

**A Reputational Theory of Two Party Competition**

Tasos Kalandrakis

Working Paper No. 41  
August 2006

W. ALLEN  
WALLIS  
*Institute of*  
POLITICAL  
ECONOMY

UNIVERSITY OF  
**ROCHESTER**

# A Reputational Theory of Two Party Competition<sup>1</sup>

by

Tasos Kalandrakis<sup>2,3</sup>

**Abstract:** We propose a reputational theory of two-party competition. We model the interaction of parties and the electorate as a stochastic game of incomplete information. The parties' preferred policies (moderate or extreme) are possibly revealed to the electorate only via their policy choices while in government, and partisan preferences change with positive probability following defeat in elections. Due to inertia within party organizations, party preferences display positive serial correlation. When partisans care sufficiently about office, extreme policies are pursued with positive probability by the government only when the ruling party is perceived relatively more extreme than the opposition. In equilibrium such policies occur when (a) both parties are perceived to be more extreme than a long-run benchmark level, and (b) neither party holds a significant advantage regarding its perceived extremism by the electorate. Equilibrium dynamics produce two qualitatively different adjustment paths: one exhibits polarized politics such that there is positive probability of non-moderate policies in the future for a protracted period of time; the other possible adjustment path produces moderation with probability one in all periods. Both adjustment paths are such that one of the two parties (possibly different over time) may win successive elections with high probability in equilibrium.

**Keywords:** Parliamentary Dynamics, Reputation, Westminster.

---

<sup>1</sup>First draft: April 6, 2005; this draft: July 30, 2006.

<sup>2</sup>Previous title: Dynamics of Westminster Parliamentarism. Thanks to Keith Krehbiel, and John Roemer, for their feedback. I have also benefited from helpful comments from audiences at the 2005 MPSA meeting, the University of Crete, the Athens University of Economics and Business, the 2005 SED conference, Budapest, and Cornell University. I am responsible for all errors.

<sup>3</sup>Assistant Professor, Department of Political Science, University of Rochester. Correspondence Address: Department of Political Science, University of Rochester, Harkness Hall 333, Rochester, NY 14627-0146. E-mail: akalandr@mail.rochester.edu. Web: <http://mail.rochester.edu/~akalandr/>.

## 1. INTRODUCTION

The theory of two party/candidate competition has been dominated by the ‘convergence to the median’ theorem of Hotelling, 1929, and Downs, 1957. Their stark model provides compelling support for the intuition that candidates endorse moderate policies in order to attain office. Despite its appeal, Downsian electoral competition has the unfortunate implication that politics do not matter in equilibrium, since candidates adopt identical equilibrium policy positions. An alternative is obtained under the assumption that candidates have policy motivations and face electoral uncertainty. Under these conditions, Wittman, 1983, 1990, (see also Calvert, 1985) deduces equilibrium policies that are located away from the median. Both models generate ideals with trivial electoral dynamics, since these models imply either persistent policy convergence or (partial) divergence over time. At least with regard to partisan competition, empirical observation suggests otherwise.

For instance, Downsian convergence seems to be a fair approximation of the world of British politics in the 50’s and 60’s. But, this era of the “politics of consensus” (Kavanagh & Morris, 1994) came to an end in the 80’s with the governments of Margaret Thatcher. Recently, after almost two decades of perceived divergence, British politics seem to be, once more, closer to the Downsian ideal. To take another two-party parliamentary system as an example, policy moderation may appear consistent with Greek politics in the mid to late 90’s, yet the preceding period since 1974 featured intense policy polarization between the two parties contesting for power in this country.

There is a second count on which static models of electoral competition are inconsistent with empirical observation. In these models both candidates are in principle able to be competitive in elections. For example, barring a non-anonymous resolution of voters’ indifference, a candidate can perform at least as well as the opponent in Downs’ model by adopting a platform at the median’s position. In table 1 we provide electoral results since the end of the second world war for three countries (the UK, New Zealand until 1996, and Greece since 1974) with parliamentary systems in which two parties alternate in power. As is evident from these data, one of the two competing parties often holds a persistent advantage in consecutive elections. Contemporary analysts of British politics speak of Labour’s domination over the Tories, but similar claims were made for the Tories in the 80s. In the case of New Zealand, the advantage of the conservative National party appears almost systemic.

[insert table 1 about here]

We take the view that a significant source for the discrepancy between empirical observation and the theoretical predictions of variants of Downs’s model of electoral competition is the assumption made in these models that parties are able to commit to a policy platform prior to the election. We argue that the assumption of pre-election platform commitment cannot be easily defended in the case of partisan competition. As a matter of fact, there do not exist actual contracts or other constitutional provisions that can literally secure commitment by political parties on pre-election platforms. Hence, existing justifications for this assumption are indirect and take two modal forms.

The first justification posits that parties essentially commit by electing a leader with known policy preferences. Of course, party leadership is not a guaranteed position for those who attain it, and leaders’ influence on the party’s actions cannot reach too far beyond the boundaries that stem from the fact that a leader is an agent of the party. In fact, even if there exist candidate leaders with ‘known’ policy preferences, in typical two-party parliamentary systems partisan rules do not prevent challengers from contesting for leadership at any point in time. Speculation about such challenges to incumbent leaders abound, even while parties are in office.<sup>4</sup> The fact that leadership replacements are not frequent is merely a reflection of the equilibrium effect that leaders yield to sufficiently strong internal party pressures in order to prevent such challenges. That means that, contrary to what is required for platform commitment, if partisan preferences dictate otherwise, a party may implement a different policy than the one it endorsed prior to the elections or the one its leader had endorsed in the past.

The second justification for assuming platform commitment, is that it constitutes a reduced form way of modelling electoral competition as a repeated game. In this repeated game, parties adhere to past platform promises in equilibrium for one of two reasons. One possibility, is that voters’ strategies in this repeated game involve ‘retrospective’ punishment threats, so that parties reneging from their campaign promise face the prospect of future punishment by the electorate. These complex punishment strategies on the part of the electorate, though, are both empirically inconsistent and theoretically problematic. Allowing for such complex strategies leads to theoretical indeterminacy in these models (e.g., Duggan and Fey, 2004). Given the typical length of the electoral cycle, it is also implausible to expect that voters will coordinate to vote against a party they *prefer* in order to punish that party for its actions in government eight or ten years in the past.

---

<sup>4</sup>Furthermore, the timing of leadership replacement is sometimes unpredictable and surprising, as, for example, in the case of Margaret Thatcher’s replacement by John Major.

On the other hand, a second, more plausible, possibility is that voters do vote against a party that took an undesirable policy action in the past because they have updated adversely regarding the true policy preferences within that party and they *no longer prefer* that party. We emphasize the important distinction<sup>5</sup> that in this second possibility, voters are forward looking. They simply evaluate the credibility of the parties and the positions they take on the basis of their past record of behavior. As a result, parties may have the incentive to remain consistent to their platforms in order to build a *reputation* that will help the party win future elections. We claim that this type of reputational incentive, instead of actual policy commitment, is central in shaping party policy choices. Yet, to date, we know of no study of dynamic partisan competition that explicitly focuses on party reputations and their role on political actors' behavior.

The objective of the present paper is to propose a reputation based dynamic theory of party competition. We combine premises of the models of Downs and Wittman, assuming partisans with a mixture of office and policy motivations. We assume that parties are populated by individuals with different policy preferences. The prevailing policy preferences within the party are not publicly known. Although a party's preferences are correlated from period to period, these preferences may change over time due to exogenous shifts in the balance of power within the party. True party preferences are (possibly) revealed to the electorate through policy consequential choices of this party when it is in government.

Since the electorate operates under incomplete information regarding prevailing policy preferences within political parties, and since partisans desire to win office, party platform declarations prior to the elections are received by the electorate with a grain of salt. Indeed, both platform declarations or even past policies may be strategic choices by partisans with the intention to please or deceive the electorate regarding the true policy preferences that prevail within the party. Thus, in our model we dispense with pre-election policy announcements altogether.<sup>6</sup> In our theory, parties enter the electoral arena with a reputation regarding the level of their policy extremism. This

---

<sup>5</sup>There is a second important distinction that, since we assume party reputations that vary over time, dynamic interaction takes the form of a stochastic game instead of a repeated game.

<sup>6</sup>This is not to say that platforms and electoral campaigns are irrelevant. Our view is that a big part of the information contained in these political activities concerns the ability of the candidates to prioritize among policy areas as well as their ability to innovate in proposing solutions to extant policy problems. The mere successful (or not) execution of the campaign under the pressure of the imminent elections conveys information about the abilities of the candidates.

reputation is endogenously formed and reflects the accumulated history of electoral outcomes and policy choices that have transpired prior to the elections. The electorate then chooses between the parties on the basis of their reputations. If the party that is elected in government is actually controlled by policy extremists these representatives face the following strategic dilemma: to pursue a moderate policy, thus preserving or enhancing the party's reputation, or to damage that reputation by pursuing a desirable extreme policy?

Revealing the party's extremism is consequential in our model because we assume that parties, like all organizations, display inertia. Thus, if a party tarnishes its good reputation by implementing an extreme policy while in government, this revelation affects the electorate's beliefs for several electoral cycles.<sup>7</sup> We formalize this idea by assuming that following electoral defeat parties undergo an internal change that probabilistically determines the prevailing group within the party, and that the outcome of this lottery displays positive serial correlation: extremists have a higher probability of prevailing within the party if the party was controlled by extremists in the previous period.

The combination of these two simple premises (the endogenous formation of party reputations along with the assumption of inertia in the determination of party preferences) produces a rich set of insights into the workings of two party competition. Our first result is that our model is inconsistent with Downsian convergence to the median, no matter how office oriented parties are. There does not exist a robust equilibrium in which parties in government implement moderate policies with probability one independent of their combined reputations. If parties are impatient or place significant emphasis on policy relative to office, the only equilibrium involves party types implementing their ideal policy independent of the electorate's beliefs about the two parties. In the more interesting case when parties assign high weight on office utility relative to policy, we ask the following three questions: for what levels of reputation for the two parties do party extremists pursue extreme policies when in government? Under what conditions are such policies observed in equilibrium? And what are the resultant electoral and policy dynamics?

We show that parties controlled by extremists pursue extreme policies with positive probability while in government when their party's reputation is relatively worse than that of the opposition (on or off the equilibrium path). In other words, extreme policies may be pursued by the government

---

<sup>7</sup>Again, British politics provide a good example. During Thatcher's term, and for consecutive elections, the Labour party attempted to present an ideologically reformed, more moderate facade, but these attempts (at least ex post) seem to have been discounted by the electorate which became convinced of the reform only recently.

if the government faces an electoral disadvantage. This occurs even though we allow the electorate to devise retrospective voting strategies such that a government that implements an extreme policy is ousted from office with probability one. This result reflects a more general principle that stems from our focus on party reputations. The principle dictates that the choices of the governing party are not independent of the reputation of the opposition party in relation to the government's own reputation.

Second, extreme policies are observed along the equilibrium path if and only if two conditions hold (a) both parties are perceived to be extreme with a probability that exceeds a given steady state reputation level, and (b) neither party holds a significant reputation advantage compared to its opponent. In other words, if both parties have relatively bad reputations, a government may pursue extreme policies following a *close* election. By 'close' we mean an election in which both parties have similar reputation levels. Policy extremism by the government along the equilibrium path occurs due to a *regression to the mean* mechanism. A party that was perceived to be relatively extreme prior to its electoral defeat enjoys an improved reputation while in opposition, since the party undergoes changes that move it closer to moderation with higher probability. Thus, a government that is elected with a bare reputation advantage, is bound to appear worse than the opposition in the election that follows its electoral victory. As we already discussed, governments with such a disadvantage may pursue extreme policies in equilibrium.

Third, with regard to policy dynamics, the equilibrium produces two qualitatively different adjustment paths depending on initial conditions. Either there is positive probability of extreme policies in the future for a protracted period of time, when parties' reputations are worse than a given steady state level; or there is zero probability of future extreme policies in the opposite case. Thus, unlike in the static models of Downs or Wittman, two-party parliamentary systems may experience spells of (relative) extremism or moderation. In the long-run, parties' perceived extremism converges to levels that guarantee moderate policies with probability one.

Finally, with respect to electoral dynamics, the equilibrium is consistent with the empirical evidence we report in table 1. Along the equilibrium adjustment path one of the two parties may hold a significant electoral advantage for protracted periods of time. In the absence of probabilistic elections, this electoral advantage is specific to one of the two parties in the case of the equilibrium adjustment path with policy moderation. In the adjustment path with policy polarization, the electoral advantage may alternate between the two parties over time.

Before we review related literature, we conclude with a remark on the applicability of our model which is confined to partisan, not individual candidacies. We believe that among instances of two party competition our analysis is best suited for parliamentary systems. In these systems, a party with a majority in parliament can exercise effective control of parliamentary procedure in order to implement a relatively unconstrained policy agenda. On the other hand, our theory would require modifications in order to be applicable for a Congressional such as that prevailing in the US, where policy choices arise as complex compromises among multiple institutional actors.

Recent reviews of the literature on two party/candidate competition appear in Duggan, 2004, and Grofman, 2004. Partisan models of two-candidate competition under complete information include those of Aldrich, 1983, and John Roemer 1999, 2000. In Aldrich, 1983, policy divergence is due to party activism and voter alienation. Roemer proposes a model of party competition in multiple issue dimensions that is premised on the idea that disagreement within parties generates party competition equilibria when none of the parties can unanimously improve on their own platform given the platform of the opposition. This model produces a range of equilibria, that subsume the predictions of the models of Downs or Wittman.

Static models of two-candidate competition in which one of the candidates has an exogenous advantage include Groseclose, 2001, and Aragonés and Palfrey, 2005. Alesina, 1988, and Duggan and Fey, 2004, study repeated versions of the models in which candidates have policy and office motivations, respectively, and the electorate can devise complex, history dependent strategies. They characterize the set of subgame perfect equilibria which are consistent with a wide range of policy platforms for the parties in Alesina, 1988, and include all possible equilibrium policy outcomes in Duggan and Fey, 2004. Gul, Dixit, and Grossman, 2000, use the same equilibrium concept characterizing efficient equilibria in a model in which parties' re-election probabilities follow an exogenous Markov process that depends on the incumbent's policy choice.

Unlike our study, all of the above models assume complete information. Related studies with incomplete information include those by Banks and Sundaram, 1993, Duggan, 2000, Banks and Duggan, 2002, and Bernhardt, Dubey, and Hughson, 2004.<sup>8</sup> These models apply to situations with individual candidacies, where challengers to the incumbent are drawn from an identical pool of

---

<sup>8</sup>Other models of incomplete information focus on the fact that the incumbent's action while in office is unobserved (hidden action). Such models include Ferejohn, 1986, Rogoff and Sibert, 1988, Rogoff, 1990, and Meirowitz, 2003. Banks and Sundaram, 1993, combine aspects of both models.



possible candidates over time. As we have discussed, the assumption of challengers drawn from a stationary distribution seems inappropriate for partisan candidacies because of inertia in party organizations. In our analysis, beliefs about the incumbent party's opposition in each period are influenced by past behavior of that opposition party.

The remainder of this paper is organized as follows. In section 2 we describe the model. We analyze this model in sections 3 and 4. Section 4 contains the main results of the paper concerning situations in which parties value office significantly compared to policy. We characterize an equilibrium (proposition 3), discuss equilibrium properties (proposition 4), and equilibrium dynamics (proposition 5). In section 5 we extend our analysis to the case the electoral outcome is not deterministic. We conclude in section 6.

## 2. MODEL

The game is played between the electorate represented by a moderate or median voter,  $M$ , and a set of partisan 'types' within each of two political parties. These players interact an infinity of periods  $t = 1, 2, \dots$ . We denote a generic party by  $P$ , which is either a left-wing party ( $P = L$ ), or a right-wing party, ( $P = R$ ). We also use  $-P$  to denote the party in opposition of party  $P$ .

Each of the two parties contains individuals with two different ideological convictions, call them *moderates* and *extremists*. These two groups/types disagree as to the optimal government policy. In each period one of the two groups holds the prevailing ideological position of the party. Thus, in period  $t$  party  $P \in \{L, R\}$  is either an extreme type,  $e$ , or a moderate type,  $m$ . We denote party  $P$ 's type in period  $t$  by  $\tau_P^t$  with  $\tau_P^t \in \{e, m\}$ .

We assume that following elections an internal shake up occurs within the losing party that may upset the balance of power between extremists and moderates within that party. Depending on the outcome of this internal battle, the type of the losing party may become moderate or extreme with positive probability. We assume the prevailing type within the party is determined probabilistically because of factors no group within the party can have full control over (*e.g.* the availability of sufficient number of competent members in positions of influence such as MPs, MP candidates, in local and national party organizations, unions, etc., as well as the existence of competent leadership for each of the two ideological groups within the party).

The only assumption we make regarding this process of party type determination (inequality (1) below), is that the new partisan type is extreme (moderate) with higher probability if the

party was controlled by extremists (moderates) in the previous period. This serial correlation is due to inertia in the manner in which partisan populations evolve, or due to the fact that the prevailing ideological group in the party commands resources and/or other institutional advantages that render it better positioned to fight the internal battle for control of the party in the next period.

Formally, we assume that if the party loses the election in period  $t$  when its type is  $\tau \in \{e, m\}$ , it is of the same type in period  $t + 1$ , with probability  $\pi_\tau$ ,  $\tau \in \{e, m\}$ . These transition probabilities satisfy

$$1 > \pi_e > 1 - \pi_m > 0, \tag{1}$$

so that the probability of an extreme party type is higher if the party's type in the previous period is extreme. On the other hand, if the party wins the election, then its type remains unaltered between periods. Parties know the realization of their own type in each period, but that information is not revealed to other players except via policy consequential choices of the party/type while in government.

**Beliefs/Party Reputations** Players hold (and rationally update) beliefs about the probability that each party is moderate or extreme. We can think of, and will refer to, these beliefs as the *reputation* of each party regarding its moderation. In particular, in each period there is a pair of probabilities  $\mathbf{b} = (b_L, b_R) \in \mathcal{B}$ , where  $\mathcal{B} \equiv [0, 1]^2$ , that represent the common beliefs of the voter about the two parties and of the parties for each other.<sup>9</sup> Thus, probability  $b_L$  represents the belief of  $M$ , (and party  $R$ ) that party  $L$  is extreme. Similarly,  $b_R$  is the corresponding belief that party  $R$  is extreme.

**Timeline of the Game** Each period in the game is a complete political cycle. A period starts with elections, in which the voter,  $M$ , chooses one of the two parties to control the government. Following the election, the party that ended up in the opposition undergoes a reorganization that may result in a change in its type (extreme or moderate). The party/type in government

---

<sup>9</sup>Note that parties know the beliefs of the electorate regarding the extremism of the two parties. If players know the initial priors of the electorate at the beginning of the game, they can trace the posterior beliefs of the electorate in each period, because the information via which these beliefs are updated is publicly available. Alternatively, we can assume parties acquire such knowledge of the electorates' beliefs directly in each period via polling and similar devices.

chooses and implements a policy  $x^t \in X$ .<sup>10</sup>

In general, there are four possible policies in each period, a left-wing policy,  $x_e^L$ , a moderate left-wing policy,  $x_m^L$ , and corresponding right-wing policies  $x_e^R$  and  $x_m^R$ . As will become evident by our assumptions on players' payoffs, we do not preclude the possibility that  $x_m^L = x_m^R$  is a common policy. This permits a 'convergence to the median' equilibrium to occur. But we allow  $x_m^L \neq x_m^R$ , *i.e.*, there may exist residual partisanship even if the moderates are the prevailing group within each party. In summary, we have  $X = \{x_e^L, x_m^L, x_m^R, x_e^R\}$ .

[insert figure 1 here]

With regard to the choice of government policies, we assume (naturally) that moderate types always implement the moderate policy  $x_m^P$ .<sup>11</sup> The strategic burden in the model is borne by the extreme partisan types. In particular, extreme types ( $\tau_P^t = e$ ) may choose either an extreme policy,  $x_e^P$ , revealing their type, or a moderate policy,  $x_m^P$ , imitating moderate types.

The policy choice by the governing party is observed by all players and the game moves to the next period. In that period, the voter elects a new government, new partisan types are realized, the governing party implements a policy, etc. The timeline of the game is represented graphically in figure 1.

**Preferences** Since moderate partisan types always pursue the same action, we only need state payoffs for the voter and the two extreme partisan types (left and right). The preferences of

---

<sup>10</sup>The exact sequencing of the incumbent government's policy choice and the realization of the opposition party's new prevailing type is not essential for the analysis as long as both occur prior to the following election. Empirically, parties that lose the election undergo a period of internal restructuring that may alter the balance of power within the party at the beginning of the interelection period. On the other hand, partisans typically exercise restraint and refrain (or are compelled to do so) from such internal fights closer to new elections or while the party is in power.

<sup>11</sup>This is the behavior that would arise endogenously in an equilibrium of the type we characterize.

these players over policies in  $X$  are summarized by the following within period (or stage) payoffs:

	Payoff from policy:			
	$x_e^L$	$x_m^L$	$x_m^R$	$x_e^R$
voter $M$	$v_e^L$	$v_m^L$	$v_m^R$	$v_e^R$
Type $e$ of Party $L$	$u_e^L$	$u_m^L$	$a_m^L$	$a_e^L$
Type $e$ of Party $R$	$a_e^R$	$a_m^R$	$u_m^R$	$u_e^R$

We assume that  $v_m^L = v_m^R > v_e^L = v_e^R$ , *i.e.*, the voter prefers moderate policies and parties are symmetrically located in each direction from the voter. For extreme partisan types  $\tau_P = e$ , we assume  $u_e^P > u_m^P \geq a_m^P > a_e^P$ ,  $P \in \{L, R\}$ .

The above preferences coincide with the intuitive interpretation of the different types: extremists of each party prefer the respective partisan policy most, moderate policies next, and they least prefer the partisan policy of the other party. To preserve the symmetry of the game, we set  $u_\tau^L = u_\tau^R$  and  $a_\tau^L = a_\tau^R$ ,  $\tau \in \{e, m\}$ . A graphic rendition of admissible configurations of policies in the classical one-dimensional spatial model is given in figure 2.

[insert figure 2 about here]

While the voter only cares about the policy outcome, parties also prefer to control the government independent of the policy that the government pursues. In particular, partisan types receive utility  $G \geq 0$  when their *party* (*i.e.*, independent of prevailing party type) is in government. We assume that the voter is strategic but cares only about the policy outcome in the current period. Partisan types are (potentially) more farsighted and care about the electoral and policy outcome in two periods, the current period  $t$  as well as period  $t + 1$ . The weight parties place in the outcome of the next period is given by a discount factor  $\delta \in [0, 1]$ .

**Party Strategies** We shall focus our attention on Perfect Bayesian Nash equilibria in strategies that are appropriately Markovian, *i.e.*, strategies that depend only on a summary of the history of the game in each period. Given the structure of the model, the payoff relevant strategic environment for the parties is summarized by the electorates' beliefs about the probability that extremists control the two parties,  $\mathbf{b} \in \mathcal{B}$ . In other words, from the perspective of a player about to choose an action in period  $t$ , the payoff relevant information is the current reputation of the two

parties given by beliefs  $\mathbf{b} \in \mathcal{B}$ , even if these beliefs may have been reached from different histories.<sup>12</sup>

Thus, a strategy for type  $e$  of party  $P$  is given by a function:

$$\sigma^P : \mathcal{B} \longrightarrow [0, 1], P \in \{L, R\}. \quad (2)$$

Accordingly,  $\sigma^P(\mathbf{b})$  is the probability that extreme type,  $e$ , of party  $P$  implements policy  $x_e^P$  when party reputations are given by  $\mathbf{b} \in \mathcal{B}$ .

In principle, we could similarly restrict the voter,  $M$ , to pursue a Markovian strategy that depends only on party reputations  $\mathbf{b} \in \mathcal{B}$ . Instead, we allow the voter's strategy to also depend on the policy choice of the party in government in the period prior to the election.<sup>13</sup> This allows us to build a *retrospective* element on voter's strategies, even though the voter is still prospective and strategic. Furthermore, this type of history dependence ensures the existence of equilibrium, which is not in general guaranteed in this class of stochastic games.

Thus, a voter strategy is given by a function

$$\Phi : X \times \mathcal{B} \longrightarrow [0, 1]. \quad (3)$$

Now,  $\Phi(x^{t-1}, \mathbf{b}) = 1$  means voter elects party  $L$  in government in period  $t$ , while  $\Phi(x^{t-1}, \mathbf{b}) = 0$  means the right-wing party is elected in government.  $\Phi(x^{t-1}, \mathbf{b}) \in (0, 1)$  means the voter randomizes accordingly.

**Evolving Reputations** In order to update their beliefs regarding the extremism of the two parties, players use all the information accumulated between two consecutive elections. This information includes both the identity of the party in opposition, as well as the policy pursued by the government.

Specifically, players use Bayes' rule to update beliefs about the incumbent party after observing its policy choice. They also use their knowledge of the structure of the game in order to update beliefs regarding the opposition party's type, which is drawn by nature according to the probabilities in (1). We emphasize that both of these pieces of information (the government's policy choice and the identity of the party in opposition) are publicly observable.

---

<sup>12</sup>For a dynamic game in which players' Markov strategies are conditioned on beliefs in a similar fashion, see Mailath and Samuelson, 2001.

<sup>13</sup>This is in addition to the indirect effect that these policies have on the voter's beliefs. In other words, the voter may choose a different voting action following two different policies, even if these two policies lead to the same posterior beliefs.

Thus, the updated beliefs of voter  $M$  in period  $t+1$  with beliefs in period  $t$  given by  $(b_L, b_R) \in \mathcal{B}$  are represented by a function  $\beta : \mathcal{B} \times \{L, R\} \times X \rightarrow \mathcal{B}$ . The coordinate of  $\beta$  that corresponds to the updated belief about the incumbent party  $P$  after it implemented a policy  $x \in X$  is obtained by Bayes' rule as

$$\beta_P(\mathbf{b}, -P, x) = \begin{cases} 1 & \text{if } x = x_e^P \\ \frac{(1-\sigma^P(\mathbf{b}))b_P}{1-\sigma^P(\mathbf{b})b_P} & \text{if } x = x_m^P, \sigma^P(\mathbf{b})b_P < 1 \end{cases}, P \in \{L, R\}. \quad (4)$$

Recall that  $\sigma^P(\mathbf{b})$  is the probability (to be determined endogenously in equilibrium) that an extreme type of party  $P$  chooses an extreme policy,  $x_e^P$ , when in government, with party  $-P$  in opposition.

Effectively, we assume a (standard) refinement on beliefs by assuming the electorate believes the government party is extreme with probability one after observing an extreme policy ( $\beta_P(\mathbf{b}, -P, x_e^P) = 1$ ) even if such a policy has zero probability in equilibrium ( $\sigma^P(\mathbf{b}) = 0$ ). Similarly, we complete (4) in the out of equilibrium event that a moderate policy is observed when  $b_P = \sigma^P(\mathbf{b}) = 1$ , by setting  $\beta_P(\mathbf{b}, -P, x_m^P) = 0$ .

Note that  $\beta_P(\mathbf{b}, -P, x_m^P) = b_P$  if  $\sigma^P(\mathbf{b}) = 0$ , *i.e.*, there is nothing learned about the governing party if both extreme and moderate types choose moderate policies with probability one. In this case (when  $\sigma^P(\mathbf{b}) = 0$ ) the updated reputation about the governing party is identical to that this party attained prior to implementing a government policy.

[insert figure 3 about here]

The reputation of the opposition party changes because the electoral loss triggers an internal shake up in the party that may alter the *status quo* within that party. Thus, the coordinate of  $\beta$  that corresponds to the opposition party  $P$  is given by:

$$\beta_P(\mathbf{b}, P, x) = \pi_e b_P + (1 - \pi_m)(1 - b_P), P \in \{L, R\}. \quad (5)$$

There is a level of belief about the extremity of the opposition party, at which that party's reputation remains unchanged between periods. Since this level is important in our ensuing discussion, we derive this 'steady state' probability of the party's reputation by solving  $\beta_P((b_{-P}, b^o), P, x) = b^o$  to get

$$b^o = \frac{1 - \pi_m}{2 - \pi_e - \pi_m}. \quad (6)$$

It is straightforward to verify using assumption (1) that

$$b_P > \beta_P((b_{-P}, b_P), P, x) \iff \beta_P((b_{-P}, b_P), P, x) > b^o \iff b_P > b^o \quad (7)$$

*i.e.*, a party with a perceived extremism above (below) the long-term steady state is moving monotonically toward that steady state from either direction. In figure 3 we depict the changes in the reputation of the opposition party following an election. Note that the direction and magnitude of change of beliefs regarding the party’s extremism differ with the direction and distance of beliefs at the time of the election from the steady state reputation level  $b^o$ .

**Equilibrium Concept** As we have already discussed in defining players’ strategies, the equilibrium concept we use is a semi-Markov version of Perfect Bayesian Nash equilibrium. We provide a formal definition of the equilibrium in the appendix. In the remainder, we explain certain refinements of this concept that we apply in our analysis.

First, the history dependence of voting strategies allows us to incorporate a retrospective element on voting behavior, in accordance with prevalent claims regarding retrospective voting behavior in the empirical literature. In particular, we say that a voting strategy is *retrospective* if the voter does not re-elect a party that pursued an extreme policy while in government in the period prior to the election. This amounts to requiring:

$$\Phi^* (x_e^P, \beta(\mathbf{b}, -P, x_e^P)) = \begin{cases} 1 & \text{if } P = R \\ 0 & \text{if } P = L \end{cases}, \text{ for all } \mathbf{b} \in \mathcal{B}. \quad (8)$$

Since parties only care about one future period, a retrospective voting strategy gives parties a strong incentive to pursue moderate policies. Thus, equilibria in which parties pursue extreme policies are significantly more credible when they are retrospective equilibria. We emphasize that we do not assume retrospective voting as a “hard-wired” behavioral trait of the electorate. In other words, retrospective voting constitutes a best response if present in a retrospective equilibrium.

Our equilibrium definition leaves room for a further refinement on voting strategies. To motivate this refinement, note that in general the voter is allowed to randomize in arbitrary fashion in the elections when indifferent between the two parties. This is not controversial in our symmetric setup if the reputation of the two parties is identical. But if the beliefs about the two parties diverge, it seems intuitive that the voter may favor the party that is perceived to be more moderate.

This intuition becomes obvious if we consider the case both parties are expected to pursue a moderate policy with probability one following elections. In that case, the voter is indifferent between the two parties. But this indifference is not robust to the possibility of a (small) exogenous probability  $\varepsilon > 0$  that extreme types may ‘tremble’ and choose an extreme policy, contrary to the

prescribed equilibrium strategy. Clearly, if such trembles are possible, the voter strictly prefers that between the two parties that is perceived less extreme.

Informally, a *robust equilibrium* is one in which the voter’s strategy is a best response even if extreme party types may ‘tremble’ away from their equilibrium strategy choice with a small probability  $\varepsilon$ . We define robust equilibria formally in the appendix. There exists an obvious connection between our requirement and standard refinement arguments dating to Selten’s trembling hand perfect notion. It is important to emphasize that the concepts are also different. First, we consider one among a (very) large range of possible trembles of partisan strategies. Furthermore, we do not consider the consequences of such trembles on the optimality or robustness of parties’ strategies, even though that is an obvious avenue to pursue. Our goal with a robust equilibrium is more limited in that we simply seek to resolve the electorate’s possible indifference in a manner that is responsive to its beliefs about the relative extremism of the two parties.

There is a more direct (and apparently more restrictive) manner to impose such a refinement. In particular, it seems intuitive in our setup to conjecture that parties are *weakly* preferred by the voter when they are perceived to be less extreme than the opposition. Thus, if we require this intuitive property *and* resolve indifference in favor of the least extreme party, we may define an *intuitive equilibrium* as one in which the voting strategy satisfies

$$\Phi^*(x, \mathbf{b}) = \begin{cases} 1 & \text{if } b_L < b_R \\ 0 & \text{if } b_L > b_R \end{cases} \quad (9)$$

for all  $x \in X$ .

Note that, because of assumption (1), condition (8) is implied by condition (9). In other words, an intuitive equilibrium must be retrospective because the posterior belief following an extreme policy,  $x_e^P$ , by party  $P$  is given by  $\beta_P(\mathbf{b}, -P, x_e^P) = 1 > \beta_{-P}(\mathbf{b}, -P, x_e^P)$ . Nevertheless, we will use the somewhat redundant combined term referring to an equilibrium as an *intuitive retrospective* equilibrium if the voter’s strategy satisfies (9) (and (8)).<sup>14</sup>

In the next two sections we proceed to an analysis of the game. First we consider the analogues of ‘pooling’ and ‘separating’ equilibria in our dynamic game. In such equilibria, extreme

<sup>14</sup>A second remark is that, in effect, condition (9) renders the surviving equilibria closer to genuine Markov Perfect equilibria. In particular, the voter is limited to (possibly) condition her action on past policy choices only in a set of payoff relevant states  $b \in B$  such that  $b_L = b_R$ . Thus, an intuitive equilibrium involves voting strategies that are Markovian, except for a set of payoff relevant states  $b \in B$  of measure zero.



partisan types pursue the same policy (moderate or extreme, respectively) independent of party reputations  $\mathbf{b} \in \mathcal{B}$ , hence we call these equilibria simple. Our main results appear in section 4, where we consider robust retrospective equilibria that are not simple and involve parties that place high weight in office (high  $G$ ) and in the future (high  $\delta$ ).

### 3. SIMPLE EQUILIBRIA

Naturally, the primary focus of our analysis is in the dynamics induced by the strategic calculus of extreme partisan types when they contemplate the trade-off between a (preferable) extreme policy in the current period and the possible utility loss in the next period due to averse electoral consequences from this extreme policy choice. In particular, we are interested in the range of possible combinations of party reputations (*i.e.*, beliefs held by the electorate about the extremism of these parties) for which the extreme partisan types pursue extreme policies (if at all), and the policy dynamics these strategies generate.

Before we move to this more interesting analysis, we consider two simple types of equilibria in which parties' strategy does not depend on party reputations  $\mathbf{b} \in \mathcal{B}$ . First, in proposition 1, we give a precise range of parameters in which extreme partisan types implement extreme policies whenever in power, independent of party reputations  $\mathbf{b} \in \mathcal{B}$ . We have:

**Proposition 1** *An equilibrium in which extreme partisan types pursue an extreme policy with probability one independent of party reputations,  $\mathbf{b} \in \mathcal{B}$ , exists if and only if*

$$\delta \leq \frac{u_e^P - u_m^P}{\pi_e (a_m^L - a_e^L) + G + u_e^P - a_e^P}. \quad (10)$$

*All such equilibria are robust, intuitive, and retrospective. If the inequality (10) is strict then all equilibria must involve such party strategies.*

Note that (except for the case of equality) when condition (10) holds all equilibria of the game involve extreme partisan types pursuing their ideal policy, independent of the electorates' beliefs. Thus retrospective voting is not sufficient to induce moderation when either (a) parties are impatient (low  $\delta$ ), or (b) parties place low value to office (low  $G$ ), or (c) the loss in utility due to the policies pursued by the opposition party controlling the government is small (low  $u_e^P - a_e^P$ ), or (d) when the ability of extremists to maintain control of the party following electoral defeat is small (low  $\pi_e$ ).

One may conjecture that when these conditions are reversed we may instead obtain a simple ‘pooling’ equilibrium in which extreme partisan types always imitate the moderate partisan types by pursuing a moderate policy. It is possible to construct such equilibria (for high enough  $G$  &  $\delta$ ) exploiting voters’ indifference, but these equilibria are not robust. Indeed we can show that there does not exist a robust ‘pooling’ equilibrium:

**Proposition 2** *There does not exist a robust equilibrium such that all governments pursue moderate policies with probability one, independent of party reputations,  $\mathbf{b} \in \mathcal{B}$ .*

Thus the analogue to a ‘convergence to the median’ result is not attainable in our game in a robust equilibrium, even with retrospective voting. The reasoning behind proposition 2 is straightforward. If all party types moderate policies independent of the electorate’s beliefs, then the electorate is indifferent between the two parties. In a robust equilibrium, the voters then will elect that between the two parties that is (strictly) perceived to be more moderate. Thus, a party that is in government, is controlled by extremists, and is perceived to be more extreme than the opposition even when implementing a moderate policy, has no incentive to pursue such a moderate policy. *This party faces electoral defeat independent of policy choice, so types in control of the party might as well pursue their ideal policy.*

Thus, in combination, propositions 1 and 2 imply that when condition (10) fails, a robust retrospective equilibrium must involve some positive probability of moderate policies pursued by extreme types who anticipate a future electoral gain from doing so, as well as some positive probability of extreme policies pursued by these types. We take the analysis of such more interesting equilibria in the next section.

#### 4. EQUILIBRIUM WITH OFFICE MOTIVATED PARTIES

Proposition 1 demonstrates that the strategic calculus of parties is trivial when partisans are primarily motivated by policy. Such parties/types simply pursue their ideal policy. Thus, the interesting strategic environment is one in which parties value office significantly compared to policy (high  $G$ ) and are patient (high  $\delta$ ). Three questions emerge in such an environment: (a) Does there exist a robust retrospective equilibrium in which extreme party types pursue extreme policies (if elected) for some beliefs? (b) Are extreme policies observed along the equilibrium path?, and (c) what are the policy and electoral dynamics that prevail? In what follows, we answer question (a),

(b), and (c) in subsections 4.1, 4.2, and 4.3 respectively.

#### 4.1 Equilibrium

Our goal in this section is to establish an equilibrium when condition (10) fails, and parties are sufficiently patient and motivated predominantly by office considerations (high  $G$ ). Proposition 3 establishes that there exists a *unique* intuitive equilibrium in these cases:

**Proposition 3** *Assume*

$$\delta > \frac{u_e^P - u_m^P}{G + u_m^P - a_m^P}. \quad (11)$$

*There exists a unique<sup>15</sup> intuitive retrospective equilibrium in which the probability of an extreme policy choice by extreme partisan types is given by*

$$\sigma^P(\mathbf{b}) = \begin{cases} \frac{b_P - \pi_e b_{-P} - (1 - \pi_m)(1 - b_{-P})}{b_P(1 - \pi_e b_{-P} - (1 - \pi_m)(1 - b_{-P}))} & \text{if } b_P > \pi_e b_{-P} + (1 - \pi_m)(1 - b_{-P}), \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

*This equilibrium is robust.*

Note that parties' equilibrium mixing probabilities are independent of  $G$ ,  $\delta$ , or of the players' payoffs  $u_\tau^P$ ,  $a_\tau^P$ ,  $v_\tau^P$ ,  $P \in \{L, R\}$ ,  $\tau \in \{e, m\}$ . Furthermore, the equilibrium in proposition 3 holds for arbitrarily large values of  $G$ , as long as parties place some weight in the future ( $\delta > 0$ ).

Thus, no matter how office oriented parties are, there exists a configuration of beliefs by the electorate about the extremism of the two parties that makes it worthwhile for extreme partisan types to pursue extreme policies when in government. As we explain shortly, this occurs when the party is disadvantaged electorally. Figures 4(i) display the equilibrium probability of an extreme policy choice by extreme partisans of party  $L$ ,  $\sigma^L(\mathbf{b})$ , for different values of the parameters  $\pi_e$ ,  $\pi_m$ .

[insert figure 4 here]

From the perspective of the electorate, the expected probability that, say, party  $L$  will pursue an extreme policy given beliefs  $\mathbf{b} \in \mathcal{B}$ , is given by  $b_L \sigma^L(\mathbf{b})$ . In figures 4(ii) we plot this probability. In both cases of figures 4(i) and 4(ii) it is straightforward to verify via calculus or visual inspection that the probability of an extreme policy increases as the party is perceived to be more extreme.

<sup>15</sup>In particular, the equilibrium uniquely determines the value of the voting strategy,  $\Phi$ , except for  $\Phi(x, b, b)$ ,  $b < 1 - \pi_m$  or  $b > \pi_e$ . Due to (5), these beliefs cannot be reached along the path of play.

Furthermore, the party must be perceived *relatively* more extreme than it's opposition in order for it to pursue extreme policies.

The reason why extreme types of party  $L$  mix between extreme and moderate policies for beliefs  $(b_L^t, b_R^t) \in \mathcal{B}$  in period  $t$  with  $b_L^t > \pi_e b_R^t + (1 - \pi_m)(1 - b_R^t)$ , becomes obvious if we consider the electoral prospects of this party by following either an extreme or a moderate policy with probability one (instead of mixing between the two). For such beliefs,  $(b_L^t, b_R^t)$ , the party loses the election in an intuitive equilibrium if it pursues a moderate policy with probability one. This is because in that case (if  $\sigma^L(\mathbf{b})$  is equal to zero), the policy of the government conveys no information to the electorate and other players in the game regarding the government's type. Thus, following a moderate policy choice, posterior beliefs in period  $t + 1$  satisfy  $b_L^{t+1} = b_L^t > b_R^{t+1} = \pi_e b_R^t + (1 - \pi_m)(1 - b_R^t)$ , and party  $L$  loses the election despite its attempt to appear moderate. Thus, pursuing a moderate policy with probability one is not an equilibrium.

Similarly, it is not an equilibrium for extremists of party  $L$  to implement an extreme policy with probability one ( $\sigma^L(\mathbf{b}) = 1$ ). If  $\sigma^L(\mathbf{b}) = 1$ , the electorate expect different party types to reveal their true preferences. Thence, extremists have an incentive to deviate and implement a moderate policy instead, in order to carry the upcoming election. Such a deviation convinces the voter that the party is moderate, when in fact it is extreme. Thus, the only possibility for an equilibrium is a mixed strategy, where the mixture probability is such that it makes the party barely competitive against its opponent at the elections when the realization of the party's randomization is a moderate policy.

Finally, when the party has an electoral advantage ( $b_L^t < \pi_e b_R^t + (1 - \pi_m)(1 - b_R^t)$ ) it has no incentive to spoil its electoral prospects by implementing an extreme policy. In particular, the party wins the election whenever it sets a moderate policy. Thus, given that partisan types care sufficiently about office, the only equilibrium choice is to set  $\sigma^L(\mathbf{b}) = 0$ , *i.e.*, extreme types of party  $L$  choose a moderate policy with probability one.

## 4.2 Extreme Policies Along the Equilibrium Path

The fact that extreme partisan types pursue their ideal policy with positive probability ( $\sigma^P(\mathbf{b}) > 0$ ) for some beliefs  $\mathbf{b} \in \mathcal{B}$  in proposition 3 is not, in general, sufficient to produce extreme policies along the equilibrium path. This is because these types pursue extreme policies only when their party is perceived (relatively) more extreme. But, parties that are perceived more

extreme are not elected in government in the first place. In other words, along the equilibrium path, the probability that an extreme policy is observed is regulated via appropriate *screening* from the electorate.

Do we obtain extreme policies along the equilibrium path despite this selection effect of elections? The answer is in the affirmative and our analysis provides a precise mechanism for this to occur. Extreme policies are observed in equilibrium following elections in which: (a) both parties are perceived to be extreme (above the steady state level of extremism,  $b^o$ ), and (b) the election is ‘close’. Specifically, the set of party reputations for which extreme policies are expected with positive probability is defined as follows:

$$\tilde{\mathcal{B}} \equiv \{\mathbf{b} \in \mathcal{B} : b_L > \pi_e b_R + (1 - \pi_m)(1 - b_R) \ \& \ b_R > \pi_e b_L + (1 - \pi_m)(1 - b_L)\}. \quad (13)$$

This set of beliefs for which extreme policies are possible in equilibrium is depicted graphically in figures 4(iii) for different values of the transition probabilities  $\pi_e$  and  $\pi_m$ . It is straightforward to verify from assumption (1) that this set is not empty, and that reputations  $(b_L, b_R)$  in set  $\tilde{\mathcal{B}}$  are such that both parties are above the steady state level of extremism, *i.e.*,  $b_L, b_R > b^o$ .

Extreme policies occur in this area of the space of possible party reputations due to a ‘*regression to the mean*’ effect on these reputations. Parties that are perceived relatively more extreme and barely lose the election to the opposition undergo internal changes, or reforms following their electoral defeat. Because these parties’ perceived extremism is above their steady state level (6), this internal shake up moves the party towards moderation closer to its long term perceived extremism (by condition (7)). As a result, a government that comes to power with a bare advantage, is bound to be perceived more extreme than the opposition immediately following the election. In these cases the government may pursue extreme policies.

We summarize our discussion of the equilibrium in the following proposition:<sup>16</sup>

**Proposition 4** *The equilibrium in proposition 3 is such that:*

- (a) *The (expected) probability that party L implements an extreme policy following victory in an election with party reputations  $\mathbf{b} \in \mathcal{B}$ , weakly increases with  $b_L$ , and weakly decreases with  $b_R$ ,*
- (b) *The expected probability of an extreme policy following an election with party reputations  $\mathbf{b} \in \mathcal{B}$ , is weakly increasing with  $b_L$  when  $b_L < b_R$ , and is weakly decreasing with  $b_L$  when  $b_L > b_R$ , and*

---

<sup>16</sup>The proof of this and the following proposition are straightforward and are omitted.

- (c) *The expected probability of an extreme policy following an election with party reputations  $\mathbf{b} \in \mathcal{B}$  is positive if and only if  $\mathbf{b} \in \tilde{\mathcal{B}}$ .*

We conclude our analysis in this section by considering the dynamics of beliefs and policies induced by the equilibrium in proposition 3.

### 4.3 Equilibrium Dynamics

Starting from any initial party reputations  $\mathbf{b} \in \mathcal{B}$ , beliefs evolve over time via Bayes' rule following government's policy, and via the electorate's anticipation of internal restructuring within parties that lose the election. It is straightforward to verify that equilibrium beliefs remain unchanged if for some reason the system rests at belief points  $(b_L, b^o) \in \mathcal{B}$  with  $b_L \leq b^o$  and party  $L$  is in government.<sup>17</sup> In these cases the party in government is pursuing a moderate policy with probability one, and there are no changes in beliefs regarding the opposition because the opposition is already at its steady state level of beliefs,  $b^o$ .

Indeed, the equilibrium in proposition 3 is such that party reputations converge with probability one in the long-run to some level  $(b_P, b^o)$  with  $b_P \leq b^o$ , for one of the two parties  $P$ , without any (endogenous) forces inducing a change in beliefs ever after. At all these possible long-run party reputation levels, there is probability zero of an extreme policy. But, both the eventual long-term beliefs and the path that leads to these beliefs *differ qualitatively depending on initial conditions*.

We summarize these dynamics in proposition 5, and elaborate on these points in the remainder of this subsection:

**Proposition 5** *The equilibrium in proposition 3 is such that when initial party reputations are  $\mathbf{b} = (b_L, b_R) \in \mathcal{B}$ :*

- (a) *If  $b_P \leq b^o$  for some  $P \in \{L, R\}$  there is zero probability of an extreme policy along the path of play and, in the long-run, equilibrium party reputations converge to  $(b_L, b^o)$  if  $b_L < b_R$ , to  $(b^o, b_R)$  if  $b_L > b_R$ , and to either  $(b_L, b^o)$  or  $(b^o, b_R)$  if  $b_R = b_L$ ,*
- (b) *If  $b_L, b_R > b^o$  equilibrium party reputations converge to  $(b^o, b^o) \in \mathcal{B}$  in the long-run, and for any period  $t$  there is positive probability of an extreme policy in some future period  $t' > t$ , i.e., beliefs in the set  $\tilde{\mathcal{B}}$  emerge infinitely often along the path of play.*

---

<sup>17</sup>Or, symmetrically if  $(b^o, b_R) \in \mathcal{B}$  with  $b_R \leq b^o$  and party  $R$  in government.

*Independent of initial party reputations, one of the two parties may enjoy a significant electoral advantage in multiple successive elections.*

The dynamics described in proposition 5 are illustrated graphically in figure 5. Following any election with beliefs satisfying the condition of case (a) of the proposition, the party that is elected in government is guaranteed to be perceived more moderate than the opposition. As a result, the government always implements a moderate policy and is re-elected with probability one. This process continues until players' beliefs about the extremism of the opposition party reach the steady state level  $b^o$  given by (6).

[insert figure 5 about here]

The situation is much different when both parties are perceived to be above their long-term steady state level of extremism,  $b^o$ . In these cases, we have one of two possibilities. Either the party in government is relatively more extreme than the opposition in which case it implements an extreme policy with positive probability; or, the party is favored electorally and pursues a moderate policy. In the latter case, the governing party wins re-election until internal adjustments in the opposition 'turn the tide,' and the opposition is perceived more moderate than the government.

Since both beliefs  $b_L, b_R$  exceed the long-term steady state level of perceived extremism  $b^o$ , such a situation will arise 'infinitely often' along the equilibrium path due to condition (7).<sup>18</sup> By that we mean that for any period in the game there is probability one of a future period in which party reputations belong in the set  $\tilde{\mathcal{B}}$ . As a consequence, *if the system starts from a situation in which both parties are perceived to be relatively extreme, extreme policies will occur in the future with strictly positive probability for every period along the path of adjustment to the long term equilibrium party reputations  $(b^o, b^o) \in \mathcal{B}$ .*

Importantly, the path to the long-run steady state may be quite long when  $b_L, b_R > b^o$ , depending on the values of transition probabilities  $\pi_e, \pi_m$ . Thus, even though in the long run the political system converges to a situation consistent with the predictions of Downsian competition, equilibrium adjustment dynamics may contain a significant number of electoral cycles away from that long-term steady state and with a positive expectation of extremism.<sup>19</sup>

---

<sup>18</sup>Of course, the probability of an extreme policy dissipates as beliefs approach the long term level  $(b^o, b^o)$ .

<sup>19</sup>In a special case, this adjustment process is immediate. This occurs when the probability that

[insert table 2 here]

Last, but not least, the equilibrium dynamics in either case (a) or (b) of proposition 5 imply that one of the two parties may enjoy an electoral advantage in successive elections. To corroborate that claim, we report calculations of equilibrium probabilities of the possible outcomes (victorious party and government policy) in a sequence of 20 electoral cycles in table 2, for certain initial conditions and parameter values.

As is immediate from table 2, the implied pattern for the sequencing of electoral victories for the two parties from these calculated probabilities is highly consistent with the observed data in parliamentary systems in which two parties alternate in power, such as New Zealand or the UK. In these systems, as evident from the data in table 1, one of the two parties has enjoyed electoral success for 3, or 4 consecutive elections for time periods spanning well over a decade. On the contrary, static theories of two party competition produce equilibrium predictions that consistently project the same probability of victory for one of the two contesting parties in equilibrium.

## 5. PROBABILISTIC ELECTIONS

The model in the previous section constitutes a clean, baseline environment from which to evaluate the consequences of introducing more complicated assumptions. In this section we consider one such extension, namely the possibility of probabilistic elections.

Even for the most tame political environments it is reasonable to assume that events out of the control of the players may influence the outcome of the electoral campaign and give a critical electoral advantage to one of the two parties contesting for power. Such exogenous events can be both favorable to the government (*e.g.* a victorious war or success in foreign policy) or the opposition, (*e.g.* scandal involving the government, etc.). They may simply represent a temporary swing on the electorate's ideological convictions.

To incorporate this possibility, we assume that in each election period there is an (exogenous) probability,  $w$ , that the incumbent government is re-elected or ousted, independent of the voter's a party is extreme following an internal shake-up is independent of the previous identity of the prevailing group in the party, i.e. when  $b^o = \pi_e = 1 - \pi_m$ . This case corresponds to the assumption in existing electoral models with incomplete information (*e.g.* Banks and Sundaram, 1993, or Banks and Duggan, 2002). In these models, extremism (or quality) is not serially correlated over time, and equilibrium is stationary along the equilibrium path.



strategy. In effect, this amounts to restricting the voter's strategy to:

$$\Phi : X \times \mathcal{B} \longrightarrow [w, 1 - w], \quad 0 < w < \frac{1}{2}, \quad (14)$$

with the obvious modifications on conditions (8) and (9).<sup>20</sup>

Of course, with this assumption a party that is perceived relatively more extreme is no longer guaranteed defeat in elections. This has two main implications for our analysis. First, substantively we obtain equilibrium outcomes that are closer to empirical observation. Just like Margaret Thatcher's Conservatives won reelections in 1983, there is positive probability that a party may remain in government and implement an extreme policy in successive periods in our analysis.

The second implication of our assumption on the probabilistic nature of elections has to do with equilibrium dynamics. In particular, it is straightforward to derive the following extension of propositions 3 and 5:<sup>21</sup>

**Proposition 6** *Assume (14) and*

$$\delta > \frac{u_e^P - u_m^P}{(1 - w)(G + u_m^P - a_m^P) - w(G + u_e^P - a_e^P)}. \quad (15)$$

- (a) *There exists a unique intuitive, retrospective equilibrium with partisan strategies given by (12). This equilibrium is robust.*
- (b) *The expected probability of an extreme policy following an election with party reputations  $\mathbf{b} \in \mathcal{B}$ , is positive independent of the identity of the victorious party if and only if  $\mathbf{b} \in \tilde{\mathcal{B}}$ ,*
- (c) *For initial party reputations  $b_L, b_R > b^\circ$ , beliefs in  $\tilde{\mathcal{B}}$  occur infinitely often along the equilibrium path,*
- (d) *For initial party reputations with  $b_P \leq b^\circ$  for some  $P \in \{L, R\}$ , beliefs in  $\tilde{\mathcal{B}}$  never occur in equilibrium,*
- (e) *For any initial party reputations  $\mathbf{b} \in \mathcal{B}$ , equilibrium party reputations converge to  $(b^\circ, b^\circ) \in \mathcal{B}$  in the long-run.*

---

<sup>20</sup>A slightly more complicated assumption in the same spirit is to assume  $w$  is an appropriate function of the electorate's beliefs  $b \in B$ . This can be implemented in the analysis to follow, without any gain in insight.

<sup>21</sup>The proof is available upon request.

Condition (15) is analogous to condition (11), adjusting for the probabilistic nature of elections. This is a sufficient (not necessary) condition that is satisfied for large enough  $G$ , *i.e.*, if parties are sufficiently office oriented.

Qualitatively, the equilibrium in proposition 6 is very similar to that in proposition 3. In particular, the strategies of extreme party types are identical, and equilibrium dynamics display similar properties. When the belief about the extremity of at least one of the two parties is less than or equal to  $b^o$ , then extreme policies are observed only when, contrary to the systematic<sup>22</sup> preference of the median voter (with probability  $w$ ), a relatively moderate party loses the election. Furthermore, along the path of adjustment at least one of the two parties is always perceived to be extreme with probability less than  $b^o$ .

On the other hand, if the reputation level of both parties exceeds the long-term level  $b^o$  there is positive probability of extreme policies by both contestants in the election along the equilibrium path of adjustment. In other words, for beliefs that are visited infinitely often along the equilibrium path there are elections in which the winner implements extreme policies with positive probability, whether the winner is the relatively moderate party or not. Thus, it is still the case that the path of adjustment for initial party reputations  $b_L, b_R > b^o$  produces more policy extremism than is the case when either  $b_L \leq b^o$  or  $b_R \leq b^o$ .

The main difference of the adjustment dynamics in proposition 6 compared to proposition 3, is that with probabilistic voting the party reputations at which the political system converges to in the long-run do not depend on initial conditions. In particular, equilibrium beliefs eventually converge to  $(b^o, b^o) \in \mathcal{B}$  from any initial level  $\mathbf{b} \in \mathcal{B}$ .

## 6. CONCLUSIONS

Stylized models of two candidate competition assume the contenders for office make truthful policy promises which voters take at face value. In this study, we have departed from these models to propose a reputation based theory of two-party competition. We have assumed that political parties enter the electoral arena with a(n endogenously formed) reputation regarding the prevailing policy preferences within the party. Instead of campaign promises, these party reputations shape the electorate's expectations about the policies that are likely to be pursued by the victorious party

---

<sup>22</sup>*i.e.*, when exogenous shocks in preferences such as scandals, foreign policy developments, etc. alter the median's ranking between the two contesting parties.

following elections. The second basic premise of our analysis is that party reputations improve or deteriorate gradually following electoral defeat, due to inertia within party organizations.

From these two simple premises, we have built a dynamic model of two party competition in which government actions are dependent not only on the incumbent party's reputation among the electorate, but also on the opposition party's reputation. These choices shape future reputation levels as well as future policy and electoral outcomes. We showed that in robust equilibria in which parties care sufficiently about office, the ruling party pursues extreme policies when it has a relatively worse reputation compared to the opposition. Extreme policies occur in equilibrium when (a) both parties' reputations are above some fixed steady state level and (b) elections are close, *i.e.*, both parties have similar reputations. We also showed that, depending on initial conditions, policy dynamics can be one of two types. One displays polarized politics that involve extreme policies with positive probability along the adjustment path. The other involves moderate policies with probability one along the adjustment path. In the long run, both parties pursue moderate policies with probability one, independent of initial conditions. The electoral dynamics associated with either adjustment path may involve one of the two parties enjoying an electoral advantage in consecutive elections.

As we discussed in the introduction, these results are consistent with empirical observation in a number of parliamentary systems in which two parties alternate in power. In particular, politics in such systems may alternate between phases of convergence and polarization. Also, different countries with the same institutional setup may experience completely different dynamics under the same equilibrium, due to different initial conditions. Furthermore, the electoral dynamics produced by our model as exemplified in the simulation reported in table 2 are consistent with actual electoral results in a number of countries such as those reported in table 1.

Of course, our setup leaves a number of open avenues for improvement. First, we have assumed a coarse policy/type space, by allowing only two policy choices and two possible party preferences. Allowing more party types generates a much more interesting strategic environment, since governments in our model condition their policy choice on the reputation of the opponent. As a result, with intermediate party policy preferences possible, extreme parties may push government policies in the preferred direction just enough in order to maintain an advantage against an opposition that has a worse reputation. This is not possible in the present study because any extreme policy choice by the government implies that the updated reputation for the governing party must

be worse than that of the opposition party.

Second, the equilibrium we characterize in this study when parties are primarily office motivated involves a long-run steady state level of reputations for the two parties such that both parties pursue moderate policies with probability one. This conclusion is qualified if we introduce additional noise in the political system in one of two forms: (a) if we assume that the preferences of the ruling party change with positive probability while the party is in government, or (b) if we assume random (exogenous) shocks on party reputations in any given period.

The second of the above two model modifications also allows us to introduce more generality in our analysis. In particular, asymmetry in party and voter preferences, as well as a larger set of possible party types and policies can be accommodated if we assume party reputations are subjected to an exogenous random shock. This source of noise in the political system can be construed as an alternative to our version of probabilistic elections. Under mild conditions, this alternative guarantees existence of Markovian equilibria in the associated stochastic games. Of course, such generality comes at the cost of the lack of analytical solutions, a very appealing feature of our maintained setup.

## REFERENCES

1. Aldrich, J. H. 1983. "A Downsian Spatial Model with Party Activism," *American Political Science Review*, 77(4): 974-990.
2. Alesina, A. 1988. "Credibility and Policy Convergence in a Two-party System with Rational voters," *American Economic Review*. 78: 796–805.
3. Aragonés, E. and T. Palfrey. 2005. "Spatial Competition Between Two Candidates of Different Quality: The Effects of Candidate Ideology and Private Information," in *Social Choice and Strategic Behavior: Essays in Honor of Jeffrey Scot Banks*. D. Austen-Smith and J. Duggan eds. Springer: Berlin.
4. Banks, Jeffrey and John Duggan. 2002. "A Multidimensional Model of Repeated Elections," mimeo.
5. Banks, Jeffrey and R. Sundaram. 1993. "Moral Hazard and Adverse Selection in a Model of Repeated Elections," in *Political Economy: Institutions, Information, Competition and Representation*, W. Barnett et al. eds., New York: Cambridge University Press.
6. Bernhardt, Dan, Sangita Dubey and Eric Hughson. 2004. "Term Limits and Pork Barrel Politics," *Journal of Public Economics*, 2383-2422.
7. Calvert, R. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence," *American Journal of Political Science*, 29: 69–95.
8. Dixit, Avinash, Gene M. Grossman, and Faruk Gul. 2000. "The Dynamics of Political Compromise," *Journal of Political Economy*, 108(3):531-68.
9. Downs, Anthony. 1957. *An Economic Theory of Democracy*. Harper Collins. New York. N.Y.
10. Duggan, John. 2000. "Repeated Elections with Asymmetric Information", *Economics and Politics*, 12: 109-136.
11. Duggan, John. 2004. "Candidate Objectives and Electoral Equilibrium" forthcoming (Oxford Handbook of Political Economy, edited by Barry Weingast and Donald Wittman).

12. Duggan, John and Mark Fey. 2002. "Repeated Downsian Electoral Competition," mimeo, University of Rochester.
13. Ferejohn, John. 1986. "Incumbent Performance and Electoral Control," *Public Choice*, 50: 5-25.
14. Grofman, Bernard. 2004. "Downs and Two-Party Convergence," *Annual Review of Political Science*, 7: 25-46.
15. Groseclose, T. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage," *American Journal of Political Science*, 45(October): 862-86. 2001.
16. Hotelling, H. 1929. "Stability in Competition," *Economic Journal*, 39: 41-57.
17. Ledyard, J. 1984. "The Pure Theory of Large Two-Candidate Elections," *Public Choice*, 44: 7-41.
18. Mailath, George J. and Larry Samuelson. 2001. "Who wants a Good Reputation?," *Review of Economic Studies*, 68: 415-41.
19. Meirowitz, Adam. 2003. "Accountability and representation in repeated elections with uncertain policy constraints," working paper, Princeton University.
20. Roemer, John. 1999. "The Democratic Political Economy of Progressive Income Taxation," *Econometrica*, 67(1), 1-19.
21. Roemer, John. 2000. *Political Competition*. Cambridge: Harvard University Press.
22. Rogoff, Kenneth. 1990. "Equilibrium Political Budget Cycles," *American Economic Review* 80(1): 21-36.
23. Rogoff, Kenneth and Anne, Sibert. 1988. "Elections and Macroeconomic Policy Cycles," *Review of Economic Studies* 55(1): 1-16.
24. Wittman, Donald. 1983. "Candidate Motivation: A Synthesis of Alternative Theories," *American Political Science Review*, 77: 142-157.

25. Wittman, Donald. 1990. "Spatial Strategies When Candidates Have Policy Preferences," *Advances in the Spatial Theory of Voting*, J. Enelow and M. Hinich, eds., New York: Cambridge Press.

## APPENDIX

In this appendix we state the proofs of propositions 1 to 3. Before we do so, we elaborate on some definitions that were discussed informally in the main body of the paper. Given strategies  $\Phi$ ,  $\sigma^P$ ,  $P \in \{L, R\}$ , we first derive expressions for players' expected payoffs. The expected utility of the voter  $M$  from choosing party  $P$  when beliefs are given by  $\mathbf{b} \in \mathcal{B}$ , is obtained as

$$V(\mathbf{b}, P) = \begin{cases} b_L \sigma^L(\mathbf{b}) (v_e^L - v_m^L) + v_m^L & \text{if } P = L \\ b_R \sigma^R(\mathbf{b}) (v_e^R - v_m^R) + v_m^R & \text{if } P = R \end{cases}. \quad (16)$$

Likewise, the expected payoff of type  $e$  of party  $L$  from implementing a policy  $x_\tau^L \in \{x_m^L, x_e^L\}$  is given by:

$$U^L(\mathbf{b}, x_\tau^L) = u_\tau^L + G + \delta \left( \begin{array}{l} \Phi(x_\tau^L, \beta(\mathbf{b}, R, x_\tau^L)) [\sigma^L(\beta(\mathbf{b}, R, x_\tau^L)) (u_e^L - u_m^L) + u_m^L + G] \\ + (1 - \Phi(x_\tau^L, \beta(\mathbf{b}, R, x_\tau^L))) [\beta_R(\mathbf{b}, R, x_\tau^L) \sigma^R(\beta(\mathbf{b}, R, x_\tau^L)) (a_e^L - a_m^L) + a_m^L] \end{array} \right). \quad (17)$$

Observe that this expression reflects party  $L$ 's updated assessment in the ensuing period regarding the extremity of party  $R$ ,  $\beta_R(\mathbf{b}, R, x_\tau^L)$ , after party  $R$ 's term in the opposition. The corresponding expression for party  $R$  is obtained in an analogous fashion.

With the above we state the definition of our equilibrium concept as follows:

**Definition 1** *An equilibrium is a pair of party strategies  $\sigma^{P*}$ ,  $P \in \{L, R\}$ , and a voting strategy  $\Phi^*$  such that:*

$$\Phi^*(x, \mathbf{b}) \begin{cases} = 1 & \text{if } V(\mathbf{b}, L) > V(\mathbf{b}, R) \\ \in [0, 1] & \text{if } V(\mathbf{b}, L) = V(\mathbf{b}, R) \\ = 0 & \text{if } V(\mathbf{b}, L) < V(\mathbf{b}, R) \end{cases}, \text{ for all } \mathbf{b} \in \mathcal{B} \quad (18)$$

and

$$\sigma^{P*}(\mathbf{b}) \begin{cases} = 1 & \text{if } U^P(\mathbf{b}, x_e^P) > U^P(\mathbf{b}, x_m^P) \\ \in [0, 1] & \text{if } U_e^P(\mathbf{b}, x_e^P) = U^P(\mathbf{b}, x_m^P) \\ = 0 & \text{if } U_e^P(\mathbf{b}, x_e^P) < U^P(\mathbf{b}, x_m^P) \end{cases}, \text{ for all } \mathbf{b} \in \mathcal{B}, P \in \{L, R\}. \quad (19)$$

We define a robust equilibrium as follows:

**Definition 2** *An equilibrium  $\Phi^*$ ,  $\sigma^{P*}$ ,  $P \in \{L, R\}$ , is robust if there exists an  $\bar{\varepsilon} \in (0, \frac{1}{2})$  such that for each  $\varepsilon$ ,  $\bar{\varepsilon} > \varepsilon > 0$ , the voting strategy  $\Phi^*(x, \mathbf{b})$  is a best response when party strategies are*



perturbed according to

$$\sigma_\varepsilon^{P^*}(\mathbf{b}) = \begin{cases} 1 - \varepsilon & \text{if } \sigma^{P^*}(\mathbf{b}) > 1 - \varepsilon \\ \varepsilon & \text{if } \sigma^{P^*}(\mathbf{b}) < \varepsilon \\ \sigma^{P^*}(\mathbf{b}) & \text{otherwise.} \end{cases} \quad (20)$$

We now proceed to the proofs. For ease of reference, we restate the propositions to be proven. We start with proposition 1:

**Proposition 1** *An equilibrium with  $\sigma^P(\mathbf{b}) = 1$ , for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$  exists if and only if*

$$\delta \leq \frac{u_e^P - u_m^P}{\pi_e(a_m^L - a_e^L) + G + u_e^P - a_e^P}. \quad (15)$$

*All such equilibria are robust, intuitive, and retrospective. If (10) is strict then  $\sigma^P(\mathbf{b}) = 1$ , for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$  in all equilibria.*

**Proof.** We start by showing that all equilibria with  $\sigma^P(\mathbf{b}) = 1$  for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$  must be robust and intuitive (hence retrospective). For perturbed party strategies  $\sigma_\varepsilon^P(\mathbf{b}) = 1 - \varepsilon$  we calculate voter's expected utility as

$$V_\varepsilon(\mathbf{b}, P) = b_P(1 - \varepsilon)(v_e^P - v_m^P) + v_m^P, \quad P \in \{L, R\}.$$

We have  $V_\varepsilon(\mathbf{b}, L) > V_\varepsilon(\mathbf{b}, R) \iff b_L < b_R$  for all  $\varepsilon \geq 0$ . Thus, an equilibrium with strategies  $\sigma^P(\mathbf{b}) = 1$  for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$  is robust and intuitive.

Next, we show that if  $\delta < \frac{u_e^P - u_m^P}{\pi_e(a_m^L - a_e^L) + G + u_e^P - a_e^P}$  we must have  $\sigma^P(\mathbf{b}) = 1$  for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$  in all equilibria. Suppose, to get a contradiction, that there exists an equilibrium with  $\sigma^P(\mathbf{b}) < 1$  for some  $\mathbf{b} \in \mathcal{B}$ . Then we must have  $U^P(\mathbf{b}, x_m^P) \geq U^P(\mathbf{b}, x_e^P)$  for these beliefs. Note that logically (independent of strategies) expected utilities must satisfy  $U^P(\mathbf{b}, x_e^P) \geq u_e^P + G + \delta(\beta_{-P}(\mathbf{b}, -P, x_e^P) a_e^P + (1 - \beta_{-P}(\mathbf{b}, -P, x_e^P)) a_m^P)$ . Since, by (1) and (5),  $\max \beta_{-P}(\mathbf{b}, -P, x_e^P) = \pi_e$ , the last inequality implies  $U^P(\mathbf{b}, x_e^P) \geq u_e^P + G + \delta(\pi_e a_e^P + (1 - \pi_e) a_m^P)$ , for all  $\mathbf{b} \in \mathcal{B}$ . Similarly, we obtain  $U^P(\mathbf{b}, x_m^P) \leq u_m^P + G + \delta(u_e^P + G)$ , for all  $\mathbf{b} \in \mathcal{B}$ . Thus, using these bounds, we deduce from  $U^P(\mathbf{b}, x_m^P) \geq U^P(\mathbf{b}, x_e^P)$  that  $u_e^P + G + \delta(\pi_e a_e^P + (1 - \pi_e) a_m^P) \leq u_m^P + G + \delta(u_e^P + G) \iff \delta \geq \frac{u_e^P - u_m^P}{\pi_e(a_m^L - a_e^L) + G + u_e^P - a_e^P}$ , a contradiction.

Next, we verify that  $\sigma^P(\mathbf{b}) = 1$  for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$  is part of an equilibrium when (10) holds. From the above arguments,  $\sigma^P(\mathbf{b}) = 1$  for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$  are (at least weak) best responses, independent of the voting strategy when  $\delta \leq \frac{u_e^P - u_m^P}{\pi_e(a_m^L - a_e^L) + G + u_e^P - a_e^P}$ . As a result, we

only need specify a voting strategy that is a best response. We have already shown that the voting strategy must satisfy  $\Phi(x, \mathbf{b}) = \begin{cases} 1 & \text{if } b_L < b_R \\ 0 & \text{if } b_L > b_R \end{cases}$ . Further set arbitrary values for  $\Phi(x, b, b)$ ,  $b \in (0, 1)$ . We have established that  $\sigma^P(\mathbf{b}) = 1$  for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$ , is part of a robust, retrospective, and intuitive equilibrium when (10) is true.

Lastly, we show that party strategies  $\sigma^P(\mathbf{b}) = 1$  for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$  cannot be part of an equilibrium when  $\delta > \frac{u_e^P - u_m^P}{\pi_e(a_m^L - a_e^L) + G + u_e^P - a_e^P}$ . Since  $(a_m^L - a_e^L) > 0$ ,  $\frac{\partial \beta_R(\mathbf{b}, R, x_e^L)}{\partial b_R} > 0$ , and  $\beta_R(\mathbf{b}, R, x_e^L) = \pi_e$  when  $b_R = 1$ , we infer (by continuity) that if  $\delta > \frac{u_e^P - u_m^P}{\pi_e(a_m^L - a_e^L) + G + u_e^P - a_e^P}$  there exists  $\hat{b}_R \in (0, 1)$  such that  $\delta > \frac{u_e^P - u_m^P}{\beta_R(\mathbf{b}, R, x_e^L)(a_m^L - a_e^L) + G + u_e^P - a_e^P}$  for all  $b_R > \hat{b}_R$ . Without loss of generality assume party  $L$  is in government. In every equilibrium with party strategies  $\sigma^P(\mathbf{b}) = 1$  for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$  we must have  $\Phi(x, \mathbf{b}) = \begin{cases} 1 & \text{if } b_L < b_R \\ 0 & \text{if } b_L > b_R \end{cases}$ . Given posterior beliefs  $\beta_L(\mathbf{b}, R, x_m^L) = 0 < \beta_R(\mathbf{b}, R, x_m^L)$  and  $\beta_L(\mathbf{b}, R, x_e^L) = 1 > \beta_R(\mathbf{b}, R, x_e^L)$ , we calculate expected utility when extremists of party  $L$  follow the prescribed strategy as

$$U^L(\mathbf{b}, x_e^L) = u_e^L + G + \delta (\beta_R(\mathbf{b}, R, x_e^L) (a_e^L - a_m^L) + a_m^L). \quad (21)$$

On the other hand, a one-period deviation by extremists of party  $L$  to a moderate policy accrues

$$U^L(\mathbf{b}, x_m^L) = u_m^L + G + \delta (u_e^L + G) \quad (22)$$

where we have obtained (21) and (22) by substituting for  $\Phi(x, \cdot, \cdot)$  in (17). Now, comparing the two expected utilities we obtain

$$U^L(\mathbf{b}, x_e^L) \geq U^L(\mathbf{b}, x_m^L) \iff \delta \leq \frac{u_e^L - u_m^L}{\beta_R(\mathbf{b}, R, x_e^L) (a_m^L - a_e^L) + G + u_e^L - a_m^L}.$$

But the latter inequality is false for  $b_R > \hat{b}_R$ . As a consequence,  $\sigma^P(\mathbf{b}) = 1$ , for all  $\mathbf{b} \in \mathcal{B}$  cannot be a robust retrospective equilibrium when (10) is violated. ■

We continue with the proof of proposition 2.

**Proposition 2** *There does not exist a robust equilibrium such that  $\sigma^P(\mathbf{b}) = 0$ , for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$ .*

**Proof.** If  $\sigma^P(\mathbf{b}) = 0$ , for all  $\mathbf{b} \in \mathcal{B}$ ,  $P \in \{L, R\}$ , every robust equilibrium must be intuitive, *i.e.* the voting strategy  $\Phi(x, \mathbf{b})$  must satisfy condition (9). This is because with perturbed strategies

$\sigma_\varepsilon^P(\mathbf{b}) = \varepsilon$ , voter's expected utility is such that

$$\begin{aligned} V_\varepsilon(\mathbf{b}, L) &< V_\varepsilon(\mathbf{b}, R) \iff \\ b_L \varepsilon (v_e^L - v_m^L) + v_m^L &< b_R \varepsilon (v_e^R - v_m^R) + v_m^R \iff \\ b_L &> b_R, \end{aligned}$$

for every  $\varepsilon > 0$ .

Recall that posterior beliefs in (4) following moderate policies satisfy  $\beta_P(\mathbf{b}, -P, x_m^P) = b_P$  when  $\sigma^P(\mathbf{b}) = 0$ . Note that there exist  $\mathbf{b} \in \mathcal{B}$  such that  $\beta_L(\mathbf{b}, R, x_m^L) = b_L > \beta_R(\mathbf{b}, R, x_m^L) \iff b_L > (\pi_e + \pi_m - 1)b_R + 1 - \pi_m$ . Thus, for such  $\mathbf{b} \in \mathcal{B}$  with party  $L$  in power, we have from (9) that  $\Phi(x_m^L, \beta(\mathbf{b}, R, x_m^L)) = 0$  in a robust equilibrium. Furthermore, since  $\beta_L(\mathbf{b}, R, x_e^L) = 1 > \beta_R(\mathbf{b}, R, x_e^L)$  for all  $\mathbf{b} \in \mathcal{B}$ , we must also have  $\Phi(x_e^L, \beta(\mathbf{b}, R, x_e^L)) = 0$  in equilibrium. Thus, substituting in the expected utility expression (17) we get

$$U^L(\mathbf{b}, x_m^L) = u_m^L + G + \delta a_m^L < U^L(\mathbf{b}, x_e^L) = u_e^L + G + \delta a_m^L.$$

Thus, there exist  $\mathbf{b} \in \mathcal{B}$  such that  $\sigma^L(\mathbf{b}) = 0$  is not optimal for party  $L$  and equilibrium condition (19) is violated. ■

Lastly we prove proposition 3.

**Proposition 3** *Assume*

$$\delta > \frac{u_e^P - u_m^P}{G + u_m^P - a_m^P}. \quad (16)$$

*There exists a unique intuitive retrospective equilibrium in which the probability of an extreme policy choice by extreme partisan types is given by*

$$\sigma^P(\mathbf{b}) = \begin{cases} \frac{b_P - \pi_e b_{-P} - (1 - \pi_m)(1 - b_{-P})}{b_P(1 - \pi_e b_{-P} - (1 - \pi_m)(1 - b_{-P}))} & \text{if } b_P > \pi_e b_{-P} + (1 - \pi_m)(1 - b_{-P}), \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

*This equilibrium is robust.*

**Proof.** We shall prove the proposition using a few lemmas. We first show that in a robust equilibrium with the party strategies in (12), the voting strategy  $\Phi$  must be intuitive.

**Lemma 1** *In a robust equilibrium with party strategies given by (12), the voting strategy satisfies condition (9).*

**Proof.** The strategies in (12) satisfy:

$$\frac{\partial \sigma^P(\mathbf{b})}{\partial b_P} = \begin{cases} \frac{\pi_e b_{-P} + (1 - \pi_m)(1 - b_{-P})}{b_P^2 (1 - \pi_e b_{-P} - (1 - \pi_m)(1 - b_{-P}))} > 0 & \text{if } b_P > \pi_e b_{-P} + (1 - \pi_m)(1 - b_{-P}) \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\frac{\partial \sigma^P(\mathbf{b})}{\partial b_{-P}} = \begin{cases} \frac{-(1 - b_P)(\pi_e + \pi_m - 1)}{b_P (1 - \pi_e b_{-P} - (1 - \pi_m)(1 - b_{-P}))^2} < 0 & \text{if } b_P > \pi_e b_{-P} + (1 - \pi_m)(1 - b_{-P}) \\ 0 & \text{otherwise.} \end{cases}$$

Hence, since  $\sigma^R(b, b) = \sigma^L(b, b)$  for all  $b \in [0, 1]$ , we have  $b_L < b_R \implies \sigma^R(\mathbf{b}) \geq \sigma^L(\mathbf{b})$ , and an intuitive voting strategy is part of a best response. To see that it is the unique best response in a robust equilibrium, note that for the perturbed (according to (20)) strategies,  $\sigma_\varepsilon^P(\mathbf{b})$ , we have  $b_L < b_R \implies b_L \sigma_\varepsilon^L(\mathbf{b}) < b_R \sigma_\varepsilon^R(\mathbf{b})$  for all  $\varepsilon > 0$  ( $\varepsilon < \frac{1}{2}$ ). But

$$b_L \sigma_\varepsilon^L(\mathbf{b}) < b_R \sigma_\varepsilon^R(\mathbf{b}) \iff V_\varepsilon(\mathbf{b}, L) > V_\varepsilon(\mathbf{b}, R)$$

where  $V_\varepsilon(\mathbf{b}, L)$ ,  $V_\varepsilon(\mathbf{b}, R)$  are voter's expected utilities from electing the left and right parties respectively given party strategies  $\sigma_\varepsilon^P(\mathbf{b})$ . Thus, concluding the proof of the lemma, we deduce that the only robust equilibrium must involve intuitive voting strategies that satisfy condition (9).   
 ■

Next, we show using two lemmas that given intuitive retrospective voting strategy satisfying (9), the unique equilibrium party strategies are given by (12).

**Lemma 2** *Assume (11) and a voting strategy that satisfies (9). Then  $\sigma^P(\mathbf{b}) = 0$ ,  $P \in \{L, R\}$ , for all  $\mathbf{b} \in \mathcal{B}$  with  $b_P < \pi_e b_{-P} + (1 - \pi_m)(1 - b_{-P})$ .*

**Proof.** Without loss of generality we consider the strategy of party  $L$ . Since  $\beta_L(\mathbf{b}, R, x_e^L) = 1 > \beta_R(\mathbf{b}, R, x_e^L)$ , it must be that  $\Phi(x_e^L, \beta(\mathbf{b}, R, x_e^L)) = 0$  when  $\Phi$  satisfies (9), hence we have

$$U^L(\mathbf{b}, x_e^L) = u_e^L + G + \delta (\beta_R(\mathbf{b}, R, x_e^L) \sigma^R(\beta(\mathbf{b}, R, x_e^L)) (a_e^L - a_m^L) + a_m^L).$$

We also have  $\beta_L(\mathbf{b}, R, x_m^L) \leq b_L < \beta_R(\mathbf{b}, R, x_m^L)$ , for all possible  $\sigma^L(\mathbf{b}) \in [0, 1]$ . Thus,  $\Phi(x_m^L, \beta(\mathbf{b}, R, x_m^L)) = 1$  from (9). So, the expected utility from pursuing a moderate policy is

$$U^L(\mathbf{b}, x_m^L) = u_m^L + G + \delta (G + \sigma^L(\beta(\mathbf{b}, R, x_m^L)) (u_e^L - u_m^L) + u_m^L).$$

We have

$$U^L(\mathbf{b}, x_m^L) > U^L(\mathbf{b}, x_e^L) \iff$$

$$u_m^L - u_e^L + \delta(G + u_m^L - a_m^L) > -\delta \left( \begin{array}{c} \sigma^L(\beta(\mathbf{b}, R, x_m^L))(u_e^L - u_m^L) \\ + \beta_R(\mathbf{b}, R, x_e^L) \sigma^R(\beta(\mathbf{b}, R, x_e^L))(a_m^L - a_e^L) \end{array} \right)$$

The right hand side of the above is less than or equal to zero, while the left hand side satisfies

$$u_m^L - u_e^L + \delta(G + u_m^L - a_m^L) > 0 \iff \delta > \frac{u_e^L - u_m^L}{G + u_m^L - a_m^L},$$

which is true by (11). We conclude that  $U^L(\mathbf{b}, x_m^L) > U^L(\mathbf{b}, x_e^L)$  is true and  $\sigma^L(\mathbf{b}) = 0$  is the unique optimal strategy when (9) holds and  $b_L < \pi_e b_R + (1 - \pi_m)(1 - b_R)$ . ■

The last lemma is:

**Lemma 3** Assume (11) and a voting strategy that satisfies (9). Then  $\sigma^P(\mathbf{b}) = \frac{b_P - \pi_e b_{-P} - (1 - \pi_m)(1 - b_{-P})}{b_P(1 - \pi_e b_{-P} - (1 - \pi_m)(1 - b_{-P}))}$ ,  $P \in \{L, R\}$ , for all  $\mathbf{b} \in \mathcal{B}$  with  $b_P \geq \pi_e b_{-P} + (1 - \pi_m)(1 - b_{-P})$ .

**Proof.** Without loss of generality we consider the strategy of party  $L$ . Following a policy  $x_e^L$  by party  $L$  we have  $\beta_L(\mathbf{b}, R, x_e^L) = 1$ , so that  $\beta_R(\mathbf{b}, R, x_e^L) < \pi_e \beta_L(\mathbf{b}, R, x_e^L) + (1 - \pi_m) \beta_L(\mathbf{b}, R, x_e^L) \implies \sigma^R(\beta(\mathbf{b}, R, x_e^L)) = 0$ , by lemma 2. Thus,  $U^L(\mathbf{b}, x_e^L) = u_e^L + G + \delta a_m^L$ , because  $\beta_L(\mathbf{b}, R, x_e^L) = 1 > \beta_R(\mathbf{b}, R, x_e^L)$  implies  $\Phi(x_e^L, \beta(\mathbf{b}, R, x_e^L)) = 0$  when  $\Phi$  satisfies (9). Let  $\sigma^L(\mathbf{b}) = \sigma$ ; the posterior belief following a moderate policy choice,  $x_m^L$ , is given from (4) as  $\beta_L(\mathbf{b}, R, x_m^L) = \frac{(1 - \sigma)b_L}{1 - \sigma b_L}$ , and satisfies  $\frac{\partial \beta_L(\mathbf{b}, R, x_m^L)}{\partial \sigma} < 0$ . Hence, due to (9), we get

$$\begin{aligned} \sigma < \frac{b_L - \pi_e b_R - (1 - \pi_m)(1 - b_R)}{b_L(1 - \pi_e b_R - (1 - \pi_m)(1 - b_R))} &\implies \beta_L(\mathbf{b}, R, x_m^L) > \beta_R(\mathbf{b}, R, x_m^L) \ \& \ \Phi(x_m^L, \beta(\mathbf{b}, R, x_m^L)) = 0 \\ \sigma = \frac{b_L - \pi_e b_R - (1 - \pi_m)(1 - b_R)}{b_L(1 - \pi_e b_R - (1 - \pi_m)(1 - b_R))} &\implies \beta_L(\mathbf{b}, R, x_m^L) = \beta_R(\mathbf{b}, R, x_m^L) \\ \sigma > \frac{b_L - \pi_e b_R - (1 - \pi_m)(1 - b_R)}{b_L(1 - \pi_e b_R - (1 - \pi_m)(1 - b_R))} &\implies \beta_L(\mathbf{b}, R, x_m^L) < \beta_R(\mathbf{b}, R, x_m^L) \ \& \ \Phi(x_m^L, \beta(\mathbf{b}, R, x_m^L)) = 1. \end{aligned}$$

Thus, by choosing a moderate policy, extreme types of party  $L$  expect:

$$U^L(\mathbf{b}, x_m^L) = u_m^L + G$$

$$+ \delta \left( \begin{array}{c} \Phi(x_m^L, \beta(\mathbf{b}, R, x_m^L)) (\sigma^L(\beta(\mathbf{b}, R, x_m^L))(u_e^L - u_m^L) + u_m^L + G) \\ + (1 - \Phi(x_m^L, \beta(\mathbf{b}, R, x_m^L))) (\beta_R(\mathbf{b}, R, x_m^L) \sigma^R(\beta(\mathbf{b}, R, x_m^L))(a_e^L - a_m^L) + a_m^L) \end{array} \right).$$

Now, we verify with straightforward algebraic manipulation from the above that if  $\sigma < \frac{b_L - \pi_e b_R - (1 - \pi_m)(1 - b_R)}{b_L(1 - \pi_e b_R - (1 - \pi_m)(1 - b_R))}$  we have

$$U^L(\mathbf{b}, x_e^L) > U^L(\mathbf{b}, x_m^L) \iff u_e^L > u_m^L$$

which is true, so that it cannot be that  $\sigma^L(\mathbf{b}) < \frac{b_L - \pi_e b_R - (1 - \pi_m)(1 - b_R)}{b_L(1 - \pi_e b_R - (1 - \pi_m)(1 - b_R))}$ . Similarly, if  $\sigma > \frac{b_L - \pi_e b_R - (1 - \pi_m)(1 - b_R)}{b_L(1 - \pi_e b_R - (1 - \pi_m)(1 - b_R))}$ , we get

$$U^L(\mathbf{b}, x_e^L) < U^L(\mathbf{b}, x_m^L) \iff u_e^L - u_m^L - \delta(G + u_m^L - a_m^L) < \delta(\sigma^L(\beta(\mathbf{b}, R, x_m^L))(u_e^L - u_m^L)),$$

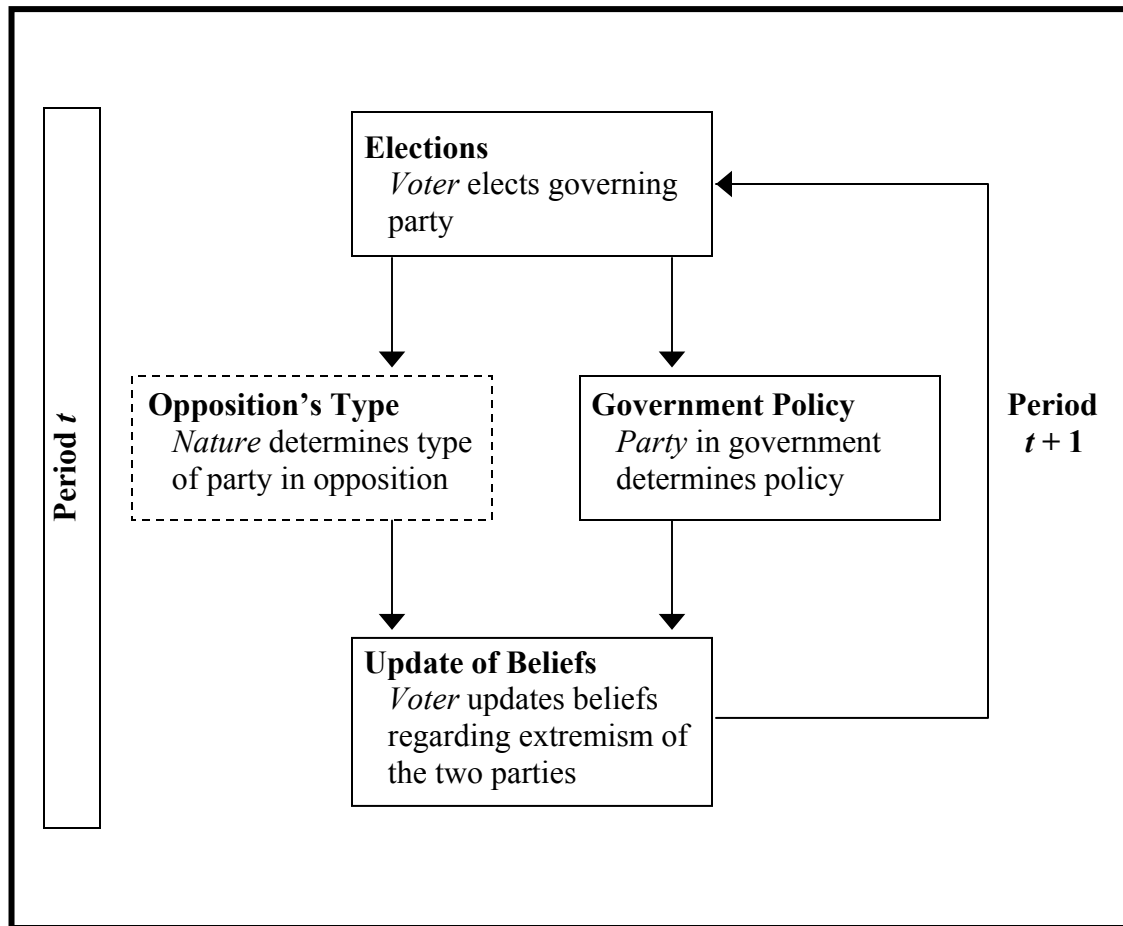
which is true because (11) implies  $u_e^L - u_m^L - \delta(G + u_m^L - a_m^L) < 0$ . Thus, we cannot have  $\sigma > \frac{b_L - \pi_e b_R - (1 - \pi_m)(1 - b_R)}{b_L(1 - \pi_e b_R - (1 - \pi_m)(1 - b_R))}$  either. As a consequence, equilibrium can only be attained when  $\sigma = \sigma^L(\mathbf{b}) = \frac{b_L - \pi_e b_R - (1 - \pi_m)(1 - b_R)}{b_L(1 - \pi_e b_R - (1 - \pi_m)(1 - b_R))}$ . Furthermore (because the inequalities above are strict) there exists unique  $\Phi(x_m^L, \beta_L(\mathbf{b}, R, x_m^L)) \in (0, 1)$  such that  $U^L(\mathbf{b}, x_e^L) = U^L(\mathbf{b}, x_m^L)$ .<sup>23</sup> ■

To summarize, in lemmas 2 and 3 we have shown that the only equilibrium with intuitive voting involves party strategies given by (12). Furthermore, by lemma 1, when party strategies are given by (12) the only robust voting strategy is intuitive. Thus, there exists an essentially (except for values of  $\Phi(x_m^L, b, b)$  when  $b \notin [1 - \pi_m, \pi_e]$ ) unique intuitive retrospective equilibrium that is also robust. ■

---

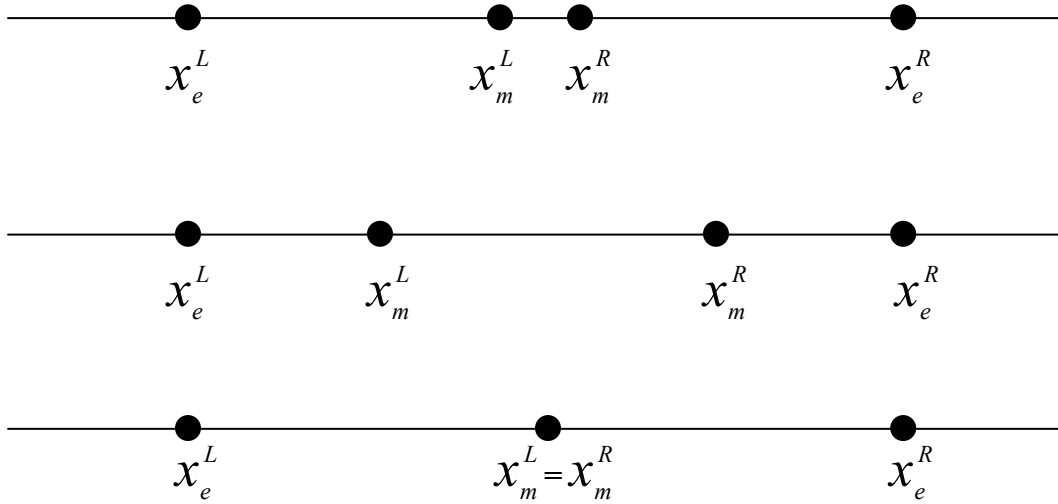
<sup>23</sup>The exact values of  $\Phi(x_m^L, \beta_L(\mathbf{b}, R, x_m^L))$  are available upon request.

**Figure 1:** Timeline of the Game



**Key:** A period in the game is a complete political cycle that includes elections, the determination of government's policy and the opposition party's type, and the formation of voter's updated beliefs regarding the extremism of the two parties. While the electorate observes the government's policy, it does not directly observe the new type of the opposition party.

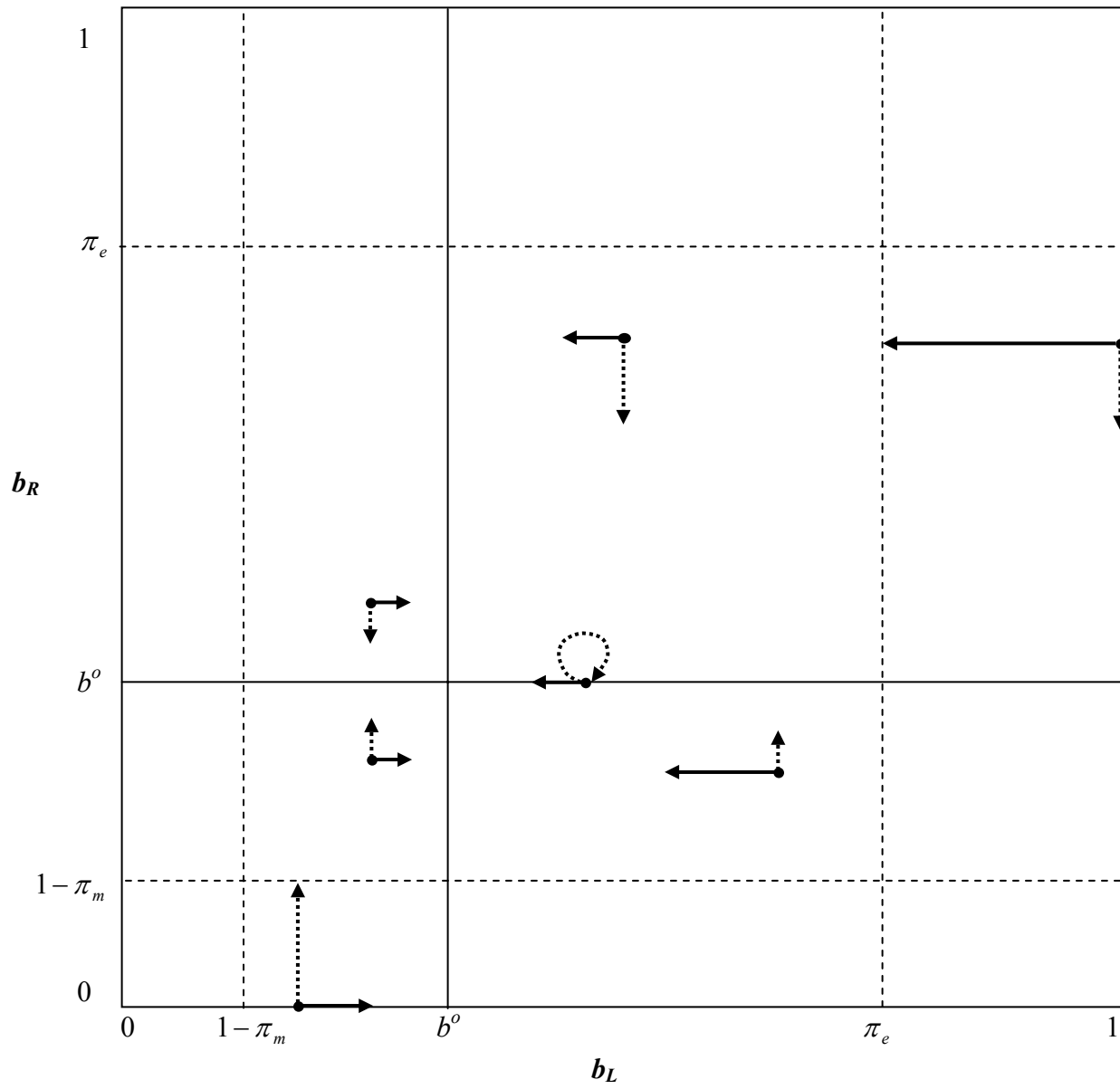
**Figure 2:** Rendition of Government Policies in the Spatial Model



**Key:** *The ideal policy of moderate partisans may be identical (the median) in one dimension, or may reflect partisan bias even when the party is controlled by moderates.*

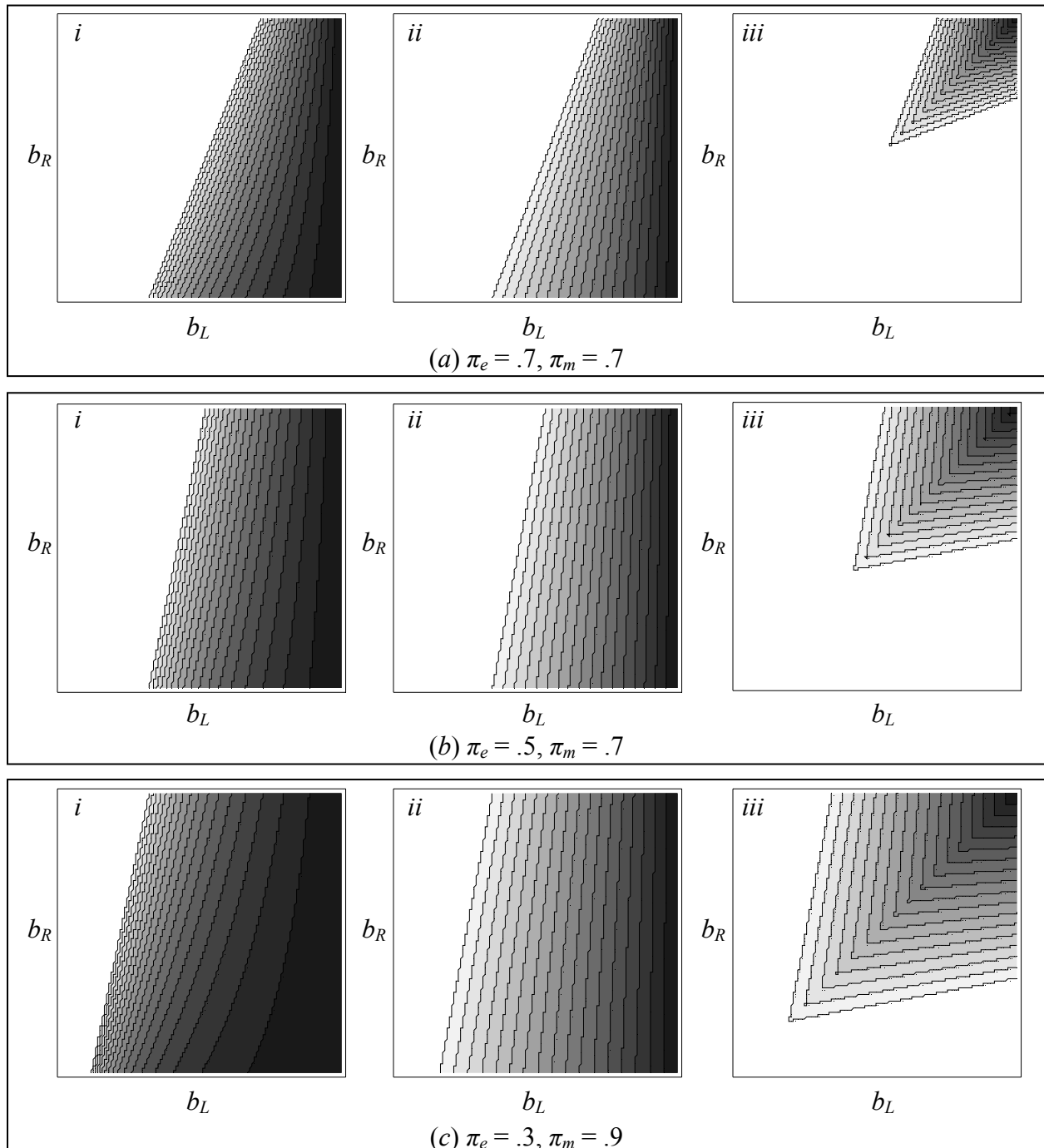


**Figure 3:** Evolution of Beliefs about Party in Opposition



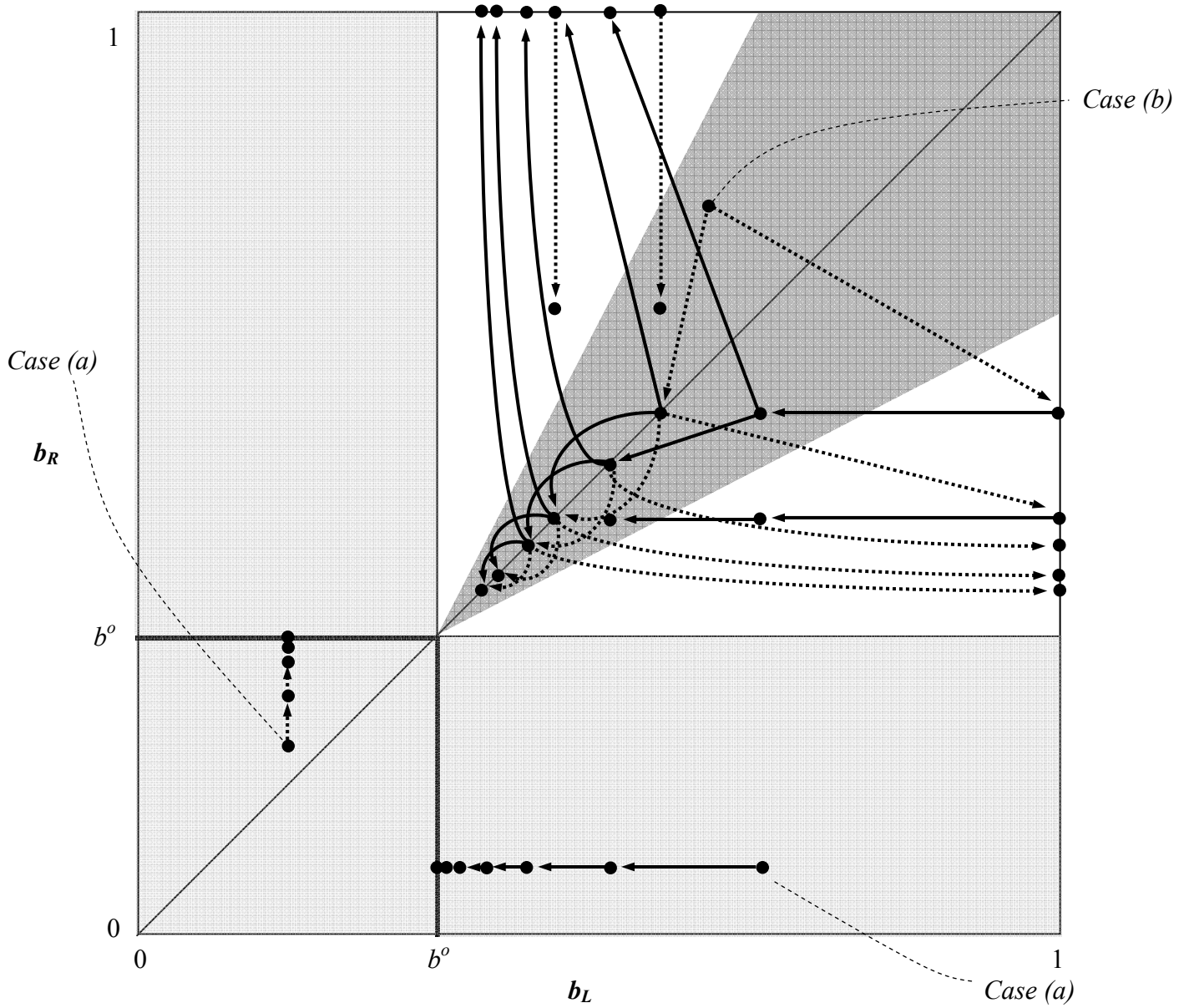
**Key:**  $\longrightarrow$  change in beliefs following defeat of the Left  
 $\cdots\cdots\longrightarrow$  change in beliefs following defeat of the Right

**Figure 4: Probability of Extreme Policies & Electorates' Beliefs**



**Key:** Contour plots of probability of an extreme policy in the space of beliefs,  $\mathcal{B} = [0,1]^2$ . For different values of parameters  $\pi_e, \pi_m$ : figure (i) depicts the probability that an extreme type of party L would implement an extreme policy, if party L is elected; figure (ii) depicts the expected probability that party L implements an extreme policy, if elected; figure (iii) indicates the equilibrium expected probability of an extreme policy choice (by a government of either party) following an election with the corresponding beliefs about the extremism of the two parties. Darker areas indicate higher probability. Probability is zero in blanc areas.

**Figure 5: Equilibrium Dynamics**



- Key:**
- Change in beliefs following a government of the Right,
  - ⋯→ Change in beliefs following a government of the Left,
  - Beliefs at election,
  - ▨ Case (a) of Proposition 5,
  - ▩ Case (b) of Proposition 5,
  - Set  $\tilde{B}$  defined in (13).

**Table 2:** Simulated Equilibrium Dynamics

<b>Government:</b>	<i>Party L</i>		<i>Party R</i>	
<b>Policy:</b>	$x_e^L$	$x_m^L$	$x_m^R$	$x_e^R$
Period 1	0.000	1.000	0.000	0.000
Period 2	0.182	0.818	0.000	0.000
Period 3	0.050	0.169	0.645	0.136
Period 4	0.024	0.276	0.632	0.068
Period 5	0.014	0.331	0.578	0.077
Period 6	0.014	0.427	0.517	0.042
Period 7	0.027	0.442	0.499	0.033
Period 8	0.026	0.663	0.297	0.014
Period 9	0.014	0.357	0.602	0.026
Period 10	0.007	0.336	0.637	0.021
Period 11	0.011	0.533	0.443	0.014
Period 12	0.006	0.330	0.650	0.014
Period 13	0.008	0.570	0.414	0.008
Period 14	0.004	0.368	0.619	0.009
Period 15	0.004	0.391	0.598	0.007
Period 16	0.003	0.389	0.602	0.006
Period 17	0.003	0.394	0.598	0.005
Period 18	0.003	0.602	0.392	0.003
Period 19	0.002	0.357	0.638	0.003
Period 20	0.002	0.393	0.603	0.002

**Key:** Table reports computed equilibrium probabilities of observing the outcome of the respective column in the period indicated by the corresponding row. Shaded cells indicate the party that wins the election with probability higher than 50% in the corresponding period. Calculations are based on the following parameter values:  $\pi_e = .8$ ,  $\pi_m = .95$ ,  $u_e^P = 1$ ,  $u_m^P = a_m^P = 0$ ,  $a_e^P = -1$ ,  $G = 4$ ,  $\delta = 1$ . Initial beliefs are  $b_L = .7$ ,  $b_R = .95$ .

**Table 1:** Electoral Outcomes in Majoritarian Parliamentary Systems

UK			New Zealand			Greece		
Election Date	Labour	Conservatives	Election Date	Labour	National Party	Election Date	ΠΑΣΟΚ	Νέα Δημοκρατία
07/05/45	47.7% (61.4%)	36.8% (31.1%)	11/27/46	51.3% (52.5%)	48.4% (47.5%)	11/17/74*	13.6% (4.0%)	54.4% (73.3%)
02/23/50	46.1% (50.4%)	43.4% (47.7%)	11/30/49	47.2% (42.5%)	51.9% (57.5%)	11/20/77	25.3% (31.0%)	41.8% (57.0%)
10/25/51	48.8% (47.2%)	48.0% (51.4%)	09/01/51	45.8% (37.5%)	54.0% (62.5%)	10/18/81	48.1% (57.3%)	35.9% (38.3%)
05/26/55	46.4% (44.0%)	49.7% (54.8%)	11/13/54	44.1% (43.8%)	44.3% (56.3%)	06/02/85	45.8% (53.7%)	40.8% (42.0%)
10/08/59	43.8% (41.0%)	49.4% (57.9%)	11/30/57	48.3% (51.3%)	44.2% (48.8%)	06/18/89	39.1% (41.7%)	44.3% (48.3%)
10/15/64	44.1% (50.3%)	43.4% (48.3%)	11/26/60	43.4% (42.5%)	47.6% (57.5%)	11/05/89	40.8% (43.0%)	46.1% (49.3%)
03/31/66	48.0% (57.8%)	41.9% (40.2%)	11/30/63	43.7% (43.8%)	47.1% (56.3%)	04/08/90	38.7% (41.7%)	47.0% (50.3%)
06/18/70	43.1% (45.7%)	46.4% (52.4%)	11/26/66	41.4% (43.8%)	43.6% (55.0%)	10/10/93	46.9% (56.7%)	39.3% (37.0%)
02/28/74	37.2% (47.4%)	37.9% (46.8%)	11/29/69	44.2% (46.4%)	45.2% (53.6%)	09/22/96	41.5% (54.0%)	38.1% (36.0%)
10/10/74	39.3% (50.2%)	35.8% (43.6%)	11/25/72	48.4% (63.2%)	41.5% (36.8%)	04/09/00	43.8% (52.7%)	42.7% (41.7%)
05/03/79	36.9% (42.4%)	43.9% (53.4%)	11/29/75	39.6% (36.8%)	47.6% (63.2%)	03/07/04	40.5% (39.0%)	45.4% (55.0%)
06/09/83	27.6% (32.2%)	42.4% (61.1%)	11/25/78	40.4% (43.5%)	39.8% (55.4%)			
06/11/87	30.8% (35.2%)	42.3% (57.8%)	11/28/81	39.0% (46.7%)	38.8% (51.1%)			
04/09/92	34.4% (41.6%)	41.9% (51.6%)	07/14/84	43.0% (60.0%)	35.9% (37.9%)			
05/01/97	43.2% (63.6%)	30.7% (25.0%)	08/15/87	48.0% (58.8%)	44.0% (41.2%)			
06/07/01	41.3% (62.7%)	32.2% (25.2%)	10/27/90	35.0% (28.9%)	48.0% (70.1%)			
05/05/05	35.4% (55.1%)	32.5% (30.7%)	11/06/93	34.7% (45.5%)	35.0% (50.5%)			

**Key:** In the bigger part of the post WWII period, the UK, New Zealand, and Greece (since 1974), have had parliamentary systems in which one of two parties controlled a majority in parliament. New Zealand interrupted this mode of single-party majority governments by adopting a more proportional electoral system since the 1996 elections, and Greece operated under a significantly more proportional electoral system briefly for the three elections in 1989, and 1990. For each election date and party, percentages reflect vote share and (in parenthesis) seat share in parliament. Shaded cells indicate the party that won the majority in parliament in the corresponding election. The sequence of victors and electoral outcomes indicate prolonged periods with an advantage for one of the two contenders for control of the parliament (and government).

\* ΠΑΣΟΚ was the 3d party in this election.