Political Competition and the Dynamics of Parties and Candidates

Stefan Krasa[*]

April 11, 2018

Abstract

The paper introduces a dynamic model of electoral competition in which parties must select candidates from their current members. The chosen candidates, and the policies that they represent, determine the future party membership and thus the set of candidates that are available in future elections. This framework allows us to address a new set of questions for which existing models of political competition do not apply: (1) When does a party "overreach," by trying to exploit a current majority to implement policies that negatively impact the party's future electoral prospects? (2) How long can parties retain majorities? (3) Under what conditions does a realignment of parties take place?

[*]University of Illinois, Department of Economics, Urbana, IL 61801 USA; skrasa@illinois.edu
I would like to thank Dan Bernhardt, Peter Buisseret, Jon Eguia, Richard Van Weelden, Myrna Wooders and seminar audiences at Vanderbilt, University of Missouri, University of Warwick, the LSE, and the PECA conference at Northwestern University for many helpful comments.

1 Introduction

Democratic party competition frequently results in long periods of dominance by one party. For instance, the Democrats controlled the presidency in the US between 1933 and 1953; Tories were prime ministers of the UK between 1970 and 1997, followed by ten years of labor control. Similar, political operatives often interpret the outcomes of one election to imply persistence of observed trends. For example, in 2001 Karl Rove predicted that George W. Bush would usher in a permanent Republican majority,[1] and in 2014 Greg Walden, the Chairman of the National Republican Committee observed that Republicans are "as back to a majority as we have seen in our lifetimes. It may be a hundred-year majority."[2]

This persistence (or perceived persistence) of electoral outcomes is hard to reconcile with the predictions of standard models of electoral competition such as (Downs, 1957) and (Wittman, 1973). These models predict that parties should adjust their positions so that every election is close and each party has the same probability of winning. We explain the inertia of electoral results by focusing on the dynamic process that relates a party's previous positions, its membership, and the policies it adopts. Our dynamic theory explains how a majority party with a moderate leadership can stay in power for a long sequence of elections. Precisely because the dominant party has gathered majority support, the opposition party is reduced to its extreme core, and it necessarily selects extreme candidates, perpetuating its minority status.

This cycle of dominance ends when the majority party "overreaches": Fiorina (2016) defines "overreaching" as governing "in a manner that alienates the marginal member of its electoral majority," the majority party overreaches when it seizes the opportunity to adopt non-median policies it favors, such as the Iraq war for Blair, the ACA for Obama, or the 2017 "tax cut and job acts" for Trump. These policies are unpopular, and they cause the party to lose moderate members, who join the opposition. As opposition strengthens with an influx of moderate members and candidates, the majority partys base of moderate support shrinks, ultimately leading to an electoral defeat, ending the period of dominance.

A party's incentive to overreach poses the ultimate limit for a party to stay in power. According to Fiorina (2016), then Speaker Nancy Pelosi may have pushed for passages of the ACA even if she would have known how severely Democrats would be punished for it in the next election. Standard models of electoral competition do not capture the rational for "overreaching" and the effect of the choice of a particular candidate on the party's future electoral prospects. Moreover,

[1] c.f., http://edition.cnn.com/ALLPOLITICS/time/2001/04/30/busy.html

[2] c.f.., http://www.cnbc.com/2014/11/05/gops-walden-predicts-a-hundred-year-majority.html

these models also predict that candidates who lose an election, and are therefore never in a position to overreach, have no impact on their party's future. However, this is certainly not true in practice as Barry Goldwater's campaign for the presidency illustrates. In 1964, Goldwater lost against Johnson by a margin of over 20%. Nevertheless, candidates inspired by Goldwater's conservative positions gained seats in Congress, Republicans won eight governorships in 1966 (see Edwards (2014)), and paved the way for Reagans' presidency 16 years later.

To investigate the interplay of party dominance and overreaching, we introduce a dynamic model that restricts parties to selecting candidates from among their existing members. However, future membership changes in response to the policies proposed by candidates in the current period, just as Goldwater influenced a new generation of candidates to run for office, and Obama's ACA led to the rise of the Tea Party Republicans.[3]

As in Osborne and Slivinski (1996) and Besley and Coate (1997), candidates are unable to commit ex-ante to policies; if elected candidates select policies that maximize their own utility. As mentioned, we assume that each party must select a candidate from the set of its members. As in Snyder and Ting (2002), Osborne and Tourky (2008), or Buisseret and Weelden (2017) we can think of parties as providing information about the candidate types that otherwise would not be known to voters, and parties have this information about their members only. Alternatively, we could assume that individuals only want to be candidates for a party, if their preferences are compatible with those of other party members. This reflects the experience of many Southern Democrats who became Republicans when they felt that their views were no longer welcome in the Democratic Party (Strom, 1990).

In the model, we assume that candidates are chosen by party leaders, whose identity can depend on the party membership. For example, if the space of citizen types is one-dimensional, and the candidate is determined in a primary in which all party members participate, then the median party member would be the leader. If, instead, partisans participate at higher rates in primaries, then party leaders would be more extreme than the party medians.[4]

Our model describes a dynamic game in which the state of the game at each point in time is given by the allocation of citizens into parties. We first provide an existence result for Markov perfect equilibria and investigate equilibria when the type and policy spaces are one dimensional,

[3]The effect of party's policy choices on the set of future candidates was also pointed out by Steny Hoyer, the Democratic Minority Whip in the House, after the House vote on May 4, 2017 to repeal and replace the ACA (see Hulse (2015)): "We have too many [candidates] wanting to run. They are just coming out of the woodwork because they smell victory in the air and they are angry about what the Republicans are doing."

[4]The mapping that selects the party leader and thus the party preferences, resembles the preference aggregation in Baron (1993) or Caplin and Nalebuff (1997).

as in most standard models. We then consider the model with two time periods, and show that policies change more than median voters' positions if different parties win, but that policy may not respond to changes of the median voter's position if the same party wins twice. Intuitively, if a party wins only in one period, the party finds it advantageous to "overreach" and win with a more partisan candidate. As a consequence, the party is weakened in the next period, and the opposing party can win with a less centrist candidate. If the party wins in both periods, then it may be optimal to smooth policies intertemporally.[5] All equilibria feature policy divergence.

Next, we show that overreaching increases if a party's leadership becomes more extreme, for example in response to an increasing number of partisans participating in the party's primary. This is because a more extreme party places more emphasis on implementing partisan policies in the current period, rather than on winning in the next period.

We then investigate how long the dominance of one party can last, when the model's time horizon is infinite. In the model, a party with an electoral advantage in the current period could obtain a permanent majority, by choosing candidates that are sufficiently moderate. However, in equilibrium parties choose more extreme candidates which means that the permanent or "hundred-year" majorities envisioned by Karl Rove or Greg Walden cannot occur. In particular, we show that a party's advantage disappears after only a few (in many cases less than three) elections, at which point each party nominates a candidate whose type is identical to that of the median voter. Moreover, convergence occurs faster if parties are more extreme, because more extreme parties overreach more in each period, thereby weakening their future strength and accelerating convergence to the median. One can also interpret this result as a dynamic extension of the classic median voter theorem. However, our result is stronger as it not only predicts that policies converge, but also that no party can dominate in the long run.

If the policy space is one-dimensional, then the strategy of a competing minority party is to nominate the most moderate candidate possible in order to constrain the other party from nominating extremists. However, this need not be true if there are multiple policy dimensions. Similar to Goldwater influencing a new generation of conservatives to run for office for the Republican party, a party in our model may choose policies that lose in the short run, but attract different types of individuals to the party, whose positions may be successful in future campaigns.

[5]Note that in a standard model with policy motivated candidates policies track the position of the median voter, and it is irrelevant for the equilibrium wether different parties win.

2 Related Literature

Downs (1957), p. 144 proposed that parties face limitations about the candidates or policies they can select: a party cannot move ideologically "beyond the nearest party toward which it is moving." However, in a static model, this assumption does not work very well. For example, consider the policy interval [0, 1], and suppose that party 1 is currently located at 0.4 and party 2 at 0.5. Then party 1 could move to 0.5, but not beyond, so both parties having candidates at 0.5 would be an equilibrium. However, if party 2 moves to 0.4 then we would reach a different equilibrium. Similarly, both parties could simultaneously move, for example to 0.45, which would also be an equilibrium. In other words, if Downs' assumption is used as stated, we would get a continuum of equilibria. Our dynamic model resolves this difficulty but retains Downs' original insight that a party's set of candidates is limited.

An important aspect of our model is the link between current candidate policies and future party membership. Poutvaara (2003) provides a dynamic model in which such a link exists, but parties are myopic and policy choice is unrestricted as in standard models. As a consequence, parties are policy motivated, and size or party power is irrelevant. Gomberg et al. (2004) and Gomberg et al. (2016) consider static models, in which parties aggregate preferences and propose policies, and individuals sort themselves into parties as a function of these policies. In contrast to our model, policy choice is not driven by political competition.

In our model, parties select candidates who implement their most preferred policy if elected. We could also assume that parties choose platforms, but that these platforms are restricted to the Pareto set of all party members, as in Levy (2004).[6] Such an alternative approach would not affect the results in the one-dimensional case, because we show that parties are intervals and therefore any policy in a party's Pareto set can be implemented as the preferred policy of one party member. Similarly, our results for discrete type and policy spaces would remain essentially unchanged.

There is a large literature that uses dynamic games to investigate how policies are determined. Most prominently, the literature on legislative bargaining extends the sequential bargaining model of Baron and Ferejohn (1989) to settings where policy changes can be enacted repeatedly over time, with both a deterministic status quo policy as in Baron (1996) and Kalandrakis (2004), or with a stochastic status quo as in Duggan and Kalandrakis (2012). There are several differences between these models and ours. First, we have restrictions on the action space that change endogenously over time, because of our assumption that candidates must be selected from their respective parties.

[6] Similarly, in Morelli (2004), parties also impose restrictions on the policies that candidates can select.

Second, we always have two simultaneous rather than sequential proposers, who are the leaders of the two parties. Third, there is no status quo policy, because in every period a new policy can be enacted. Finally, the state is the party allocation, rather than the previous policy, and it depends deterministically on the actions of players in previous periods. Buisseret and Bernhardt (2017) consider a two-period model in which the policy proposer faces uncertainty about the political control in the future, but unlike our model current policies do not affect a candidates probability to become the policy proposer.

Gul and Pesendorfer (2016) consider a continuous time model with two competing parties that can only adjust policy gradually overt time. Their paper shares with ours the feature that parties are unable to immediately adopt median positions. In their main result they show that their model's steady steady state solution corresponds to that of Wittman (1973).

The paper is organized as follows. Section 3 describes the model. The definition and existence of equilibria is discussed in section 4. Section 5 analyzes the case where the policy space is the interval $[0, 1]$, proves existence of equilibria and provides comparative static results. Section 6 derives a dynamic median voter theorem and considers a situation with two policy dimensions. Section 8 concludes. Proofs are found in the text and in the Appendix.

3 Model

Time is indexed by $t = 0, \ldots, T$, where T can be either finite or infinite. Let X be the policy space at each time t. The preferences of a citizen of type $\theta \in \Theta$ are described by a utility function $u_\theta \colon X \to \mathbb{R}$. We assume that there is a unique policy $x(\theta) \in X$ that maximizes the citizen's utility, i.e., $x(\theta)$ solves $\max_x u_\theta(x)$. The population of citizens at time t is described by a probability distribution ϕ_t on Θ. Each citizen discounts future utility at a rate β, where $0 < \beta < 1$.

There are two parties indexed by $i = 1, 2$. The set of party members or equivalently the party's potential candidates at time t are given by $S_{1,t}, S_{2,t} \subset \Theta$, where $S_{1,t}, S_{2,t} \neq \emptyset$.

Each party has a leader at time t who selects the candidate for the election. As in Buisseret and Weelden (2017) we can think of the leader as the party elite, or alternatively as the median voter in a primary. The identity of the party leader could change over time as the party changes, but this additional degree of freedom in the model is not needed to derive any of the results. Formally, let $\mathscr{S}$ be the set of all subsets of Θ. Then the leader of party i is determined by a function $m_i \colon \mathscr{S} \to \Theta$. If $\Theta \subset \mathbb{R}$, then the party leader could, for example, be the median of S_i. We require that the party leader is always a member of the party, i.e., $m_i(S) \in S$ for all $S \in \mathscr{S}$.

At each time t, the leader of party i selects a candidate $\theta_{i,t} \in S_{i,t}$. All citizens vote for one of the candidates, and the candidate with the majority of votes wins. Winning candidates select policies that maximizes their utility, i.e., a type θ candidate implements policy $x(\theta)$.[7]

In the subsequent periods, party membership may change, depending on policies chosen in each period. Party membership in period $t+1$ is described by a function $\psi\colon \Theta^2 \times \mathscr{S}^2 \to \mathscr{S}^2$, mapping the candidate types θ^i, $i = 1, 2$ and party structures at time t into a party structure at time $t+1$.

Note that the future party composition depends on the candidates selected by the parties in the current period, but not on the actual outcome of the election. In particular, this allows for the possibility that losing candidates, such as Barry Goldwater in 1964, can influence the future of their party. As we explain in the following section it also means that voters do not have to be concerned about affecting a party's future composition through their vote. In contrast, party leaders must be forward looking, because function ψ depends on their actions.

4 Equilibrium: Definition and Existence

A subgame at time t is determined by the sets $S_{i,t}$, $t = 1, 2$ that represent party membership. Because party membership in period $t+1$ is determined by the candidates representing the parties at time t, and not by the identity of the winning candidate, the election outcome only determines current but not future payoffs. Thus the voting stage reduces to a one-period problem. Given a set of candidates $\theta^i_t \in S_{i,t}$ at t, and assuming that citizens use weakly dominant strategies, it follows that type θ votes for candidate i if $u_\theta(x(\theta_{i,t})) > u_\theta(x(\theta_{-i,t}))$.

Let $x(\theta_i)$ be the policy chosen by candidate θ_i, $i = 1, 2$. Then the probability that party i wins the election at time t is given by $\pi_{i,t}(\theta_1, \theta_2)$ where

$$\pi_{i,t}(\theta_1, \theta_2) = \begin{cases} 1 & \text{if } \phi_t\Big(\big\{\theta \,\big|\, u_\theta(x(\theta_i)) > u_\theta(x(\theta_{-i}))\big\}\Big) > 0.5; \\ 0 & \text{if } \phi_t\Big(\big\{\theta \,\big|\, u_\theta(x(\theta_i)) > u_\theta(x(\theta_{-i}))\big\}\Big) < 0.5; \\ q \in [0,1] & \text{otherwise;} \end{cases} \tag{1}$$

and $\pi_{1,t}(\theta_1, \theta_2) + \pi_{2,t}(\theta_1, \theta_2) = 1$

In contrast to voting by citizens, the selection of candidates by party leaders impacts payoffs in future time periods, because it influences future party composition. We consider Markov strategies.

[7]In particular, this means that candidates are not strategic players. Only party leaders take the impact of the current policy choice on the party's future into account. Thus, as mentioned above, it is equivalent to assume that party leaders can choose a policy that is in the Pareto set of the party's supporters as in Levy (2004).

The state at time t is given by the party composition (S_1, S_2), and Markov strategies for the game would have to be defined for all divisions of the type space into parties, although it may be impossible to attain many of them. To simplify the analysis we can therefore restrict attention to party allocations that are feasible.

Definition 1 *A set of party allocations $\mathscr{K} \subset \mathscr{S} \times \mathscr{S}$ is feasible if and only if $\psi(\theta_1, \theta_2, S_1, S_2) \in \mathscr{K}$ for all $(S_1, S_2) \in \mathscr{S} \times \mathscr{S}$ and $\theta_i \in S_i$, $i = 1, 2$.*

Let $\mathscr{K}$ be a feasible set of party allocations, and $\mathscr{P}(\Theta)$ the set of all probabilities on Θ. Then a (mixed) Markov strategy for party i is given by a transition probability $\mu_i \colon \mathscr{P}(\Theta) \times \mathscr{K} \to [0, 1]$. That is, if $K = (S_1, S_2) \in \mathscr{K}$, then $\mu_i(\cdot, S_1, S_2)$ is a probability distribution over candidate i's type $\theta_i \in \Theta$. Because candidates are selected from parties, $\theta_i \in S_i$, the support of $\mu_i(\cdot, S_1, S_2)$ must be contained in S_i, i.e., $\mu_i(S_i, S_1, S_2) = 1$.

The payoff of a party leader of type $\theta \in \Theta$ from strategies $\mu_{i,t} = 1, 2$ at time t, given party allocation $K = (S_1, S_2) \in \mathscr{K}$ is

$$\begin{aligned} U_t(K, \mu_{1,t}, \mu_{2,t}, \theta) = \int \Big(& \sum_{i=1}^{2} \pi_{i,t}(\theta_1, \theta_2) v_\theta(x(\theta_i)) \\ & + \beta U_{t+1}(\psi(\theta_1, \theta_2, K), \mu_{1,t+1}, \mu_{2,t+1}, \theta) \Big) d\mu_{1,t}(d\theta_1, K) d\mu_{2,t}(d\theta_2, K) \end{aligned} \tag{2}$$

In a Markov perfect equilibrium, the candidate choice by the leader $m_i(S_i)$ of party i at time t, must be optimal given the candidate chosen by the leader $m_j(S_j)$ of the rival party j.

Definition 2 *$(\mu_{1,t}, \mu_{2,t})$, $\pi_{i,t}$, $t = 1, \ldots, T$ is a Markov perfect equilibrium if and only if*

1. *$\pi_{i,t}$ satisfies (1) for all $t = 1, \ldots, T$;*
2. *For all $t = 1, \ldots, T$, and for all $K \in \mathscr{K}$ and all probabilities ν with support on S_i, where $K = (S_1, S_2)$,*
$$U_t(K, \mu_{i,t}, \mu_{-i,t}, \theta_i) \geq U_t(K, \mu_{i,t}, \nu, \theta_i), \tag{3}$$
for party leader $\theta_i = m_i(S_i)$ and parties $i = 1, 2$.

If the type space Θ is finite and there are finitely many time periods T, then existence of Markov perfect equilibria is standard (c.f., Theorem 1.3.1 in Fudenberg and Tirole (1991)). The existence result can be generalized to the case with infinitely many time periods.

Proposition 1 *Let Θ be finite and time T finite or infinite. Then there exists a Markov perfect equilibrium in mixed strategies.*

In the final sections of this paper, we consider a model with finitely many types, where Proposition 1 is applicable. However, we also want to investigate party dynamics when the type space and policy space are the interval $[0, 1]$ to compare our result to classic settings.

5 The Model with Two Time Periods

5.1 Existence of Pure Strategy Equilibria

Proposition 1 does not apply if the type space is continuous. In fact, it is well known that subgame perfect equilibria may not exist even in a two-stage game, when actions spaces are no longer finite (Harris et al., 1995). One way to derive general existence results it to introduce noise in the process that affects the future state (the allocation of parties) in the next period. More specifically, to get existence with a finite time horizon, the transition rule for states must be setwise continuous (Rieder, 1979). Alternatively, if the time horizon is infinite, norm-continuity is required (see Theorem 2 of Jaśkiewics and Nowak (2016)). Neither property holds if the transition function is deterministic as in our case, but the properties can be satisfied if sufficient noise is added (Duggan, 2012). For example, we could assume that after ψ selects an interim party allocation $K \in \mathscr{K}$, a random shock, modeled as a transition probability $q_t\colon \mathscr{P}(\mathscr{K}) \times \mathscr{K} \to [0, 1]$, determines the final party allocation according to the probability distribution $q_t(\cdot, K)$.

We choose to work with a deterministic setting instead, because this allows us to get a sharper results on convergence of policies. In exchange, we have to make more specific assumption on preferences and on the mapping, ψ, that determines party affiliations. In particular, we assume that each type is affiliated in the next period with the party whose policy they prefer in the current period. That is,

$$u_\theta(x(\theta_{i,t})) > u_\theta(x(\theta_{j,t})) \text{ implies } \theta \in S_{i,t+1} \text{ and } \theta \notin S_{j,t+1}, \tag{4}$$

where $S_{i,t+1} = \psi_i(\theta_1, \theta_2, S_{1,t}, S_{2,t})$, $i = 1, 2$.

If, in addition, a single-crossing property holds, then we can show that parties can be represented as intervals.

Lemma 1 *Let $X = \Theta = [0, 1]$ and suppose that ψ_1 and ψ_2 are compact valued and satisfy (4). Further, suppose utility satisfies the single-crossing property that $\frac{\partial u_\theta(x)}{\partial \theta}$ is strictly monotone (increasing or decreasing) in x. Let $\theta_1 \neq \theta_2$ be the previous period's policies. Then there exists $s \in [0, 1]$ such that $\psi_i(\theta_1, \theta_2, S_1, S_2) = [0, s]$ and $\psi_{-i}(\theta_1, \theta_2, S_1, S_2) = [s, 1]$.*

For example, suppose that agents have quadratic utility $u_\theta(x) = -(x-\theta)^2$, and that the previous period's candidates are $\theta_1 < \theta_2$. Then the next period party cutoff is $s = 0.5(\theta_1 + \theta_2)$, and parties will be $S_1 = [0, s]$ and $S_2 = [s, 1]$. [8]

Finally, we describe the functions m_i that determine the party leader who strategically selects a candidate. In view of Lemma 1 we can restrict attention to parties that are given by intervals of the form $[0, s]$ or $[s, 1]$. As a consequence, with a slight abuse of notation we can describe the party leader by functions m_1 and m_2 that only depend on s. That is, $m_1(s)$ is the party leader if the party is given by $[0, s]$, while $m_2(s)$ is the party leader if the party is given by $[s, 1]$.

In the remainder of this section we assume that the position of the party leader does not change as long as that party has an electoral advantage. This allows us to focus solely on the tradeoff between implementing more extreme policies and losing more moderate supporters. We discuss the extension to the case where the leadership changes in section 7.

Assumption 1 *Suppose that parties are given by $S_1 = [0, s]$ and $S_2 = [s, 1]$. Let $\theta_{m,t}$ be position of the median voter at time t. Then there exist $\bar{m}_i$, $i = 1, 2$ such that for all time periods t:*

1. *If $s \geq \theta_{m,t}$, then $m_1(s) = \bar{m}_1 < \theta_{m,t}$.*
2. *If $s \leq \theta_{m,t}$, then $m_2(s) = \bar{m}_2 > \theta_{m,t}$.*

Note that the assumption also indicates the leader of the majority party must be more extreme than the median voter.

Next, we start by analyzing the two-period setting, and prove existence of pure strategy subgame perfect equilibria. Even if there are multiple equilibria, payoffs are always the same.

Proposition 2 *Let $T = 2$. Suppose that condition* (4) *and assumption 1 hold. If preferences satisfy the single crossing property of Lemma 1 then there exists a Markov perfect equilibrium in pure strategies. All pure strategy subgame perfect equilibria are payoff equivalent.*

The existence result holds for any initial allocation of citizens into parties, and below we will analyze the interesting case where parties differ. However, note that if $S_1 = S_2 = [0, 1]$ then there

[8] More generally, we can allow for the possibility that there are independents, i.e., some types that are not affiliated with one of the two parties. For example, suppose that currently parties are given by $S_1 = [0, s_1]$ and $S_2 = [s_2, 1]$ with $s_1 < s_2$, and that policies are $\theta_1 < \theta_2$. Party membership is slower to adjust to changes, if $\psi_1(\theta_1, \theta_2, S_1, S_2) = [0, \alpha s_1 + (1-\alpha)0.5(\theta_1 + \theta_2)]$ and $\psi_2(\theta_1, \theta_2, S_1, S_2) = [\alpha s_2 + (1-\alpha)0.5(\theta_1 + \theta_2), 1]$, where $0 \leq \alpha \leq 1$. If $\alpha = 0$ then we have again the case without independents, where adjustments are immediate, while $\alpha = 1$ means that party composition never changes.

also exists a trivial equilibrium in which both parties select the median voter as their candidate in both periods, i.e., $\theta_{1,t} = \theta_{2,t} = \theta_{m,t}$. In this special case the equilibrium of our model corresponds to that of Wittman (1973). Note, however, that this equilibrium would not be robust with respect to adding uncertainty about the position of the median voter if party leaders' preferences differ. The uncertainty about the median voter would lead to some policy divergence, which means the parties would nominate different candidate types. According to Lemma 1 we would then get parties of the form $[0, s]$ and $[s, 1]$ in the second period.

5.2 Non Centrist Policies and Policy Divergence

We now use the result in Proposition 2 that equilibrium payoffs are unique, to analyze how equilibrium policies change in response to movements of the median voter's position. We show that if different parties win in the two periods, then equilibrium policies can move by more than the position of the median voter. In contrast, if the same party wins twice, policies may not respond at all to changes of the electorate. This is substantially different from the standard one-dimensional case without uncertainty, where policy positions shift exactly by the same amount as the median voter. The results are summarized in Proposition 3 below.

To solve for equilibria, we proceed by backward induction, following the argument in the proof of Proposition 2.

Consider a subgame starting at $t = 2$ and suppose that the party cutoff is s_2. If the position of the median voter, $\theta_{m,2}$, at $t = 2$ is to the left of the party cutoff s_2, then party 1 wins as long as its candidate, x_2, is at least as close to the median voter as party 2's candidate, i.e., $x_2 \geq 2\theta_{m,2} - s_2$. If $\bar{m}_1 > 2\theta_{m,2} - s_2$, then party 1 can select a candidate at the leader's ideal point $\bar{m}_1$.

Similarly, if $\theta_{m,2} > s_2$ then party 2 wins. In general, the winning policy at $t = 2$ is therefore

$$x_2 = h(s) = \begin{cases} \bar{m}_2 & \text{if } s < 2\theta_{m,2} - \bar{m}_2; \\ 2\theta_{m,2} - s & \text{if } 2\theta_{m,2} - \bar{m}_2 \leq s \leq 2\theta_{m_2} - \bar{m}_1; \\ \bar{m}_1 & \text{if } s > 2\theta_{m,2} - \bar{m}_1. \end{cases} \tag{5}$$

Now consider the first period of the game, and suppose that $\theta_{m,1} \leq s_1$ (the case where $\theta_{m,1} \geq s_1$ is symmetric). Let x_1 and y_1 be the positions of the candidates of party 1 and 2, respectively.

Note that $y_1 \geq s_1$, because party 2's candidate must be a member of party 2. Further, decreasing y_1 lowers the second-period party cutoff $s_2 = (x_1 + y_1)/2$. Because (5) shows that the winning policy in the second period is a non-increasing function of s, this implies that $y_1 = s_1$ is an optimal candidate choice for party 2.[9]

If party 2 chooses a candidate at $y_1 = s_1$, then the party cutoff in the second period is $s_2 = (x_1 + s_1)/2$. Party 1's candidate must be a member of party 1, and hence $x_1 \leq s_1$. Further, party 1 wins if x_1 is closer to the median voter, $\theta_{m,1}$ than party 2's candidate $y_1 = s_1$. In other words, it is never optimal for party 1 to choose a candidate, x_1, with $\theta_{m,1} - x_1 > s_1 - \theta_{m_1}$, because by raising x_1 party 1 could win in period 1. Further, (5) implies that raising x_1 (and hence s_1) cannot make the winning policy at $t = 2$ less favorable to party 1.

The optimal candidate, x_1, selected by party 1 must therefore solve the following optimization problem.

$$\max_{x_1 \in [0,s_1]} u_\theta(x_1) + \beta u_\theta\left(h\left(\frac{x_1 + s_1}{2}\right)\right) \text{ s.t. } x_1 \geq 2\theta_{m,1} - s_1, \tag{6}$$

where $\theta = \bar{m}_1$ is the position of the leader of party 1.

If the constraint of problem 6 binds, then it is immediate that policies move more than the median voter if party 2 wins in the second period. In particular, if the constraint binds then the winning policy in the first period is $x_1 = 2\theta_{m,1} - s_1$, resulting in a party cutoff $s_2 = \theta_{m,1}$ in the next period. If $\theta_{m,2} > \theta_{m,1}$ then party 2 wins in the next period, and the winning policy $x_2 = h(s_1) \geq \theta_{m,1}$. Thus, $|x_2 - x_1| > |\theta_{m,2} - \theta_{m,1}|$.

If $\theta_{m,2} < \theta_{m,1} < s_1$ then party 1 wins in both periods, and if the constraint of Problem 6 binds, then the winning policy in both periods is strictly to the left of the median voter (assuming that $m_1(1) < \theta_{m,2}$). In this scenario it is possible that the policy shifts by less than the median voter.

For example, suppose that $\theta_{m,1} = 0.5$, $\theta_{m,2} = 0.45$ and $s_1 = 0.6$. If the constraint of Problem 6 binds then $x_1 = 2\theta_{m,1} - s_1 = 0.4$. In the next period, the party cutoff is $s_1 = 0.5$. The winning policy at $t = 2$ is therefore $2\theta_{2,m} - s_2 = 0.4$ and hence the policy does not respond to changes in the electorate's preferences. Note that in both periods policies are strictly to the left of the median voters' ideal points.

We now provide more general conditions (including the case where the constraint of problem 6 is slack), under which the shift in policies exceeds the shift of the median voter if different parties win, but policies shift less if the same party wins both times.

Proposition 3 *Suppose that utility takes the form $u_\theta(x) = -|x - \theta|^\gamma$, $\gamma \geq 1$ and that $\bar{m}_1 < 2\theta_{m,1} - s_1$. Let $\theta_{m,1} < s_1$ and $\beta > 0$. Then:*

1. *Policies change more than the median voters, i.e., $|x_1 - x_2| > |\theta_{m,1} - \theta_{m,2}|$; party 1 wins in*

[9] If (5) is locally constant, then $y_1 > s_1$ can also be optimal. However, the payoffs to both parties remain the same. In other words, equilibria are not always unique, but equilibrium payoffs are unique, as shown in Proposition 2.

period 1; party 2 in period 2, and there is policy divergence in both periods if one of the following two conditions holds:

(a) $s_1 < (\theta_{m_1} + \theta_{m,2})/2$*;*

(b) $(\theta_{m_1} + \theta_{m,2})/2 \leq s_1 < \theta_{m,2}$ *and* $\beta < \min\left\{2^{2-\gamma}, 1\right\}$.

2. *Let* $\theta_{m,1} < \theta_{m,2} < s_1$*. Then party 1 wins in both periods. If* $\gamma \to \infty$ *then in the limit winning policies do not change in response to changes of the median voter's position, i.e.,* $|x_1 - x_2| \to 0$ *as* $\gamma \to \infty$.

The proposition is written for the case where $\theta_{m,1} < s_1$. The case $\theta_{m,1} > s_1$ is analogous. Further, the assumption that $\bar{m}_1 < 2\theta_{m,1} - s_1$ excludes the case where party 1's electoral advantage is so large that it can nominate a candidate at its ideal position $\bar{m}_1$.

Intuitively, why does the restriction that candidates are selected from parties increase the volatility of policies? Suppose that party 1 wins in the first period and party 2 wins in the second period. Party 1 could constrain party 2 at $t = 2$ by nominating a more moderate candidate at $t = 1$. The new party cutoff s_2 would then be closer to the median voter, which in turn means that party 1's candidate is more competitive, preventing party 2 from nominating a candidate who is too extreme. However, the new party cutoff is $s_2 = (s_1 + x_1)/2$. Hence nominating a candidate who is more moderate by "one unit" only results in the party cutoff moving to the right by 1/2 of a unit. For example, if utility is $u_\theta(x) = -|x - \theta|$, it follows that moderation at $t = 1$ is not worthwhile.

If the curvature of utility and if β are increased, parties may lower the volatility of policies, but when $\beta < 2^{2-\gamma}$ policies still move by more than the median voter. For example, if utility is $u_\theta(x) = -(x - \theta)^2$, then the result holds for any $\beta < 1$.

However, even when γ, and hence relative risk aversion go to infinity, volatility does not go to zero when different parties win, and we can still find cases in which policies move more than the median voters as condition (a) of the proposition indicates. Policy x_1 must be in $[0, s_1]$, and if s_1 is closer to $\theta_{m,1}$ than to $\theta_{m,2}$, party 2 uses its electoral advantage in the second period to nominate a more partisan candidate, resulting in a large policy change $|x_1 - x_2|$.

Proposition 3 also shows that policies may change very little if the same party wins in both periods. If the curvature of utility is sufficiently large, and the same party has an electoral advantage in both periods, the winning party will attempt to minimize policy changes.

5.3 Extreme Parties and "Overreaching"

Consider a situation in which one party selects an extremist candidate. Such a candidate will have a negative impact on the future composition of the party, resulting in a tradeoff between winning with a more extreme candidate and improving the party's future electoral prospects. This tradeoff is at the core of the discussion of "overreaching" by parties. As mentioned in the introduction, Fiorina (2016) defines "overreaching" as selecting more extreme policies today at the cost of alienating moderates.

To better understand this tradeoff, consider the following example. Suppose that the current party cutoff is $s_1 = 0.6$: All types $\theta \leq 0.6$ belong to party 1, and all types $\theta \geq 0.6$ to party 2. Because parties can only select one of their members as a candidate, party 2's most competitive candidate is located at 0.6. If the median voter is at $\theta_{m,1} = 0.5$, party 1 can win by nominating any candidate θ between 0.4 and 0.6. The new party cutoff becomes $s_2 = (\theta + 0.6)/2$. Hence, if party 1 nominates a more extreme candidate (i.e., chooses a lower θ), it alienates more party members, moving s_2 to the left. If s_2 remains to the right of the median voter, $\theta_{m,2}$, then party 1 can still win at $t = 2$, but it must choose a more centrist candidate, given that party 2 now has access to more competitive candidates. If $\theta_{m,2} > s_2$ then party 1 will lose in the next round, because party 2 can select a candidate who is more attractive to the median voter. Thus, party leaders face the tradeoff of selecting more partisan candidates in the current period versus retaining the party's size and dominance in the future.

We next show that overreaching increases if a party's leadership becomes more extreme. Before we provide the formal argument, we illustrate the result via an example.

Suppose that utility is given by $u_\theta(x) = |x - \theta|$, that $\theta_{m,1} = \theta_{m,2} = 0.5$ and $s_1 = 0.6$. Let the party leader's position, $\bar{m}$ be strictly between 0.4 and 0.5. Then problem 6 implies that party 1 nominates a candidate at $x_1 = \bar{m}_1$. In particular, if the winning policy in the second period (i.e., function h defined in (5)) is locally independent of the next period's party cutoff, s_2, then this is immediate. Otherwise, the winning policy is $x_2 = h(s_2) = 2\theta_{m,2} - s_2 = 1 - s_2$. Then the fact that $s_2 = (s_1 + x_1)/2$ implies that $x_1 = \bar{m}_1$ is optimal. Thus, a more extreme party, characterized by a lower $\bar{m}_1$, chooses a more extreme policy in the first period.

If $x_1 = \bar{m}_1$ then the party cutoff in period 2 becomes $s_2 = (s_1 + \bar{m}_1)/2 = 0.3 + 0.5\bar{m}_1$. The winning policy in the next period is $x_2 = h(s_2) = 2\theta_{m,2} - s_2 = 0.7 - 0.5\bar{m}_1$ (see (5)). The distance between x_2 and the median voter is $\left|\theta_{m,2} - x_2\right| = 0.5\bar{m}_1 - 0.2$. Thus, if party 1 is more extreme, i.e., $\bar{m}_1$ is lower, then the second period policy must be close to the position of the median voter. The reason is that the party has lost some of its electoral advantage by choosing an extreme policy in the first period.

The result is similar if party 2 wins in the second period. For example, suppose that $\theta_{m,2} = 0.55$ and that $0.4 < \bar{m}_1 < 0.5$. The party cutoff in the second period is $s_2 = 0.3 + 0.5\bar{m}_1$, but now $s_2 > \theta_{m,2}$. Thus party 2 wins, and the winning position is $x_2 = 0.7 - 0.5\bar{m}_1$. The distance of this policy from the median voter is $\left|\theta_{m,2} - x_2\right| = 0.15 - 0.5\bar{m}_1$. Thus, a more moderate party 1 will now result in a more moderate winning position by party 2. The reason is that a more moderate party 1 can nominate more competitive candidates at $t = 2$, thereby limiting party 2's ability to win with more extreme candidates in the second period.

Proposition 4 show that the insights from this example hold more generally.

Proposition 4 *Suppose that utility takes the form* $u_\theta(x) = |x - \theta|^\gamma$ *for* $\gamma \geq 1$ *and that* $\theta_{m,1} < s_1 < 4\theta_{m,2} - 3\bar{m}_1$.

Let x_t *denote the winning policies in periods* $t = 1, 2$. *Suppose that the winning party* i *at* $t = 1$ *becomes more extreme, i.e., the distance between the party leader,* $\bar{m}_1$, *and the median voter,* $\theta_{m,1}$, *increases, resulting in winning policies* y_t, $t = 1, 2$. *Then the new winning policy is more extreme in period 1, i.e.,* $y_1 \leq x_1$; *it becomes more moderate in period 2 if party 1 wins again, i.e.,* $x_2 \leq y_2 \leq \theta_{m,2}$, *and it becomes more extreme if party 1 loses at* $t = 2$, *i.e.,* $\theta_{m,2} \leq x_2 \leq y_2$.

This result show that if the decision makers in a party become more extreme, for example, because partisans start participating more in primaries than moderates, then these decision makers become more concerned about electing a partisan than about damaging the party's future electoral prospects. They will therefore overreach. A "weaker" party i in the second period will have to nominate a more moderate candidate and be less able to prevent the opposing party from winning.

6 The Model with Infinitely Many Time Periods

6.1 Long-term Party Dominance

In the previous section we analyzed the tradeoff between a party's electoral dominance and its ability to implement non-median policy in a two-period model. We now extend the time horizon to address the following question: how long can a party extend its dominance? As discussed in the introduction, parties can dominate elections over several periods, a phenomenon that the Downs and Wittman models do not explain. However, unlike predictions of political pundits, such periods of dominance do not last very long.

Without loss of generality suppose that party 1 dominates at the outset, i.e., the party cutoff is $s > \theta_m$. In principle, party 1 could win in all future time periods, with left-of-center candidates.

For example, let $x_{t+1} = \theta_m - (s - \theta_m)/(2(4)^t)$. Then $s_{t+1} = \theta_m + (s - \theta_m)/4^t$ and $x_{t+1} > 2\theta_m - s_{t+1}$, which implies that party 1 would win in all periods using candidates at x_t, no matter which candidates party 2 nominates.

However, we have seen that parties have the incentive to overreach by nominating more extreme candidates to the detriment of the party's future electoral prospects. We show that this overreaching is strong enough that a party's dominance disappears after finitely many periods, and the party cutoff becomes $s_t = \theta_m$.

Let $x_1(s)$ be the policy selected by the winning party's candidate, and let $V(s)$ be the continuation utility of the leader of party 1 if the current state is s. Let $x_2(s)$ be the policy of party 2. As in proposition 4 we assume that utility is $u_\theta(x) = |x - \theta|^\gamma$

We show that there is a unique Markov perfect equilibrium in which strategies $x_i(s)$ are continuous. Moreover, in this equilibrium, party 2, the minority party, selects the most moderate policy that is feasible, i.e., $x_2(s) = s$. Further, we show that party 1's equilibrium optimal policy solves

$$x_1(s) \in \arg\max_{x \in [0,s]} - |x - \bar{m}_1|^\gamma + \beta V\left(\frac{x+s}{2}\right) \text{ s.t. } x \geq 2\theta_m - s, \tag{7}$$

where

$$V(s) = - |x - \bar{m}_1|^\gamma + \beta V\left(\frac{x_1(s)+s}{2}\right). \tag{8}$$

Lemma 2 *Suppose that utility is of the form $u_\theta(x) = |x - \theta|^\gamma$, for $\gamma \geq 1$. Then there exists a unique Markov perfect equilibrium $x_i(s)$, $i = 1, 2$ in pure strategies. If $s \geq \theta_m$ then $x_1(s)$ solves Problem* (7)*, and $x_2(s) = s$. Further, $x_1(s)$ is continuous and strictly decreasing in s for $\gamma > 1$, i.e., a party with a larger electoral advantage chooses more extreme policies.*

To prove the Lemma we show that party 2's strategy is $x_2(s) = s$ in every equilibrium. First, if party 2 chooses this strategy, then party 1's best response must solve the recursive optimization problem (7), and the policy function $x_1(s)$ itself is unique. Thus, to show that we have an equilibrium it is sufficient to verify that party 2 has no incentive to deviate.

Suppose, for example, that $s = 0.6$ and that the median voter is at 0.5. Further, suppose that party 2 decides to deviate in period $t = 1$ and chooses policy $x_2 = 0.62$ instead of 0.6. Because policies are chosen simultaneously there is no effect on the current period and party 1's policy remains at $x_1(0.6)$. However, the party cutoff at $t = 2$ changes from $s' = 0.5(x_1(0.6) + 0.6)$ to

$\tilde{s}' = 0.5(x_1(0.6) + 0.62) = s' + 0.01$, i.e., party 1 becomes more dominant. At first glance it looks like a bad choice by party 2 to increase party 1's dominance, because party 1 will now select a more extreme policy (in fact we show that $x_1'(s) < 0$). However, there is one possible upside for party 2: If party 1's uses its increased dominance to overreach and adopts very extreme policies, the speed of convergence to the median may be increased.

For example, suppose that $x_1(0.6) = 0.46$. If party 2 chooses the equilibrium policy $x_2 = 0.6$, the party cutoff at $t = 2$ is 0.53. If party 2 chooses $x_2 = 0.62$, then the party cutoff would be 0.54. Suppose that $x_1(0.53) > 0.47$ and $x_1(0.54) = 0.46$. If no deviation occurs the party cutoff at $t = 3$ is strictly greater than 0.5 and party 1 still dominates. In contrast, if party 2 deviated at $t = 1$, then the party cutoff becomes 0.5, and the median voter's policy 0.5 wins in all future periods.

In the proof of Lemma 2 we show that this cannot happen, because $x_1'(s) > -1$, which implies that increasing party 2's policy results in a higher party cutoff in all future periods.

To show uniqueness, it suffices to show that $x_2(s) = s$ must hold in any equilibrium. For a general strategy $x_2(s)$ the constraint of problem (7) becomes $x \geq 2\theta_m - x_2(s)$. However, if $x_2(s) > s$ then this constraint cannot bind. Otherwise, party 2 could win by lowering the policy marginally. If s is close to θ_m, however, the constraint must bind, and thus $x_2(s) = s$ close to θ_m. Thus, $x_2(s)$ can only start to differ from s for some $\hat{s} > \theta_m$, i.e., $x_2(s) = s$ for $s \in [\theta, \hat{s}]$. We then show that party 2 would be strictly better off lowering $x_2(s)$ for some $s \geq \hat{s}$, which shows that $x_2(s) = s$ in equilibrium.

In view of Lemma 2 we now analyze equilibria by solving problem (7). Without loss of generality we consider the case where party 1 starts with an electoral advantage, i.e., where the party cutoff s is to the right of the median voter. Proposition 5 provides a sense in which convergence to the median voter cutoff occurs very quickly, and it explicitly details party 1's policy function, whenever s is sufficiently close to the median, and convergence to the median occurs within two periods.

Proposition 5 *Suppose that the median voter θ_m is the same in all periods, and that party 1 has an electoral advantage, i.e., $s > \theta_m$. Suppose that utility is of the form $u_\theta(x) = |x - \theta|^\gamma$, for $\gamma > 1$.*

1. *In any pure-strategy Markov the electoral advantage disappears after finitely many periods, n. That is, starting from any arbitrary party cutoff s, after n periods, the party cutoff becomes θ_m and the equilibrium policies are $x_1 = x_2 = \theta_m$. If N satisfies*

$$\left(2^N - 1\right)(\theta_m - \bar{m}_1)\left(1 - \left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}}\right) \geq s - \theta_m, \tag{9}$$

then $n \leq N$.

2. *Let*

$$\bar{s}_1 = \theta_m + (\theta_m - \bar{m}_1)\left(1 - \left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}}\right); \tag{10}$$

$$\bar{s}_2 = \bar{s}_1 + \left(1 - \left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}}\right)(\bar{s}_1 - \bar{m}_1) + 2\left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}}(\bar{s}_1 - \theta_m). \tag{11}$$

If $\theta_m < s \leq \bar{s}_1$ then the party cutoff converges to θ_m in the next period. If $\bar{s}_1 < s \leq \bar{s}_2$, then convergence to θ_m takes two periods.

3.

$$x_1(s) = \begin{cases} 2\theta_m - s & \text{if } s \in [\theta_m, \bar{s}_1]; \\ 2\theta_m - \bar{s}_1 - \frac{\left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}}(s - \bar{s}_1)}{2 + \left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}}} & \text{if } s \in [\bar{s}_1, \bar{s}_2]. \end{cases} \tag{12}$$

The proposition provides conditions for the speed of convergence of the party cutoff s to θ, i.e., the number of periods after which party 1's electoral advantage has disappeared. As (9) indicates, convergence is faster if (a) the party is more extreme, i.e., the party leader's position $\bar{m}_1$ is closer to zero; or if (b) β is smaller, or if (c) γ is closer to 1. Intuitively, if party 1 is more extreme, it will implement more extreme policies while it has an electoral advantage, which in turn means that the party loses its electoral advantage more quickly.

Similarly, a lower value of β increases speeds up convergence, because party 1 cares less about the future and therefore implements more extreme policies in the current period. A lower parameter γ implies that the party is less concerned about having to moderate in future periods and cares less about smoothing payoffs intertemporally. Again, this leads to more extreme policies in the current period and therefore quicker convergence.

We can also determine the policy function $x_1(s)$ for all states s for which convergence occurs in at most two periods. These functions are piecewise linear between θ_m and $\bar{s}_2$, with a "kink" at $\bar{s}_1$.

If $\gamma = 1$ then it is easy to solve problem (7) and determine the policy function $x_1(s)$ for all s. In particular, $x_1(s) = \max\{2\theta_m - s, \bar{m}_1\}$. As stated in the proposition, $\bar{s}_1 = 2\theta_m - \bar{m}_1$. For $n > 1$, $\bar{s}_n$ satisfies $\bar{s}_{n-1} = (x_1(\bar{s}_n) + \bar{s}_n)/2 = (\bar{m}_1 + \bar{s}_n)/2$, i.e., $\bar{s}_n = 2\bar{s}_{n-1} - \bar{m}_1$ which implies $\bar{s}_n = \theta_m + (2^n - 1)(\theta_m - \bar{m}_1)$. The smallest index n for which $\bar{s}_n \geq s$ determines the number of periods it takes for parties to converge to equal size starting from party cutoff s. This corresponds exactly to condition (9) in the Proposition, if we take the limit for $\gamma \downarrow 1$. For example, if $\theta_m = 1/2$

and $\bar{m}_1 = 1/4$ then $\bar{s}_1 = 3/4$ and $\bar{s}_2 = 2\bar{s}_1 - \bar{m}_1 = 5/4 > 1$. Thus, for all party cutoffs s, it takes at most two periods two converge when $\gamma = 1$.

For quadratic utility, $\gamma = 2$, the policy function $x_2(s)$ is also piecewise linear, and closed form solutions for $x_2(s)$ and $\bar{s}_n$ for $n > 2$ can be determined iteratively.

Consider again the case $\theta = 1/2$ and $m = 1/4$. Then the first statement in proposition 5 reveals that convergence is reached in at most three periods. Further, $\bar{s}_1 = \frac{3}{4} - \frac{1}{8}\beta$ and $\bar{s}_2 = \frac{5}{4} - \frac{1}{4}\beta - \frac{1}{16}\beta^2$. Using the formula for $x_1(s)$ for $s \leq \bar{s}_2$ from the proposition, we can determine $x_2(s)$ for $s \geq \bar{s}_2$ by solving

$$\max_{x\in[0,s]} -(x-\bar{m}_1)^2 - \beta\left(x_1\left(\frac{x+s}{2}\right) - \bar{m}_1\right)^2 - \beta^2\left(2\theta_m - \frac{x_1\left(\frac{x+s}{2}\right) + \frac{x+s}{2}}{2} - \bar{m}_1\right)^2 - \frac{\beta^3}{1-\beta}(\theta - \bar{m}_1)^2$$

where $x_1(s)$ is given by (12), and where the constraint $x \geq 2\theta_m - s$ is slack. We therefore get

$$x_1(s) = \begin{cases} 1 - s & \text{for } \frac{1}{2} \leq s \leq \bar{s}_1; \\ 1 - \bar{s}_1 - \frac{\beta(s-\bar{s}_1)}{4+\beta} & \text{for } \bar{s}_1 \leq s \leq \bar{s}_2; \\ 2\bar{s}_1 - \bar{s}_2 - \frac{(4-\beta)\beta^2(s-\bar{s}_2)}{(4+\beta)(16+\beta^2)} & \text{for } s \geq \bar{s}_2. \end{cases}$$

Note that $x_1(s)$ is strictly decreasing in s as stated in Lemma 2.

6.2 Political Realignment

If there is only one policy dimension, the optimal strategy of the minority party is to select the most moderate candidate available, in order to win back the center and regain the support of the single median voter type. In contrast, if policies cannot be condensed into one dimension, then there is no single voter type who is decisive. As a consequence, the objective of both parties is to assemble a coalition of voter types that can win the election. A party with an electoral disadvantage may find it in its interest not to pursue policies that may be the most competitive in the current period, but instead attracts different types of individuals to the party, who may form a winning coalition in the future.

As mentioned earlier, while Barry Goldwater's policy positions were unpopular with a large majority of the electorate, they attracted social conservatives to the Republican party, leading to electoral successes in future elections. As a consequence, the composition of parties themselves changed in a fundamental way, leading to a political realignment as defined in Key (1955). In this section we show that such a realignment can occur in our model. Further, we show that it can be optimal for a party to select a candidate who does worse in the general election than another type of candidate that may be available. The party may select such a candidate to appeal to new voter types that can form a winning coalition in the future.

We consider two different policy issues $i = 1, 2$. For example, we can think of issue 1 as pertaining to economics and issue 2 relating to social issues. For simplicity suppose that there are two policy positions on each issue,[10] a liberal, L, a and a conservative, C. There are four types of citizens, characterized by their ideal point on each position. Citizens' utilities are given by

$$u_\theta(x) = \begin{cases} 0 & \text{if } x = \theta; \\ -1 & \text{if } x_1 \neq \theta_1,\ x_2 = \theta_2; \\ -d & \text{if } x_1 = \theta_1,\ x_2 \neq \theta_2; \\ -e & \text{if } x_1 \neq \theta_1,\ x_2 \neq \theta_2; \end{cases} \tag{13}$$

where $e > d$ and $e > 1$ reflects that a policy that differs from the voter's ideal point on all issues is worse than one that differs only on one issue. The utility normalization of the first two cases in (13) is without loss of generality.

As before we assume that citizens join the party whose policy they preferred in the previous period (condition (4)). Given that preferences in this model are strict, if parties nominate different candidate types then parties in the next period are a partition of the type space, i.e., $S_{1,t} \cap S_{2,t} = \emptyset$, and $S_{1,t} \cup S_{2,t} = \Theta$. If $S_{i,t} = \{(i, 1), (i, 2)\}$ then parties are differentiated with respect to economics issues and aligned on social issues. The reverse is true if $S_{i,t} = \{(1, i), (2, i)\}$. The objective is to understand under what circumstance we can observe a switch from one party structure to the other.

Suppose that $d > 1$, which means that voters put more emphasis on differences on issue 2, the social issue. Assume we start with a situation in which parties are originally only differentiated with respect to economic issues. That is, party 1 consists of all economic liberals, while party 2 consists of all economic conservatives. Suppose that party leaders are types (L, L) and (C, C). If there are more social liberals than social conservatives, i.e., if $\phi_t(\{(L, L), (C, L)\}) > \phi_t(\{(L, C), (C, C)\})$, then it is an equilibrium for both candidates to select a social liberal. In particular, party 1 selects a candidate (L, L), and party 2 a candidate (C, L). If in some periods economic conservatives are in the majority, and in other periods economic liberals are in the majority, then this is in fact the unique equilibrium. If, for example, party 2 deviated and nominated a candidate at (C, C), then this candidate loses and in all remaining periods, party 1 consists of all social liberal and party 2 of all social conservatives. If $\phi_t(\{(L, L), (C, L)\}) > \phi_t(\{(L, C), (C, C)\})$ for all periods t, then party 2 never wins, while party 1 can nominate its most preferred candidate, (L, L), which makes party 2 strictly worse off.

[10]See Krasa and Polborn (2010) for a general description of models where policy choices are discrete.

Now suppose that beginning with periods $\bar{t}$, there are always more social conservatives than social liberals, i.e., $\phi_t(\{(L, L), (C, L)\}) > \phi_t(\{(L, C), (C, C)\})$ for $t \geq \bar{t}$. In all these periods it is no longer an equilibrium for both parties to selects social liberals. In fact, there are two possible equilibria. First, parties remain the same, i.e., differentiated with respect to economic issues, but both candidates are social conservatives: (L, C) for party 1, and (C, C) for party 2. Alternatively, we can have a realignment of parties. That is, party 1 consists of all social liberals, party 2 of all social conservatives and party 2 wins with a candidate at (C, C) — party 1's candidate selection is irrelevant. The question is which of these subgames will be reached.

Suppose that economic liberals outnumber economic conservatives in all periods. Then party 2 would want to select a candidate (C, C) at $t = 1$. This candidate would lose, because social conservatives are in the minority, but it attracts social conservatives to party 2. As a result party 2 consists of all social conservatives and at $\bar{t}$ we would reach the subgame where party 2 wins with candidate (C, C) in all remaining periods. To prevent this from happening, party 1 could fight for the support of social conservatives by nominating a social conservative at $t = 1$, i.e., candidate (L, C). Party 1 would remain the majority party and (L, C) would therefore win in all periods. However, party 1's leader prefers a candidate who is liberal on both positions, and thus loses $-d$ units of utility in each period $t < \bar{t}$. As long as β is not too close to l, this loss dominates party 1's gain of appealing to social conservatives.

If economic conservatives outweigh economic liberals starting with period $\bar{t}$, then party 2 does not have a strict incentive of appealing to social conservatives before time $\bar{t}$ if party 1 keeps nominating liberal candidates, (L, L). Thus, if it is optimal for party 1 to nominate candidates (L, L), then an equilibrium without party realignment exists. However, if β is not too small, party 1 has now the incentive of appealing to social conservatives before party 2. This, in turn, provides incentives for party 2 to preempt this move by party 1. We now summarize the results and provide conditions on the discount factor β under which party realignments always occurs.

Proposition 6 *Suppose that at $t = 1$ party 1 consists of all economic liberals and party 2 of all economics conservatives, and that party leaders are types (L, L) and (C, C), respectively. Suppose that (a) there are more social liberals than social conservatives in periods $t < \bar{t}$,*[11] *while (b) the reverse is true when $t \geq \bar{t}$, and (c) that economic liberals outnumber economic conservatives in periods $t < \bar{t}$. Then*

1. *When economic liberals outnumber economics conservatives in all periods, a party realignment occurs in every equilibrium if and only if $\beta < (d/e)^{1/\bar{t}}$. After the realignment, party 1*

[11]That is, $\phi_t(\{(L, L), (C, L)\}) > \phi_t(\{(L, C), (C, C)\})$ for all $t < \bar{t}$.

consists of all social liberals, and party 2 of all social conservatives,

2. *When economic conservatives outnumber economic liberals for $t \geq \bar{t}$. a party realignment occurs in every equilibrium if and only if $\beta > 1/(1 + e - d)$. After the realignment either party 1 or party 2 may become the social-conservative party.*

The first statement of the proposition considers the case where economic conservatives are always in the minority. The only option for party 2 to win in the future is to select a socially conservative candidate in order to appeal to social conservatives. In the current period this may lead to a larger loss in the election.

For example, suppose that the distribution of types is $\phi_t(\{(L, L)\}) = 0.4$, $\phi_t(\{(C, L)\}) = 0.3$, $\phi_t(\{(L, C)\}) = 0.2$, and $\phi_t(\{(C, C)\}) = 0.1$. If party 2 selects a candidate (C, L) against party 1's candidate (L, L), then party 2 receives 40% of the vote. In contrast, by selecting a candidate (C, C), party 2's candidate only gets the vote of all social conservatives, i.e., of 30% of the population. However, this short run loss is outweighed by the fact a realignment of parties takes places. Party 2 now consists of all social conservatives, and when social conservatives form a majority at $\bar{t}$, party 2 can win with candidate (C, C), even though a minority of the population supports economic conservative positions. In contrast, if no realignment occurs, then party 1 can nominate a candidate (L, C) in periods $t \geq \bar{t}$ who wins, because economic liberals are in the majority.

The second statement considers the case where economic liberals become the minority at $\bar{t}$. Now party 2 has no longer a strict incentive to support socially conservative positions before period $\bar{t}$, because (C, C) would win agains (L, C). In contrast, it may now be optimal for party 1 to switch to a socially conservative position before period $\bar{t}$. This would lose the election for party 1. However, the resulting realignment would allow party 1 to win with an economic liberal position in periods $t \geq \bar{t}$. In other words, the long-run benefit of a realignment to party 1 can be large enough to outweigh losing some elections in the short run.

Both statements of the proposition show that when the distribution of voter types changes, parties that are currently in the minority may have a strong incentive to reach out to new groups of voters, in order to change or increase the talent pool for future elections.

For example, when Tony Blair become leader of the Labour Party in 1994, the party had not won an election since 1974. Blair's goal was a to redefine the party's message away from its more traditional socialist roots: "Above all, in a society in which fewer and fewer people thought of themselves as traditional working class, we needed to build a new coalition between the aspirant up-and-coming and the poorest and most disadvantaged" (Blair, 2015).

Although standard models of electoral competition also predict that candidates would adjust a policy platform in response to shifts in voter preferences, our model captures the impact of these policies on party membership and future potential candidates. In fact, in response to Blair's policy shift, membership in the Labour Party increased from 300,000 to 400,000 between 1994 and 1997 (Keen and Apostolova, 2017), presumably because the individuals who joined found Labour's new platform appealing. We can also argue that this lead to the election of more moderate Labour MPs. Thus, when Jeremy Corbyn was running for the leadership of the Labour party in 2015 emphasizing more traditional, pre-Blair Labour positions, only 14 MPs supported his bid.

However, it is interesting that this new realignment of Labour is taken place again when the party is out of power. Further, as in 1994 party membership is again increasing. In the two months after Corbyn's election 50,000 individuals joined the party (Stone, 2015), and in 2017 party membership is 517,000 up from 194,000 in 2014 (Keen and Apostolova, 2017). In other words, Corbyn's policy positions have changed the composition of the Labour party. In turn, this will influence the type of candidates who will run for office and represent the party in the future.

7 Changes of Party Leadership

We have assumed that the ideologies of party leaders are fixed. For example, this would be the case if party control rests with an elite as in Buisseret and Weelden (2017). We now discuss what happens in the one-dimensional case when leadership changes in responses to changes in the set of party members.

Consider a two period model, and suppose that party 1 has an electoral advantage in the first period, i.e., $s > \theta_{m,1}$. If s increases, and party members determine candidates in a primary, we can imagine a situation in which increasing s increases the median primary voter's type, i.e., the party's leader becomes more moderate. As a result, party 2 may be better off to drop out of the race if party 1 is sufficiently dominant. If both parties compete in the election then the equilibrium of the subgames starting in the second period if both parties compete is the same as in (5), replacing $\bar{m}_i$ by $m_i(s)$. However, if $s > 2\theta_{m,2} - m_1(s) = 0.5$ then candidate $\theta = m_1(s)$ wins. In this case party 2 would be better off not competing, if this would induce members of the now inactive party 2 to participate in the primary of party 1, thereby raising $m_1(s)$. A similar argument can be made in the first period, if s is too large.

Note that if we consider a party whose leadership is more extreme, i.e, where $m_1(s)$ is lower for all s, then party 2 is more likely to compete because $2\theta_{m,2} - m_1(s)$ is lowered. In other words, party 2 faces the following tradeoff. By competing, party 2 can discourage party 1 from nominating an extreme candidate. However, if party 2 does not have any competitive candidates, then this effect may be very limited. Instead, if the main election is not contested, more moderate voters may participate in party 1's primary, resulting in the selection of a more moderate candidate. In practice, many elections at the state level in the U.S. are not contested. For example, in 2016 one third of all House and Senate elections in North Carolina in 2016 had only a single candidate on the general election ballot. In practice a party's inability to find a qualified candidate is often mentioned as the reason for this problem, and this precisely what drives this result.

8 Conclusion

The paper introduces a dynamic model of electoral competition in which parties must select candidates from among their members, linking the identity of current candidates with the set of candidates that are available in future elections. This setup allows us to address a new set of questions for which existing models of political competition do not apply: (1) Do parties "overreach," by trying to take advantage of a current majority to implement policies that negatively impact the parties future electoral prospects? (2) How long can parties retain majorities? (3) If there is more than one policy issue, under what conditions do parties realign, for example, by moving from a situation in which parties represent economic liberal and conservatives, respectively, to a new one in which social issues become the fault line?

We consider a single election in each period, but it would be interesting to extend the analysis to multi-district elections. For example, suppose that there is one policy dimension, and parties are given by $S_1 = [0, 0.5]$, $S_2 = [0.5, 1]$, where 0.5 is the position of the national median. If median voters in legislative districts differ from the national median, then one of the parties has an electoral advantage, and can therefore win with a more extreme candidate. An increase in heterogeneity between legislative districts, for example through gerrymandering, would then result in an increased polarization of candidates that run in the individual elections.

For example, if we start in a situation where the median voters in all districts are all located at 0.5, then there is no polarization because all candidates are located at 0.5. If districts become more heterogenous, and medians move to 0.4 in some districts and to 0.6 in other districts, then we get a more polarized outcome. That is, in the districts with median 0.4, party 1 wins with a candidate $\theta < 0.4$, while in districts with median 0.6, candidates $\theta > 0.6$ win.

Similarly, we can also consider the case where candidates for federal offices, must be selected from the group of individuals that are successful in state races. Then one party dominance in a state, will lead to a more extreme set of possible candidates, which in turn increases polarization at the federal level. Such an extension would allow us to consider spillovers between state and federal elections.

9 Appendix

Proof of Proposition 1. For $T < \infty$ the result follows immediately. Suppose that time is infinite. Because Θ is finite, the number of players in the game remains finite. Theorem 2 of Jaśkiewics and Nowak (2016) or the related theorem in Mertens and Parthasarathy (2003) implies existence if the transition rule satisfies norm continuity. The transition rule in our game is deterministic and given by function ψ. However, since the action space of the game, i.e., the set of candidate types Θ, is finite, norm continuity is immediate. ■

Proof of Lemma 1. Let x and y denote the policies that would be implemented by the candidates of parties 1 and 2, respectively. We can assume that $x \neq y$.

Suppose that there exist $\theta_1 < \theta_2 < \theta_3$ with $u_{\theta_i}(x) - u_{\theta_i}(y) \geq 0$ for $i = 1, 3$ and $u_{\theta_2}(x) - v_{\theta_2}(y) < 0$. Then the function $f(z) = u_z(y) - v_z(x)$ assumes a local minimum at some point $\bar{z}$ in the open interval (θ_1, θ_3). As consequence, $0 = f'(z) = \frac{\partial u_z(y)}{\partial \theta} - \frac{\partial u_z(x)}{\partial \theta}$. This, however, is a contradiction because $\frac{\partial u_z(x)}{\partial \theta}$ is strictly monotone in x, which together with the fact that $x \neq y$ implies that $\frac{\partial u_z(y)}{\partial \theta} \neq \frac{\partial u_z(x)}{\partial \theta}$.

As a consequence, there are three cases:

1. $u_\theta(x) - u_\theta(y) > 0$ for all θ,
2. $u_\theta(x) - u_\theta(y) < 0$ for all θ,
3. There exits a unique $\bar{\theta}$ such that $u_{\bar{\theta}}(x) - u_{\bar{\theta}}(y) = \alpha$.

In the first case, all types prefer party 2, and in the second case all types prefer party 1, and we clearly have the desired structure. In the third case, define $s = \bar{\theta}$. Because $\bar{\theta}$ is unique if follows that the set $\{\theta | u_\theta(x) - u_\theta(y) \geq 0\}$ is an interval of the form $[0, s]$ or $[s, 1]$, and similar for $\{\theta | u_\theta(x) \geq u_\theta(y)\}$. ■

Proof of Proposition 2. By assumption each type θ has a unique most preferred policy $x(\theta)$. Further, if $\theta \neq \theta'$ then the single crossing property implies that $x(\theta) < x(\theta')$ for $\theta < \theta'$.

In particular if $y = x(\theta)$ is the individual's optimal policy then $\frac{\partial u_\theta(x)}{\partial x} = 0$. By assumption $\frac{\partial^2 u_\theta(x)}{\partial x \partial \theta} > 0$. As a consequence, $\frac{\partial u_{\theta'}(x)}{\partial x} > 0$. Since utility is strictly concave it follows that $\frac{\partial u_{\theta'}(y)}{\partial x} = 0$ for some $y > 0$.

Without loss of generality we can therefore rescale Θ such that type θ's most preferred policy is $\theta = x(\theta)$. After this rescaling it is possible that Θ is a strict subset of $[0, 1]$. We can define arbitrary preferences on $[0, 1] \setminus \Theta$, because types in $[0, 1] \setminus \Theta$ are never picked by party leaders and therefore do not affect the game.

Suppose that at $t = 2$ parties are given by $S_1 = [0, s]$ and $S_2 = [s, 1]$, and let $\theta_{m,2}$ be the median type at $t = 2$. There are two cases to consider:

Case 1: $\theta_{m,2} \leq s$.

Assumption 1 implies that party 1's leader is at $\bar{m}_1$. If party 2 competes, then party 1 selects candidate $\theta_1 = \max\{\bar{m}_1, 2\theta_{m,2} - s\}$; party 2 selects candidate $\theta_2 = s$. The median voter at $\theta_{m,2}$ and all types $\theta < \theta_{m,2}$ prefer candidate 1. Hence, electing candidate 1 is an equilibrium. The winning candidate is therefore a continuous, decreasing function of s.

Case 2: $\theta_{m,2} \geq s$.

In this case party 2 wins. The analysis is analogous to that of case 1. Specifically, the winning candidate is given by $\bar{\theta}_2 = \min\{\bar{m}_2, 2\theta_{m,2} - s\}$.

If follows immediately that the equilibrium at each subgame at $t = 2$ is unique.

Now consider the first stage of the game. Suppose that parties are given by $[0, s]$ and $[s, 1]$. Assume without loss of generality that the median voter at time $t = 1$ satisfies $\theta_{m,1} \leq s$. Then party 1 wins if $\theta_1 \geq 2\theta_{m,2} - s$. As a result, the party leader, type $\theta = m_1(s)$ solves

$$\max_{x \in [0,s]} u_\theta(x) + \beta u_\theta\left(2\theta_{m,2} - \frac{x + s}{2}\right) \text{ s.t. } x \geq 2\theta_{m,1} - s. \tag{14}$$

The second derivative of the object with respect to x is

$$u''_\theta(x) + \frac{\beta}{4} u''_\theta\left(2\theta_{m,2} - \frac{x + s}{2}\right) < 0.$$

Hence, (14) has a unique solution, x_1, which is party 1's equilibrium strategy. It is not optimal for party 2 to deviate, because this would raise the party cutoff at $t = 2$ which, as shown above, lowers or at best keeps the policy at $t = 2$ the same.

We next prove uniqueness.

Without loss of generality suppose that $s > \theta_{m,1}$. If party 2 selects a candidate at $\theta_2 = s$ then the fact that that the solution of (14) is unique implies that the equilibrium is unique. We can therefore assume that $\theta_2 > s$.

If party 2 deviates by changing θ_2, then as show above the winner in the second period changes unless $\bar{\theta} = \bar{m}_1$ or $\bar{\theta} = \bar{m}_2$, depending on whether party 1 or 2 wins in the second period. Given that local changes of θ_2 do not change the identity of the winning candidate at $t = 2$, the same is true for small changes of the type of party 1's candidate.

The constraint of optimization problem 14 must be therefore be slack. Else, lowering θ_2 would result in party 1 winning the election. Thus, the leader of party 1's optimal candidate is locally unconstrained, and therefore $\theta_1 = \bar{m}_1$ is optimal, independent of the choice of θ_2. Hence equilibrium payoffs are unique.

Finally, note that the only case in which strategies are not unique is when one of the parties is indifferent between competing and not competing. ■

Proof of Proposition 3. First, suppose that $\theta_{m,1} < s_1 < (\theta_{m_1} + \theta_{m,2})/2$, which implies that $s_1 < \theta_{m,2}$. The equilibrium policy solves problem 6. Thus, party 1 wins in the first period, and policy is $x_1 \leq s_1$. The winning policy in the next period is $x_2 = h(x_1) = 2\theta_{m,2} - 0.5(x_1 + s_1) \geq 2\theta_{m,2} - s_1$. Because $s_1 < \theta_{m,2}$ it follows that $x_2 \geq x_1$. Thus, $|x_1 - x_2| = x_2 - x_1 \geq 2(\theta_{m,2} - s_1) > \theta_{m,2} - \theta_{m,1}$, because $s_1 < (\theta_{m_1} + \theta_{m,2})/2$.

To consider the remaining case, we determine the first-order condition for policy x_1.

Suppose that $s_2 > 2\theta_{m,2} - \bar{m}_1$. Then $\frac{\partial h(0.5(x_1+s_1))}{\partial x_1} = 0$. The objective or problem 6 is strictly concave because $u'' < 0$.

If the constraint is slack then $u'_\theta(x_1) = 0$, which implies $x_1 = \theta = \bar{m}_1$. The constraint must be satisfied and hence $\bar{m}_1 \geq 2\theta_{m,1} - s_1$. However, this contradicts the assumption that $\bar{m}_1 < 2\theta_{m,1} - s_1$.

Thus, $2\theta_{m,2} - \bar{m}_2 \leq s_1 \leq 2\theta_{m_2} - \bar{m}_1$ and hence $(\partial/\partial x_1)h(0.5(s_1 + x_1)) = -0.5$. Further, we must have $\theta = \bar{m}_1 < x_1$. The derivative of the object of problem 6 is therefore

$$-\gamma(x_1 - \theta)^{\gamma-1} + \frac{\gamma\beta}{2}\left|2\theta_{m,2} - \frac{x_1 + s_1}{2} - \theta\right|\left(2\theta_{m,2} - \frac{x_1 + s_1}{2} - \theta\right)^{\gamma-2}. \tag{15}$$

If follows immediately that the second derivative is strictly negative.

Note that if $x_1 < \frac{1}{3}(2\theta_{m,2} + 2\theta_{m,1} - s_1)$, then $x_2 > \frac{1}{3}(5\theta_{m,2} - \theta_{m,1} - s_1)$ and $|x_1 - x_2| > |\theta_{m,2} - \theta_{m,1}|$. It is therefore sufficient to prove that (15) is strictly negative at $x_1 = \frac{1}{3}(2\theta_{m,2} + 2\theta_{m,1} - s_1)$, i.e., that

$$-\gamma\left(\frac{2}{3}\theta_{m,2} + \frac{2}{3}\theta_{m,1} - \frac{1}{3}s_1 - \theta\right)^{\gamma-1} + \frac{\gamma\beta}{2}\left(\frac{5}{3}\theta_{m,2} - \frac{1}{3}\theta_{m,1} - \frac{1}{3}s_1 - \theta\right)^{\gamma-1} < 0. \tag{16}$$

Note that (16) holds if and only if

$$\left(\frac{2}{3}\theta_{m,2} + \frac{2}{3}\theta_{m,1} - \frac{1}{3}s_1 - \theta\right) > \left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}}\left(\frac{5}{3}\theta_{m,2} - \frac{1}{3}\theta_{m,1} - \frac{1}{3}s_1 - \theta\right). \tag{17}$$

Next, $\theta = \bar{m}_1 < 2\theta_{m,1} - s_1$. Thus, (17) holds if

$$\left(\frac{2}{3}\theta_{m,2} - \frac{4}{3}\theta_{m,1} + \frac{2}{3}s_1 - \theta\right) \geq \left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}} \left(\frac{5}{3}\theta_{m,2} - \frac{7}{3}\theta_{m,1} - \frac{1}{3}s_1 - \theta\right). \tag{18}$$

Finally, if (18) holds for $s = 0.5(\theta_{m,1} + \theta_{m,2})$ then it holds for any $s \geq 0.5(\theta_{m,1} + \theta_{m,2})$. As a consequence, (18) holds if

$$(\theta_{m,2} - \theta_{m,1}) \geq \left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}} (2\theta_{m,2} - 2\theta_{m,1}). \tag{19}$$

Thus, (19) holds if $\beta \leq 2^{2-\gamma}$, in which case $|x_1 - x_2| > |\theta_{m,1} - \theta_{m,2}|$.

Finally, suppose that $\theta_{m,1} < \theta_{m,2} < s_1$. If we set (15) equal to zero, then

$$x_1 - \theta = \left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}} \left(2\theta_{m,2} - \frac{x_1 + s_1}{2} - \theta\right). \tag{20}$$

If $\gamma \to \infty$ then (20) implies that $x_1 \to (4\theta_{m,2} - s_1)/3$.

Next, we show that the constraint of problem 6 is slack for large γ. It is sufficient to show that $(4\theta_{m,2} - s_1)/3 > 2\theta_{m,1} - s_1$. This inequality is equivalent to $2\theta_{m,2} > 3\theta_{m,1} - s_1$, which holds because $s_1 > \theta_{m,1}$ and $\theta_{m,1} < \theta_{m,2}$. Because $s_1 > \theta_{m,2}$ it also follows that $(4\theta_{m,2} - s_1)/3 < s_1$. Thus, for large γ the solution to problem 6 is characterized by (20).

Finally, it follows that $h((4\theta_{m,2} - s_1)/3) = 4(\theta_{m,2} - s_1)/3$. Therefore $|x_1 - x_2| \to 0$ as $\gamma \to \infty$. ■

Proof of Proposition 4. By assumption $s_1 > \theta_{m,1}$ and therefore party 1 wins in the first period. If function h defined in (5) is locally independent of s, then problem 6 implies that $x_1 = \theta = \bar{m}_1$, i.e., the party chooses its most preferred candidate in the current period. In the second period, party 1 again chooses a candidate at $\bar{m}_1$. Thus, $\bar{m}_1 \geq 2\theta_{m,2} - s_2$, where $s_2 = (\bar{m}_1 + s_1)/2$. This implies, $s_1 \geq 4\theta_{m,2} - 3\bar{m}_1$, a contradiction to the assumption of the Proposition.

Now let $h(s) = 2\theta_{m,2} - s$. If the constraint of problem 6 binds, then $x_1 = 2\theta_{m,1} - s$, which implies $x_1 = y_1$. The party cutoff in the next period is $s_2 = \theta_{m,1}$. If $\theta_{m,2} < \theta_{m,1}$ then $x_2 = \min\{\bar{m}_2, 2\theta_{m,2} - \theta_{m,1}\}$ and hence $x_2 = y_2$. If $\theta_{m,2} \geq \theta_{m,1}$ then party 1 wins again and $x_2 = \max\{\bar{m}_1, 2\theta_{m,2} - \theta_{m,1}\}$, which is non-decreasing in $\bar{m}_1$. Thus, $\theta_{m,2} \geq y_2 \geq x_2$.

Next, suppose that there exists an interior solution. Let $\theta = \bar{m}_1$. Then the first order condition of problem 6 with a slack constraint is given by

$$v'(x - \theta) = \frac{\beta}{2} v' \left(2\theta_{m,2} - \frac{x + s_1}{2} - \theta\right), \tag{21}$$

where $v'(z) = \gamma z |z|^{\gamma-2}$. In order for (21) to have a solution, $x - \theta$ and $2\theta_{m,2} - \frac{x+s_1}{2} - \theta$ must have the same sign. Both cases yield the same solution

$$x = \frac{\left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}} (4\theta_{m,2} - 2\theta - s_1) + 2\theta}{\left(\frac{\beta}{2}\right)^{\frac{1}{\gamma-1}} + 2}. \tag{22}$$

Equation (22) implies that a more extreme party leader, i.e., a lower θ results in a more extreme policy in the current period. This, however, means that in the next period, the winning candidate's policy is more moderate. If party 1 wins again in the next period, which is the case if $\theta_{m,2} < (s_1 + x_1)/2$, then this means that party 1 must nominate a more moderate candidate. If party 2 wins at $t = 2$, then they will have a more partisan candidate.

■

Proof of Lemma 2. Let $C([\theta_m, 1])$ be the set of continuous function on $[\theta_m, 1]$ (endowed with the sup norm). Define $T\colon C([\theta_m, 1]) \to C([\theta_m, 1])$ by

$$T(V) = \max_{x\in[0,s]} -|x - \bar{m}_1|^\gamma + \beta V\left(\frac{x+s}{2}\right) \text{ s.t. } x \geq 2\theta_m - s. \tag{23}$$

Then T satisfies Blackwell's sufficient conditions (c.f., Theorem 3.3 in Stokey et al. (1989)). Thus, T is a contraction mapping on $C([\theta_m, 1])$. Let $\tilde{V}$ denoted the fixed point of T. Theorems 4.7 and 4.8 of Stokey et al. (1989) imply that $\tilde{V}$ is strictly increasing and strictly concave in s. The Berge maximum theorem therefore implies that the policy function $x_1(s)$ is continuous. Further, theorem 4.10 of Stokey et al. (1989) implies that $\tilde{V}$ is differentiable at all x that are interior, i.e., where $2\theta_m - s < x < s$.

We next show that $x_1(s)$ is strictly decreasing in s. This is immediate if the constraint $x \geq 2\theta_m - s$ binds. Further, the constraint $x_1 \leq s$ must be slack when $s > \theta_m$. Otherwise, if $x_1(s) = s$ then next period's state is again s and the winning policy would remain at s. Party 1 could strictly improve by choosing policy $x_1 = \theta_m$ in all future time periods.

Because $\tilde{V}$ is strictly increasing, $x_1(s) \geq \bar{m}_1$. Otherwise, if $x_1(s) < \bar{m}_1$ the objective of (7) can be increased by raising x_1. Thus, $x_1(s)$ satisfies the first order condition of (7) given by

$$-\gamma(x_1(s) - \bar{m}_1)^{\gamma-1} + \frac{\beta}{2}\tilde{V}'\left(\frac{x_1(s)+s}{2}\right) = 0. \tag{24}$$

Taking the derivative of (24) with respect to s yields

$$-\gamma(\gamma-1)x_1'(s)(x_1(s) - \bar{m}_1)^{\gamma-2} + \frac{\beta}{4}\tilde{V}''\left(\frac{x_1(s)+s}{2}\right)(x_1'(s)+1) = 0. \tag{25}$$

Suppose that $x_1'(s) \geq 0$. Then the left-hand side of (25) would be strictly negative, a contradiction. Thus, $x_1'(s) < 0$ and hence $x_1(s)$ is strictly decreasing in s. Further, $x_1'(s) \geq -1$. Otherwise, if $x_1'(s) < -1$ then the left-hand side of (25) is strictly positive, a contradiction.

We have shown that there exist $x_1(s)$ that is a best response to $x_2(s) = s$ in every state $s \geq 2\theta_m$. To prove that $x_1(s)$, $x_2(s)$ is a Markov perfect equilibrium, it is sufficient to show that there does not exist a one-step deviation for party 2.

Suppose by way of contradiction that it is optimal for party 2 to deviate at some state s from $x_2(s) = s$ to $\hat{x}_2 > s$. Because parties move simultaneously, this deviation does not affect the current policy, and thus only increases next period's state from $s' = (x_1 + s)/2$ to $\bar{s}' = (x_1 + \hat{x}_2)/2$. Because $x_1(s)$ is strictly decreasing, this implies that next period's winning policy is strictly further to the left, lowering the period utility of party 2's leader. $x_1'(s) \geq -1$ implies that $s'' = (s' + x(s'))/s \geq (\hat{s}' + x(\hat{s}'))/2 = \hat{s}''$, where s'' and $\hat{s}''$ are the states (i.e., party cutoffs) in two periods. Hence, in all future periods, the winning policy (i.e., party 1's policy) is moved weakly to the left after the deviation. The deviation therefore makes party 2 strictly worse off, which proves that $x_i(s)$, $i = 1, 2$ is a Markov perfect equilibrium.

Next, we prove uniqueness of the equilibrium. To do this, we first list necessary conditions for an equilibrium.

If $x_i(s)$, $i = 1, 2$, is a Markov perfect equilibrium then $x_1(s)$ must satisfy

$$x_1(s) \in \arg\max_{x \in [0,s]} -|x - \bar{m}_1|^\gamma + \beta V\left(\frac{x + x_2(s)}{2}\right) \quad \text{s.t.} \quad x \geq 2\theta_m - x_2(s), \tag{26}$$

where $V(s)$ is given by (8). Similarly, in order for $x_2(s)$ to be a best response,

$$x_2(s) \in \arg\max_{x \in [s,1]} -|x_1(s) - m_2(s)|^\gamma + \beta V\left(\frac{x_1(s) + x_2}{2}\right) \quad \text{s.t.} \quad x_2(s) \geq s, \tag{27}$$

where

$$V(s, \theta) = -|x_1(s) - \theta|^\gamma + \beta V\left(\frac{x_1(s) + x_2(s)}{2}, \theta\right). \tag{28}$$

Further, it should not be possible for party 2 to win, by a selecting a more moderate policy, i.e., $x_1(s) \geq 1 - s$. In other words, if $x_2(s) > s$ then the constraint $x \geq 2\theta_m - x_2(s)$ in (26) must be slack.

We show that $x_2(s) = s$ in order for these necessary conditions for a Markov perfect equilibrium to be satisfied.

Suppose by way of contradiction that there exists an equilibrium with $x_2(s) > s$ for some $s > \theta_m$. Let $\hat{s} = \inf\left\{s \mid x_2(s) - s > 0, \ s \geq \theta_m\right\}$.

Case 1: $\hat{s} = \theta_m$.

Note that $x_2(\theta_m) = \theta_m$. Suppose by way of contradiction that $x_2(\theta_m) > \theta_m$. As noted above, in equilibrium, $x_1(s) \geq 2\theta_m - s$ and therefore $x_1(\theta_m) = \theta_m$. Party 1 could deviate to x_1 with $2\theta_m - x_2(\theta_m), \bar{m}_1 < x_1 < \theta_m$, which would make the leader of party 1 at $\bar{m}_1$ strictly better off.

Similarly, it follows that $\lim_{s\downarrow\theta_m} x_2(s) = \theta_m$. Otherwise, there exists a $\varepsilon > 0$ and s that are arbitrarily close to θ_m for which $x_2(s) - x_2(\theta_m) \geq \varepsilon$ (recall that $x_2(s) \geq s$). In equilibrium $x_1(s) \geq 2\theta_m - s$. If s is sufficiently close to θ_m it would be optimal for party 1 to deviate to some $x_1 < x_1(s)$ with $2\theta_m - x_2(s), \bar{m}_1 < x_1 < x_1(s)$, a contradiction.

Define

$$V(s) = -\left|2\theta_m - x_2(s) - \bar{m}_1\right|^\gamma - \frac{\beta\left|\theta_m - \bar{m}_1\right|^\gamma}{1-\beta}. \tag{29}$$

Let T^n be operator T applied n times, where T is defined in (23). Because T is a contraction mapping, it follows that there exists a fixed point $\tilde{V}$ of T such that $\lim_{n\to\infty} T^n(V) = \tilde{V}$. Because $x_1(s)$ is a best response, it must solve (26) for value function $\tilde{V}$.

We next show that there exists $\varepsilon > 0$ such that $\tilde{V}(s) = V(s)$ for all $s \in [\theta_m, \theta_m + \varepsilon]$. Given that T is a contraction mapping, it is sufficient to prove that there exists $\epsilon > 0$ such that $T(V)(s) = V(s)$ for all all $s \in [\theta_m, \theta_m + \varepsilon]$.

In order to determine $T(V)$ we solve

$$\max_{x\in[0,s]} -\left|x - \bar{m}_1\right|^\gamma - \beta\left|2\theta_m - \frac{x + x_2(s)}{2} - \bar{m}_1\right|^\gamma - \frac{\beta^2\left|\theta_m - \bar{m}_1\right|^\gamma}{1-\beta} \quad \text{s.t} \;\; x \geq 2\theta_m - s. \tag{30}$$

It is immediate that the objective of (30) is differentiable and strictly concave in x. Further, if s is close to θ_m then $x > \bar{m}_1$. Thus, the solution is given by the first order condition

$$-\gamma(x - \bar{m}_1)^{\gamma-1} + \frac{\gamma\beta}{2}\left(2\theta_m - \frac{x + x_2(s)}{2} - \bar{m}_1\right)^{\gamma-1} + \lambda = 0, \tag{31}$$

where λ is the Largrange multiplier of the constraint.

We have shown that $\lim_{s\downarrow\theta_m} x_2(s) = \theta_m$. Further, $x_1(s) \leq \theta_m$ and $x_1(s) \geq 2\theta_m - x_1(s)$. Thus, $\lim_{s\downarrow\theta_m} x_1(s) = \theta_m$. Taking the limit for $s \downarrow \theta_m$ in (31) implies

$$-\gamma\left(1 - \frac{\beta}{2}\right)(\theta_m - \bar{m}_1)^{\gamma-1} + \lambda = 0, \tag{32}$$

which implies that $\lambda > 0$. Thus, there exists $\varepsilon > 0$ such that $\lambda > 0$ and the constraint therefore binds for $s \in [\theta_m, \theta_m + \varepsilon]$. Hence, $T(V)(s) = V(s)$ for all such s.

We have shown that the optimal response for all $s \in [\theta_m, \theta_m + \varepsilon]$ is $x_1(s) = 2\theta_m - x_2(s)$. However, if $x_2(s) > s$ this contradicts the necessary condition for an equilibrium that $x_1(s) \geq 2\theta_m - s$. Hence, $x_2(s) = s$ for all $s \in [\theta_m, \theta_m + \varepsilon]$, which contradicts the assumption of case 1.

Case 2: $\hat{s} > \theta_m$.

Let $C([\theta_m, \hat{s}])$ be the set of continuous functions on $[\theta_m, \hat{s}]$. We define T as in the existence proof, which establishes that there is an optimal response $x_1(s)$ which is strictly decreasing in s. This also implies that $x_1(s) < \theta_m$ for all $s < \theta_m$. Let s be marginally to the right of $\hat{s}$ with $x_2(s) > s$. Then next periods state $s' < \hat{s}$. However, because $x_1(s)$ is strictly decreasing, party 2 would be better off deviating to $x_2(s) = s$, a contradiction. Thus $x_2(s) = s$ for all $s \geq \theta_m$. Hence, the equilibrium is unique. ■

Proof of Proposition 5. The proof of lemma 2 shows that

$$T(V) = \max_{x \in [0,s]} -|x - \bar{m}_1|^\gamma + \beta V\left(\frac{x+s}{2}\right) \text{ s.t. } x \geq 2\theta_m - s. \tag{33}$$

is a contraction mapping. Define

$$V(s) = -|2\theta_m - s - \bar{m}_1|^\gamma - \frac{\beta |\theta_m - \bar{m}_1|^\gamma}{1-\beta}. \tag{34}$$

In order to determine $T(V)$, we substitute (34) into the maximization problem in (33), i.e.,

$$\max_{x \in [0,s]} -|x - \bar{m}_1|^\gamma - \left|2\theta_m - \frac{x+s}{2} - \bar{m}_1\right|^\gamma - \frac{\beta |\theta_m - \bar{m}_1|^\gamma}{1-\beta} \text{ s.t. } x \geq 2\theta_m - s. \tag{35}$$

The first order condition of (35) for $x > \bar{m}_1$ is

$$-\gamma(x - \bar{m}_1)^{\gamma-1} + \frac{\beta\gamma}{2}\left(2\theta_m - \frac{x+s}{2} - \bar{m}_1\right)^{\gamma-1} + \lambda = 0, \tag{36}$$

where λ is the Lagrange multiplier of the constraint. It follows immediately that the objective of (35) is strictly concave in x for $\gamma > 1$. Substituting $x = 2\theta_m - s$ into (36) implies that the constraint is slack if

$$-\gamma(2\theta_m - s - \bar{m}_1)^{\gamma-1} + \frac{\beta\gamma}{2}(\theta_m - \bar{m}_1)^{\gamma-1} \geq 0, \tag{37}$$

It follows immediately that the (37) holds with equality for $s = \bar{s}_1$, where $\bar{s}_1$, defined in (10), is the value of s at which (37) holds with equality. Clearly, $\bar{s}_1 > \theta_m$. Note that the left-hand side of (37) is strictly increasing in s. Thus, (37) is violated for all s with $\theta_m \leq s < \bar{s}_1$, i.e., the constraint binds and $x_1(s) = 2\theta_m - s$. Thus, $T(V)(s) = V(s)$ for all $s \in [\theta_m, \bar{s}_1]$. Similarly, it follows that $T^n(V)(s) = V(s)$ for all $s \in [\theta_m, \bar{s}_1]$

Let V^* be the fixed point of T. Because T is a contraction mapping and thus $\lim_{n\to\infty} T^n(V) = V^*$ it follows that $V^*(s) = V(s)$ for all $s \in [\theta_m, \bar{s}_1]$. Hence, in equilibrium it takes one period for the party cutoff to be $s = \theta_m$ if $s \in [\theta_m, \bar{s}_1]$.

Lemma 2 shows that in equilibrium $x_2(s) = s$ and that $x_1(s)$ is decreasing. Hence, $x_1(s) \leq 2\theta_m - \bar{s}_1$ for all $s \geq \bar{s}_1$. Let $s_1 = \bar{s}_1$ and define s_{i+1} by $(s_{i+1} + 2\theta_m - \bar{s}_1)/2 = s_i$. If $s \leq s_i$, it takes at most i periods until the party cutoff converges to θ_m. It follows that

$$s_i = \theta_m + \left(2^i - 1\right)(\bar{s}_1 - \theta_m). \tag{38}$$

(38) immediately implies the upper bound in (9).

Because the constraint $x \geq 2\theta_m - s$ binds for $s \in [\theta_m, \bar{s}_1]$, it follows that $x_1(s) = 2\theta_m - s$. It remains to determine $\bar{s}_2$ and $x_1(s)$ for $s \in [\bar{s}_1, \bar{s}_2]$. Consider again the first order condition of (36). For $s \geq \bar{s}_1$ the constraint is slack and $x_1(s)$ as defined in (12) solves the first order conditions. Because $-1 < x_1'(s) < 0$, there exists a unique value $\bar{s}_2 > \bar{s}_1$ such that $(x_1(\bar{s}_2) + \bar{s}_2)/2 = \bar{s}_1$. This value is given by (11). If $\bar{s}_1 < s \leq \bar{s}_2$ then it follows by construction that the party cutoff becomes θ_m in two periods. ■

Proof of Proposition 6. Consider the first case where the number of economic liberals outweigh social conservatives in all period. Suppose that party 1 selects candidates of type (L, L) at $t < \bar{t}$. Then party 2 is strictly better off nominating a candidate at (C, C) before party 1 switches to (L, C).

If party 1 selects a candidate (C, C) at $t < \bar{t}$ and party 2 selects (C, C), then the payoff to party 1's leaders is $-e \sum_{k=\bar{t}}^{\infty} \beta^{k-\hat{t}} = -e\beta^{\bar{t}-\hat{t}}/(-\beta)$. If, instead, party 1 nominates a social conservative at t, then the payoff is $-d \sum_{k=t}^{\infty} \beta^{k-\hat{t}} = -d/(1-\beta)$. Thus, selecting (L, L) at t is optimal for party 1 if $e\beta^{\bar{t}-t} < d$. Note that this inequality is weakest if t is minimal, i.e., $t = 1$. Thus, if $e\beta^{\bar{t}-1} \geq d$, then there exists an equilibrium in which party 1 selects (L, C) in the first period, and no party realignment occurs. Otherwise, if $e\beta^{\bar{t}-1} < d$ then it is always strictly better for party 1 to nominate a candidate (L, L) in some periods $t < \bar{t}$, and party 2 selects (C, C) at least in the last period where the opposing party chooses (L, L), leading to a party realignment.

Now suppose that there are more economic conservatives than economic liberals in periods $t \geq \bar{t}$. If no party realignment occurs, then the equilibrium in periods $t \geq \bar{t}$ is for party 1 to select a candidate (L, C) and for party 2 to select (C, C). Candidate (C, C) wins because there are more economic conservatives when $t \geq \bar{t}$.

If parties are realigned, and party 1 contains all social conservatives, then by the assumption of the proposition party 1's leader is (L, C). The resulting equilibrium is for party 1 to nominate a candidate (L, C) who would win in all periods $t \geq \bar{t}$. If party 2 contains all social conservatives, then the party leader (C, C) will select a candidate of type (C, C) and win.

Consider period $\bar{t} - 1$.

Suppose parties choose candidates (L, C) and (C, C). Then party 1 wins because there are more economic liberals in periods $t < \bar{t}$. If party 1 deviates to (L, L) then party 1 would win, because there are more social liberals in period $t < \bar{t}$, resulting in a utility increase of d. Parties realign, i.e., party 2 contains all social conservative, which means that (C, C) wins in all remaining periods, i.e., payoffs in periods $t \geq \bar{t}$ are unaffected by party 1's deviation. Thus, party 1's deviation to (L, L) is profitable, and (L, C) and (C, C) cannot be supported as an equilibrium at $\bar{t} - 1$.

Next, suppose that candidates are (L, L) and (C, L). Party 2's payoff remains unchanged when deviating to (C, C). In contrast, if party 1 deviates to (L, C), party 2 would win in the current period, but the winning candidate in all periods $t \geq \bar{t}$ would be changed from (C, C) to (L, C). Thus, this deviation by party 1, leading to a party realignment is optimal if

$$-1 - \frac{d\beta}{1-\beta} > -\frac{e\beta}{1-\beta}, \tag{39}$$

i.e., if $\beta > 1/(1 + e - d)$. As a consequence, if this inequality holds and parties do not realign, then the only candidate equilibrium at $\bar{t} - 1$ is (L, C), (C, C). However, it is optimal for party 1 to deviate to (L, L) leading to party realignment. Hence if (39) holds then a party realignment will take place at $\bar{t} - 1$.

Proceeding by backward induction consider a time period $\hat{t} < \bar{t}$. Then candidates selecting (L, C), (C, C) in all periods $\hat{t} \leq t < \bar{t}$ is not an equilibrium because party 1 would be better off deviating. Thus, suppose that parties select candidates at (L, L) and (C, L) until $\bar{t} - 1$. Party 1 is strictly better off deviating to (L, C) at $\hat{t}$ if

$$-\sum_{k=\hat{t}+1}^{\bar{t}-1} \beta^{k-\hat{t}} - \sum_{k=\bar{t}}^{\infty} d\beta^{k-\hat{t}} > -\sum_{k=\bar{t}}^{\infty} e\beta^{k-\hat{t}}. \tag{40}$$

One can check that if (40) holds then (39) must hold. Thus, if an equilibrium without realignment exists then (39) cannot hold. In this equilibrium both parties nominate candidates (L, L), (C, L) until period $\bar{t}$ when they both switch to (L, C) and (C, C), respectively.

Suppose that (39) holds but (40) is not satisfied for any $\hat{t} < \bar{t} - 1$. If no realignment of parties has occurred, then (L, L), (C, C) are the equilibrium strategies at $\bar{t} - 1$. If (40) holds for all $t \geq \hat{t}$ where $\hat{t} < \bar{t} - 1$ and no realignment has previously occurred then the equilibrium starting hat $\hat{t}$ is again (L, L), (C, C). Finally, if (39) does not hold, then candidate choices (L, L), (C, L) in periods $t < \bar{t}$ and (L, C), (C, C) in periods $t \geq \bar{t}$ are an equilibrium, i.e., no realignment occurs. ■

References

Baron, D. P. (1993). Government formation and endogenous parties. *American Political Science Review 87*(01), 34–47.

Baron, D. P. (1996). A dynamic theory of collective goods programs. *American Political Science Review 90*(2), 316–330.

Baron, D. P. and J. A. Ferejohn (1989). Bargaining in legislatures. *AmericanPolitical Science Review 83*(4), 1181–1206.

Berman, R. (2017). Republicans grapple with the risk of overreach. The Atlantic, `https://www.theatlantic.com/politics/archive/2017/01/the-era-of-republican-overreach-begins/512087/`.

Besley, T. and S. Coate (1997). An economic model of representative democracy. *Quarterly Journal of Economics 112*, 85–114.

Blair, T. (2015). Tony Blair: What I got right and labour now is getting wrong. The Spectator, `https://www.spectator.co.uk/2015/12/tony-blair-what -i-got-right-and-labour-now-is-getting-wrong`.

Buisseret, P. and D. Bernhardt (2017). Dynamics of policymaking: Stepping back to leap forward, stepping forward to keep back. forthcoming, American Journal of Political Science.

Buisseret, P. and R. V. Weelden (2017). Crashing the party? elites, outsiders, and elections. Working Paper, University of Chicago.

Caplin, A. and B. Nalebuff (1997). Competition among institutions. *Journal of Economic Theory 72*(2), 306–342.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy 67*(2), 135–150.

Duggan, J. (2012). Noisy stochastic games. *Econometrica 80*(5), 2017–2045.

Duggan, J. and T. Kalandrakis (2012). Dynamic legislative policy making. *Journal of Economic Theory 147*(5), 1653–1688.

Edwards, L. (2014). Barry M. Goldwater: The most consequential loser in american politics. First Principles, The Heritage Foundation `http://thf_media.s3.amazonaws.com/2014/pdf/MAPT11.pdf`.

Fiorina, M. P. (2016). The temptation to overreach. Essays on Contemporary American Politics Series 5, Hoover Institution.

Fudenberg, D. and J. Tirole (1991). *Game Theory, 1991*. Cambridge, Massachusetts.

Gomberg, A., F. Marhuenda, and I. Ortuño-Ortín (2004). A model of endogenous political party platforms. *Economic Theory 24*(2), 373–394.

Gomberg, A. M., F. Marhuenda, and I. Ortuño-Ortín (2016). Endogenous party platforms: stochastic membership. *Economic Theory 62*, 839–866.

Gul, F. and W. Pesendorfer (2016). Policy competition in real-time. Working Paper, Princeton University.

Harris, C., P. Reny, and A. Robson (1995). The existence of subgame-perfect equilibrium in continuous games with almost perfect information: A case for public randomization. *Econometrica 63*, 507–544.

Hulse, C. (2015). G.O.P. cheers a big victory. but has it stirred a hornet's nest? The New York Times: May 5, 2017.

Jaśkiewics, A. and A. S. Nowak (2016). Non-zero-sum stochastic games. unpublished manuscript.

Kalandrakis, T. (2004). A three-player dynamic majoritarian bargaining game. *Journal of Economic Theory 116*(2), 294–314.

Keen, R. and V. Apostolova (2017). Membership of UK political parties. House of Commons Library: Briefing Paper SN05125.

Key, V. O. (1955). A theory of critical elections. *The Journal of Politics 17*(1), 3–18.

Krasa, S. and M. Polborn (2010). The binary policy model. *Journal of Economic Theory 145*, 661–688.

Levy, G. (2004). A model of political parties. *Journal of Economic Theory 115*(2), 250–277.

Mertens, J. and T. Parthasarathy (2003). Equilibria for discounted stochastic games. In A. Neyman and S. Sorin (Eds.), *Stochastic Games and Applications*, pp. 131–172. Kluwer, Dordrecht.

Morelli, M. (2004). Party formation and policy outcomes under different electoral systems. *Review of Economic Studies 71*(3), 829–853.

Osborne, M. and A. Slivinski (1996). A model of political competition with citizen candidates. *Quarterly Journal of Economics 111*, 65–96.

Osborne, M. J. and R. Tourky (2008). Party formation in single-issue politics. *Journal of the European Economic Association 6*(5), 974–1005.

Poutvaara, P. (2003). Party platforms with endogenous party membership. *Public Choice 117*(1-2), 79–98.

Rieder, U. (1979). Equilibrium plans for non-zero-sum markov games. In O. Moeschlin and D. Pallaschke (Eds.), *Game theory and related topics*, Chapter 1, pp. 91–101. North-Holland Publishing Company.

Snyder, J. M. and M. M. Ting (2002). An informational rationale for political parties. *American Journal of Political Science 46*(1), 90–110.

Stokey, N. L., R. E. Lucas, and E. C. Prescott (1989). *Recursive Methods in Economic Dynamics*. Harvard University Press.

Stone, J. (2015). More people have joined labour since the election than are in the entire conservative party. The Independent, `http://www.independent.co.uk/news/uk/politics/more-people-have-joined-labour-since-the-election-than-are-in-the-entire-conservative-party-a6686001.html`.

Strom, K. (1990). A behavioral theory of competitive political parties. *American Journal of Political Science 34*(2), 565–598.

Wittman, D. A. (1973). Parties as utility maximizers. *American Political Science Review 67*(2), 490–498.