (or lower cost of effort) improves electoral control, i.e. higher effort by elected officials (e.g. Ferejohn (1986) and Ferraz and Finan 2009).

In the next section, we define the benchmark model. Section 3 characterizes the optimal effort choice of public officials. Section 4 discusses the implications of the salience of extraneous events and voter memory. Section 5 explores some extensions of our model. It sheds light on the implication of correlated events, the frequency of elections, and recurring candidates. Section 6 concludes. All proofs are relegated to the Appendix.

2 The Model

The model considers a infinite series of periods. Dates are denoted by \( t \in \mathbb{N} \), although sometimes we will omit the date subscripts to simplify notions if no confusion arises from doing so. We assume for now that an election is held in every period. The candidates for the date \( t \) election are the incumbent from period \( t - 1 \), denoted \( \theta_{t-1} \), and a challenger \( \gamma_t \). \( \theta_1 \) is determined exogenously. The electorate is comprised of a continuum (measure 1) of infinitely lived voters. We will refer to a voter as "she" and the incumbent as "he". Voter \( i \) votes for \( \theta_{t-1} \) at the date \( t \) election with probability (propensity) \( p_{i,t} \in [0,1] \). We denote the vote share for \( \theta_{t-1} \) at the date \( t \) election as \( P_t = \int_i p_{i,t} \). We assume that \( \theta_{t-1} \) is reelected if \( P_t > \frac{1}{2} \) (i.e. majority rule). Otherwise the challenger \( \gamma_t \) becomes the new incumbent \( \theta_t \).

The winner of the date \( t \) election chooses effort \( a_t \in \{h,l\} \subset \mathbb{R}_+ \), where one should interpret \( h \) as high effort (working for the electorate) and \( l \) as low effort (shirking). The incumbent’s effort level at date \( t \) influences the date \( t \) payoff, \( \pi_{i,t}^a \in \mathbb{R} \) for voter \( i \), which in turn determines \( i \)’th propensity to reelect the incumbent (to be specified below). Date \( t \) utility for \( \theta_t \) is \( w - a_t \), where \( w \) is the value of holding office. The incumbent chooses a (possibly finite) sequence of efforts to maximize his discounted utility with a discount factor \( \delta < 1 \). We assume that \( w > h > l \) so the office benefits compensate for the incumbent’s cost of effort.

\[12\] The relationship between incentives, effort, and performance has also been studied in the context of term limits. Term limits are associated with lower effort, higher levels of corruption, and lower performance. (Besley and Case 1995, Ferraz and Finan 2008, Ferraz and Finan 2011, Alt, de Mesquita, and Rose 2011).
While politicians are assumed to act as forward looking utility maximizers, voters respond to (past) payoffs in a myopic fashion. More specifically, we assume that the voters follow an adaptive learning heuristic consistent with the Law of Effect. This heuristic is viewed as "the most important principle in learning theory" (Hilgart and Bower, 1966; p. 481), and its key axiom is intuitive: agents increase the propensity for an action if that action has produced satisfactory feedback, they decrease propensity if feedback was negative. Feedback is satisfactory if current utility $\pi_{i,t}^a$ exceeds a threshold $\pi^*$. Formally,

$$
\begin{align*}
\pi_{i,t+1} &> \pi_{i,t} \quad \text{if } \pi_{i,t}^a > \pi^*
\pi_{i,t+1} &< \pi_{i,t} \quad \text{if } \pi_{i,t}^a < \pi^*
\pi_{i,t+1} &= \pi_{i,t} \quad \text{if } \pi_{i,t}^a = \pi^*
\end{align*}
$$

We assume that $i$'th payoff is correlated with the incumbent’s effort by imposing that $\pi_{i,t}^a = \pi^a + \epsilon_i$ where $\pi^h > \pi^* > \pi^l$ are constants and $\{\epsilon_i\}$ are iid random variables with median zero.\(^{13}\) The random component $\epsilon_i$ embodies extraneous events that are outside of incumbent’s control but nonetheless affect voters decision (e.g. shark attacks, global oil price). For simplicity, we assume $\epsilon_i$ contains no atoms and therefore the case of $\pi_{i,t}^a = \pi^*$ can be ignored.

We assume that the propensities evolve according to the Bush-Mosteller rule (Bush and Mosteller 1955), which is one of the cornerstones of the behavioral learning literature (see Borgers and Sarin 2000, Karandikar et al. 1998) and its applications in political science (Bendor et al. 2003, 2011). The Bush-Mosteller rule specifies that:

$$
\pi_{i,t+1} = \begin{cases} 
(1 - \beta)\pi_{i,t} + \beta & \text{if } \pi_{i,t}^a > \pi^* \\
(1 - \beta)\pi_{i,t} & \text{if } \pi_{i,t}^a < \pi^*
\end{cases}
$$

where $\beta \in [0, 1]$. Under the Bush-Mosteller rule, the next-period propensity is a convex combination of the current-period propensity and one if the payoff is satisfactory, and zero if the payoff is not satisfactory. It is straightforward to see (1) are satisfied. We assume that the propensity in the initial period is $\frac{1}{2}$ i.e. $p_{i,1} = \frac{1}{2}$.

Note that a higher $\beta$ means that $p_{i,t+1}$ depends less on $p_{i,t}$, which is associated with payoffs prior to date $t$, and more on the date $t$ payoff. For example, in the case of $\beta = 1$, a voter’s decision at date $t + 1$ depends only on the date $t$ payoff, i.e. voters are “satisficers”\(^{13}\)

\(^{13}\)It will be clear in what follows that assuming $\{\epsilon_i\}$ being identical is without loss of generality (see Appendix). The independence assumption is relaxed in Section 5.1.
as in Simon (1955). Thus, $\beta$ captures the voter’s forgetfulness, or more generally their tendency to place greater weight on more recent experiences. By varying $\beta$, we can explore the relationship between forgetfulness and electoral control.

It is clear from equations (2) that two (probabilistic) events are of importance: $\pi_{i,t}^a > \pi^*$ and $\pi_{i,t}^a < \pi^*$. We refer to the former as $G$(ood experience) and the latter $B$(ad experience). Define $\Psi^a = \Pr(\pi_{i,t}^a > \pi^*)$. From the incumbent’s point of view, $\{\Psi^h, \Psi^l\}$ are the only relevant properties of $\{\pi_{i,t}^a\}$, and the assumptions on $\pi_{i,t}^a$ imply that $\Psi^h > \frac{1}{2} > \Psi^l$. Treating $\{\Psi^h, \Psi^l\}$ as parameters, they can be interpreted as a measure of the amount of influence of various types of extraneous events, or their salience. For example, an increase in the salience of negative events (i.e. events that hurt the incumbent) leads to a decrease in both $\Psi^h$ and $\Psi^l$. Intuitively, the probability of a good experience for any level of effort would decrease if the voters pay greater attention to negative events. Alternatively, if both positive and negative extraneous events become more salient, then both $\Psi^h$ and $\Psi^l$ would be close to $\frac{1}{2}$. Intuitively, as voters pay attention to all types of extraneous events, the incumbent’s effort choice becomes increasingly irrelevant and the difference between $\Psi^h$ and $\Psi^l$ should decrease.

3 Electoral Control

We shall first define a few notions and terms for expositional purposes. Given the date $t$ vote share $P_t$ and effort level $a_t$, $\theta_t$’s vote share at the date $t+1$ election is:

$$Q(a_t, P_t) = \Psi^a \int [(1 - \beta)p_{i,t} + \beta] + (1 - \Psi^a) \int (1 - \beta)p_{i,t} = (1 - \beta)P_t + \beta\Psi^a$$

(3)

Note that the vote share dynamics share the same recursive structure as the underlying Bush-Mosteller process. Furthermore, when voter payoff is deterministic (i.e. $\epsilon_i$ is degenerate at 0), the vote share dynamic is exactly the Bush-Mosteller process (i.e. $\Psi^h = 1$ and $\Psi^l = 0$). Observe that our assumption of $p_{i,1} = \frac{1}{2}$ implies that $P_1 = \frac{1}{2}$.

---

14 Originally developed in the context of search behavior, it has also been applied to models of voting (e.g. Bendor 2010, Bendor, Diermeier, Siegel, and Ting 2011).

15 Intuitively, if negative events are more salient, then the distribution of $\epsilon_i$ is more spread out to the left of 0.
Observe that if \( P \) is sufficiently large, then \( \theta_t \) can shirk and still be reelected. It is useful to distinguish whether an effort level ensures reelection or not. This leads to the following definition.

**Definition 3.1.** An effort level \( a \) is adequate given \( P \) if \( Q(a, P) > \frac{1}{2} \). That is, \( a \) is adequate if under current-period vote share \( P \), the incumbent will be reelected given effort \( a \). Given current-period vote share \( P \), a (possibly finite) sequence of action is adequate if effort at each subsequent date is adequate.

The following result is a direct consequence of (3):

**Lemma 3.1.** \( Q(a, P) \) is strictly increasing in both arguments. Consequently, if \( \{a_s\}_{s=1}^{n} \) is a adequate sequence of efforts, then so is any alternative sequence \( \{a'_s\}_{s=1}^{n} \) where \( a'_s \geq a_s \forall s \).

In many political systems an incumbent who loses a reelection rarely gets nominated by his own party to run for office in the future. Therefore, we assume that \( \theta_t \neq \gamma_t + s \forall s \geq 0 \). That is, the incumbent cannot become a challenger in the future after he is voted out. This essentially rules out strategic interaction between the incumbent and challengers. We will relax this assumption in Section 5.3. We will also assume that if the challenger wins the election, the propensity for the new incumbent is reset at \( \frac{1}{2} \). In other words, the electorate is neutral towards a new incumbent. This is a reasonable assumption because in a context of moral hazard, reelection is a disciplining mechanism rather than a mechanism to select a competent leader as in an adverse selection context.

We can now characterize the optimal behavior of the incumbent. Lemma 3.2 below is a simple observation that at the optimum, the incumbent either exert sufficient effort to remain in office forever, or shirks and is voted out.

**Lemma 3.2.** At the optimum, \( \theta_1 \) either shirks in the first period and is voted out, or will remain in office forever. Furthermore, if it is optimal for \( \theta_1 \) to shirk in the first period, then for all \( t > 1 \), \( \theta_t \) shirks as well.

From now on, we shall refer to the case where the incumbent stays in office forever as "permanent incumbency". Lemma 3.2 implies that permanent incumbency is necessary for electoral control because otherwise the incumbent shirks every period.\(^{16}\) We will henceforth take as given the optimality of permanent incumbency unless otherwise stated. Lemma 3.3 below states that permanent incumbency is optimal if the incumbent is sufficiently patient.

---

\(^{16}\)As we show in section 5.1 electoral control can be maintained in the absence of permanent incumbency in a more general setting.
Lemma 3.3. Permanent incumbency is optimal if $\delta > 1 - \frac{w-h}{w-l}$.

Note that the critical value $1 - \frac{w-h}{w-l}$ is decreasing in the reward from holding office $w$ and increasing in the cost of effort $h$ and $l$.

To properly discuss electoral control, which is typically some measure of voters’ long-run payoff, we need to characterize the optimal sequence of efforts. Proposition 3.1 is an important step towards this goal. It shows that the optimal effort at date $t$ follows a cut-off rule on $P_t$. Moreover, we can explicitly compute the cut-off value $P^*$ from the parameters $\beta$ and $\Psi^l$. Interestingly, the cut-off $P^*$ is the lowest possible value of $P_t$ for which the incumbent can be reelected even if he shirks.

**Proposition 3.1.** The incumbent chooses $l$ at date $t$ if and only if $P_t > P^*$ where $P^*$ satisfies $(1 - \beta)P^* + \beta\Psi^l = \frac{1}{2}$. Furthermore, if $P^* \geq \Psi^h$, then the optimal effort in every period is $h$; otherwise, $l$ will be chosen infinitely often at the optimum.

Since the vote share aggregates the voters’ propensities, one may interpret $P_t$ as a measure of the incumbent’s "political capital" at date $t$. Empirically, it may correspond to the incumbent’s approval rating, as measured in public opinion surveys. Proposition 3.1 implies that incumbents do not seek to maintain public approval higher than necessary for reelection. Rather, if public approval is sufficiently high, the incumbent will exploit his political capital and shirk.

Observe that the cut-off rule implies for $P^* > \Psi^h$, the incumbent exerts high effort every period; while for $P^* < \Psi^l$, efforts have a cyclical structure: a sequence of high effort followed by a sequence of low effort. The cut-off rule by itself, however, does not provide any characteristics of the effort cycles. This is accomplished by Proposition 3.2 below. Somewhat surprisingly, the length cycles are almost stationary, and we can derived bounds on the length of effort cycles. For notational simplicity, let $Q_a(P) = Q(a, P)$ and $Q_a^n(P) = Q_a \circ Q_a \circ \ldots \circ Q_a(P)$.

**Proposition 3.2.** Suppose that $P^* < \Psi^h$:

- If $Q_h(\frac{1}{2}) > P^*$, then at the optimum, the incumbent chooses $l$ for $\sigma$ or $\sigma + 1$ consecutive periods each time after $h$ is chosen, where $\sigma = \min\{n : Q_h^n(Q_h(\frac{1}{2})) \leq P^*\} \geq 1$.

- If $Q_h(\frac{1}{2}) \leq P^*$, then at the optimum, the incumbent chooses $h$ for $\sigma$ or $\sigma + 1$ consecutive periods each time after $l$ is chosen, where $\sigma = \max\{n : Q_h^n(\frac{1}{2}) \leq P^*\} \geq 1$. 


Note that the length of consecutive high efforts (or low efforts) can differ between cycles. For example, the optimal sequence of effort may be something like \{h,h,l,h,l,...\} or \{h,l,l,h,l,l,...\}, but the length of the cycles cannot differ by more than one.

Given effort level \(a\), the one-period aggregate voter payoff is \(\int_t \pi^a_{i,t} dt\). For simplicity, assume that \(\epsilon_i\) has mean zero and therefore \(\int_t \pi^a_{i,t} dt = \pi^a\). Given a sequence of effort from the incumbent, one can compute the long-run aggregate voter payoff, defined as:

\[
\Pi(\{a_t\}) = \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \pi^{a_t}
\]

We defined the long-run payoff as the limit of means for notational simplicity. The results would not change if we used discounted payoffs instead. Note that since the length of effort cycles is bounded at the optimum, the corresponding long-run voter payoff, denoted \(\Pi^*\), is bounded as well. For example, if \(Q_h(\frac{1}{2}) \leq P^*\), then

\[
\frac{\sigma}{\sigma + 1} \pi^h + \frac{1}{\sigma + 1} \pi^l \leq \Pi^* \leq \frac{\sigma + 1}{\sigma + 2} \pi^h + \frac{1}{\sigma + 2} \pi^l
\]

A similar bound can be derived for the case of \(Q_h(\frac{1}{2}) > P^*\). Note that the width of the bound is small for large \(\sigma\), and therefore the bounds can be a good approximation for \(\Pi^*\).

We use the lower bound of \(\Pi^*\) as the measure of electoral control. All of our comparative statics would follow if we used the upper bound instead.

**Definition 3.2.** Let electoral control \(e\) be the lower bound of \(\Pi^*\), that is:

\[
e = \begin{cases} 
\frac{\sigma}{\sigma + 1} \pi^h + \frac{1}{\sigma + 1} \pi^l & \text{if } Q_h(\frac{1}{2}) \leq P^* \\
\frac{1}{\sigma + 2} \pi^h + \frac{\sigma + 1}{\sigma + 2} \pi^l & \text{if } Q_h(\frac{1}{2}) > P^*
\end{cases}
\]

where \(\sigma\) is defined in Proposition 3.2.

Observe that \(e\) is maximized (i.e. \(e = \pi^h\)) if the incumbent never shirks. Results in the next section shows that this can be accomplished under certain conditions. Therefore, electoral control with behavioral voters can be greater than the electoral control with rational voters as in Ferejohn (1986), where there is always a strictly positive probability that the incumbent shirks. This observation is in contrast to the view that the public control of officials does not function in an environment with an ignorant and uninterested public that relies on simple
feed-back mechanisms.\textsuperscript{17}

4 Election Characteristics and Electoral Control

In the previous section, we showed that public officials will only accumulate sufficient political capital to ensure reelection. From the voters' point of view, accumulating political capital should not be too easy. Otherwise, elected officials will tend to exercise less effort.

With the characterization of effort choice, we can now explore the relationships between various model primitives and electoral control. We provide some novel insights. First, we show how the effect of some extraneous event on electoral control depend on whether the event is positive or negative. Second, we examine the consequence of voter's (lack of) memory. Finally, we show that some of the known empirical regularities of electoral control also hold with behavioral voters. In particular, increasing in the value of office helps electoral control. Throughout this section, since $e$ is a strictly monotonic function of $\sigma$, we will often treat $\sigma$ as electoral control to simplify notations in the proofs.

4.1 Extraneous Events

Extraneous events are events beyond the control of elected officials that nevertheless impact public approval and reelection prospects. Recent empirical studies have identified a growing list of such events. For example, negative events like shark attacks (Achen and Bartels 2004a) and natural disasters (Gasper and Reeves 2011) depress a voter's propensity to reelect the incumbent. Other events such as global oil prices (Wolfers 2007) and local sports team's performance (Healy, Mo and Malhotra 2010) may either increase or decrease voter propensities, depending on the outcomes. In this section, we show that the salience of extraneous events have different implications depending on their type: electoral control is increasing in the salience of negative events and decreasing in the salience of positive events.

Recall that the parameters $\{\Psi^h, \Psi^l\}$ are interpreted as a measure the salience of extraneous events, or equivalently, the level of "attention" the voters pay to those events. As the salience

\textsuperscript{17}The popular press has engaged this issue in the context of "low information" voters, a term originally due to Popkin (1991). As an example see the opinion piece by Berkeley cognitive linguist George Lakoff (Lakoff 2012). Some political scientists have expressed similar opinions (e.g. Achen and Bartels 2004b; p. 38 ).
of negative events increases, one should expect $\Psi^h$ and $\Psi^l$ to decrease. Intuitively, if voters pay more attention to negative extraneous events, then it depresses the probability of having a good experience for any given level of effort. Proposition 4.1 below shows that electoral control is decreasing in $\Psi^h$ and in $\Psi^l$. Thus, it implies that electoral control is increasing in the salience of negative events, and decreasing in the salience of positive events. The effect of events that can be both positive and negative is ambiguous.

**Proposition 4.1.** $e$ is decreasing in $\Psi^h$ and $\Psi^l$.

The intuition behind the result is straightforward. If $\Psi^h$ is high, then it is easy for the politicians to build up political capital. This hurts electoral control because political capital encourages shirking. If $\Psi^l$ is low, then shirking is not very costly for the politicians, which clearly is harmful for electoral control. It follows that voters are better off, if they treat negative and positive events asymmetrically, giving more weight to negative rather than positive events. There is extensive evidence in the psychological literature that individuals tend to pay more attention to negative rather than positive events (Baumeister et al. 2001, Rozin and Royzman 2001). In the context of electoral control, such "negative bias" is beneficial for the electorate.

### 4.2 Voter Memory

In this section, we explore a different dimension of electoral control: the connection between voter memory and electoral control. Recall our interpretation of $\beta$ as a measure of voter's forgetfulness: a higher $\beta$ means the current experience at time $t$ has a greater weight on determining the date $t+1$ propensity than prior experiences. In the extreme case of satiscing ($\beta = 1$), the date $t+1$ propensity is solely determined by the date $t$ experience.

The following result indicates that a high level of electoral control is obtained when the voters are sufficiently forgetful. Intuitively, for a high level of forgetfulness, the incumbent’s action at date $t$ has a large impact on the outcome of election at $t+1$, but its impact on the outcome of election at $t+2$, $t+3$, ... is small (since the outcome in those elections is affected mostly by actions in $t+1$, $t+2$, ...). Consequently, the degree of voter forgetfulness determines how myopic or farsighted the incumbent is when choosing effort. And since for sufficiently large $\beta$, high effort is needed for reelection, a myopic electorate creates high incentives for the incumbent to exert high effort. It follows that electoral control is maximized in this case.
Proposition 4.2. \( e = \pi^h \) if and only if \( \beta \geq \frac{\Psi^h - \frac{1}{2}}{\Psi^h - \Psi^l} \), and \( e \to \pi^h \) as \( \beta \to \frac{\Psi^h - \frac{1}{2}}{\Psi^h - \Psi^l} \). Furthermore, \( \exists \mu \) such that \( e \) is increasing in \( \beta \) for \( \beta \in (\mu, \frac{\Psi^h - \frac{1}{2}}{\Psi^h - \Psi^l}) \).

It then follows immediately that electoral control is maximized if voters are satisficers, as expressed in the following Corollary.

Corollary 4.1. \( e = \pi^h \) if the voters are satisficers, i.e. \( \beta = 1 \)

In sum, we have shown that highly forgetful voters as identified by Achen and Bartels (2004a) can induce high levels of electoral control. However, as we will show in Section 5.2, this conclusion does depend on the frequency of elections. Highly forgetful voters induce the incumbent to behave myopically. This is beneficial for electoral control when elections are frequent (i.e. held in every period). However, when elections are held once every \( k > 1 \) periods, a myopic incumbent would have little incentive to exert high effort except in periods close to an election. This suggests that highly forgetful voters decreases electoral control when elections are infrequent. We will discuss this issue in detail in Section 5.2.

4.3 The Value of Office

One of the main empirical findings consistent with the rational voting theory of electoral control is that a higher value of office or lower cost of effort improves electoral control (Besley and Case 1995, Ferraz and Finan 2008, Ferraz and Finan 2009, Ferraz and Finan 2011, Alt, de Mesquita, and Rose 2011). These observations continue to hold in our framework as a corollary to Lemma 3.3. In particular, high value of office and low (relative) cost of high effort helps sustaining permanent incumbency, which is necessary for electoral control.

Proposition 4.3. If \( w \) is sufficiently low or \( h - l \) sufficiently high, \( e \) is minimal and equal to \( \pi^l \). For high values of \( w \) or low values of \( h - l \), \( e > \pi^l \).

The proof is a straightforward given Lemma 3.3 and is omitted. Observe that Proposition 3.1 and Proposition 3.2 imply that electoral control given permanent incumbency does not depend on \( w, h \) nor \( l \). Thus, \( w \) and \( h - l \) affect electoral control only in determining the optimality of permanent incumbency. Finally, if there is a finite term limit, then the continuation value of exerting high effort is less than without term limits. It follows that the incumbent has a higher incentive to shirk, and electoral control suffers.\(^{18}\)

5 Extensions

5.1 Common Shocks

We have so far assumed that the realization of shock $\epsilon_i$ is independent across voters (i.e. idiosyncratic shocks). In reality, however, the experiences of voters can be correlated (e.g. in the case macro-economic shocks). We can introduce common shocks into our model in the following manner. Suppose that in every period, all voters receive good experiences with probability $p$ irrespective of the incumbent’s effort, similarly, let $q$ be the probability of all voters receiving a bad experience. The common shock is realized after the incumbent takes the effort.

Proposition 5.1 below shows that the incumbent adopts similar strategy as in the benchmark model with idiosyncratic shocks. Consequently, our main observations would continue to hold in this case.

Proposition 5.1. Suppose that $q + p$ is sufficiently small and $\delta$ sufficiently large, and define $P^{**} = \frac{1}{2(1-\beta)}$:

1. if $P^{**} \geq 1$, then the incumbent follows the same strategy as in the case without common shocks. (i.e. Proposition 3.1 holds).

2. if $P^{**} < 1$, then the incumbent chooses $h$ if the reputation is such that $Q(l, P) \leq P^{**}$.

Interpreted literally, $p + q$ is a measure of the frequency of the shock. An alternative interpretation of $p+q$ is that it measures the proportion of voters affected by the shock. Thus, commonly experienced events that are sufficiently rare or effect a sufficiently small proportion of the electoral do not undermine electoral control. This applies to natural disaster (rare events) but also frequent events (e.g. football games) provided they are local in nature (in the sense that the outcomes are correlated among a comparatively small segment of the electorate). Indeed, as the second result in Proposition 5.1 shows, the presence of common shocks can improve electoral control. This follows from the fact that in the baseline model, high effort is exerted if and only if $Q(l, P) \leq \frac{1}{2}$, while here the threshold for high effort is $P^{**}$, which may be greater than $\frac{1}{2}$. Intuitively, the incumbent wishes to insulate himself against a bad shocks by maintaining higher political capital than in an environment without common shocks. Thus, he needs to exert high effort more often.
So far, we have shown that common shocks of limited impact have no effect on incumbent’s behavior. However, there are some macro-events, for example oil price shocks in times of high volatility, that both reach a large portion of the electorate and can potentially occur with high frequency. In our current setting, it is straightforward to see that if \( p + q \) is sufficiently large, then the incumbent would always shirk. Thus, these shocks can have a pernicious effect on electoral control. The intuition of this observation is as follows. If common shocks occur with high probability, then the incumbent’s effort choice has only negligible effect on the outcome of the election. Therefore the incentive to undertake costly effort is small.

5.2 Multi-period Incumbency

In this section, we examine the incumbent’s behavior when an election is held every \( k > 1 \) periods. That is, there is an election at date 1, \( k + 1 \), \( 2k + 1 \) and so on. Note that \( k = 1 \) corresponds to the baseline model in Section 2. We show that the optimal effort is still determined by a threshold rule even though the incumbent no longer needs to maintain a vote share of \( \frac{1}{2} \) in every period. This rule is stationary in the sense that threshold rule for date \( t \) will be the same as that for date \( t + k \). To simplify notations, we will refer to dates \( t \in \{\tau, k + 1 - \tau, 2k + 1 - \tau, \ldots\} \) collectively as \( \tau \). That is, \( \tau \leq k \) is the number of period until the next election (e.g. \( \tau = 1 \) denotes dates prior to an election, and \( \tau = k \) denotes election dates.)

**Proposition 5.2.** At date \( \tau \), the incumbent maintains \( P_\tau > P_{\tau}^{**} \) and he shirks if and only if \( P_\tau > P_{\tau}^* \). where \( P_{\tau}^{**} \) and \( P_{\tau}^* \) are defined as follows:

If \( \beta < 1 \),

\[
P_{\tau}^{**} \quad \text{is such that} \quad Q^*(h, P_{\tau}^{**}) = \frac{1}{2}
\]

\[
P_{\tau}^* \quad \text{is such that} \quad Q(l, P_{\tau}^*) = \frac{1}{2}
\]

\[
P_{\tau}^* \quad \text{is such that} \quad Q(l, P_{\tau}^*) = P_{\tau-1}^{**} - 1
\]

\[\text{Note that oil prices fluctuations, though reaching a large portion of electorate, can still be considered of limited impact if the fluctuation is of small magnitude or occurs with low probability.}\]
If $\beta = 1$,

\begin{align*}
P^*_1 &= 1 \\
P^*_\tau > 1 &= 0
\end{align*}

Because $P^*_{\tau}$ is decreasing in $\tau$, $P^*_{\tau}$ is also decreasing in $\tau$. Since $P^*_{\tau}$ is the threshold for shirking, this means that the incentive to shirk decreases as the election draws near (i.e. $\tau$ small). This is intuitive because when voters are forgetful, efforts early in the term have less impact on the outcome of the upcoming reelection. Furthermore, due to discounting, the cost of high effort is higher earlier in the term. These two factors imply that it is more profitable for the incumbent to shirk early in the election cycle than late. Corollary 5.1 formalizes this intuition by showing that in the optimum, the incumbent shirks prior to some point in the election cycle and exerts high effort thereafter.

We also provide a bound on the proportion of high efforts to low efforts within an election cycle, and this bound is independent of $k$ and decreasing in $\beta$. One implication of the result is that ceteris paribus, high frequency of elections is good for electoral control. A second implication is that when elections are infrequent, highly forgetful voters are no longer a boon for electoral control. For example, when voters are satisfiers i.e. $\beta = 1$, electoral control is minimized if $k > 1$ while maximized if $k = 1$.

**Corollary 5.1.** There exists some $1 \leq m \leq k$, such that for dates $\tau > m$, the incumbent shirks and for dates $\tau \leq m$, the incumbent exerts high effort. Furthermore, there is an upper bound for $m$, denoted $\bar{m}$, which is independent of $k$, decreasing in $\beta$, and $\bar{m} = m = 1$ when $\beta = 1$.

Observe that the upper bound $\bar{m}$ can be translated to an upper bound on electoral control i.e. $e \leq \frac{m}{k} \pi^h + \frac{k-m}{k} \pi^l$. Now, since $\bar{m}$ is independent of $k$, the bound on $e$ is decreasing monotonically to $\pi^l$ as $k$ increases to infinity. This suggests that low frequency of elections is detrimental to electoral control. Moreover, the fact that $\bar{m}$ is decreasing in $\beta$ suggests that high forgetfulness is bad for electoral control when elections are infrequent.

### 5.3 Recurring Candidates

We have so far assumed that the incumbent cannot reenter a future election once he loses. Consequently, the incumbent faces essentially a single agent optimization problem. In this
section, we assume that there are two long-lived candidates (i.e. $D$ and $R$) who are \textit{ex ante} identical and run against each other in every election (i.e. $\theta_{t-1}, \gamma_t \in \{D, R\}\forall t$). This is a reasonable assumption if we think of the candidates as political parties.

Thus, we have defined a non-cooperative game between two long-lived actors. Characterizing the Nash Equilibria of this game, however, is difficult, because the game lacks a recursive structure (the game is not a repeated game). Furthermore, the game may admit multiple equilibria (see Proposition 5.4).

We start with a straightforward observation:

**Proposition 5.3.** If $\delta > 1 - \frac{w-h}{w-l}$, then there is an equilibrium where permanent incumbency is achieved.

Clearly, if permanent incumbency is optimal in equilibrium, then $\theta_1$'s optimal efforts on the equilibrium path are characterized as in the baseline model.

Uniqueness of equilibrium with permanent incumbency is possible under some parameter values.

**Proposition 5.4.** If $\delta > \frac{w-l}{w-h} - 1$, then it is a dominant strategy for both players to follow the optimal sequence of efforts prescribed in Proposition 3.2.

Observe that the condition on $\delta$ in Proposition 5.4 is not the same in the condition in Proposition 5.3. In particular, the condition in Proposition 5.3 can always be satisfied by taking $\delta$ sufficiently large, the same cannot be said for Proposition 5.4 since $\frac{w-l}{w-h}$ is not bounded above. The following is a corollary of Proposition 5.3 and Proposition 5.4.

**Corollary 5.2.** If $\beta \geq \frac{\psi h - \frac{1}{2}}{\psi - \psi h}$, and $\delta \leq \frac{w-l}{w-h} - 1$, then there is an equilibrium where both $D$ and $R$ always shirk. If, in addition, $1 - \frac{w-h}{w-l} < \delta$, then there is also a permanent incumbency equilibrium.

Note that $\beta \geq \frac{\psi h - \frac{1}{2}}{\psi - \psi h}$ is a sufficient and necessary condition for maximal electoral control in the baseline model. Thus, Corollary 5.2 shows that in some instances, there can be two extremal equilibria: one in which the incumbent exercises high efforts in every period, and one in which the incumbent shirks in every period. This multiplicity of equilibria makes it difficult to conduct comparative statics.\textsuperscript{20} However, it is straightforward to see that the

\textsuperscript{20}One problem is that for some equilibria, the measure of electoral control defined earlier (i.e. $e$) may not apply since the sequence of efforts on an equilibrium path may not be well-behaved.
incumbent’s payoff under permanent incumbency must be the lower bound of the equilibrium payoff, since the incumbent can always deviate to permanent incumbency. The fact that the incumbent’s equilibrium payoff is higher than under permanent incumbency suggests that there is (weakly) more shirking in equilibrium than under permanent incumbency.\textsuperscript{21} This can be seen as evidence that allowing the incumbent to reenter the race after losing is harmful for electoral control. This is in accordance with a corresponding result in Ferejohn (1986), where electoral control is decreasing in the probability that an incumbent returns to a race after losing office.

6 Conclusion

Critics of democracy have frequently argued that democratic forms of governance require a well-informed and rational electorate (e.g. Dahl 1989). If voters are found to lack these qualities, so the argument continues, a proper justification for democratic governance is lacking (e.g. Achen and Bartels 2004b). Much of the existing debate on electoral accountability has centered on the question whether actual voter behavior is consistent with rational choice theory (e.g. Ashworth 2012) or not (e.g. Achen and Bartels 2004a). In this paper, we set this debate aside and examine the implications of an electorate that follows simple behavioral rules rather than rational choice assumptions. That is, we assume that voters act as the behavioralist critics of the rational choice models have argued: they are forgetful, uninformed, biased and care little about politics and policy. We then build a model that captures these assumptions formally. The model yields a dynamic process which we analyze with respect to its predictive and normative properties.

Our model provides several novel insights. We find that electoral control is increasing in the salience of negative events. In other words, electoral control is greater when voters “pay greater attention” to negative events. We find that extraneous events that affect large proportion of voters and occur frequently are the most harmful to electoral control. We also show that the effect of voter forgetfulness depends crucially on the frequency of elections. Voter forgetfulness is beneficial to electoral control if and only if elections are frequent.

The upshot of our analysis is that one should be cautious about making broad statements regarding the implications of a behavioral electorate. Electoral control can function well even

\textsuperscript{21}Note that the incumbent obtains $w$ in every period under permanent incumbency. Thus if in a given equilibrium both players were to obtain higher payoffs, that must mean $l$ is chosen more often.
with voters that vote how they feel, who are forgetful and influenced by extraneous events. But the details matter. Voter forgetfulness, for example, may be a boon or bane for electoral control, depending on other institutional factors such as election frequency. Our modeling approach can also be applied to other questions. In addition to the factors identified in our analysis, one may consider the issue of candidate quality as in adverse selection models or the implications of policy choice under ideological candidates. Such questions, we hope, will be the subject of future research.

Appendix

Heterogeneous $\Psi^a$

In our model, we assumed that individual shocks $\{\epsilon_i\}$ are drawn from the same distribution; this has the consequence that $\Psi^a$ is the same for all voters. The following result shows that this is without the loss of generality.

Lemma 6.1. Let $\Psi^a_i$ be the probability of a good experience occurring for voter $i$ given effort level $a$, then the optimal incumbent behavior will be the same as in the environment where every voter is endowed with $\Psi^a = \int_i \Psi^a_i$.

Proof. Note that the operator $Q$ (i.e. next period’s vote share given current period’s vote share and effort) can be reformulated as:

$$Q(a, P) = \int \Psi^a_i (1 - \beta) p_i + \beta) + (1 - \Psi^a_i)(1 - \beta)p_i$$

$$= \int \Psi^a_i \beta + (1 - \beta)p_i$$

$$= (1 - \beta)P + \beta \int \Psi^a_i$$

Since the operator $Q$ is the only relevant information for the incumbent’s problem, we see that electoral control in an environment with heterogeneous $\Psi^a_i$ is equivalent to an environment with homogenous voters.
Proof for Lemma 3.2

Proof. Let \( v(P) \) be the incumbent’s maximal discounted utility given vote share \( P \geq \frac{1}{2} \) and him exerting high effort. Observe that \( v(P) \) is increasing in \( P \) since an adequate sequence of effort under \( P \) will remain adequate under \( P' > P \). Suppose it is optimal for the incumbent to stay in the office until date \( N > 1 \) and then quit, it must be that \( v(P_N) < w-l \). By the fact that the incumbent is reelected at date \( N, P_N \geq P_1 = \frac{1}{2} \). This means \( v(P_N) \geq v(P_1) \geq w-l \). This is a contradiction. Finally, if the vote share is \( \frac{1}{2} \), \( l \) is not adequate. Therefore, if \( \theta_1 \) shirks in the first period, he will be voted out. Now, recall the assumption that when the challenger wins the election, a voter’s propensity for the new incumbent is set at \( \frac{1}{2} \). Thus the new incumbent faces the same decision problem as \( \theta_1 \). It follows that if it is optimal for \( \theta_1 \) to shirk, then it is optimal for \( \theta_{t>1} \) to shirk.

\[ \square \]

Proof for Proposition 3.1

Proof. It is straightforward to see that if \( P_t \leq P^* \), then the incumbent has to choose \( h \), otherwise he will be voted out. We will now show that \( l \) will be chosen if \( P_t > P^* \) (note \( l \) is adequate if \( P_t > P^* \) at date \( t \)). Suppose at date \( t \), the incumbent finds \( h \) optimal even though \( l \) is adequate. We can construct an alternative sequence of efforts that gives a higher payoff. There are two cases to consider. First, if a one stage deviation (i.e. play \( l \) at date \( t \) and then go back to the prescribed efforts) does not violate permanent incumbency, then clearly the incumbent is better off deviating. Suppose the one stage deviation violates permanent incumbency at \( t + s + 1 \) (i.e. the prescribed effort at date \( t + s \) after the deviation at \( t \) is no longer adequate), then we shall construct a two stage deviation where the incumbent takes effort \( l \) at date \( t \), and takes effort \( h \) (instead of \( l \) ) at \( t + s \). This two stage deviation gives a higher payoff because of discounting. We need to show that this two stage deviation is adequate. The following observations are important for the proof.

Recall that Bush-Mosteller implies that given vote share \( P \) and effort level \( a \), the vote share for the next election is \( Q(a, P) = (1-\beta)P + \beta \Psi^a \). The difference in vote share given the same \( P \) but different effort level is:

\[ Q(h, P) - Q(l, P) = \beta(\Psi^h - \Psi^l) \] (4)
and the difference in next-period vote shares between two different initial vote shares \((P' < P)\) but the same effort is:

\[
Q(a, P) - Q(a, P') = (1 - \beta)(P - P')
\]

by inductive reasoning, the difference in vote share given \(P' < P\) and taking the same vector of efforts \(\{a_1, a_2, \ldots, a_j\}\) is:

\[
(1 - \beta)^j(P - P')
\]

Now, we are ready to show the two stage deviation is adequate. Note that after the first deviation, \(l\) is prescribed at date \(t + s\) but is not adequate. By (4), the vote share at date \(t + 1\) after the deviation is \(\beta(\Psi^h - \Psi^l)\) less than under the prescribed sequence of efforts. Now, the one stage deviation is adequate at date \(t + s - 1\), and by (5), the difference in vote share for the date \(t + s\) election between the one stage deviation and prescribed efforts at \(t + s - 1\) is \((1 - \beta)^{s-1}\beta(\Psi^h - \Psi^l)\), with \(s \geq 1\). Now, the second deviation has the incumbent taking \(h\) instead of \(l\) at the date \(t + s\). By (4), \(\beta(\Psi^h - \Psi^l) \geq (1 - \beta)^{s-1}\beta(\Psi^h - \Psi^l)\). Thus, the vote share at date \(t + s + 1\) after the second deviation is higher than under the prescribed sequence of efforts. That means following the prescribed efforts from date \(t + s + 1\) onward will not violate permanent incumbency.

Now, since \(P_1 = \frac{1}{2} < \Psi^h\), incumbent’s vote share at any given point cannot exceed \(\Psi^h\). Thus if the threshold is larger than \(\Psi^h\), the incumbent never shirks. If \(P^* < \Psi^h\), then if incumbent exert high effort for many periods, the vote share will converge to \(\Psi^h\) and therefore exceed \(P^*\) at some point. It follows that the incumbent will shirk infinitely often at the optimum.

\[
\square
\]

**Proof for Proposition 3.2**

*Proof.* Assume first that \(Q_h(\frac{1}{2}) > P^*\). Since \(Q_a(P)\) is increasing in \(P\), the incumbent’s vote share following a high effort is larger than \(P^*\). Consequently, the incumbent shirks following a high effort according to Proposition 3.1. Now, the incumbent’s vote share after exerting high effort is between \(Q_h(\frac{1}{2})\) and \(Q_h(P^*)\). Therefore, the length of shirking is between \(\min\{n : Q^n_l(Q_h(\frac{1}{2})) \leq P^*\} = \sigma\) and \(\min\{n : Q^n_h(P^*) \leq P^*\} \geq \sigma\). We will show
\[
\min\{n : Q^n_l(Q_h(P^*)) \leq P^*\} \leq \sigma + 1 \text{ by arguing that:}
\]
\[
Q_{l+1}^n(Q_h(P^*)) \leq Q^n_l(Q_h(\frac{1}{2})) \leq Q^n_l(Q_h(P^*)) \tag{6}
\]

In particular, if we can show
\[
Q_l(Q_h(P^*)) < Q_h(\frac{1}{2}) < Q_h(P^*)
\]
then Lemma 3.1 will imply (6). We know \(Q_h(\frac{1}{2}) < Q_h(P^*)\). To see \(Q_l(Q_h(P^*)) < Q_h(\frac{1}{2})\) holds, first note that \(P^* < \Psi^h\) implies \(Q_h(P^*) > P^*\). Since \(P - Q_l(P) = \beta(P - \Psi^l)\) is increasing in \(P\),
\[
Q_h(P^*) - Q_l(Q_h(P^*)) > P^* - Q_l(P^*) \tag{7}
\]
Now, observe that \(P^* - Q_l(P^*) = P^* - \frac{1}{2}\) and \(Q_h(P^*) - Q_h(\frac{1}{2}) = (1 - \beta)(P^* - \frac{1}{2})\). Thus,
\[
P^* - Q_l(P^*) > Q_h(P^*) - Q_h(\frac{1}{2}) \tag{8}
\]
Equation (7) and (8) imply that \(Q_l(Q_h(P^*)) < Q_h(\frac{1}{2})\).

Now, suppose that \(Q_h(\frac{1}{2}) \leq P^*\). Since at the optimum, the vote share is no greater than \(Q_h(P^*)\), the vote share after \(l\) is between \(\frac{1}{2}\) and \(Q_l(Q_h(P^*)) = q\). Observe that since \(Q_h(P) - P\) is decreasing in \(P\),
\[
Q_h(P^*) - P^* < Q_h(\frac{1}{2}) - \frac{1}{2} \tag{9}
\]
Equation (7) and (9) implies that \(q < Q_h(\frac{1}{2})\) (note that \(P^* - Q_l(P^*) = P^* - \frac{1}{2}\)). Thus, the incumbent cannot shirk two periods in a row. The length of consecutive \(h\) that follows shirking is between \(\max\{n : Q^n_h(\frac{1}{2}) \leq P^*\} + 1\) and \(\max\{n : Q^n_h(q) \leq P^*\} + 1\). Similar to the argument for the first result, we will show that
\[
\max\{n : Q^n_h(q) \leq P^*\} \geq \max\{n : Q^n_h(\frac{1}{2}) \leq P^*\} - 1
\]
by proving \(q < Q_h(\frac{1}{2}) < Q_h(q)\) and then applying Lemma 3.1. We have already shown the first part of the inequality. \(Q_h(\frac{1}{2}) < Q_h(q)\) follows from the fact that \(q > \frac{1}{2}\).
Proof for Proposition 4.1

Proof. Because $Q^n_a(P)$ and $P^*$ are functions of $\Psi^h$ and $\Psi^l$, we will explicitly write $\Psi^h$ and $\Psi^l$ as arguments at various points in the following proof. Observe that for an arbitrary $n$ and $P$, $Q^n_a(P, \Psi^h)$ is increasing in $\Psi^h$, but $Q^n_a(P)$ and $P^*$ are unaffected by $\Psi^h$. We will first show that $e(\Psi^h, \Psi^l)$ is decreasing in $\Psi^h$. Suppose $\Psi^h > \Psi^h$, there are three cases to consider: If $Q_h(\frac{1}{2}, \Psi^h) \leq P^* < Q_h(\frac{1}{2}, \Psi^h)$, then by Proposition 3.2, we see that $e(\Psi^h, \Psi^l) < 1 \leq e(\Psi^h, \Psi^l)$.

1. If $Q_h(\frac{1}{2}, \Psi^h) < Q_h(\frac{1}{2}, \Psi^h) \leq P^*$, then it is straightforward to see that:

$$e(\Psi^h, \Psi^l) = \max\{n : Q^n_h(\frac{1}{2}, \Psi^h) \leq P^*\} \leq \max\{n : Q^n_h(\frac{1}{2}, \Psi^h) \leq P^*\} = e(\Psi^h, \Psi^l)$$

2. If $P^* < Q_h(\frac{1}{2}, \Psi^h) < Q_h(\frac{1}{2}, \Psi^h)$, then by Lemma 3.1, $Q^n_l(Q_h(\frac{1}{2}, \Psi^h)) < Q^n_l(Q_h(\frac{1}{2}, \Psi^h))$ for arbitrary $n$. It follows that:

$$\min\{n : Q^n_l(Q_h(\frac{1}{2}, \Psi^h)) \leq P^*\} < \min\{n : Q^n_l(Q_h(\frac{1}{2}, \Psi^h)) \leq P^*\}$$

Denote $\min\{n : Q^n_l(Q_h(\frac{1}{2}, \Psi^h)) \leq P^*\} = A$ and $\min\{n : Q^n_l(Q_h(\frac{1}{2}, \Psi^h)) \leq P^*\} = B$, then we see that $e(\Psi^h, \Psi^l) = \frac{1}{B+1} \leq \frac{1}{A+1} = e(\Psi^h, \Psi^l)$.

We now follow similar steps to show that $e(\Psi^h, \Psi^l)$ is decreasing in $\Psi^l$. Keep in mind that $Q^n(a, \Psi^l)$ is increasing in $\Psi^l$ and $P^*(\Psi^l)$ is decreasing in $\Psi^l$, but $Q_h(P)$ is unaffected by $\Psi^l$. Given $\Psi^h > \Psi^h$:

1. If $P^*(\Psi^l) < Q_h(\frac{1}{2}) \leq P^*(\Psi^l)$, then $e(\Psi^h, \Psi^l) < 1 \leq e(\Psi^h, \Psi^l)$.

2. If $P^*(\Psi^l) < P^*(\Psi^l) < Q_h(\frac{1}{2})$, then observe that:

$$\min\{n : Q^n_l(Q_h(\frac{1}{2}, \Psi^l)) \leq P^*(\Psi^l)) \leq \min\{n : Q^n_l(Q_h(\frac{1}{2}, \Psi^l)) \leq P^*(\Psi^l)}$$

Denote $A = \min\{n : Q^n_l(Q_h(\frac{1}{2}, \Psi^l)) \leq P^*(\Psi^l)}$ and $B = \min\{n : Q^n_l(Q_h(\frac{1}{2}, \Psi^l)) \leq P^*(\Psi^l)}$, we see $e(\Psi^h, \Psi^l) = \frac{1}{B+1} \leq \frac{1}{A+1} = e(\Psi^h, \Psi^l)$.

3. If $Q_h(\frac{1}{2}) \leq P^*(\Psi^l) < P^*(\Psi^l)$, then it is straightforward to see that:

$$e(\Psi^h, \Psi^l) = \max\{n : Q^n_h(\frac{1}{2}) \leq P^*(\Psi^l)} \leq \max\{n : Q^n_h(\frac{1}{2}) \leq P^*(\Psi^l)} = e(\Psi^h, \Psi^l)$$
Proof for Proposition 4.2

Proof. Observe that \( P^* = \left( \frac{1}{2} - \beta \Psi^l \right) \frac{1}{1-\beta} \) and the derivative

\[
\frac{dP^*}{d\beta} = \frac{1}{2(1-\beta)^2} - \Psi^l(1 - \beta) + \beta \Psi^l \frac{1}{(1-\beta)^2} = \frac{\frac{1}{2} - \Psi^l}{(1-\beta)^2}
\]

is strictly positive. Therefore, \( P^* \) is increasing in \( \beta \) and \( P^* \to \infty \) as \( \beta \to 1 \). Let \( \beta^* = \frac{\Psi^h - \frac{1}{2}}{\Psi^l - \Psi^h} \)

solve \( \left( \frac{1}{2} - \beta \Psi^l \right) \frac{1}{1-\beta} = \Psi^h \), then \( \beta \geq \beta^* \iff P^* \geq \Psi^h \). Therefore, by Proposition 3.1, \( e = \pi^h \) if and only if \( \beta \geq \beta^* \).

Note that \( Q_n^* \) is a function of \( \beta \) as well. For the rest of the proof, let \( Q_n(\beta) = Q_n^*(\frac{1}{2}, \beta) \). Note that \( \forall \beta < \beta^*, Q_n(\beta) < Q_n(\beta^*) < \Psi^h \), and \( P^* \to \Psi^h \) as \( \beta \to \beta^* \). Thus for \( \beta \) sufficiently close to \( \beta^* \), \( Q(\beta) < Q(\beta^*) < P^* \) and by Proposition 3.2:

\[
e(\beta) = \max\{ n : Q_n(\beta) \leq P^* \} \geq \max\{ n : Q_n(\beta^*) \leq P^* \}
\]

Since \( \max\{ n : Q_n(\beta^*) \leq P^* \} \to \infty \) as \( P^* \to \Psi^h \), \( e \to \pi^h \) as \( \beta \to \beta^* \).

Now, we will argue that \( e \) is weakly increasing in \( \beta \) in an interval around \( \beta^* \). The idea is to show that for \( \beta' \) sufficiently close to \( \beta^* \), \( \frac{dP^*}{d\beta'}(\beta') \) is greater than \( \frac{dQ^\kappa}{d\beta}(\beta') \), where \( \kappa = e(\beta') \).

This ensures that \( e(\beta) = \max\{ n : Q_n(\frac{1}{2}) \leq P^* \} \) is not decreasing in \( \beta \). Writing \( Q_n(\beta) \) explicitly, we have

\[
Q_n(\beta) = (1 - \beta)^n \frac{1}{2} + \beta \Psi^h \sum_{j=0}^{n-1} (1 - \beta)^j = \Psi^h + \left( \frac{1}{2} - \Psi^h \right)(1 - \beta)^n
\]

Thus:

\[
\frac{dQ^n}{d\beta} = n(\Psi^h - \frac{1}{2})(1 - \beta)^{n-1}
\]

Note that \( \frac{dQ^n}{d\beta} \) is decreasing in \( \beta \) and that \( \frac{dQ^n}{d\beta} \to 0 \) as \( n \to \infty \). We know from above that \( \frac{dP^*}{d\beta} \) is bounded away from zero, and \( \kappa \) can be made arbitrarily large by taking \( \beta' \) sufficiently close to \( \beta^* \). Therefore, for sufficiently large \( \beta' \), \( \frac{dQ^\kappa}{d\beta}(\beta') < \frac{dP^*}{d\beta}(\beta') \).

\( \square \)
Proof for Proposition 5.1

Proof. Note that there is always a positive probability of a series of consecutive bad shocks occurring, which implies that the incumbent is voted out even if he consistently exerts high efforts. Thus, the notion of permanent incumbency is no longer useful in the presence of common shocks.

For the first result, note that if \( P^{**} \geq 1 \), then the incumbent will be kicked out with probability \( q \) each period regardless of his efforts. Let \( P^* \) be as defined in Proposition 3.1. If \( P \leq P^* \), then the difference in probability of reelection between high effort and low effort is \( 1 - p - q \). Let \( v \) be the continuation value given high effort, the incumbent should choose high effort if \( \delta \cdot (1 - p - q) \cdot v \) is greater than \( h - l \). This is true if \( \delta \) is large and \( q \) small, since \( v \) would be large in that case. Finally, it is straightforward to see that if \( P > P^* \), then the two stage deviation argument applies, and therefore \( l \) is optimal if \( P > P^* \).

For the second result, we have to argue that a sufficiently patient incumbent has a strict incentive to maintain his reputation above \( P^{**} \). In particular, we want to show that the benefit of exerting high effort when \( P \) is below \( P^{**} \) (i.e. an greater of probability of being reelected in the future) outweighs the loss of utility due to high effort. Let \( v_P \) be the optimal discounted utility given reputation level \( P \) (note that \( v_P \) is increasing in \( P \)). Suppose \( P \) is such that \( Q(l, P) \leq P^{**} \leq Q(h, P) \), then if the incumbent chooses \( h \), he will have a higher probability of being reelected two periods from now than if he were to choose \( l \) (since the vote share next period will be above \( P^{**} \)). Let this difference in probability be \( \Delta \), then he should choose \( h \) if \( \delta \cdot \Delta \cdot v_P > h - l \) where \( v_P \) is the next period’s vote share given high effort (and no shocks occurring). Now, since \( w - h > 0 \), \( v_P \to \infty \) as \( \delta \to 1 \) and \( p + q \to 0 \). Therefore for sufficiently large \( \delta \) and small probability of common shocks, high effort will be optimal.

Now suppose \( P < P^{**} \) is such that \( Q(h, P) = Q_h(P) < P^{**} < Q^*_h(P) \), if the incumbent exert high effort, then given the step above, he will exert high effort again next period and his reputation will be above \( P^{**} \) after two periods (given no shocks occurring). Therefore, the incumbent’s probability of being reelected three periods into the future will be higher if he exert high effort now rather than low effort. Again, for sufficiently patient incumbent and low probability of common shock, this difference in probability is enough to induce high effort now. We can iterate the same argument for \( P \) where \( Q^*_h(P) < P^{**} < Q^*_h(P) \) for any \( n \).
Proof for Proposition 5.2

Proof. For $\beta = 1$, it is easy to see that the decision rule defined for maximizes the incumbent’s utility. For $\beta < 1$, note first that $P_{\tau}^{**}$ represents the lower bound on reputation needed for permanent incumbency (i.e. if $P_{\tau} \leq P_{\tau}^{**}$, then the incumbent cannot win the upcoming reelection even if he exerts high effort every period until the election). By the definition of $P_{\tau}$, we see that if $P_{\tau}^{**} < P_{\tau} \leq P_{\tau}^*$, then the incumbent must exert high effort at $\tau$, otherwise $P_{\tau-1} \leq P_{\tau-1}^{**}$ (define $P_{0}^{**} = \frac{1}{2}$). Now, to show that for $P_{\tau} > P_{\tau}^*$ the incumbent would shirk, we can use a two-stage deviation argument as in the proof of Proposition 3.1 (keep in mind that $Q(l, P_{\tau}) > P_{\tau-1}^{**}$ for $P_{\tau} > P_{\tau}^*$, so shirking is adequate in such a case). We omit the details for brevity.

Proof for Corollary 5.1

Proof. First it is straightforward to see that when $\beta = 1$, the incumbent only has to exert effort in the period immediately preceding the election i.e. $\tau = 1$. Thus, $m = 1$. Now, for $\beta < 1$, we will first argue that if the incumbent exert high effort at date $\tau \leq 2$, then he will exert high effort at date $\tau - 1$. In particular, we want to show that

$$Q(h, P_{\tau}^*) \leq P_{\tau-1}^*$$

since this will imply that for any $P_{\tau}^{**} < P_{\tau} \leq P_{\tau}^*$, $Q(h, P_{\tau}) \leq P_{\tau-1}^*$, and by our characterization of the optimal action, the incumbent will exert high effort at $\tau - 1$.

We shall show the following inequality holds:

$$Q(h, P_{\tau}^*) - P_{\tau}^* \leq P_{\tau-1}^* - P_{\tau}^*$$

Now, recall that $(1 - \beta)P_{\tau}^* + \beta \Psi^l = P_{\tau-1}^{**}$, and $(1 - \beta)P_{\tau-1}^* + \beta \Psi^l = P_{\tau-2}^{**}$ (where $P_{0}^{**} = \frac{1}{2}$). Subtract the two equations, we get:

$$P_{\tau-1}^* - P_{\tau}^* = \frac{P_{\tau-2}^{**} - P_{\tau-1}^{**}}{1 - \beta} > P_{\tau-2}^{**} - P_{\tau-1}^{**}.$$
Now, since \( P_{\tau - 2} = Q(h, P_{\tau - 1}^{**}) \), \( P_{\tau} > P_{\tau - 1}^{**} \) and \( P_{\tau - 1}^{**} < \frac{1}{2} < \Psi^h \), it must be that

\[
Q(h, P_{\tau}) - P_{\tau} \leq P_{\tau - 2}^{**} - P_{\tau - 1}^{**} < P_{\tau - 1}^{**} - P_{\tau}^*
\]

We will define \( \bar{m} = \inf \{ n : Q^h_n(0) > \frac{1}{2} \} \). That is, \( \bar{m} \) is the (smallest) number of consecutive high efforts that can guarantee reelection when initial reputation is 0. Thus \( \bar{m} \) is an upper bound for \( m \), i.e \( m \leq \bar{m} \). Observe that \( \bar{m} \) can be greater than \( k \) but does not depend on \( k \). It is straightforward to see from the definition of \( Q(\cdot, \cdot) \) that \( \bar{m} \) is decreasing in \( \beta \) and is equal to 1 when \( \beta = 1 \).

\[\square\]

**Proof for Proposition 5.3**

*Proof.* Without the loss of generality, assume that \( \theta_1 = D \). Suppose for now that \( R \) adopts as his strategy the optimal sequence conditional on permanent incumbency. Given this, \( D \) faces the same problem as in the baseline model, which means that for \( \delta \) large enough, it is optimal for \( D \) to stay in office forever. This in turn justifies the assumption of \( R \)'s strategy.

\[\square\]

**Proof for Proposition 5.4**

*Proof.* Suppose \( \theta_1 = D \), we will show that permanent incumbency is the dominant strategy for \( D \). By symmetry, permanent incumbency is the dominant strategy for \( R \) as well. Observe that \( D \)'s payoff of an arbitrary strategy \( S \) and an arbitrary \( R \)'s strategy is less than the payoff of an appropriately chosen alternative strategy \( S' \) and \( R \) shirking always. Since the \( D \)'s payoff under permanent incumbency is independent of \( R \)'s strategy, showing that permanent incumbency dominates all other strategies when \( R \) always shirks is sufficient to prove the dominance of permanent incumbency.

First, we will show that \( D \)'s best response to \( R \) shirking always is either permanent incumbency or shirking always. The argument for this is similar to the proof of Lemma 3.2. Let \( v_P \) be \( D \)'s maximal discounted payoff given initial distribution \( P \) and choosing \( h \). Let \( \tilde{v} \) be the \( D \)'s maximal discounted payoff for choosing \( l \) and losing the reelection. Note that \( \tilde{v} \) does not depend on the initial distribution of propensities because the propensities for a newly
elected incumbent is reset to $\frac{1}{2}$. Suppose the $D$'s best response is staying until period $N > 1$ and then quit, it must be that $v_{P_N} < \hat{v}$. However, the fact that $P_N > P_1$ and $D$ chose $h$ at date 1 means $v_{P_N} \geq v_{P_1} > \hat{v}$. A contradiction.

Observe that if it is optimal for $D$ to shirk at date 1, then it must be optimal to shirk whenever $D$ is elected to office. Thus, $D$’s best response to $R$’s strategy involves either permanent incumbency or shirking always. The payoff associated with permanent incumbency is at least $\frac{w-h}{1-\delta}$, while the payoff associated with shirking always is $\frac{w-l}{1-\delta}$. Therefore, if $\frac{w-h}{1-\delta} > \frac{w-l}{1-\delta}$, then permanent incumbency is the dominant strategy.

Proof for Corollary 5.2

Proof. If $\beta \geq \frac{\Psi^{h-\frac{1}{2}}}{w-h\Psi}$, then $l$ is never adequate. That means the payoff under permanent incumbency is exactly $\frac{w-h}{1-\delta}$. It follows from the proof of Proposition 5.4 that if $\frac{w-h}{1-\delta} \leq \frac{w-l}{1-\delta}$, then $D$’s best response to $R$ shirking always is to shirk always. This in turn justifies $R$ shirking always.
References


