Rational Parties and Retrospective Voters

Jonathan Bendor [*]

Sunil Kumar [†]

David A Siegel [‡]

March 12, 2007

Abstract

Many elections specialists take seriously V.O.Key's hypothesis (1966) that much voting is retrospective: citizens reward good performance by becoming more likely to vote for the incumbent and punish bad performance by becoming less likely. Earlier (Bendor, Siegel, and Kumar 2005) we formalized Key's verbal theory. Our model shows that people endogenously develop partisan voting tendencies, even if they lack explicit ideologies. However, that paper depicts parties as passive payoff-generating mechanisms. Here we make parties active, rational players with conventional goals: they either are pure office-seekers or have the usual mix of goals (office and policy preferences). The parties' optimal strategies reflect the incentives produced by retrospective voting. These incentives are powerful: for a wide range of parameter values they induce parties to select policies that differ not only from the median of the distribution of voter

---

[*]Professor, Graduate School of Business, 518 Memorial Way, Stanford University Stanford, CA 94305; E-mail: bendor_jonathan@gsb.stanford.edu; Phone: 650-723-4480.

[†]Professor, Graduate School of Business, 518 Memorial Way, Stanford University Stanford, CA 94305; E-mail: skumar@stanford.edu; Phone: 650-725-9220.

[‡]Assistant Professor, Department of Political Science, Florida State University, Tallahassee, FL 32306-2230; E-mail: dsiegel@fsu.edu; Phone: 850-645-0083.

ideal points, but also from the *mean*. Further, by analyzing the complex dynamics of voter adaptation and party response, we can derive and characterize the endogenous incumbency advantage enjoyed by the party in power. We establish these properties both analytically and computationally.

1 Introduction

Most models of party competition assume fully rational voters. Many also assume that citizens are well-informed. These are shaky premises. Decades of empirical research have shown that few American voters have coherent, detailed ideologies and few know much about politics. Donald Kinder summarized what we know about citizens' thinking about politics: "Precious few Americans make sophisticated use of political abstraction. Most are mystified by or at least indifferent to standard ideological concepts, and not many express consistently liberal, conservative, or centrist positions on government policy" (1999, p.796). Regarding information he reports that "the depth of ignorance demonstrated by modern mass publics can be quite breathtaking" and "the number of Americans who garble the most elementary points is... impressive" (p.785). Luskin's summary is harsher: most voters "know jaw-droppingly little about politics" (2002, p.282; see also Delli Carpini and Keeter 1996).

Years ago V.O.Key (1966) sketched out a theory of *retrospective voting* that appears to be consistent with the empirical regularities described by Kinder. His basic idea was elegantly simple. Citizens don't need well-worked out ideologies or realistic theories about how programs generate outcomes. Instead, they can decide how to vote for by assessing the performance of incumbents. Incumbents who have done well are rewarded by electoral support; those who have done poorly get fewer votes.

This is a plausible idea, but like many verbal theories it is somewhat vague and incomplete. In particular, how do voters evaluate governmental performance? How do they decide that an incumbent has performed poorly or well? And what are the effects of retrospective voting, either microscopic (e.g., the voting trajectories of individual citizens) or macroscopic (e.g., electoral outcomes)? In other work (Bendor, Kumar, and Siegel 2005; henceforth BKS 2005) we have addressed these and related questions by developing a deductive model of

retrospective voting.[1]

However, in BKS 2005 the parties are represented as passive, payoff-generating mechanisms. This assumption makes the model more tractable but is clearly only a way-station to a more plausible one. In this chapter we construct a model of party competition given retrospective voting.[2] Parties understand how citizens behave and they respond rationally to the incentives created by retrospective voting, as in Achen and Bartels (2002).[3]

In most of the chapter we allow the aspirations of citizens to be endogenous, responding to experience (realized payoffs) in a way that is conventional in the emerging literature on aspiration-based models of behavior. (See Bendor, Mookherjee and Ray 2001 for a review

---

[1]By focusing on how voters respond to realized payoffs, the present chapter is similar to work on retrospective voting stimulated by Ferejohn's (1986) seminal paper. However, our model differs in several fundamental ways from such principal-agent formulations. Most notably, (1) we don't replace the empirically-unreasonable Downsian informational assumptions by equally heroic assumptions about voters' rationality, (2) we consider a substantial population of voters, and (3) we examine dynamics away from the steady-state.

[2]For clarity and simplicity we assume purely retrospective voting: the citizens' votes are based totally on politicians' past performances. Fortunately, one can easily prove that most of our results are robust: if electoral choice is a weighted average of retrospective and prospective voting then they continue to hold if most (but not all) of the weight is on the past.

[3]The present chapter and that of Achen and Bartels (A-B) have the same goal: to analyze how retrospective voting affects the behavior of incumbents. The models differ in their assumptions about voters. In A-B's multiple elections, individuals vote deterministically for the party that is expected to produce lesser losses when in office, as deduced via Bayesian updating from their priors and the payoffs received in previous periods. Though they get only a noisy signal of parties' behavior in office, voters make the best of it, collating information gleaned from payoffs under both parties over time into two separate measures of their respective expected performances in the future. Asymptotically, the voters become perfect Bayesians (p. 23). In contrast, voters in the BKS 2005 model are less sophisticated cognitively: they are assumed only to update a single aspiration level, and view repeated elections less as an opportunity to develop better expectations on each parties' behavior than as a series of mostly independent payoff draws about which they can be more or less satisfied. What learning there is occurs more indirectly, as aspirations adjust to match realizations over time, and the frequency of being satisfied drives their probabilistic voting behavior, which never asymptotes to optimality.

of this literature.) Models with endogenous aspirations are notoriously hard to solve analytically; hence, we construct a computational model as well as an anyalytical one. In both we focus on the policies implemented by incumbents and on how long they stay in office. In Section III we derive several analytical results, at the price of some ruthless simplification. In Section IV we turn to computation, complemented by simple analytic examples. This combined approach enables us to derive several key—and observably distinct—findings. Specifically, we find that voters' retrospective behavior often induces rational, office-seeking incumbents to locate away from not only the median voter (MV) but also from the mean of the voters' ideal point distribution. (Models with probabilistic voting tend to exhibit the former property but not the latter.). We characterize the length of time incumbents can expect to stay in office—an incumbency advantage derived endogenously from the basics of voter psychology rather than determined by exogenous factors. And we show why a party might want to play to its base even when it is not worried about turnout.

The rest of the chapter is organized as follows. Section II presents the general ideas. It provides a free-standing introduction to retrospective voting[4] and explains how candidates optimize in light of voters' behavior. Section III gives the analytical model; section IV, the computational one. Section V concludes.

2   General Ideas

Retrospective Voting.   This type of voting is based on voters' evaluating the performance of an incumbent—either the party in power or a specific office-holder. The heart of Key's theory is that voters reward good performance by becoming more inclined to vote for the incumbent and punish bad performance by becoming less inclined to support the incumbent.

However, Key did not clarify the meaning of 'good' and 'bad' performance. To make these notions more precise, BKS (2005) posited that voters have *aspirations* (Simon 1955,

---

[4]It presumes no knowledge of either V.O.Key's verbal theory or BKS 2005's formal model.

1956): internal evaluation-thresholds which code an incumbent's performance as good or bad, satisfactory or unsatisfactory. Once an incumbent's performance has been assessed in this manner, the *direction* of the voter's stance toward the incumbent official or party is determined: good performance is rewarded with a higher propensity (probability) of voting for the incumbent; bad, with reduced support. These properties are formalized by the following axioms, which define a class of *adaptive voting rules* (AVoRs).

In what follows, we assume that there are n voters. Let $\pi_{i,t}$ denote voter i's payoff in period t and a_{t-1}, his current aspiration level, inherited from the previous period. (We assume that elections are held at the beginning of a period, so today's realized payoff is compared to the inherited aspiration level.) There are two parties, $\{D, R\}$ and a citizen's propensity to vote for party D at the start of period t is denoted $p_{i,t-1}(D)$. We assume that $p_{i,t-1}(D) = 1 - p_{i,t-1}(R)$. That is, everyone votes for one of the two parties with probability one. Moreover the actual votes and thus the outcome of the election depends only on the propensities. Let W_t denote the winner of the election in t, determined stochastically from these propensities. Hence, the winning party is the incumbent both during period t (when it generates payoffs for voters), as well as during the election at the beginning of period $t+1$. We posit that the voters adjust their propensities to vote for the incumbent using rules that satisfy the following assumptions.

(A1) (positive feedback): If $\pi_{i,t} \geq a_{i,t-1}$ then $p_{i,t}(W_t) \geq p_{i,t-1}(W_t)$, and this conclusion holds strictly if $\pi_{i,t} > a_{i,t-1}$ and $p_{i,t-1}(W_t) < 1$.

(A2) (negative feedback): If $\pi_{i,t} < a_{i,t-1}$ then $p_{i,t}(W_t) \leq p_{i,t-1}(W_t)$, and this conclusion holds strictly if $p_{i,t-1}(W_t) > 0$.

Voters adjust aspirations via rules that satisfy (A3).

(A3) Each agent i has an aspiration level, $a_{i,t}$, which is updated so that the following conditions hold for all i, t, and all histories leading up to t:

1. If $\pi_{i,t} > a_{i,t-1}$ then $a_{i,t} \in (a_{i,t-1}, \pi_{i,t})$.

2. If $\pi_{i,t} = a_{i,t-1}$ then $a_{i,t} = a_{i,t-1}$.

3. If $\pi_{i,t} < a_{i,t-1}$ then $a_{i,t} \in (\pi_{i,t}, a_{i,t-1})$.

For simplicity we have made a specific modeling decision about what happens when payoffs exactly equal aspirations. Since this concerns a knife-edge circumstance, it isn't very important.

For tractability's sake we impose two other assumptions on the kinds of AVoRs voters may use.

(1) We restrict attention to AVoRs that are *deterministic*: given a particular history and a current state of affairs—in particular, a voter's current vote-propensity and his aspiration-payoff comparison—an AVoR must determine a unique new vote-propensity. For example, if $W_t = D$, $\pi_{i,t}$ is some specific value above $a_{i,t-1}$ and $p_{i,t-1}(D) = 0.8$, then $p_{i,t}$ must, with probability one, be some unique propensity value in $(.8, 1]$. (Bush-Mosteller rules, often used in psychological learning theories, are deterministic in this sense.)

(2) We examine only Markovian AVoRs: those in which adjustment of both voting propensities and aspirations in period t depend only the values of the state variables $(p_{i,t-1}(W_{t-1})$, $a_{i,t-1})$ at the beginning of the current period and on what happened in that period ($\pi_{i,t}$, W_t).

Finally, to avoid hardwiring any results, we confine attention to AVoRs that are *party-neutral*.[5] Hence, citizens must learn which party to support; such tendencies are not hard-wired by their adaptive rules.

Because all the AVoRs examined in this chapter are deterministic, Markovian and party-neutral, we will not mention these properties as specific assumptions in the results that

[5]For a formal definition of party-neutrality see BKS (2005). The following example illustrates the idea. Suppose citizens i and j, who are in different electorates, use the same retrospective AVoR. In t the incumbent in i's district is D; in j's, R. If $p_{i,t-1}(D) = p_{j,t-1}(R)$ and $\pi_{i,t} = \pi_{j,t}$, then party-neutrality requires that i and j respond identically to D and R, respectively: $p_{i,t}(D) = p_{j,t}(R)$. (Note: this presumes a deterministic AVoR.)

follow.

Optimal Responses to Retrospective Voting. Although below we consider candidates with different objectives (e.g., office-seeking versus ideological motives), it helps to fix ideas by sketching out the optimal behavior of just one type, the classical, purely office-oriented politician.

Knowing that citizens vote purely retrospectively, politicians understand that elections are referenda on the incumbent; challengers' actions don't matter. Hence, we focus on the former.

The decision problem confronting an office-oriented incumbent is simple to state: what policy should she implement in order to maximize the probability of winning the current election? (Carrying out this optimization can be quite involved, of course.) This unpacks as follows. Suppose voters have ideal points in R^n, with payoffs decreasing the further the incumbent's implemented policy is from one's bliss point.[6] In the benchmark context of complete information, an incumbent knows all this, and so can determine for any voter i the probability that implementing policy x will induce i to vote for him. So any contemplated policy produces a vector of such probabilities. The candidate then selects the policy that produces the best vector.[7]

In effect, then, the incumbent, as an agent of n adaptively rational principals, selects the policy that maximizes the probability that a majority of his bosses are satisfied with his performance.

We make optimization easier to attain (hence more plausible) by assuming throughout

---

[6]Per Stokes' critique (1963), our model of retrospective voting does not presume that this is how voters think about elections. It merely represents the relation between policies and payoffs. Voters get realized payoffs, compare these to aspirations, and so forth.

[7]As is common in electoral models, we will often assume that an office-oriented incumbent is maximizing expected vote share rather than the probability of winning the election. This is a conventional move, driven by tractability demands. It is well-established (Aranson, Hinich and Ordeshook 1974) that in some situations the two objectives are not equivalent.

that the incumbent is concerned only with the present election. Thus, his policy-selection is a myopic best response to the electoral environment created by retrospective voters.

3   Analytical Results

We begin by stipulating the class of payoff functions that we consider in this chapter, via (A4). Although (A4) is stronger than necessary for some of the analytical results obtained in this section, it is needed for the computational model. So we assume it here.[8]

(A4):   Each voter i has an ideal point x_i^* in R^d. If the incumbent implements a policy x_t, also in R^d, in period t, then the payoff for voter i is $\pi_{i,t} = f_i(||x_i^* - x_t||) + \theta_i$, where θ_i is a R-valued random variable and $f_i : R \to R$ is strictly decreasing, and $||x_i^* - x_t||$ denotes the Euclidean distance between x_i^* and x_t. We further assume that θ_i is non-degenerate and has finite mean and variance. Payoffs for the same voter in different periods, as well as the payoffs for different voters in the same period, are independent of each other. That is, the shocks θ_i are obtained from i.i.d draws for each i and t.

Proposition 1:   Suppose (A4) holds. Vote-propensities are adjusted by some mix of AVoRs that satisfy (A1)-(A3). The incumbent knows the above, and wants to maximize the probability of winning the current election. Then, picking any policy outside the convex hull of the set of voters' ideal points is a weakly dominated strategy.

Corollary 1: In addition to (A1-A4) suppose that there exists a uniquely optimal policy for the incumbent, x_t^*, in each period t. Then x_t^* is in the the convex hull of the set of voters' ideal points in every t.

The proofs of this and all other statements can be found in the appendix. Proposition 1 allows the incumbent to be ignorant of many facts: exactly how voters adjust vote-propensities

---

[8]Note that (A4) assumes stochastic payoffs. This is to enhance the empirical content of the model. Given deterministic payoffs, models of aspiration-based decision making say that almost anything can happen; i.e., folk theorems hold for such models (Bendor, Diermeier and Ting 2003a).

or aspirations, the shape of their utility functions, and so forth. Despite this uncertainty, an office-seeking incumbent knows that there's no reason to locate outside the Pareto efficient set. Here's the intuition. For any policy, x_o, that's outside the efficient set there exists another one, x', that is closer to all the voters' ideal points. So by (A4) x' delivers a vector of payoffs that, for *each* voter, first-order stochastically dominates the payoffs generated by x_o. Hence, no matter what is the value of (say) voter i's aspiration level, policy x' is at least as likely to satisfy it as x_o is.

Proposition 1 implies that the more similar are the voters (i.e., the closer their bliss points) the more tightly constrained is their agent (i.e. the smaller the set of policies from which a rational party will choose). At the extreme—voters share the same ideal point—and part (ii) holds, the office-seeking incumbent's behavior is completely constrained: he'll implement the common ideal point.[9]

However, no good deed goes unpunished. The next result shows that even a perfect

---

[9]Although this result holds in a very general setting, the assumption that parties optimize myopically is necessary here. To see why, consider the following simple example of an incumbent maximizing a sum of discounted future payoffs. Assume that $n = 1$ and the single voter (even for one voter, the algebra of the general case gets messy fast) initially has extremely low aspirations. Further, assume that he immediately resets his aspiration level to within epsilon of his most recent payoff, and that his payoff shock is symmetrically distributed around zero. If the voter's aspirations are sufficiently low then the incumbent party of the first period is almost guaranteed to win the next election, regardless of the position enacted. If this position is the voter's ideal point, however, then the probability of winning the election in period two falls all the way down to a fair coin flip in expectation: the party expects that the voter's new aspiration level would be satisfied exactly half the time if the incumbent were to again implement his ideal point. In contrast, enacting a policy some distance from the voter's ideal point would still produce a win in period one, but now the voter's aspiration after this election would stay at a more easily achieved level. Locating at the voter's ideal point during the second period would thus yield a considerably higher probability of victory than fifty percent, and so in period one a fully rational incumbent maximizing a discounted utility stream will not locate at the voter's ideal point (if future payoffs are not discounted too much, of course). But because solving such maximization problems gets extremely difficult even for a few voters, we think that assuming myopic optimization is eminently reasonable.

agent, who always does what's best for her principals, may be thrown out of office in every election.

Remark 1: Suppose the assumptions of proposition 1 hold. If the supports of the voters' payoff-shocks θ_i are not bounded below then every incumbent may be fired with positive probability, in any period t.

Assuming that payoff shocks are not bounded below is convenient but it doesn't drive the conclusion, as the next result shows. To sharpen this result, we assume that voters have the same ideal point. Even in this context—i.e., even when the incumbent is implementing what is unambiguously the best policy for all citizens—retrospective voting makes getting fired an ongoing risk for the agent. We use the notation $\underline{\pi}_i$ to denote citizen i's min payoff if the politician implements x_i^* and θ_i's support is bounded.

Remark 2: Suppose the assumptions of proposition 1 hold; further, the voters have the same ideal point and politicians don't play weakly dominated strategies. Suppose that for all i, each θ_i has a continuous density that is strictly positive over some bounded interval and zero elsewhere. If there is a date T such that $a_{i,T} > \underline{\pi}_i$ for a majority of the voters then in every date after T every incumbent may be fired with positive probability.

Remarks 1 and 2 suggest that even a politician who does exactly what s/he should be doing will be fired eventually *with certainty*. The next result shows that this is true, for a wide range of stochastic environments, if aspirations are the simple average of payoffs and citizens don't become arbitrarily sluggish in adjusting their vote propensities in response to negative feedback. The latter property is formalized by **(A2′)**, which strengthens **(A2)**.

(A2′) (negative feedback): If $\pi_{i,t} < a_{i,t-1}$ then with probability one $p_{i,t}(W_t) \leq p_{i,t-1}(W_t)$. Further, there exists an $\epsilon > 0$ such that for all t and all histories leading up to t if $p_{i,t-1}(W_t) > 0$ then $p_{i,t} \leq (1 - \epsilon)p_{i,t-1}(W_t)$.

Replacing (A2) by (A2') yields the following result.

Proposition 2: Suppose the assumptions of proposition 1 hold, as does **(A2′)**. Voters have the same ideal point and politicians don't play weakly dominated strategies. Then every incumbent is thrown out of office eventually with probability one if either (i) or (ii) obtains.

- **(i)** The payoff shocks (the θ_i's) have continuous densities, and aspirations are formed by a simple averaging rule: $a_{i,t} = \frac{a_{i,0} + \pi_{i,1} + \cdots + \pi_{i,t-1}}{t}$ for all i.

- **(ii)** All the θ_i's are discrete random variables with finitely many possible values. The aspiration adjustment rules satisfy (A3) and $a_{i,0}$ lies strictly between the minimum and maximum value that the payoffs can take.

Thus, even an incumbent who implements the electorate's common ideal point will be fired eventually.[10]

Proposition 2 does not tell us anything about the expected duration of an incumbent in office. Remark 3 allows us to estimate this, albeit under restrictive assumptions.[11]

Remark 3: Let n, the number of voters, be odd. Suppose the voters have the same bliss point $x_i^* \equiv x^*$ and identical loss functions f_i in (A4). Further suppose that they are simple satisficers: $p_{i,t}(W_t) = 1$ if $\pi_{i,t} \geq a_{i,t-1}$ and $p_{i,t}(W_t) = 0$ otherwise. If each θ_i is a continuous

---

[10]Assuming that voters have the same bliss point is analytically useful: it and the assumption that politicians avoid weakly dominated strategies together imply that the incumbent uses a stationary policy. This makes analyzing the long-run properties of aspirations tractable. Further, it sharpens the point that an even unambiguously best agent cannot forever satisfy principals who respond retrospectively. However, common ideal points is not a *necessary* condition for the conclusion to hold. One can show that even if every citizen has a distinct bliss point, any incumbent who after finitely many periods settles down on a stationary policy will be fired eventually with probability one. (We also suspect that any politician who settles down on any finite *set* of policies will eventually be thrown out of office, but this remains a conjecture.)

[11]For analytical convenience, remark 3 assumes that aspirations adjust immediately to payoffs: $a_{i,t} = \pi_{i,t-1}$. Strictly speaking this doesn't belong to the set of aspiration adjustment rules that satisfy (A3). However, continuity ensures that if citizens' aspirations adjust almost all the way to their most recent payoffs, then the incumbent will be fired with a probability that is very close to $\frac{1}{2}$.

r.v. with a density and aspirations adjust immediately ($a_{i,t} = \pi_{i,t-1}$), then in every election after $t = 1$ the incumbent will be fired with probability $\frac{1}{2}$.

Remark 3 tells us that the expected duration of an incumbent in office is two periods.[12] Under the assumptions of Remark 3, the electorate votes incumbents out quite frequently, with long runs being quite unlikely.

Although retrospective voters are ungrateful they are well-served when they are ideological clones of each other. Though the new incumbent realizes that the electorate does not understand the situation and will eventually fire him even though he is doing as well as is humanly possible, being office-seeking and rational he makes the best of a difficult situation: he implements the voters' common ideal point.

This underscores the value of representative democracy. Given informed agents, retrospectively voting principals can't do too much damage. Office-seeking politicians protect the public from itself.[13] But in direct democracy *the* voters' displeasure falls on a policy rather than a candidate. This is bad.

Remark 4: Suppose the assumptions of remark 2 hold, except that citizens vote directly for policies. If the status quo policy doesn't receive a majority of votes then it is replaced by some other policy.[14] Then in every period after T, every status quo policy is overthrown with positive probability.

Thus, the combination of direct democracy and retrospective voting can lead the citizenry to take up suboptimal policies repeatedly. In representative democracy, informed and rational office-oriented candidates protect the voters from themselves. They'll do so because pleasing the voters is the best way to continue enjoying the perks of office, per Adam Smith's

---

[12]We assume n is odd to avoid ties, of course. But if we use the convention that for n even the winner of a tied election is selected by a fair coin toss, then Remark 3's conclusion will still hold.

[13]Compare to Proposition 2 of Achen and Bartels (2002), which has a similar message under a different set of assumptions.

[14]For present purposes there is no need to be explict about how policies are replaced.

famous remark: "it is not from the benevolence of the butcher, the brewer or the baker, that we expect our dinner, but from their regard to their own interest" (quoted in Downs 1957, p.28). Hence, a central part of the Smith-Downs argument about representative democracy does not require that voters be fully rational, at least not in all contexts. Instead, what is critical is that the incentives facing an office-oriented politician line up with the citizens' interests.

These can diverge, of course, if there are agency problems. A rational and informed incumbent has the capacity to do what's best for the electorate; one who is office-oriented also has the motivation to do so. But if the incumbent also has policy preferences, and these diverge from those of, say, the median voter, then agency problems can arise. To analyze such issues we turn to the computational model.

4   Computational Results

The previous section provides analytical results that give us some insights into how rational, office-seeking parties will behave when facing retrospective voters. But many questions remain unanswered: Where, exactly, do parties locate? What are the advantages of incumbency in our setting? Do policy-motivated parties behave differently from those driven purely by a desire for the perks of office? To answer these questions we must give up a level of generality, so in this section we make specific assumptions about utility functions. Further, we abandon the requirement of analytic tractability, and turn instead to computation, bolstering the intuition so derived with simple analytic examples when appropriate.[15]

Assumptions

---

[15]We focus in this chapter on phenomenology (i.e., the range of outcomes observed in the system) rather than on understanding the effect of every model parameter on every other one. The latter analysis awaits future work; here we content ourselves with specifying behavior either common across model parameterizations, or dependent upon the same in a simple way.

The basic setting of the computational model is as described in Section II. Each voter maintains an independent propensity to vote for each party, and an aspiration level. Each period begins with a majority rule election, which produces a new incumbent. This party's position yields a payoff for each voter, which causes all voters to update their propensities either upward or downward if these payoffs are either greater than or less than aspirations, respectively. Voters also update their aspirations based on these payoffs. Finally, parties choose positions for the next election and the cycle continues.[16]   Exact specifications for behavioral rules are given by the following computational analogues to earlier assumptions:

(A1c) (positive feedback for D/negative feedback for R) $(1-\lambda)*p_{i,t}(D)+\lambda$, with $\lambda \in (0,1)$.

(A2c) (negative feedback for D/positive feedback for R) $(1-\lambda)*p_{i,t}(D)$, with $\lambda \in (0,1)$.

(A3c) (aspiration adjustment) $a_{i,t} = (1-\nu)*a_{i,t-1}+\nu*\pi_{i,t}$, with $\nu \in (0,1)$.

(A4c) (voter utilities) $\pi_{i,t} = -.5*((x_{i,t}-blissx_i)^2+(y_{i,t}-blissy_i)^2+\theta_{i,t}$, where $\theta \sim N(0,\sigma^2)$ and $(blissx_i, blissy_i)$ is voter i's ideal point.

(A5c) (party utilities) $pay_{j,t} = (1-partyIdeoLevel)*perq*voteshare(x_{j,t},y_{j,t})+partyIdeoLevel*$ $(voteshare(x_{j,t},y_{j,t})*(-.5*((x_{j,t}-blissx_j)^2+(y_{j,t}-blissy_j)^2))+(1-voteshare(x_{j,t},y_{j,t}))*$ $(-.5*((x_{\sim j,t}-blissx_{\sim j})^2+(y_{\sim j,t}-blissy_{\sim j})^2))$, where $partyIdeoLevel \in [0,1]$ determines how much the parties care about the policies enacted relative to the benefit of holding office, which is defined as $perq$. $voteshare(x_{j,t},y_{j,t})$ is the expected share of the votes received when choosing the policy $(x_{j,t}, y_{j,t})$.

Each voter begins with a 50-50 propensity to vote for either party, and an aspiration level comfortably between her minimum and maximum payoffs (though see below). Voters are purely adaptive: they respond to payoffs according to (A1c)-(A4c). While not strictly

---

[16]Note that this choice is made to maximize the utilities arising from the outcome not of the next election, but rather the one after it. No action a party takes just before an election can affect the outcome of that election, as our voters are purely retrospective. (In the future we will extend the model to citizens who vote both retrospectively and prospectively, which explains the slightly odd ordering. This choice only affects what happens in the first two periods, and has no impact on future outcomes.)

rational, such behavior does tend over time to cause them to vote for the party closest to their own interests, as measured by their ideal points (BKS 2005). Parties are myopically rational, maximizing their utility in every period. This maximization is accomplished via a simple iterated line search. Vote share is used as a proxy for the probability of winning for reasons of both tractability (exact solutions for the victory probability become unreasonable for even reasonable sizes of the electorate, and Monte Carlo methods do not yield sufficient variability for maximization at even a thousand draws) and empirical reasonableness (maximizing votes seems to approximate well the actual behavior of parties in the face of great uncertainty). Parties enact their own ideal point in the first election, which occurs before they have had a chance to act.

Convergence, but Not to the Median (and Not Always to the Mean, Either)

We begin by assuming that *partyIdeoLevel* (the degree to which parties care about the enacted position) is zero. Thus all parties are purely office-motivated, driven by the desire to obtain the perks conferred by holding power. A natural question in such a setting is: do the parties obey the median voter theorem? That is, do they converge, and, if so, is it to the median voter's position? As seen in other contexts,[17] it turns out that convergence is a strong outcome of voting processes such as ours, occurring in every period for every parameterization of the model. Though it is interesting that this fundamental result continues to hold in a very different context than the standard Downsian model of prospective voting, it is not surprising that it does. After all, with two parties that are effectively clones of each other, each having the same utility function, as long as the vote share is maximized at a single point each party should be expected to locate at that point in every period.

What is more interesting is where parties locate, an issue that the analytical model of the previous section could not completely resolve. While the exact point in policy space depends upon the distribution of voters and the parameterization of the model, certain

---

[17]See Duggan (2005) for a survey of equilibrium analyses across a variety of spatial electoral models.

aspects of party location are common across these. The first aspect is that, except for perfectly symmetric voter distributions in which the mean and the median are the same, parties do not locate at the median. Thus, the main part of the median voter theorem does not hold.

Where they do locate is more complex, and requires an understanding of the aspiration-driven dynamics involved. As voters begin with equal probabilities of voting for either party, the first election is determined by an unbiased flip of a coin. Since the position chosen by each party in the first period is its own ideal point, this first winner does not necessarily receive a benefit from the win: enacting its own ideal point is unlikely to be optimal. Parties choose optimal positions (i.e., they maximize the expected vote share arising from the implementation of their announced policy) before the second election. This has no effect upon the outcome of the second election, as voters can only respond retrospectively to the payoffs received once one of these new positions is implemented. However, whichever party wins the second election has an immediate advantage in the future: the position on which it chose to run before the second election yields, by construction, the maximal expected vote share for that party in the third election.

In a large electorate, this advantage will be sufficient to guarantee electoral victory for a substantial number of periods. (How many is a function of the model's parameters, a property that will be discussed in the next subsection.) Thus, the model inherently includes a strong incumbency advantage: since voters only react to the incumbent's actions, a (myopically) rational incumbent can influence how voters will respond, and so force a victory for large electorates.

Yet, this advantage is not absolute, and the same dynamic which makes it possible also renders it ever more difficult. At first the party's chosen position is such that voters are often satisfied but, as (A3c) indicates, this implies that their aspirations also steadily increase. At some point they begin to expect the higher payoffs; hence there is less and

less the incumbent can do to satisfy them with certainty.[18] Individuals' aspirations settle near the expected payoff, and so whether they are satisfied in any period turns on a purely random draw from the exogenous shock term in (A4c). At this point it's only a matter of time before expected vote share falls enough so that the incumbent loses, per the analytic results in the previous section. After this occurs, elections become more or less random.

Parties choose the positions on which they run in the presence of this dynamic. Consider for ease of explication an electorate divided cleanly into two groups: a left one located at -1 and a right one located at +1. The parties' initial choice of position places them either to the right or to the left of the mean, based on the quasi-random events after the first election. As aspirations adjust, however, parties must strive always to be one step ahead of the curve, and so they inch backward toward the mean in an effort to make more voters happy in every period. Eventually they reach the mean and, in this example, stay close to it forever after as the incumbent's fortunes gradually decay.

That the mean rather than the median is chosen is no accident. Once aspirations have settled down, it is the point that best balances the various probabilities of success. Moving to one direction likely means increased propensities from those voters in that direction, since their chances of receiving a success will improve, but only at the cost of lowered propensities in the other direction. Since the normal PDF is symmetric, as are propensity adjustment rules around one-half, the vote-share benefit generally cancels out the cost.[19]

This is the general pattern when the voters' bliss points are distributed exactly symmetrically, and it mirrors the outcomes observed in some probabilistic voting settings (e.g. Duggan 2005). When it is asymmetric, however, more complicated behavior arises. Now the vote share function exhibits two local maxima, and either could become the global maximum at any time in response to changes in propensities and aspirations. One local maximum is still at the mean of the asymmetric distribution, but the other lies closer to the greater

---

[18]This dynamic, fundamental in our model, is absent in Achen and Bartels (2002).

[19]This need not hold if parties maximize something other than voter share.

concentration of voters. Locating at the mean is still often the best course of action, but a brief foray much closer to the clustered voters can satisfy them with near certainty, leading to a nearly uniform increase in their propensities to vote for the incumbent party. When propensities are near one-half this boost can outweigh the loss in expected vote share from locating away from the smaller group of voters, and the two local maxima switch relative sizes. This action immediately changes the distribution of propensities, though, and as a result both parties move quickly back toward the mean. Thus, we see cycles of the following sort: the parties move toward the mean of the voter distribution, get close to it, then jump back to a point closer to the greater concentration of voters, and then repeat the cycle.[20] Aspiration-based behavior therefore produces the convergence that one expects of any reasonable theory when the two parties are clones of each other, but yields a much richer array of behavior than is seen with Downsian voters.

Analytical examples Kollman and Page (forthcoming), Klochko and Ordeshook (2006) and others have cogently argued that computational modeling can aid our understanding of a phenomenon: instead of giving up when we fail to derive results analytically, we can turn to simulation. We fully agree with this position, as our use of a computational model of elections indicates. But the help can go the other way too: because the patterns generated by a computational model are often very complex, it is helpful to supplement them with examples that are simple enough to be solved analytically.[21] Then we can see, starkly and directly, some underlying tendencies of the computational model.

Hence, here we provide some simple analytical results that we believe illuminate the harder-to-grasp patterns of these electoral simulations. These examples parallel the preceding computational subsection by focusing on the locational strategies of the incumbent party. In

---

[20]In Achen and Bartels (2002) parties can also locate away from the mean/median, though this occurs due to the influence of voters' priors rather than asymmetric voter distributions.

[21]For a discussion of how simple analytical examples can help illuminate a complex computational model, see Bendor, Diermeier and Ting (2003b).

order to do this, we make two simplifying assumptions. First, we assume binary payoffs, l and h, with $h > l$. Second, we invoke the following result, established in BKS 2005, regarding the long-run behavior of endogenous aspirations.

Proposition 0: Consider a decision-theoretic problem in which the payoffs are either l or h, and every feasible action produces either payoff with positive probability. If aspirations adjust via (A3) then the following conclusions hold.

- (i) If $a_{t'} \in (l, h)$ then with probability one $a_t \in (l, h)$ for all $t > t'$.

- (ii) Suppose aspirations start outside (l, h): either $a_0 \leq l$ or $a_0 \geq h$. If additionally aspiration-adjustment is bounded away from zero uniformly in t, then a_t moves monotonically toward (l, h) and is absorbed into that interval with probability one as $t \to \infty$.

In the context of our model, Proposition 0 implies that all voters will come to regard h's as satisfying and l's as dissatisfying. This result enables us to have our analytical cake and eat it too: we may reasonably suppress aspirations, yet it remains true that vote-propensities are modified as if aspirations explicitly guided propensity-changes. Hence, here we replace (A1) and (A2) by simpler counterparts that depend only on whether a voter got a high or low payoff.

(A1') (positive feedback): If $\pi_{i,t} = h$ then with probability one $p_{i,t}(W_t) \geq p_{i,t-1}(W_t)$.

(A2') (negative feedback): If $\pi_{i,t} = l$ then with probability one $p_{i,t}(W_t) \leq p_{i,t-1}(W_t)$.

The essence of spatial payoffs extends easily to the binary payoff setting.

(A4'): The probability that voter i gets a payoff of h is strictly decreasing in the distance between i's bliss point and the incumbent's policy.

We will call the function that represents (A4') the probability loss function, or plf. This in effect parallels standard spatial utility functions, with expected utility replacing utility.

To complete the parallel with the computational model, we assume that voters differ only in the location of their bliss points, i.e., they have the same plf.

The first two examples show why office-seeking incumbents often will not locate at the median voter's (MV's) ideal point. We can make the point crisply by assuming a very simple form of retrospective voting: a citizen votes for the incumbent if and only if she is satisfied, i.e., she received an h-payoff in the current period. This will be called simple satisficing.[22]

There are two distinct causal mechanisms that produce non-MV behavior. The first one turns on voters' being very fussy, i.e., the incumbent is unlikely to satisfy citizen i unless she implements a policy very close to i's point. Visually, a voter is fussy if her plf looks like a sunken tent, mathematically, if the magnitude of her plf drops off quickly away from her ideal point.[23]

Example 1: Suppose (A4') holds and voters are simple satisficers. If they are sufficiently fussy and the distribution of ideal points is unimodal, then an office-seeking incumbent will locate at that mode.

[Figure 1 about here]

The intuition of Example 1 is conveyed visually by Figure 1. The dots correspond to the distribution of ideal points, with bigger dots implying a greater concentration of voters sharing that point. Thus, in the example there are more Liberal voters (L) than Moderate (M) or Conservative (C) voters. Overlaid on these points is the plf for voters at each point, which is almost maximally fussy: if the incumbent chooses any policy that differs from a voter's bliss point, then the probability of an h-payoff for that voter quickly drops to zero. (Strictly speaking the plf as drawn doesn't satisfy (A4'), but obviously the relevant kind of continuity allows us to draw a plf such that if the policy distance is more than δ then

---

[22]We use this name because the rule described in the text is, in fact, the simplest variant of the satisficing heuristic that one could imagine. Not all satisficing models assume that dissatisfaction triggers search for a new alternative with probability one (e.g., BKS 2005).

[23]The limiting case of a fussy voter's plf would be a Dirac delta: a pure spike of zero width and infinite magnitude.

the probability of an h is less than ϵ, etc.) Since the incumbent has almost no chance of satisfying anyone if she locates between ideal points, she must locate at one of the bliss points. And since there are more liberals than moderates or conservatives, she locates at the liberals' ideal point, even though the moderate is the MV.[24]

Of course, this example is very simple. In the more complex computational model, aspirations explicitly adjust to experience; hence, what voters regard as satisfactory payoffs varies over time. This produces much more varied behavior. However, this example can still help us understand the more complete model: as a snapshot of a complex dynamic, the example highlights the properties of voters that induce incumbents to locate closer to a mode than to the median or the mean. The cyclic pattern of party location described earlier, in which parties temporarily locate much closer to the greater concentration of voters, illustrates this. In both the full model and in the stylized example, the immediate benefits of locating closer to the greater number of voters (closer, i.e., to the mode) outweigh the votes lost by moving away from the median or the mean. The differences between the example and the computational model are clear. The former is simple enough so that we can say with certainty where the parties will locate in every period. The model, with endogenous aspirations coupled to a full range of payoffs, is sufficiently complex so that incumbents locational choices are only tendencies, not starkly identified certainties.

This argument applies equally well to a second cause of non-MV behavior, a sufficiently asymmetric distribution of ideal points. Again, suppose that there are three types of voters, liberal, moderate and conservative, with ideal points l^*, m^* and c^*, respectively. This time we assume that they're equally numerous, but that the distribution of ideal points is skewed

---

[24]Since we are assuming that the median is differs from the mode, there must be less than half the population in the liberal camp, and so the chance of any party's winning with this strategy is identically zero. Thus, the strong dominance of this strategy relies upon the assumption of vote share maximization rather than maximization of the probability of winning. However, the strategy remains weakly dominant in the latter case, and we can recover our stronger result by instituting lexicographic preferences: a party first maximizes the probability of winning and then maximizes vote share if still indifferent between policies.

sharply to the right: c^* is much farther from m^* than is l^*. It is clear from Figure 2 that if the C's are sufficiently extreme then no policy will satisfy both them and the M's, let alone C's and L's. This leads to Example 2, displayed in Figure 2.

Example 2: Suppose (A4') holds and voters are simple satisficers. There are three types of voters. For any plf that asymptotes to zero, there exists a distribution of voters asymmetric enough so that an office-seeking incumbent will not locate at the MV.

[Figure 2 about here.]

Applying Example 2 to our example, for any given plf that asymptotes to zero and any fixed $\epsilon > 0$, if the C's are sufficiently far from the M's and if the incumbent locates at $\frac{m^*+c^*}{2}$ then the probability of giving an h to either type is less than ϵ. *A fortiori*, the chance that an L gets an h will also be less than ϵ. Hence, a rational office-seeking party gives up on extremist principals: trying to please them is too costly.

Given this, and given that all three voter-types are equally common, it's easy to show that it cannot be optimal for an office-seeker to locate to the right of m^*. (Here's a sketch of why that must be true. Locating at m^* maximizes the chance of satisfying the M's, there's a greater chance of getting L's, compared to any $x > m^*$, and given that the chance of satisfying C's was negligible at the midpoint between m^* and c^*, the loss in C-support incurred by locating at m^* must also be negligible.) So x^* must be in $[l^*, m^*]$. The only remaining question is whether it is always optimal for the incumbent to pick this interval's right boundary, m^*. We know from Example 1 that this *would* be best if the voters are sufficiently fussy. But if the plf is sufficiently concave over a big enough interval (over (l^*, m^*) will do) then an office-seeking party will locate to the left of m^*. The reason is that the plf's concavity makes attracting the support of L's an attractive possibility: moving a bit to the left of m^* increases the probability of producing h's for L's faster than it reduces the chance of satisfying M's. And the C's are so extreme that the chance of satisfying them is already negligible for $x^* = m^*$, so there's virtually no cost on that front.

In effect, then, a rational office-seeking politician simply ignores the extremist C's. They

23

are so far from everyone else that it is not worthwhile to try to satisfy them. Hence, the incumbent's future lies entirely in the hands of the other two blocs, the liberals and moderates. And under plausible conditions, it is optimal to avoid corner solutions, i.e., not to do exactly what is best for either bloc of voters. So again non-MV behavior prevails.[25] This illustrates the cyclic dynamics of the computational model, in which it sometimes becomes beneficial to downgrade the potential votes of the extremists in one's location calculations.

In light of work on location at the mean of the ideal-point distribution (e.g., Enelow and Hinich 1984), it is worth noting that in the present example the incumbent does *not* locate at the distribution's mean either. Indeed, the distance between the mean and x^* grows arbitrarily large as $c^* \to \infty$. Retrospective voting, which is driving the present result, does not necessarily entail mean-seeking behavior by an office-seeking incumbent; the policy on which parties collocate depends fundamentally on the probability that voters are satisfied in any period. In these simple examples this probability is driven by only one factor: the distribution of voters' ideal points. If they are too asymmetric either in concentration or in location, parties locate at neither the mean nor the median of the distribution. In the more complex computational model, changing aspirations play a role as well. Though the payoff function, (A4c), typically is such that for the voter distributions we examine parties collocate at or near the mean, they do not simply sit there as they would in a probabilistic voting model. Instead, at times the adjustment of aspirations alters the situation to match up more closely with these simple examples, resulting in policies chosen (albeit briefly) away from both mean and median.

Incumbency Advantage

In the previous subsection we spoke of the incumbent's natural advantage, given retrospective voting: since citizens react only to outcomes driven by the incumbent's behavior

---

[25]This behavior will persist indefinitely. Hence, whereas parties asymptotically converge to the MV in the Achens-Bartel model (2002, p.24), this is not true in general of our model.

in office, the latter can influence, especially early on, how the former will respond.[26] This is an important point. Methodologically, it means we have recovered some of the near-invulnerability of many incumbents without invoking additional modeling structure. Substantively, it suggests that something far more central to human decision-making than a reaction to transient advertising bought with an incumbent's extra resources can produce extended terms in office. Given this, it bears further examination. In particular, our previous discussion does not examine just how long an incumbent can be expected to stay in office. While this is dependent on several parameters, the focus of this chapter is on the effects aspiration-based behavior has on elections, and we leave most analysis to future work. Here we concentrate on the impact the parameters of propensity and aspiration adjustment, λ and ν, have on the length of time in office. (We briefly examine the impact of the initial aspiration level at the end of this subsection.)[27] Since the first period's occupant is random, as discussed previously, we calculate this length as the difference between the second election and the last election won by the incumbent before the party in power changes. The setting will be the same bipolar electorate described in the previous section: half of the electorate has an ideal point of -1, the other half at +1.

[Figure 3 about here.]

Figure 3a illustrates the effect that increasing the rate of aspiration adjustment has on the length of time in office, for a relatively slow propensity adjustment rate of 0.1. As the graph plainly shows, increasing ν decreases time in office. Slow updating of aspiration levels produces a powerful incumbency advantage, as it takes a long time for voters' aspirations to settle down to their mean payoff level. In contrast, as aspirations update faster, voters more

---

[26]Achen and Bartels (2002) also find an incumbency advantage, though under different assumptions. Further, as learning progresses in their model, the advantage tends to increase over time, while ours eventually decreases as aspirations adjust.

[27]We don't discuss the impact of the number of voters on the incumbency advantage, though that is fairly clear: the more numerous the voters, the less stochastic are electoral outcomes, so incumbency advantage rises.

quickly become accustomed to their typical payoffs ("what have you done for us lately?") and it becomes more difficult for the incumbent party to satisfy its constituents for any stretch of time.

If propensities adjust more quickly, the problem for incumbents only worsens. Figure 3b displays the same information as in Figure 3a, save that $\lambda = 0.5$. Again we find a weakening of the incumbency advantage as aspirations adjust more rapidly, but now the curve as a whole is lower than it was for slower propensity updating. Individuals respond more rapidly in their behavior (propensities) to changes in their beliefs (aspirations) here, and so the movement of their aspirations to the mean payoff level catches up to the incumbent more quickly. However, this effect is of lesser magnitude than the decrease driven by slowing down the rate of aspiration adjustment, indicating again that it is really the dynamic of aspiration adjustment that is driving the show.

In Figure 3c we look at the limiting case: propensities adjust immediately. A satisfied voter will with certainty vote for the incumbent party, while a dissatisfied voter will most assuredly vote against it. This behavioral extremum yields some real differences from more gradual updating—for one, the distribution of propensities at any time is bimodal at zero and one rather than tightly clustered around one-half—but the trend is the same. Again, increasing the rate of propensity adjustment decreases the incumbent's average time in office, and again this effect is less than that produced by faster aspiration adjustment.

We see then that the general trend appears to be that faster updating reduces incumbency advantage. This is certainly true for the middling initial aspirations assumed in the above analysis, and also holds when initial aspirations are low. If initial aspirations are high, however, we see a reversal of the effect. Figure 3d is identical to Figure 3a, save that voters' initial aspirations are now high: just below the maximum. Voters thus expect a great deal, far more than any incumbent can deliver, and incumbency becomes a *disadvantage*. When aspirations adjust slowly, this disadvantage persists for a long time. But as aspirations update

faster, voters learn more quickly to expect less, which rapidly eliminates this disadvantage.[28] As with any default choice, incumbents benefit from low aspirations.

Analytical Examples Though the complexity of endogenous aspiration adjustment implies that we cannot analytically derive the graphs in Figure 3, a simplification can illustrate the computational model's major points regarding how initial aspiration levels and the speed of aspiration-adjustment together impact how long incumbents stay in power. We consider the two extreme possibilities: that the speed of adjustment is minimal—$\nu = 0$ so $a_{i,t} = a_{i,t-1}$—and that it is maximal—$\nu = 1$ so $a_{i,t} = \pi_{i,t}$. We start with the former. To keep the analytical example simple, we assume that the voters share the same ideal point and payoff shock, θ.

In this environment the ability of an office-seeking incumbent to stay in power depends entirely on the voters' initial aspirations. If θ's support is bounded and $a_{i,0} < \min(\pi)$ for all i, then whichever party is elected first will stay in office forever after: the incumbent will optimize by implementing the voters' common ideal point, which by (A4) and the assumption of bounded shocks must give everyone more than their minimal payoff. On the other hand, if everyone's initial (and fixed) aspiration level exceeds $\max(\pi)$, then in every election the incumbent is thrown out of office with probability one.

Now suppose that aspirations adjust immediately to payoffs: $\nu = 1$. If θ has a continuous density over a bounded support, then Remark 3 applies, so incumbents only have a 50-50 chance of staying in office, for all elections after the first two.

These simple analytical examples highlight properties of the computational model: speeding up aspiration-adjustment has different effects on incumbency advantage, depending on the level of initial aspirations. If initial aspirations are low, then incumbency advantage falls as aspirations adjust more rapidly. But if initial aspirations are high then the incumbent's duration-in-office rises as aspirations adjust more rapidly.

---

[28]Incumbency advantage is a consequence of the dynamics of the model, and so the fact that initial conditions alter it does not imply non-ergodicity.

Policy-Motivated Parties

Thus far we have considered only parties motivated purely by the perks of office, so parties are effectively clones of each other. But what if this is not the case, and politicians, like other citizens, care about the policy enacted as well (so $partyIdeoLevel > 0$)? We will discuss two aspects of this question here.

The first, and the most natural, question to ask in this context is: Do the parties still converge? The answer to this is simply no. The objective functions each party maximizes are no longer identical, and so the maxima are generically no longer identical as well. How much the parties diverge depends on the relative values of the parameters *perq* and *partyIdeoLevel*. The greater the perks enjoyed by the party in office, or the less parties care about the policy enacted, the closer they will locate, *ceteris parabis*. This is perhaps unsurprising; probabilistic voting introduces an element of chance just like uncertainty in position does, and, as Calvert (1985) showed, this along with policy-motivated parties suffices to induce policy divergence. Because we ground the notion of uncertainty in a specific behavioral model, however, we can go further and explore consequences of policy-motivation beyond convergence.

One such consequence is that the policies parties select exhibit more interesting dynamics than is the norm in Downsian analyses, which leads to our second question: where do the parties locate? Earlier we saw cyclic behavior as voters updated their aspirations. The same general dynamic holds in the policy-motivated regime, with some important differences. Consider again the bimodal voter distribution, with each party's ideal point matching the location of one of the modes.

As stated, the winning party locates further toward its ideal point than in the purely office-motivated case. At first the other party does the same, taking up the mirror image position of the first, due to the symmetry of the setup. This, however, does not last. Once one party begins winning, earning the incumbency advantage described above, the other party is forced to move closer to the center in an attempt to sway a greater proportion of the voters,

were the challenging party to win. Of course, for reasonably slow aspiration adjustment the incumbency advantage is too strong; voters react only to the party in power, and the challenger's approach to the mean affects neither their aspirations nor their propensities. Still, the challenger tries to increase his chance of winning by moving closer to the other candidate (or, more precisely, its objective function dictates optimal positions closer to the other candidate), in some cases actually crossing the mean while doing so.

This behavior continues while the incumbent party keeps trying to increase the average propensity of the electorate to vote for it. After awhile, however, aspirations begin to reach mean payoffs, and the average propensity to vote for the incumbent begins to fall. Now the other local maximum for the challenger–its ideal point–begins to look more attractive. If the challenger were to win, after all, not only would at least some voters (those closer to the challenger's ideal point than the incumbent's position) be likely to be satisfied with its performance at its own ideal point, but it would receive the maximal payoff from the policy-related component of its utility. Eventually this option becomes optimal, and the challenger adopts its own ideal point, in a maneuver that looks to outside observers much like quitting: with no apparent chance of winning, the challenger stops trying and just runs on its beliefs (e.g., the British Labor party in the Thatcherite era).

Unfortunately for the challenger, this behavior, though myopically optimal, is *not* dynamically so. As detailed above, as aspirations continue to move toward the mean of voters' payoffs, the expected vote share of the incumbent party declines. It moves closer to the mean in an attempt to recapture votes, and this works briefly, but inexorably there comes a time when the incumbent is removed from office. Were the challenger located at the same point as the incumbent, this would result in a more-or-less random sequence of subsequent electoral outcomes, but instead the challenger has retreated to its ideal point. It wins a single election–a moral victory, perhaps–and will win the occasional victory in the future as well, but on the whole is relegated to an existence of being a largely irrelevant opposition party. The incumbency advantage in this context is thus even greater than without policy

motivation.[29] The challenger's stubborn desire to be happy with the policy outcome were it to win prevents it from winning often, while the incumbent, willing to cater partially to the masses, reaps the benefit.[30]

5   Conclusions

How citizens vote is a central part of the electoral environment for parties. Therefore, if the latter are rational but the former are only boundedly so, then optimizing parties will take citizens' limitations into account. This chapter has explored how retrospective voting—which in general is not perfectly rational behavior—affects parties' strategies. We show that rational, office-motivated parties will rarely find it optimal to locate at the median voter's ideal point, and often will choose policies away from the mean as well. This deviation from the mean—atypical in models of probabilistic voting—gives a reason why a rational party would play to its base independent of turnout considerations. We also find an endogenous incumbency advantage, deriving from the way in which voters adapt, that offers an alternative explanation to standard exogenous factors—such as funding differentials—for the well-established empirical norm of robust incumbency.

On the methodological front, this chapter adopts the view that analytical and computational methods are complementary. This is not meant as a way to make peace between warring academic factions; we mean 'complementary' in the standard dictionary sense.[31]

---

[29]Note, however, that this is somewhat dependent on parameterization. In settings where the incumbent also locates near its own ideal point, the challenger party will be far more likely to win in the future after winning once.

[30]One can argue that this is merely a result of myopic optimization, and that a more full characterization of the objective would rectify this. While quite likely true, as discussed in footnote 9, the choice problem becomes extremely complex very quickly has the size of the electorate increases, and we believe that full optimization of a discounted stream of utilities is empirically untenable.

[31]The third definition of 'complementary' in the Webster's Ninth New Collegiate Dictionary is "mutually

On the one hand, it is very difficult to figure out analytically the dynamics of party behavior when (a) aspirations are endogenous, (b) one posits retrospective voting rules that are more complex than simple satisficing, and (c) there are many different ideal points in the population. For such contexts, computation is extremely useful. On the other hand, simulation results can be hard to interpret, and supplementing them with analysis of simple, highly stylized contexts can often help us figure out exactly what's driving what in a complex computational model.

When problems are hard, it makes sense to tackle them with a variety of methods. Natural scientists have known this home truth for a long time;[32] it is time for social scientists to adopt this pragmatic perspective.

Appendix

Proof of Proposition 1

Consider x_o not in the convex hull of the ideal points. Then there exists a point x_r in the convex hull such that for every voter i, $||x_i^* - x_r|| \leq ||x_i^* - x_o||$. From (A4), for each voter i, the probability that the payoff exceeds a threshold $P(\pi_{i,t} > a \mid x_r) \geq P(\pi_{i,t} > a \mid x_o)$ for any level a. From (A1) and (A2) it follows that $P(p_{i,t}(W_t) > p_{i,t-1} \mid x_r) \geq P(p_{i,t}(W_t) > p_{i,t-1} \mid x_o)$ regardless of $p_{i,t-1}$ or $a_{i,t-1}$, for every voter. Therefore the probability that each voter votes for the incumbent (weakly) increases when x_r is implemented rather than x_o. Therefore, the probability of the incumbent winning the election (weakly) increases. QED.

Proof of Remark 1

For *any* given vector of aspiration levels $a_{i,t-1}$ every voter can get a payoff such that $P(\pi_{i,t} < a_{i,t-1}) > 0$ since all payoff shocks are not bounded below. Hence, by (A2) the

---

supplying each other's lack" (1989, p.269).

[32] As the Nobel prize winning physicist Percy Bridgman provocatively remarked, "there really is no scientific method...[just people doing their] damndest with [their] minds, no holds barred" (1947, p.144-45). Had any of his colleagues cavalierly dismissed a problem-solving method such as computation, Bridgman's reaction probably would have been unprintable.

probability that everyone has a vote-propensity less than 1 is strictly positive for every election. So the probability that a majority votes against the incumbent is strictly positive, in every given period. QED.

Proof of Remark 2

Let $N(T)$ be the subset of $\{1, 2, \ldots, n\}$ such that for any $i \in N(T)$, $a_{i,T} > \underline{\pi}_i$. Consider any i in $N(T)$.

Since θ_i's density is continuous, the event $\pi_{i,t} = \underline{\pi}_i$ occurs with probability zero. Hence, (A3) implies that $a_{i,T+1} > \underline{\pi}_i$ almost surely, and by induction this inequality holds at every finite date.

Given this, and given that θ_i's density is strictly positive over its support, $\Pr(\pi_{i,t} \in [\underline{\pi}_i, a_{i,t}) > 0$ for each $t > T$ and all $i \in N(T)$.

So with positive probability i will be disappointed, whence (A2) implies that $p_{i,t+1}(W_t) < 1$. Since this holds for all i in $N(T)$ and the latter constitutes a majority of the electorate, the incumbent will lose the election with positive probability in $t + 1$. QED.

Proof of Proposition 2

We begin by showing the result under (i). We will then indicate how the proof under (i) can be modified to obtain the result under (ii). Fix an $\delta_1 > 0$. From the law of large numbers, we expect $a_{i,t}$ to converge to $\bar{a}_i = E[\theta_i] + f_i(0)$. To be precise, there exists a time T such that $P(a_{i,t} > \bar{a}_i - \delta_1) \geq \frac{1}{2}$ for all $t > T$. Given that θ_i satisfy a continuous density and the arbitrary choice of δ_1, for all $t > T$ we must have $P(\pi_{i,t+1} < a_{i,t}) > \frac{1}{2}\delta_2$ for some $\delta_2 > 0$. That is, with positive probability each voter can be dissatisfied in every period after T.

By (A2') this implies that for all $t > T$, $P(p_{i,t+1}(W_t) < 1 - \epsilon) > \frac{1}{2}\delta_2$. Given that the actual votes only depend on the propensities, we have

$P(\text{voter i does not vote for incumbent in t+1}) > \frac{1}{2}\delta_2\epsilon$ for all $t > T$. Therefore

$P(\text{every voter does not vote for incumbent in t+1}) > \left(\frac{1}{2}\delta_2\epsilon\right)^n > 0$. Therefore

$P(W_{T+1} = W_{T+2} = \cdots = W_{T+k}) < \left(1 - \left(\frac{1}{2}\delta_2\epsilon\right)^n\right)^k$ for every k.

Letting $k \to \infty$ yields the result.

To mimic this argument under (ii), we note that it is sufficient to establish that $P(\pi_{i,t+1} < a_{i,t}) > \frac{1}{2}\delta_2$ for some $\delta_2 > 0$ for all t beyond some T as above. This actually holds with $T = 0$ given the choice $a_{i,0}$ and (A3). That is, the aspiration can never reach the minimum value of the payoff and hence at every period each voter is dissatisfied with positive probability. The rest follows as in (i). QED.

Proof of Remark 3

In each election after $t = 1$ the probability that any given voter votes for the incumbent is $P(\pi_{i,t} > \pi_{i,t-1})$. Since both parties situate themselves at x^*, the f_i are identical and the payoff shocks θ_i are i.i.d., this probability equals $\frac{1}{2}$. Since each voter is equally likely to vote for either party, we have $P(n - k\text{vote for incumbent}) = P(k\text{vote for incumbent})$ for each $k = 0, 1, \ldots, n$. In particular we have

$\sum_{k=(n+1)/2}^{n} P(k\text{vote for incumbent}) = \sum_{k=0}^{(n-1)/2} P(k\text{vote for incumbent})$. But the former is simply the probability that the incumbent wins and the latter is the probability that he does not. Hence, the result follows. QED.

Proof of Remark 4

This follows exactly along the lines of Remark 2. We simply replace the party with a policy, and since the party identity or the number of parties do not affect the proof of Remark 2, the result follows. QED.

References

Achen, Christopher, and Larry Bartels. 2002. "Ignorance and Bliss in Democratic Politics: Party Competition with Uninformed Voters." Presented at the 2002 Annual Meeting of the Midwest Political Science Association.

Aranson, Melvin Hinich and Peter Ordeshook. 1974. "Election Goals and Strategies: Equiva-

lent and Nonequivalent Campaign Objectives." *American Political Science Review* 58:135-52.

Bendor, Jonathan, Daniel Diermeier and Michael Ting. 2003a. "The Empirical Content of Adaptive Models." Presented at the 2003 Annual Meeting of the Midwest Political Science Association.

Bendor, Jonathan, Daniel Diermeier and Michael Ting. 2003b. "A Behavioral Model of Turnout." *American Political Science Review* 97:261-80.

Bendor, Jonathan, Sunil Kumar and David Siegel. 2005. "V.O.Key Formalized: Retrospective Voting as Adaptive Behavior." Presented at the 2005 Annual Meeting of the American Political Science Association.

Bendor, Jonathan, Dilip Mookherjee and Debraj Ray. 2001. "Aspiration-Based Reinforcement Learning in Repeated Games: An Overview." *International Game Theory Review* 3: No.2-3.

Bridgman, Percy. 1947. "New Vistas for Intelligence." In E. Winger (ed.), *Physical Science and Human Values*. Princeton: Princeton University Press.

Calvert, Randall. 1985. "Robustness of the Multidimensional Voting Model." *American Journal of Political Science* 29:69-95.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper Row.

Duggan, John. 2005. "A Survey of Equilibrium Analysis in Spatial Models of Elections." mimeo.

Enelow, James, and Melvin Hinich. 1984. "Probabilistic Voting and the Importance of Centrist Ideologies in Democratic Elections." *Journal of Politics* 46:459-78.

Delli Carpini, M., and S. Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven: Yale University Press.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50:5-25.

Key, V.O. 1966. *The Responsible Electorate: Rationality in Presidential Voting 1936-1960*.

Cambridge MA: Belknap Press.

Kinder, Donald. 1998. "Opinion and Action in the Realm of Politics." In D.Gilbert, S.Fiske and G.Lindzey (eds.), *The Handbook of Social Psychology*, 4th ed. Boston: McGraw-Hill.

Klochko, Marianna, and Peter Ordeshook. 2006. "Endogenous Time Preferences and Social Complexity." [this volume]

Kollman, Ken, and Scott Page. "Computational Methods and Models of Politics." Forthcoming.

Luskin, Robert. 2002. "From Denial to Extenuation (and Finally Beyond): Political Sophistication and Citizen Performance." In James Kuklinski (ed.), *Thinking about Political Psychology*. New York: Cambridge University Press.

Simon, Herbert. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69(February):99-118.

Simon, Herbert. 1956. "Rational Choice and the Structure of the Environment." *Psychological Review* 63:129-38.

Stokes, Donald. 1963. "Spatial Models of Party Competition." *American Political Science Review* 57(2):368-77.
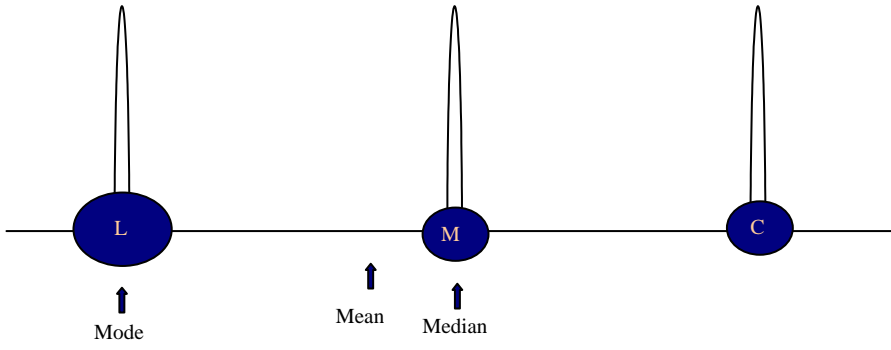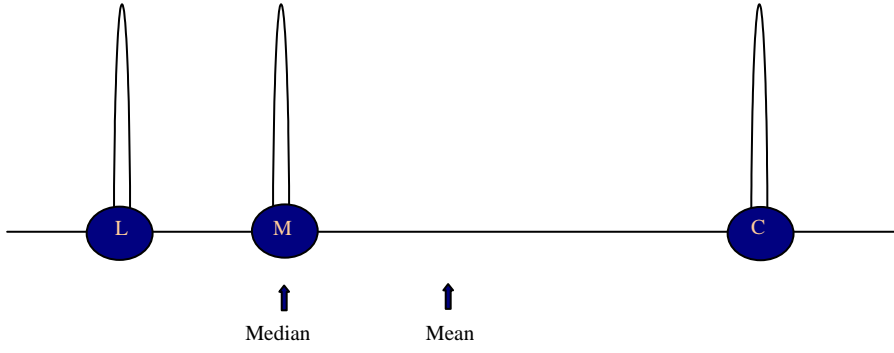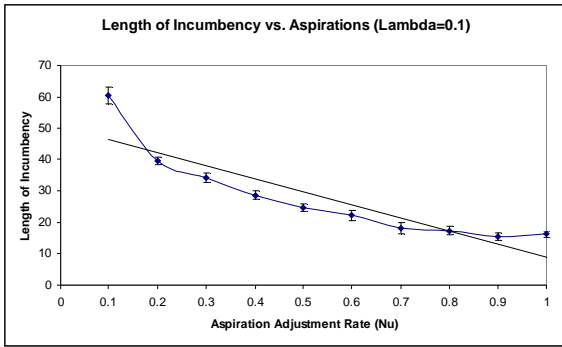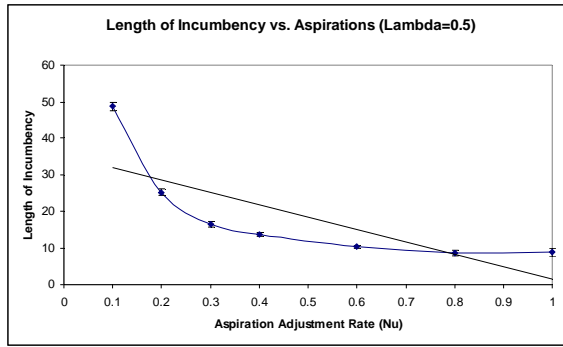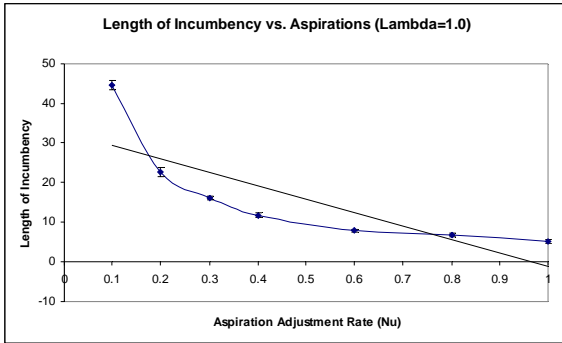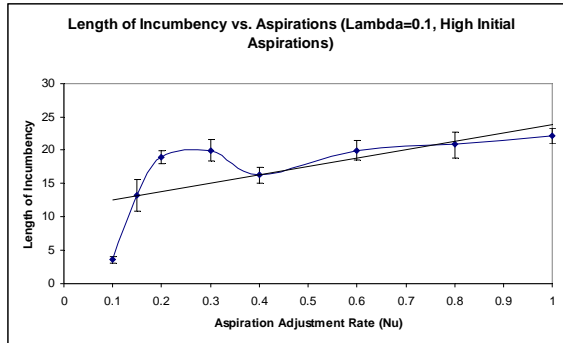
Figure 1: Example 1

Figure 2: Example 2

Figure 3: Length of Incumbency as a Function of the Speed of Voters' Updating