Electoral Control with Behavioral Voters

Daniel Diermeier* and Christopher Li†

Abstract

We present a model of electoral control with behavioral voters. The model is intended to capture the main regularities of voting behavior found in empirical studies. Specifically, the voters' propensity to keep an incumbent in office is governed by a stochastic reinforcement process instead of strategic reasoning. The likelihood of a positive feed-back for a voter depends on the effort level exercised by the public official. We show that despite the lack of rational responses by voters, the electoral control of public officials can be substantial. Indeed, electoral control is the highest when voters are most forgetful. Moreover, our model generates comparative statics that are consistent with the main empirical regularities of electoral accountability.

JEL classification: D03, D72, D83
Keywords: electoral control, adaptive behavioral voting, electoral accountability

*IBM Professor of Regulation and Competitive Practice, Department of Managerial Economics and Decision Sciences (MEDS), Kellogg School of Management, Northwestern University, 2001 Sheridan Rd., Evanston, IL. 60208, d-diermeier@kellogg.northwestern.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 as Bayesians and act strategically, and the game form 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 is bad for 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 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 are 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 candidates but nevertheless influence voters' attitudes. Studies have shown that the voters' decision is 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

[4]These 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).

[5]Some 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 reflect 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]

The issues regarding voter sophistication have potentially important normative consequences for the assessment of democratic governance structures. Arguments going as far back as Plato's *Republic* have proclaimed that an ill-informed and irrational public renders democracy unsuitable as a form of government.[7] A recent version of this criticism was formulated by Achen and Bartels (2004b; p. 38):

> Democratic government as practiced in the United States, then, is a form of limited, random oligarchy. It is an *oligarchy* because, year to year, the voters are not paying any attention and thus have no say in what the government does. Only at election time do they assess how they feel. At that point, the oligarchy becomes *random*, because the choice between alternative governing teams often comes down to accidental and arbitrary criteria such as droughts and recessions,

[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 (irrational) voting rule 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).

which no government can affect, and aspects of the candidates' personal histories and personalities with no relevance to the job.

In other words, the public control of officials does not function in an environment of an ignorant and uninterested public; electoral accountability depends on an educated and engaged populace.[8]

In this paper, we assess whether such conclusions are valid. Much of the existing debate has focused on the empirical validity of the rational voting models. Behaviorally oriented researchers have provided evidence that contradicts the assumption of voter rationality. Rational choice oriented scholars have either tried to undermine the existing evidence against rational voters or tried to argue that rational choice models of electoral control do a satisfactory job of accounting for empirical data (e.g. Ashworth 2012, Ashworth and Bueno de Mesquita 2013a).[9]

In this paper, we take a different approach. We do not engage in a debate on the degree of rationality, information, and engagement found in the general public. Instead, we will explore whether electoral accountability can exist even when the voters are not rational. 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 by the empirical studies.[10]

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

[8] 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).

[9] In a recent paper Healy and Lenz (2014) provide evidence that the myopia exhibited by voters is based on the misinterpretation of information by voters, not rational disregard of an incumbent's early performance. They show first that voters base they voting choices on election year performance alone as suggested by the economic voting literature. But if they are presented with cumulative economic performance over the entire period of incumbency voters use the more comprehensive information instead.

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

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 as opposed to rational utility maximization. The reinforcement process we use satisfy intuitive properties and is well founded in the psychological learning literature. 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.[11] 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 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 show that electoral control of public officials, as measured by voter welfare, may exceed that of an environment with rational voters. We also obtain some of the comparative statics that are known in the electoral control literature and have been supported by empirical studies. For example, higher pay-offs (or lower cost of effort) improves electoral control, i.e. higher effort by elected officials (e.g. Ferejohn (1986) and Ferraz and Finan 2009).[12] There is also evidence that more informative signals of effort increase electoral control (e.g. Berry and Howell 2007, Snyder and Stromberg 2010). In our model, we show that electoral control is enhanced if the voters pay attention to factors that are informative of low effort but uninformative of high effort.

Because of the behavioral interpretation of our model, we can explore issues that are novel to the electoral control literature. One such issue is the relationship between voter memory and electoral control. The empirical literature on electoral behavior has argued that incumbent performances close to the date of the the election have a disproportionately large effect on

[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.

[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).

voter decision. One explanation is that voters have limited memory. The behavioral framework allows us to parametrize voter forgetfulness. We find that when elections are frequent, higher levels of forgetfulness may be beneficial to electoral control. Indeed, when election occurs 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 (see Section 5.3)

In the next section we define the baseline model. Section 3 characterizes the optimal behavior of the public officials. Section 4 discusses the implications of our model in light of comparative statics. Section 5 explores some generalizations of our model. In particular, we show the robustness of our results to heterogeneous feed-back and common shocks as well as discuss the implications of recurring candidates and multi-period incumbency. Section 6 concludes.

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 is the incumbent from period $t-1$, denoted θ_{t-1}, and a challenger γ_t (θ_1 is given) 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 θ_{t-1} at the date t election with probability (propensity) $p_{i,t} \in [0,1]$. We denote the vote share for θ_{t-1} at the date t election as $P_t = \int_i p_{i,t}$. We assume that θ_{t-1} is reelected if $P_t > \frac{1}{2}$ (i.e. majority rule). Otherwise the challenger γ_t becomes the new incumbent θ_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, $\pi^a_{i,t} \in \mathbb{R}$ for voter i, which in turn determines i'th propensity to reelect the incumbent (to be specified below). Date t utility for θ_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.

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 follows an adaptive learning heuristic often called 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 feed-back. In the current context, this means that the voters increase the propensity of reelecting the incumbent if the payoff is satisfactory. Formally, let π^* be the threshold for satisfactory payoffs, then:

$$\begin{array}{ll} p_{i,t+1} > p_{i,t} & \text{if } \pi^a_{i,t} > \pi^* \\ p_{i,t+1} < p_{i,t} & \text{if } \pi^a_{i,t} < \pi^* \\ p_{i,t+1} = p_{i,t} & \text{if } \pi^a_{i,t} = \pi^* \end{array} \tag{1}$$

We assume that i'th payoff is correlated with the incumbent's effort by imposing that $\pi^a_{i,t} = \pi^a + \epsilon_i$ where $\pi^h > \pi^* > \pi^l$ are constants and $\{\epsilon_i\}$ are iid random variables with median zero.[13] ϵ_i embodies noisy events that are outside of incumbent's control but nonetheless affect voters decision (e.g. shark attacks). For simplicity, we assume ϵ_i contains no atoms and therefore the case of $\pi^a_{i,t} = \pi^*$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:

$$p_{i,t+1} = \begin{cases} (1-\beta)p_{i,t} + \beta & \text{if } \pi^a_{i,t} > \pi^* \\ (1-\beta)p_{i,t} & \text{if } \pi^a_{i,t} < \pi^* \end{cases} \tag{2}$$

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 (zero) if the payoff is (not) satisfactory. It is straightforward to see (1) are satisfied.

Note that a higher β 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. they are "satisficers" as in Simon (1955).[14] Thus, β captures the voter's forgetfulness, or more generally their tendency to place

[13]It will be clear in what follows that assuming $\{\epsilon_i\}$ being identical is without the loss of generality (see Section 5.1). The independence assumption is relaxed in Section 5.2.

[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).

greater weight on more recent experiences. By varying β, we can explore the relationship between forgetfulness and electoral control.

It is clear from equations (2) that two (probabilistic) events are of importance: $\pi^a_{i,t} > \pi^*$ and $\pi^a_{i,t} < \pi^*$. We refer to the former as G(ood experience) and the latter B(ad experience). Define $\Psi^a = \Pr(\pi^a_{i,t} > \pi^*)$. From the incumbent's point of view, $\{\Psi^h, \Psi^l\}$ are the only relevant properties of $\{\pi^a_{i,t}\}$, and the assumptions on $\pi^a_{i,t}$ imply that $\Psi^h > \frac{1}{2} > \Psi^l$. One can interpret $\{\Psi^h, \Psi^l\}$ as a measure of the impact of noisy events (e.g. the weather, the global oil prices, or local sporting events) on voter behavior. Suppose such random events are rare, i.e. ϵ_i has low variance, or the incumbent's effort has great impact on payoff relevant events, i.e. $\pi^h - \pi^l$ is large, then Ψ^h would be close to 1 and Ψ^l to 0. On the other hand, if irrelevant events have a large impact on voter's behavior, then both Ψ^h and Ψ^l would be close to $\frac{1}{2}$.

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, θ_t's vote share at the date $t+1$ election is [[Chris, formula below looks wrong. Please check. May be my word processor]] [[Fixed, I think sometimes some latex-based programs does weird things when import/export latex files]

$$\begin{aligned} Q(a_t, P_t) &= \Psi^{a_t} \int_i [(1-\beta)p_{i,t} + \beta] + (1 - \Psi^{a_t}) \int_i (1-\beta)p_{i,t} \\ &= (1-\beta)P_t + \beta\Psi^{a_t} \end{aligned} \tag{3}$$

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. ϵ_i is degenerate at 0), the vote share dynamic is exactly the Bush-Mosteller process (i.e. $\Psi^h = 1$ and $\Psi^l = 0$). Interestingly, adding noise to voter payoff is not necessarily harmful for electoral control (see the discussion in Section 4.2).

Observe that if P_t is sufficiently large, then θ_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 5.4. 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.

Lemma 3.2 below is a simple observation that in 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, θ_1 either shirks in the first period and is voted out, or will remain in office forever. Furthermore, if it is optimal for θ_1 to shirk in the first period, then for all $t > 1$, θ_t shirks as well.*

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 $v(P_N) \geq v(P) \geq w - l$. This is a contradiction. Finally, if the vote share is $\frac{1}{2}$, l is not adequate. Therefore, if θ_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 θ_1. It follows that if it is optimal for θ_1 to shirk, then it is optimal for $\theta_{t>1}$ to shirk.

$\square$

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.[15] 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.

Lemma 3.3. *Permanent incumbency is optimal if* $\delta > 1 - \frac{w-h}{w-l}$.

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 any given period is a function of the vote share at that period. In particular, the function is a cut-off rule.

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.*

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:

$$Q(h,P) - Q(l,P) = \beta(\Psi^h - \Psi^l) \tag{4}$$

[15] Lemma 3.2 is not a crucial part of our insight as electoral control can be maintained in the absence of permanent incumbency in a more general setting (see section 5.2)

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:

$$(1-\beta)^j(P-P') \tag{5}$$

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 Ψ^h. Thus if the threshold is larger than Ψ^h, the incumbent never shirks. If $P^* < \Psi^h$, then if incumbent exert high effort for many periods, the vote share will converge to Ψ^h and therefore exceed P^* at some point. It follows that the incumbent will shirk infinitely often at the optimum.

□

Since the vote share aggregates the voters' propensities, one may interpret it as a measure of the incumbent's "political capital". [[Chris, I think we can take the footnote out if we just drop "reputation", as I did, which may create confusion anyway.]][[Agreed]] 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). If public approval is sufficiently high, the incumbent will exploit his political capital and shirk.

Observe that given initial propensities, one can compute the sequence of optimal efforts following procedure in Proposition 3.1. In the case of $P^* < \Psi^h$, it is difficult to infer much about the optimal efforts except that efforts are cyclical. Proposition 3.2 below gives a finer characterization of the optimal sequence of efforts. In particular, the effort 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 $Q_a^n(P) = \underbrace{Q_a \circ Q_a \ldots \circ Q_a}_{n}(P)$.

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

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

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 $\min\{n : Q_l^n(Q_h(\frac{1}{2})) \leq P^*\} = \sigma$ and $\min\{n : Q_l^n(Q_h(P^*)) \leq P^*\} \geq \sigma$. We will show $\min\{n : Q_l^n(Q_h(P^*)) \leq P^*\} \leq \sigma + 1$ by arguing that:

$$Q_l^{\sigma+1}(Q_h(P^*)) \leq Q_l^{\sigma}(Q_h(\frac{1}{2})) \leq Q_l^{\sigma}(Q_h(P^*)) \tag{6}$$

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 $\max\{n : Q_h^n(\frac{1}{2}) \leq P^*\} + 1$ and $\max\{n : Q_h^n(q) \leq P^*\} + 1$. Similar to the argument for the first result, we will show that

$$\max\{n : Q_h^n(q) \leq P^*\} \geq \max\{n : Q_h^n(\frac{1}{2}) \leq P^*\} - 1$$

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 $q > \frac{1}{2}$.

□

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

Given effort level a, the one-period aggregate voter payoff is $\int_i \pi_{i,t}^a di$. For simplicity, assume that ϵ_i has mean zero and therefore $\int_i \pi_{i,t}^a di = \pi^a$. Given a sequence of effort from the incumbent, one can compute the long-run aggregate voter payoff, defined as:

$$\Pi\left(\{a_t\}\right) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \pi^{a_t}$$

We defined the long-run payoff as the limit of means for notational simplicity. The results would not change if we use discounted payoff instead. Note that since the length of effort cycles is bounded at the optimum, the corresponding long-run voter payoff, denoted Π^*, 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 σ, and therefore the bounds can be a good approximation for Π^*.

We use the lower bound of Π^*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* Π^*, *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 σ *is defined in Proposition 3.2.*

Observe that e is maximized (i.e. $e = \pi^h$) if the incumbent never shirks. Proposition 3.1 shows that this is the case for some specifications of the parameters. Therefore, electoral control with behavioral voters can be *greater* than the electoral control with rational voters as in the Ferejohn model, where there is always a strictly positive probability that the incumbent shirks.

4 Comparative Statics

In this section, we explore the relationship between various primitives of our model and electoral control. In particular, we look at novel issues such as when voter ignorance and forgetfulness can improve electoral control. Note that since e is a strictly monotonic function of σ, we will often treat σ as electoral control to simplify notations in the proofs.

4.1 The Value of Office

One of the fundamental insights provided by 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.

Corollary 4.1. *If* w *is sufficiently low or* $h - l$ *sufficiently high,* e *is minimal and equal to* π^l. *For high values of* w *or low values of* $h - l$, $e > \pi^l$.

One can see from Proposition 3.1 and Proposition 3.2 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.[16]

4.2 Informativeness of Efforts

A second important finding of the existing literature pertains to the informativeness of the signal observed by the electorate: electoral control improves if observed outcomes become more informative of incumbent's effort (e.g. Berry and Howell 2007, Snyder and Strömberg 2010, Ashworth 2012). Recall the interpretation of Ψ^h and Ψ^l: Ψ^h is large (Ψ^l small) if the impact of noisy events on voter behavior is small vis-a-vis factors within incumbent's control. Thus, one may interpret Ψ^h and Ψ^l as measures the informativeness of effort even though our voters are not Bayesians.

Unlike in standard models, the effect of informativeness of high and low effort on electoral control is asymmetric. In particular, electoral control improves when the electorate is influenced by factors that are informative of low effort but uninformative of high effort. This is intuitive because if the electorate is particularly sensitive to high effort, then it is easier for the politicians to build up political capital, which allows the politicians to shirk; while if the electorate is sensitive to low effort, then shirking is very costly for the politicians. Thus, to improve electoral control voters should be skeptical, but unforgiving. This is stated formally in the following proposition.

Proposition 4.1. *e is weakly decreasing in Ψ^h and Ψ^l.*

Proof. Because $Q_a^n(P)$ and P^* are functions of Ψ^h and Ψ^l, we will explicitly write Ψ^h and Ψ^l as arguments at various points in the following proof. Observe that for an arbitrary n and P, $Q_h^n(P, \Psi^h)$ is increasing in Ψ^h, but $Q_l^n(P)$ and P^* are unaffected by Ψ^h. We will first show that $e(\Psi^h, \Psi^l)$ is decreasing in Ψ^h. Suppose $\hat{\Psi}^h > \tilde{\Psi}^h$, there are three cases to consider:

If $Q_h(\frac{1}{2}, \tilde{\Psi}^h) \leq P^* < Q_h(\frac{1}{2}, \hat{\Psi}^h)$, then by Proposition 3.2, we see that $e(\hat{\Psi}^h, \Psi^l) < 1 \leq e(\tilde{\Psi}^h, \Psi^l)$.

[16] For the impact on term-limits see Besley and Case (1995), Ferraz and Finan (2008), Ferraz and Finan (2011), Alt, de Mesquita, and Rose (2011).

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

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

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

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

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

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

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

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

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

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

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

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

□

It follows immediately from the above proposition that e is weakly decreasing in π^h and π^l. However, the effect of greater noise (say an increase in the variance of ϵ_i) on electoral control is ambiguous. For illustration, suppose the noise ϵ_i is a normal random variable. An increase in the variance would increase Ψ^h while decrease Ψ^l. The former effect is harmful for electoral control while the latter is beneficial. The net effect would depend on other parameters of the model. For example, if $\pi^h - \pi^*$ is greater than $\pi^* - \pi^l$, then due to the nature of normal distribution, an increase in the variance of ϵ_i would have a greater marginal effect on Ψ^l than on Ψ^h. In this case, more noise is likely to be beneficial for electoral control.

4.3 Voter Memory

In this section, we explore a new dimension of electoral control: the connection between voter memory and electoral control. Recall our interpretation of β as a measure of voter's forgetfulness: a higher β means the date t experience has a greater weight on determining the date $t+1$ propensity. In the extreme case of satisficing ($\beta = 1$), 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 $t+2$, $t+3, \ldots$ is small (since the outcome in those elections is affected mostly by actions in $t+1, t+2, \ldots$). Consequently, the degree of voter forgetfulness determines how myopic or farsighted the incumbent is when choosing effort. And since for sufficiently large β, high effort is needed for reelection, a myopic incumbent would have very high incentive 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 β for $\beta \in (\mu, \frac{\Psi^h - \frac{1}{2}}{\Psi^h - \Psi^l})$.*

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

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

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

Note that Q_a^n is a function of β as well. For the rest of the proof, let $Q^n(\beta) = Q_h^n(\frac{1}{2}, \beta)$. Note that $\forall \beta < \beta^*, Q^n(\beta) < Q^n(\beta^*) < \Psi^h$, and $P^* \to \Psi^h$ as $\beta \to \beta^*$. Thus for β sufficiently close to β^*, $Q(\beta) < Q(\beta^*) < P^*$ 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 β in an interval around β^*. The idea is to show that for β' sufficiently close to β^*, $\frac{dP^*}{d\beta}(\beta')$ 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 β. Writing $Q^n(\beta)$ explicitly, we have

$$Q^n(\beta) = (1-\beta)^n \frac{1}{2} + \beta \Psi^h \sum_{j=0}^{n-1} (1-\beta)^j = \Psi^h + (\frac{1}{2} - \Psi^h)(1-\beta)^n$$

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 β 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 κ can be made arbitrarily large by taking β' sufficiently close to β^*. Therefore, for sufficiently large β', $\frac{dQ^\kappa}{d\beta}(\beta') < \frac{dP^*}{d\beta}(\beta')$.

□

Corollary 4.2. *$e = \pi^h$ if the voters are satisficers i.e. $\beta = 1$*

Proof. $1 > \frac{\Psi^h - \frac{1}{2}}{\Psi^h - \Psi^l}$ by our assumption that $\Psi^h > \frac{1}{2} > \Psi^l$.

□

Note that the incumbent being farsighted (low β) does not necessarily imply less incentive for high effort because both high and low effort have long run affects on outcomes. This explain why we can establish monotonicity of e with respect to β for high levels of β. In sum, we have shown that highly forgetful voters as identified by Achen and Bartels (2004a) can induce high levels of electoral control, but, as we will show in Section 5.3, this conclusion does depend on the frequency of elections. Recall that 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 is likely to be bad for electoral control when elections are infrequent.

5 Generalizations

5.1 Heterogeneous Ψ^a

So far, we have assumed that the probability of a good experience occurring given effort a is the same across voters. This is without loss of generality, as the following result demonstrates.

Lemma 5.1. *Let Ψ_i^a 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 $\Psi^a = \int_i \Psi_i^a$.*

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:

$$\begin{aligned} Q(a, P) &= \int_i \Psi_i^a \left((1-\beta)p_i + \beta\right) + (1-\Psi_i^a)(1-\beta)p_i \\ &= \int_i \Psi_i^a \beta + (1-\beta)p_i \\ &= (1-\beta)P + \beta \int_i \Psi_i^a \end{aligned}$$

Since the operator Q is the only relevant information for the incumbent's problem, we see that electoral control in an environment with heterogeneous Ψ_i^a is equivalent to an environment with homogenous voters. □

5.2 Common Shocks

We have so far assumed that the realization of experience is independent across voters. In reality, however, it is possible that the experiences of voters are correlated, e.g. in response to macro-economic shocks. One 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 action. Note that there is always a positive probability of a series of consecutive bad shocks occuring, 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 Proposition 5.1 below, we define $P^{**} = \frac{1}{2(1-\beta)}$. If the incumbent has political capital greater than P^{**}, then he will be reelected in the next period even if a bad shock occurs (i.e. $(1-\beta)P^{**} = \frac{1}{2}$). We show that the presence of common shocks does not significantly change the incumbent's optimal behavior if such shocks occur with small probability.

Proposition 5.1. *Suppose that $q+p$ is sufficiently small and δ sufficiently large:*

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 $Q(l,P) \leq P^{**}$.*

Proof. 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 δ 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 $Q(l,P) \leq P^{**} < Q(h,P)$, 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 $\triangle$, then he should choose h if $\delta \cdot \triangle \cdot v_{P'} > h-l$ where P' is the next period's vote share given high effort (and no shocks occurring). Now, since $w-h>0$, $v_P \rightarrow \infty$ as $\delta \rightarrow 1$ and $p+q \rightarrow 0$. Therefore for sufficiently large δ and small probability of common shocks, high effort will be optimal. Now suppose $P < P^{**}$ is such that $Q(h,P) = Q_h(P) < P^{**} < Q_h^2(P)$, 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 $Q_h^{n-1}(P) < P^{**} < Q_h^n(P)$ for any n. $\square$

The second result in Proposition 5.1 suggests that 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 threhold for high effort is P^{**}, which may be greater than $\frac{1}{2}$. Intuitively, the incumbent wishes to insulate himself against a bad shock by maintaining higher political capital than in an environment without common shocks. Thus, he needs to exert high effort more often.

5.3 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 τ. 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 τ, the incumbent maintains $P_\tau > P_\tau^{**}$ and he shirks if and only if $P_\tau > P_\tau^*$. where P_τ^{**} and P_τ^* are defined as follows:*

If $\beta < 1$,

$$
\begin{aligned}
P_\tau^{**} & \quad \textit{is such that} \quad Q^\tau(h, P_\tau^{**}) = \frac{1}{2} \\
P_1^* & \quad \textit{is such that} \quad Q(l, P_1^*) = \frac{1}{2} \\
P_\tau^* & \quad \textit{is such that} \quad Q(l, P_\tau^*) = P_{\tau-1}^{**}
\end{aligned}
$$

If $\beta = 1$,

$$
\begin{aligned}
P_1^* &= 1 \\
P_{\tau>1}^* &= 0
\end{aligned}
$$

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_τ^{**} 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_τ^*, we see that if $P_\tau^{**} < P_\tau \leq P_\tau^*$, then the incumbent must exert high effort at τ, 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.

□

Because P_τ^{**} is decreasing in τ, P_τ^* is also decreasing in τ. Since P_τ^* is the threshold for shirking, this means that the incentive to shirk decreases as the election draws near (i.e. τ 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 β. 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 satisficers 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 β, and $\bar{m} = m = 1$ when $\beta = 1$.*

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 $\bar{m} = \inf\{n : Q^n_h(0) > \frac{1}{2}\}$. 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 β and is equal to 1 when $\beta = 1$.

$\square$

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

5.4 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 *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.*

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 δ large enough, it is optimal for D to stay in office forever. This in turn justifies the assumption of R's strategy. □

Clearly, if permanent incumbency is optimal in equilibrium, then θ_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.*

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 $v_{P_N} < \tilde{v}$. However, the fact that $P_N > P_1$ and D chose h at date 1 means $v_{P_N} \geq v_{P_1} > \tilde{v}$. 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 $\frac{w-l}{1-\delta^2}$. Therefore, if $\frac{w-h}{1-\delta} > \frac{w-l}{1-\delta^2} \iff \delta > \frac{w-l}{w-h} - 1$, then permanent incumbency is the dominant strategy.

□

Observe that the condition on δ 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 δ 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 $\beta \geq \frac{\Psi^h - \frac{1}{2}}{\Psi^h - \Psi^l}$, 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.*

Proof. If $\beta \geq \frac{\Psi^h - \frac{1}{2}}{\Psi^h - \Psi^l}$, 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 $\frac{w-h}{1-\delta} \leq \frac{w-l}{1-\delta^2} \iff \delta \leq \frac{w-l}{w-h} - 1$, then D's best response to R shirking always is to shirk always. This in turn justifies R shirking always.

$\square$

Note that $\beta \geq \frac{\Psi^h - \frac{1}{2}}{\Psi^h - \Psi^l}$ 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.[17] However, it is straightforward to see that the 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.[18] 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 democracy is lacking (e.g. Achen and Bartels 2004b). In political economy, much of the debate has centered on whether in reality, voter behavior is consistent with the rational choice foundation (e.g. Ashworth 2012) or not (e.g. Achen and Bartels 2004a) . In this paper, we set this debate aside and investigate the underlying premise of the controversy: electoral control of public officials requires a well-informed and rational electorate. To do so 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.

Our model can account for many identified empirical regularities such as the detrimental effect of lowering office benefits, e.g. by imposing term limits. We can also investigate the

[17]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.

[18]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.

importance of new factors such as voter forgetfulness. Surprisingly, if elections are frequently held, electoral control is highest if voters are most forgetful and ignore all but the most immediate past.

Our results suggest that the institution of electoral control of public officials may be far less dependent on the reasoning ability of the electorate than previously believed. Indeed, behavioral voters may achieve higher levels of electoral control than rational voters if elections occur frequently. This conclusion may no longer hold, however, for less frequent elections.

We also explore variois extensions and variations of the model, such as heterogeneity, common shocks, recurring candidates and multi-period incumbency. Most extensions confirm the main results of the paper. Novel issues, however, are raised by multi-period incumbency, which poses new questions about the interplay of attention, term length, and voter memory. Potential extensions include candidate quality as in adverse selection models or issues of policy choice and ideological candidates. Such questions, we hope, will be the subject of future research.

References

[1] Achen, C. & Bartels, L. (2004a). “Blind Retrospection: Electoral Responses to Drought, Flu, and Shark Attacks.â Unpublished manuscript. Princeton University.

[2] Achen, C., & Bartels, L., (2004b) “Musical Chairs: Pocketbook Voting and the Limits of Democratic Accountability.” Unpublished manuscript. Princeton University.

[3] Achen, C., & Bartels, L., (2002) “Ignorance and Bliss in Democratic Politics: Party Competition with Uninformed Voters.” Unpublished manuscript. Princeton University.

[4] Alt, J., Bueno de Mesquita E., & Rose S. (2011) “Disentangling Accountability and Competence in Elections: Evidence from US Term Limits.â *Journal of Politics* 73(1): 171-86.

[5] Andonie, C & Diermeier, D. (2012). “A Behavioral Model of Multi-Candidate Elections”. working paper.

[6] Ashworth, S. (2012). “Electoral Accountability: Recent Theoretical and Empirical Work”. *Annual Review of Political Science* 15:183-201

[7] Ashworth, S. (2005). “Reputational Dynamics and Political Careersâ *Journal of Law Economics and Organization* 21:441-66

[8] Ashworth, S., and Ethan Bueno de Mesquita. (2008). “Electoral Selection, Strategic Challenger Entry, and the Incumbency Advantage.â Journal of Politics 70 (4): 1006–25.

[9] Ashworth, S., and de Mesquita. (2013a). "Disasters and Incumbent Electoral Fortunes: No Implications for Democratic Competence." Mimeo. University of Chicago

[10] Ashworth, S., and de Mesquita. (2013b). "Is Voter Competence Good for Voters?: Information, Rationality, and Democratic Performance." Mimeo. University of Chicago

[11] Ashworth, S., de Mesquita, E., & Friedenberg, A. (2012). “Accountability Trapsâ. mimeo. University of Chicago.

[12] Bailenson, J., Iyengar, S., Yee, N., & Collins, N. (2009). “Facial Similarity between Candidates and Voters Causes Influence.â*Public Opinion Quarterly* 72: 935-61.

[13] Barro, R. (1973). “The Control of Politicians: An Economic Model.â *Public Choice* 14 (1): 19–42

[14] Bartels, L.M., (2008). *Unequal Democracy.* Princeton: Princeton University Press.

[15] Bartels, L. M., & Zaller, J. (2001) “Presidential Vote Models: A Recount.â *PS: Political Science and Politics* 34: 9-20.

[16] Bellew, C. C., II. & Todorov, A. (2007) “Predicting Political Elections from Rapid and Unreflective Face Judgments,â *Proceedings of the National Academy of Sciences* 104: 17948-53.

[17] Bendor, J., Kumar, S.,&Siegel,D. (2010). “Adaptively Rational Retrospective Voting”. *Journal of Theoretical Politics*, 22: 26-63

[18] Bendor, J. Diermeier, D., Siegel, D., & Ting, M. (2011). “A Behavioral Theory of Elections” Princeton University Press

[19] Bendor, J., Diermeier, D.,& Ting, M. (2003). “A Behavioral Model of Turnout”. *American Political Science Review*, 97(2).

[20] Berelson, B.R., Lazarsfeld, P.F., & McPhee, W.N. (1954). *Voting: A Study of Opinion Formation in a Presidential Campaign.* Chicago: University of Chicago Press.

[21] Berry, C. R., & Howell, W. G. (2007) “Accountability and Local Elections: Rethinking Retrospective Voting.â *Journal of Politics* 69(3): 844-58.

[22] Besley, T. & Case, A. (1995) “Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatiorial Term Limits.â *Quarterly Journal of Economics.* 110(3): 769-98.

[23] Bisin, A., Lizzeri, A., & Yariv, L., (2011) “Government Policy with Time Inconsistent Voters.â Working paper.

[24] Board, S., and Moritz Meyer-Ter-Vehn. (2013). “Reputation for Quality.â *Econometrica* 81 (6): 2381-2462.

[25] Borgers, T., Sarin, R. (2000). “Naive Reinforcement Learning with Endogenous Aspirations”. *International Economic Review*, 41: 921-950.

[26] Bush, R., Mosteller, F. “Stochastic Models of Learning,” Wiley, New York, 1955.

[27] Campbell, A., Converse, P., Miller, W., & Stokes, D. (1960). *The American Voter*. New York: John Wiley & Sons.

[28] Cole, S., Healy, A., & Werker, E. (2012). "Do Voters Demand Responsive Governments? Evidence from Indian Disaster Relief." *Journal of Development Economics* 97: 167-181.

[29] Dahl, R. A. (1989). *Democracy and Its Critics*. New Haven: Yale University Press.

[30] Duch, R. M., & Stevenson, R. D. (2008) *The Economic Vote*. New York: Cambridge University Press.

[31] Erikson, R. S. (1990) "Economic Conditions and the Congressional Vote: A Review of the Macro-Level Evidence.â *American Journal of Political Science* 34: 373-99.

[32] Erikson, R.S. (1989). "Economic Conditions and the Presidential Vote.â*American Political Science Review* 83: 568 -73.

[33] Fair, R. (1978). "The Effect of Economic Events on Votes for President.â *Review of Economics and Statistics* 60: 159-72.

[34] Fearon, J. (1999). "Electoral Accountability and the Control of Politicians.â In *Democracy, Accountability and Representation*, eds. Adam Przeworski, Susan C. Stokes, and Bernard Manin. New York: Cambridge University Press, 55–97

[35] Ferejohn, J. (1986). "Incumbent Performance and Electoral Control". *Public Choice* 50:5-25

[36] Ferraz, C., & Finan, F. (2011) "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments.â *American Economic Review* 101(4): 1274-311.

[37] Ferraz, C., & Finan, F. (2009) "Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance.â Working Paper.

[38] Ferraz, C., & Finan, F. (2008) "Exposing Corrupt Politicians: The Effect of Brazil's Publicly Released Audits on Electoral Outcomes.â*Quarterly Journal of Economics* 123(2): 703-45.

[39] Gasper, J.T. & Reeves, A. (2011). "Make it Rain? Retrospection and the Attentive Electorate in the Context of Natural Disasters." *American Journal of Political Science* 55(2): 340-355.

[40] Harrington Jr, J. (1993) "Economic Policy, Economic Performance and Elections" *American Economic Review* 83:27-42

[41] Healy, A., Mo, C., & Malhotra, N. (2010). "Irrelevant Events Affect Voters' Evaluations of Government Performance.â *Proceedings of the National Academy of Sciences* 107(28): 12506-12511.

[42] Healy A, and Gabriel S. Lenz (2014) "Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy" *American Journal of Political Science* 58 (1):31-47

[43] Hibbs, D. A. Jr. (2006). "Voting and the Macroeconomy.âIn Weingast, B. R. and Wittmen, D., eds., *The Oxford Handbook of Political Economy*, 565-586. Oxford: Oxford University Press.

[44] Hilgard, E.R. & Bower, G.H. (1966). *Theories of Learning*. New York: Appleton-Century-Crofts.

[45] Huber, G.A., Hill, S.J., Lenz, G. S. (2012) "Sources of Bias in Retrospective Decision Making: Experimental Evidence on Voters' Limitations in Controlling Incumbents.â *American Political Science Review* 106(4): 720-41.

[46] Karandikar, R., Mookherjee, D.,& Ray, D. (1998). "Evolving Aspirations and Cooperation". *Journal of Economic Theory* 80: 292-331

[47] Kramer, G. (1971). "Short-Term Fluctuations in U.S. Voting Behavior.â *American Political Science Review* 65: 131-43.

[48] Lakoff, G. (2012). "Dumb and Dumber.â *Foreign Policy.com.*

[49] Lau, R.R., & Redlawsk, D.P. (2006). *How Voters Decide: Information Processing during Election Campaigns*. New York: Cambridge University Press.

[50] Lewis-Beck, M. (1988). *Economics and Elections: The Major Western Democracies*. Ann Arbor: University of Michigan Press.

[51] Lizzeri, A., amd Yariv, L., (2012) "Collective Self-Control.â working paper.

[52] Lupia, A., and McCubbins, M. D., (1998). *The Democratic Dilemma: Can Citizens Learn What They Need to Know*? New York: Cambridge University Press.

[53] Popkin, S.L. (1991). *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.

[54] Riker, W.H. (1982). “The Two-Party System and Duverger’s Law.â*American Political Science Review* 76: 753-66.

[55] Rogoff, K. (1990) .“Equilibrium Political Budget Cycles”. *American Economic Review* 80:21-36

[56] Simon, H. (1955). “A Behavioral Model of Rational Choice.â *Quarterly Journal of Economics* 69: 99-118.

[57] Stimson, J. (2004). *Tides of Consent: How Public Opinion Shapes American Politics.* New York: Cambridge University Press.

[58] Snyder, J.M. & Stromberg, D. (2010) “Press Coverage and Political Accountability.â *Journal of Political Economy* 118(2): 355-408.

[59] Todorov, A., Mandisodza, A. N., Goren, A., and Hall, C.C. (2005) “Inferences of Competence from Faces Predict Election Outcomes,â*Science* 308: 1623-1626.

[60] Tufte E. (1978). *Political Control of the Economy.* Princeton : Princeton University Press.

[61] Wolfers, J. (2007). “Are Voters Rational? Evidence from Gubernatorial Elections.â Unpublished manuscript.