A Higher Calling: Career Concerns and the Number of Political Parties*

Nicolas Motz†

Department of Economics, Universidad Carlos III de Madrid

First Version: 10/2014 This Version: 08/2018

Abstract

In a number of US states, elections are dominated by a single party. Given this lack of competition, why are no third parties able to enter successfully? This paper proposes the career concerns of politicians as an explanation and demonstrates the underlying logic in a novel model of party formation: any two-party equilibrium requires that politicians care sufficiently strongly about career opportunities at the federal level. In addition, the paper provides existence conditions for a particular equilibrium with two parties, one centre-left and one centre-right.

Keywords: Political parties, Duverger's law, electoral competition.
JEL Classification: D72.

*An earlier version of this paper was circulated under the title "How Political Parties Shape Electoral Competition".

†nmotz@eco.uc3m.es; This paper is based on a chapter of my doctoral thesis submitted at University College London. I am grateful to my supervisors Ian Preston and Guy Laroque for their support. The paper also benefited greatly from the help of and discussions with Omer Ali, Philipp Denter, Boris Ginzburg, and Antoine Loeper. In addition, I would like to thank Marco Bassetto, Antonio Cabrales, Benoît Crutzen, Anders Jensen, Gilat Levy, Aureo de Paula, Michael Ting, and Lukas Wenner.

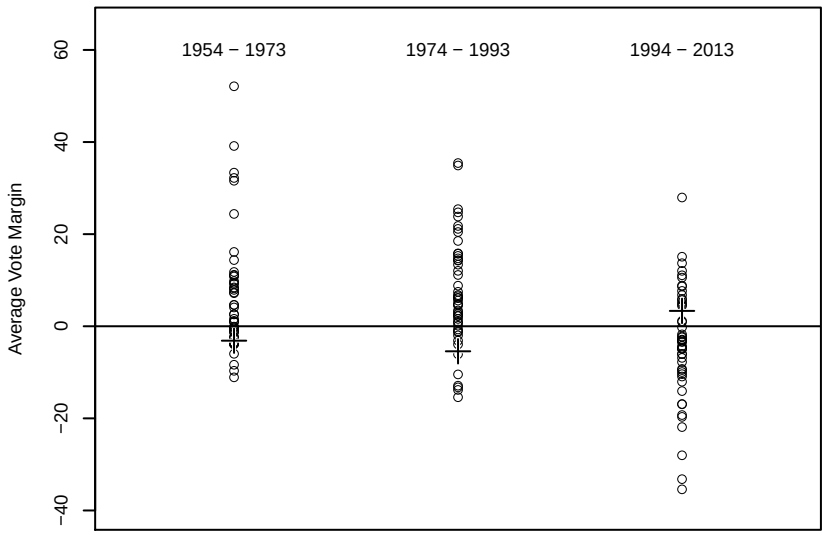
1 Introduction

Duverger's famous law states that first-past-the-post (FPTP) elections in combination with single member districts should lead to competition among two parties, and is frequently cited to explain the persistent dominance of the Democratic and the Republican Party in US politics. The logic underlying this claim is that losing parties will either be abandoned by voters or will decide by themselves to drop out of the race until only two parties remain. As has been recognized (Cox 1994), this line of reasoning applies to a single election, but not to elections held across separate districts or at multiple levels of government. Applying Duverger's law to the US, for example, we should expect to see two parties competing for the governorship of California and two parties competing for the Presidency, but there is no reason why the same two parties should be competing in both of these elections. In fact, it seems surprising that there are only two effective parties observable once one takes a look at some broad patterns in election outcomes. Figure 1 displays average differences in the vote share of Democratic and Republican candidates in presidential and gubernatorial elections across three 20-year periods.[1] The figure clearly shows that the outcomes of presidential elections are generally fairly close and do not consistently favour one party. Under these conditions the chances of any third party successfully contesting the presidential election indeed seem slim. In contrast, there is no lack of states where the candidates of one party consistently win elections with margins of victory of above 20 percent. In some extreme cases, margins of victory in gubernatorial elections approach 40 percent *on average*. Why are no third parties able to exploit this lack of competition?

In this paper I propose a new determinant of the number of parties that can explain the persistent duopoly in US politics, namely the career concerns of politicians: Federally successful parties can prevent members at the state level from defecting by offering career prospects at the federal level. I illustrate the strength of this logic in a novel model of party formation that puts no ad-hoc limits on the number of parties existing in equilibrium. The main result of the paper shows that equilibria with two parties exist only if politicians' career concerns are sufficiently strong. If politicians are mostly motivated by opportunities at the state level, in contrast, any equilibrium must feature three or more parties.

[1] Considering these elections has the advantage that they are not influenced by gerrymandering.

Figure 1: Average Differences in Vote Shares of Democratic and Republican Candidates



Notes: Each circle represents gubernatorial elections in a given state, while crosses stand for presidential elections. In the latter case, the numbers are based on popular vote shares.
Sources: Presidential elections - www.ropercenter.uconn.edu/elections/common/pop_vote.html; Gubernatorial elections up to 1990 - ICPSR (1995); Gubernatorial elections after 1990 - library.cqpress.com/elections/

In order to explain the logic underlying my results more clearly, let me first provide some details about the model. Given the question at hand, the model naturally features elections for state governments as well as a federal election. Candidates for all of these elections are nominated by political parties. Politicians standing at the beginning of their career join these parties in order to signal their policy preferences to voters. Parties enable politicians to do so by allowing only certain types of politicians to join. Parties thus serve as "informative labels" (Snyder & Ting 2002) that provide information about their members to voters. A politician who has won a state election then has a chance to become their party's candidate for the federal election.

A crucial feature of the model, which is also consistent with the data pre-

sented in Figure 1, is that there is a minimum amount of heterogeneity in voter preferences across states. This forces parties to adopt a broad ideological profile if they want to cater to voters' diverse tastes and prevent entry of additional parties. But if parties allow a wide range of politicians to join, this creates intense internal competition for the party nomination in the run-up to state elections. Politicians can be willing to put up with this competition if they see the state election as a mere stepping stone towards more attractive positions at the federal level. However, if politicians do not value such opportunities much, they will be willing to join smaller, more ideologically target parties that feature less internal competition. The key to the main result is then to show that any constellation of two parties is vulnerable to entry of such smaller parties, in which case two parties can only be an equilibrium if career concerns are sufficiently strong.

A second assumption that is crucial for the main result is that parties are not able to target the candidates they select at the state level with great precision to the preferences of voters. If such targeting was possible, the state politicians who appeal most strongly to the state median voter would be nominated with high probability and have no incentive to defect and join a smaller party, even if they had no aspirations towards higher offices. As a consequence a two-party equilibrium would always exist, independently of the strength of career concerns. However, the assumption that candidate selection at the state level is a noisy process is realistic. The potential candidates are largely unknown to voters, making it difficult for parties to commit to nominating a particular type of politician. Different factions within the party will be in disagreement about the ideal candidate. In addition, voters may be inattentive and base their expectations about the behaviour of a candidate partially on the observed behaviour of members of the same party in other states.

Beyond the general result that any two-party equilibrium requires sufficiently strong career concerns, I also provide existence conditions for a particular equilibrium with a centre-left and a centre-right party that looks very similar to what we observe in the US. In the context of the model, this equilibrium seems natural. A wider overlap among the sets of politicians allowed to join each party would create stronger internal competition and might lead to defections. A gap between parties, on the other hand, could create an opportunity for entry of a centrist party. In addition, this equilibrium recreates the pattern in Figure 1: States with extreme median voters vote overwhelmingly in favour of one party, resulting in wide vote margins. States with centrist median voters,

in contrast, are more competitive. The federal election, finally, is competitive as both parties generate symmetric candidate pools centred around the federal median voter. While a simpler model with a fixed number of parties might also be able to replicate this pattern in election results, such a model would clearly not be able to answer the deeper question about why we see the same parties competing across different geographic levels. The main contribution of the paper is therefore to establish the career concerns of politicians as an explanation for the lack of entry at the state level and as a driver of party formation more generally.

In line with the theoretical argument, being elected at the local or regional level appears to be a prerequisite for access to higher offices in the US. Diermeier et al. (2005) collect a sample of members of Congress in the period 1947 to 1994. They find that 78 percent of these politicians held a different local, state, or federal elected office before joining Congress. About ten percent of representatives in their sample run for a Senate seat. Of those who leave Congress, 35 percent stay in politics. In addition, political careers typically take place within one party and politicians who change their party affiliation are rarely observed (Yoshinaka 2015).

A concurrent and independent paper by Aldrich & Lee (2016) also highlights the importance of political ambitions in explaining why only two parties exist in the US. To make this point, these authors specify a utility function for politicians and explain how the utility of joining a party that offers the highest probability of winning a state election can be lower than joining a national party as long as the national party offers a sufficiently high probability of winning elections at the federal level. This utility function is not embedded in an equilibrium model and there is no explanation why the chances of winning the state election should be lower as a member of the national party in the first place. In my paper national parties are less attractive due to intense internal competition for nominations, which arises endogenously. In addition, my model makes clear that heterogeneity in voter preferences is a crucial ingredient: If the median voter had the same position in all elections, an equilibrium with a single party would exist even if politicians do not care about winning federal elections at all.

This paper is related to the literature on political competition with entry (Palfrey 1984, Osborne 1993, 2000, Callander 2005), which analyses the effect that the threat of entry has on the equilibrium behaviour of two parties. Perhaps closest to the current paper is Callander (2005), who studies competition between two parties in multiple single-member districts with threat of entry at

the district level. Parties, which are not explicitly modelled, are free to choose any platform. Callander (2005) finds that the threat of entry leads to the divergence of party platforms, similar to this paper. The mechanism through which entry is deterred is different though. In addition, the equilibrium presented by Callander requires very specific assumptions on the distribution of voters across districts, while the restrictions imposed on voter distributions in this paper are mild. Eyster & Kittsteiner (2007) also present a model that features multiple districts, but take the number of parties as fixed. Neither of these papers mentions career concerns.

Citizen candidate models as introduced by Osborne & Slivinski (1996) and Besley & Coate (1997) have previously been used to investigate the determinants of the number of parties competing in elections (See, for example, Dickson & Scheve 2010). In these models parties are identical to individual candidates. The current paper therefore requires a different approach, as parties have to be organisation that span multiple levels of government. Few papers have modelled parties as organisations comprising multiple politicians while endogenising the number of parties existing in equilibrium (Jackson & Moselle 2002, Levy 2004, Morelli 2004, Osborne & Tourky 2008, Eguia 2011) and I am the first to do so employing the concept of parties as informative labels.[2] My model is tractable and naturally offers itself for the purpose of investigating other questions, such as the relationship between social diversity and the number of political parties (Dickson & Scheve 2010, Milazzo et al. 2018). Indeed, the results presented here suggest that a theoretical analysis of the number of parties competing in a particular district may be misleading if linkages across levels of government are not taken into account.

The rest of the paper is organized as follows: Section 2 explains the details of the model, while Section 3 gives the theoretical results. Robustness of the results to relaxing some of the assumptions made in the basic version of the model is discussed in Section 4. Section 5 concludes.

2 The Model

A federal state consisting of $S \geq 4$ states selects federal and state governments through FPTP elections. Candidates for these elections are nominated by po-

[2] In contrast, Snyder & Ting (2002) as well as other contributions building on their approach (Ashworth & Bueno de Mesquita 2008, Bernhardt et al. 2009) consider the behaviour of a given number of parties.

litical parties. Initially a large number of parties exists, but only those that attract members can compete in elections. The timing is as follows: In the beginning of the game politicians decide which party to join. Once affiliation decisions have been made, parties nominate candidates in each state and state elections are held. Each winner of a state election then has a chance to become their party's candidate for the federal election. After the federal election the game ends. The following sections describe the elements of the model in detail.

2.1 Policy Space

The policy space is given by a grid of real numbers denoted by $\mathcal{T} \subset [-2, 2]$ such that $\{-1, 0, 1\} \subseteq \mathcal{T}$. Policies are evenly distributed such that there is the same distance between any two adjacent policies and the set $\mathcal{T}$ is symmetric around zero. For example, if there were five policies the two possible cases are $\mathcal{T} = \{-2, -1, 0, 1, 2\}$ and $\mathcal{T} = \{-1, -0.5, 0, 0.5, 1\}$.

2.2 Players

The strategic players of the game are voters and politicians.

2.2.1 Voters

Each state $s \in \{1, ..., S\}$ contains a set of voters that is large, finite, and odd. Voters are identified by their ideal policy $i \in \mathcal{T}$. The set of voters in state s is described by a measure V_s that assigns to any subset of $\mathbb{R}$ the number of voters whose ideal policies lie in this subset. Let m_s denote the ideal policy of the median voter of state s. Similarly, let V_f be the measure of voters participating in the federal election with median m_f where

$$V_f(D) = \sum_{s=1}^{S} V_s(D)$$

for any $D \subseteq \mathbb{R}$. It is assumed that m_f is equal to zero. It will often be important to know what share of voters in some region $r \in \{1, .., S, f\}$ is located in some interval $[a, b]$. I will therefore define

$$\Lambda_r([a, b]) \equiv \frac{V_r([a, b])}{V_r(\mathbb{R})} \ .$$

Apart from the normalisation $m_f = 0$ introduced in the previous paragraph,

the only assumptions imposed on the set of voters specify that there is some minimum amount of heterogeneity in voter distributions across states: let there be at least one state s such that $m_s \leq -1$, at least one state s' such that $m_{s'} = 0$ and $\Lambda_{s'}((-0.5, 0.5)) > 0.5$, and at least one state s'' such that $m_{s''} \geq 1$. As the labels of states are arbitrary it is without loss of generality to denote these states as states 1, 2, and 3, respectively. Assuming the existence of some states with relatively extreme voter preferences increases the incentives of politicians to join parties targeted at particular states and therefore ensures that the model provides a non-trivial answer to the question of why such parties fail to compete successfully in reality.

2.2.2 Politicians

Each state s has a finite set of politicians. Every politician is endowed with a platform, which is an element of the policy space $\mathcal{T}$. Each state has $|\mathcal{T}|$ politicians, none of which share the same platform. Put differently, there is one politician located at each possible platform in each state.

2.3 Political Parties

Parties are not strategic actors, but form part of the environment. Following Snyder & Ting (2002), a party is basically a subset of the policy space and only politicians whose platforms fall within this subset can join. One way to think of this is that parties can screen their members and exclude those whose platforms do not agree with the party line. As will become clear in the following sections, observing that a politician is a member of a certain party allows voters to update their beliefs about the platform of the politician. Formally, the set of politicians who can join a party is given by an interval $[a, b]$ with $\{a, b\} \subset \mathcal{T}$. If $a = b$ I simply write $[a]$. Individual parties will be denoted by capital letters and for any such party P the interval representing the party is given by I_P. Let $\mathcal{I} = \{[a, b] | a, b \in \mathcal{T}\}$ be the collection of all possible shapes parties can have. The set of parties that exists in the beginning of the game is denoted by $\mathcal{P}$. Since any party can only compete in elections if it is joined by at least one politician, the set $\mathcal{P}$ is referred to as the set of potential parties. Any party that does attract members is referred to as an active party. The set of potential parties will be described in more detail below.

In addition to the set of politicians that is allowed to join a particular party, it also needs to be specified how parties nominate candidates for elections. At

the state level, any politician who has joined a party within a state is eligible to become the party's candidate for the state election and each one of them is nominated with equal probability. Denoting by $\mathcal{M}_{P,s}$ the set of politicians who have joined party P in state s, each member of $\mathcal{M}_{P,s}$ is thus nominated with probability $|\mathcal{M}_{P,s}|^{-1}$.

The mechanism for candidate selection at the federal level is modelled in a very general way. Denote by $\mathcal{M}_{P,f}$ the set of politicians who have won a state election as a member of party P. The probability that a politician with platform p who belongs to $\mathcal{M}_{P,f}$ is nominated is given by a function $\eta_P(p|\mathcal{M}_{P,f})$. The only restrictions placed on this function are those needed to ensure that η_P yields well-defined probabilities, namely that the codomain of η_P is equal to $[0,1]$ and that for any $\mathcal{M}_{P,f}$ the nomination probabilities across all members of $\mathcal{M}_{P,f}$ sum to one.

2.4 Timing

The game proceeds as follows: In a first step, all politicians decide whether to join a party or to stay passive. After politicians have made their affiliation decisions, any party nominates a candidate in any state where it has been joined by at least one politician. Once candidates have been nominated, each voter casts a vote at the election in their state and the winner is the candidate who receives the highest number of votes with ties broken randomly. Winners of state elections then implement their platform as the state policy. Subsequently, each party that has won at least one state election nominates one of their winning candidates as their candidate for the federal election. All voters vote in the federal election. The winner is again the candidate who receives the highest number of voters, who then implements their platform as the federal policy.[3]

2.5 Information

A crucial feature of the concept of political parties employed here is that voters have limited information about politicians. Specifically, it is assumed that the electorate cannot distinguish between different politicians at the beginning of

[3]Even in a system of FPTP elections the implementation of policies requires a majority in parliament. With more than two parties competing the choice of policy may therefore require a process of coalition formation. I abstract from such issues here. At least the two-party equilibria presented below do not depend on what is assumed about the process of policy formation when no party achieves a majority. This is because voters will be allowed to vote strategically, which implies that there always exists a voting equilibrium with one party winning with a strict majority, even off the equilibrium path when a third party has entered.

the game and only knows how their platforms are distributed. As there is one politician for each possible platform in each state, the prior belief of voters over the platform of a randomly selected politician thus assigns probability $|\mathcal{T}|^{-1}$ to each platform. Furthermore, voters can see which parties have nominated a candidate in their state, but not how many politicians have joined each party. Voters do know, however, how candidates are selected. This knowledge combined with a belief about which politicians have joined a particular party allows voters to update their beliefs about the platform of a party's candidate prior to casting their vote at the state-level election. Suppose, for example, that a candidate in a certain state is a member of a party of shape $[0.5, 1]$ and voters believe that only a politician with platform 1 has joined this party. Voters will consequently believe that the candidate of the party must have platform 1. If the electorate instead believes that two politicians with platforms 0.5 and 1 have joined, they are aware that either of these will be nominated with probability one-half. Accordingly their belief over the platform of the candidate of the party will assign probability one-half to each of the platforms 0.5 and 1.

The winner of a state election implements her platform at the state level, thus revealing it to voters. Voters accordingly have full information about candidates at the federal level. All agents are also fully informed about the distribution of voters in all states and at the federal level.

2.6 Strategies

Voters are not allowed to abstain from voting and the action space of a voter in the election in their state and in the federal election is therefore equivalent to the set of nominated candidates in each respective election. The only strategic choice that politicians make is their choice of party affiliation. As parties can only compete in elections if they attract members, the affiliation choices of politicians determine the number of effective parties that is formed. In order to not put any ad-hoc restrictions on the possible constellations of parties that can emerge, it is important to let politicians choose from a sufficiently broad set of parties. I therefore assume that the set of potential parties $\mathcal{P}$ is "large". In particular, for any possible shape $I \in \mathcal{I}$ there exists at least one party $P \in \mathcal{P}$ such that $I_P = I$.

Denote by $\mathcal{P}(p)$ the set of potential parties that allow politicians with plat-

form p to join. Formally,

$$\mathcal{P}(p) \equiv \{P \in \mathcal{P} |\ p \in I_P\}\ .$$

The action space of a politician with platform p is given by $\mathcal{P}(p) \cup \{\emptyset\}$, where $\emptyset$ represents the choice not to join a party.

2.7 Payoffs

Let p_s and p_f denote the policies that are implemented in state s and at the federal level, respectively. Given beliefs over the platforms of nominated candidates and the behaviour of other voters, the objective of a voter with ideal policy $i \in \mathbb{R}$ in an election in region $r \in \{1, ..., S, f\}$ is to maximize

$$\mathbb{E}[u(|p_r - i|)]\ ,$$

where $u : \mathbb{R}_+ \to \mathbb{R}$ is continuous, decreasing, and concave.[4]

The payoffs of politicians are determined by the their individual electoral success as well as that of their party. The utility of a passive politician is normalised to zero. The utility of a politician who joins a party that does not win a single election is equal to $y_z < 0$. If a party wins at least one election, each member receives a payoff $y_w > 0$. These payoffs can be thought of as the psychological costs or benefits of being on a losing/winning team and may be arbitrarily close to zero. Of much greater importance for the analysis are the payoffs associated with personally winning elected office. The winning candidate in an election at the state level receives a payoff of $y_s > 0$, while the utility of the winning candidate at the federal election further increases by $y_f > 0$. These payoffs subsume the material and immaterial benefits of holding office. For example, purely local concerns that might motivate a politician could form part of y_s. The strength of politicians' career concerns is captured by the ratio of the payoffs y_f and y_s. The larger this ratio, the more politicians are driven by the pursuit of success at the federal level.

In order to clearly define the utility of a politician who joins a party that wins at least one election, let π_s be the probability that a politician is nominated for and wins the election in her state. Conditional on doing so, let π_f give the

[4] Elections at the state level determine who becomes a candidate at the federal level, but it is assumed that voters do not take this interdependence into account when voting in a state election.

probability that a politician is nominated for and wins the federal election. Both these probabilities are determined in equilibrium. The expected utility of a politician who has joined a party that wins at least one election is then given by

$$y_w + \pi_s(y_s + \pi_f y_f) \ .$$

2.8 Equilibrium

The party-formation game described in the previous sections features incomplete information and requires a corresponding equilibrium concept. I will focus on the set of sequential equilibria. Without further restrictions, this choice entails a huge number of equilibria due to the fact that voters are allowed to vote strategically instead of assumed to vote sincerely (See, for example, the discussion in Acemoglu et al. 2009). In order to rule out at least some of the most implausible equilibria, I impose the following restriction: if a candidate in some election is the unique most preferred option of a strict majority of voters based on their beliefs at the point when the election is held, then a voting equilibrium where this candidate wins the election is selected. While such an equilibrium always exists under the stated conditions, there are typically additional equilibria where a different candidate gets elected. Nevertheless, it seems plausible that voters will be able to coordinate on electing a candidate who is favoured by a strict majority.[5]

I focus on pure strategy equilibria. The following definition summarises the equilibrium concept:

Definition 1 (Party-Formation Equilibrium)**.** *A party-formation equilibrium is a sequential equilibrium of the party-formation game in pure strategies that satisfies the following condition: If a candidate in some election is the unique most preferred option of a strict majority of voters based on their beliefs at the point when the election is held, then this candidate wins the election.*

Equilibrium objects are indicated by stars. In particular, $\mathcal{P}^*$ will denote the set of active parties in equilibrium, while $N^* \equiv |\mathcal{P}^*|$. The expected platform of the candidate of party P in state s in equilibrium is given by $\bar{p}^*_{P,s}$. Conditional on a particular equilibrium it is also possible to compute the unconditional probability that a politician with platform p belonging to party P will be nominated

[5] For example, Myatt (2007) models a three-candidate election as global game. While coordination failure is generally a feature of equilibrium in this model, a candidate favoured by a majority of voters wins with certainty.

for the federal election:

$$\eta_P^*(p) = E_{\mathcal{M}_{P,f}^*}\left[\eta_P(p|\mathcal{M}_{P,f}^*)\right] \ .$$

Finally, denote by $\omega_{P,f}^*(p)$ the probability that party P wins the federal election in equilibrium if it nominates a politician with platform p.

2.9 Discussion

In the following I will discuss some of the features of the model. The assumption that politicians are endowed with a fixed platform that determines their policy choices is not standard. One possible interpretation is that the platform represents the ideal policy of the politician. A number of empirical studies indicates that preferences over policies are the main driver of the choices that politicians make in office (Levitt 1996, Chattopadhyay & Duflo 2004, Lee et al. 2004, Bhalotra & Clots-Figueras 2014). Of course it would be preferable that this behaviour emerges as part of an equilibrium, rather than being imposed from the outset. I will allow politicians to be more flexible in their policy choices in Section 4.2.

Another feature of the model that certainly deserves to be discussed in more detail is the process of candidate selection. As I argued in the introduction, the assumption that candidate nomination at the state level is a noisy process is crucial but also realistic. However, the more specific assumption that candidates are selected uniformly at random is needed for tractability. While the assumption of random candidate selection also features in previous contributions (Snyder & Ting 2002, Ashworth & Bueno de Mesquita 2008, Bernhardt et al. 2009), I consider the possibility of relaxing this assumption in Section 4.1. In contrast, very little is assumed about the functions that determine candidate selection at the federal level. One possibility is that

$$\eta_P(p|\mathcal{M}_{P,f}) = 1/|\mathcal{M}_{P,f}| \ ,$$

in which case nominees for the federal election are chosen at random just like at the state level. Another possibility is that parties always nominate one of the politicians whose platforms are located closest to the federal median voter. This latter case might seem like the most natural choice, but may actually lead to costly defections by members who stand little chance of being nominated, as will become clear below. The party thus has an incentive to ensure that such politicians are nominated with sufficiently high probability. It would be

interesting to make the nomination mechanism a strategic choice, but this is beyond the scope of this paper.

Finally, it would also be possible to introduce an electoral college at the federal level to fit the model more closely to the US. In this case the results go through unchanged if the median voter of the state with the median electoral vote is assumed to be located at zero.[6]

3 Results

The central insight of this paper is to show how the number of parties in a system of FPTP elections depends on the career concerns of politicians. In particular, the main result demonstrates that two parties jointly dominating elections across all levels of government is a possible outcome only if politicians care sufficiently strongly about winning elections at the federal level. To gain some intuition for why this is the case, note that the politicians in the model are generally happier the fewer members their party has: a higher number of members entails more competition for the party nomination. Suppose for a moment that all that politicians cared about was being elected at the state level. In this case, if any politicians had a chance of joining a party with fewer members that nevertheless allows them to win the state election, surely they would take it. An equilibrium where only two parties attract members could then exist only if these parties are positioned in a way that makes it impossible for third-party candidates to win any elections. Two features of the model aid parties in doing so: First of all, coordination failure among voters was not ruled out and this can make entry of third parties difficult. Secondly, the types of politicians who join a party are not necessarily the same across all states. This enables parties to have a different ideological profile in different states. Nevertheless, Proposition 1 below demonstrates that in any two-party equilibrium there are politicians who could compete successfully at the state level after joining a smaller, more ideologically targeted third party. As a consequence, any two-party equilibrium ceases to exist as the payoff from winning a state election becomes large relative to the payoff of winning the federal election (Proposition 2).

In deriving the results, it will be useful to be able to state concisely that

[6]The median electoral vote can be calculated as follows: Create a distribution of electoral votes by assigning the electoral college votes of the state to the ideal policy of the median voter of the state. Then find the median of this distribution. When there are two parties competing at the federal election, the party closest to the median voter of the state with the median electoral vote wins a majority of electoral votes.

a politician has the opportunity to win a state election by joining a previously passive party. I will say that such a politician can contest the state election.

Definition 2 (Contestable States)**.** *Consider any equilibrium. Let $\mathcal{C}_{p,s}$ denote the set of ideal points $i \in \mathcal{T}$ such that*

$$u(|p-i|) \; > \sum_{p' \in \mathcal{M}^*_{P,s} \setminus p} \frac{1}{|\mathcal{M}^*_{P,s} \setminus p|} \; u(|p'-i|) \tag{1}$$

is satisfied for all $P \in \mathcal{P}^*$ such that $\mathcal{M}^*_{P,s} \setminus p \neq \emptyset$. It is said that a politician with platform p in state s can contest the state election if

$$\Lambda_s(\mathcal{C}_{p,s}) > 0.5 \; .$$

A state is contestable in an equilibrium if there exists a politician who can contest the state election.

Should a politician with platform p deviate and join a party P' with shape $[p]$, then she is strictly preferred over any other candidate in the state election by any voter whose ideal point i satisfies Condition (1).[7] If these voters form a strict majority of the state electorate, then the restrictions on equilibrium require that the deviating politician wins the state election.[8]

I will now present two lemmas that partially characterize equilibrium behaviour. The first one gives necessary conditions for the behaviour of politicians to be consistent with equilibrium.

Lemma 1. *Consider any equilibrium. Then*

i. any politician who is eligible to join a party that wins at least one state election must have done so, while all other parties attract no members. Furthermore,

ii. any politician in any state s who is eligible to join the party that wins the election in state s must have done so if $N^* = 2$.

[7] Lemma 3 in the appendix implies that in this case voters would hold correct beliefs about which parties politicians have joined and accordingly equation (1) is appropriate to determine that a voter strictly prefers the candidate of party P' over any other candidate.

[8] If the conditions stated in Definition 2 fail, then it is possible to construct the equilibrium such that no politician can contest the state in question. This is true since a violation of the stated conditions implies that there exists an equilibrium of the subgame reached after the deviation such that the deviating politician does not win the state election. In particular, there exists an equilibrium where all voters vote for one particular other candidate, since there is no strict majority in favour of the deviating politician.

The first part of Lemma 1 is a simple consequence of the payoffs y_z and y_w, which imply that politicians are better off as a member of a party that wins at least one election compared to staying passive, which in turn is preferred to being a member of a party that doesn't win any elections. Similarly, politicians prefer to be a member of the party that wins the election in *their* state over being a member of a party that wins an election in *some* state, as any member of a party in a state has the chance of being nominated for the state election. This logic underpins the second part of the lemma.[9]

An implication of the preceding lemma is that a party that is successful in a state will always feature internal competition for the party nomination in that state. The only exception is when a party allows only one type of politician to join. However, when there are only two parties at least one of them must allow more than one type politician to join; otherwise the heterogeneity in voter preferences across states would allow a third party to enter successfully. This is demonstrated by the following lemma:

Lemma 2. *In any equilibrium it cannot be the case that* $I_P \cap I = \emptyset$ *for all* $P \in \mathcal{P}^*$ *and for any* $I \in \{[-2,-1],(-1,1),[1,2]\}$.

In words, Lemma 2 states that in any equilibrium there must be at least one active party that allows at least some left-wing politicians with a platform no larger than -1 to join, at least one active party that allows some politicians with a platform somewhere between -1 and 1 to join, as well as at least one active party that allows at least some right-wing politicians with a platform no smaller than 1 to join. This result is a consequence of the assumed heterogeneity in voter preferences across states, which forces parties to cater sufficiently to the preferences of voters in different parts of the policy space. Suppose, for example, that none of the active parties allowed politicians with a platform no larger than -1 to join. Then the politician with platform -1 in a state with a very left-leaning median voter could join a party instead of remaining passive. If this party has a narrow ideological profile, voters believe that the party's candidate is more left-wing than any competitor. The newly-formed party would therefore win the state election.

In combination, Lemmas 1 and 2 imply that in any two-party equilibrium there must be at least one party where members are competing for the nomina-

[9] The reason that part *ii* of Lemma 1 only applies to two-party equilibria is that deviations by politicians can potentially affect which party wins the election. As the proof of Lemma 1 shows, such changes in the outcome of the election can be ruled out as long as there are only two active parties.

tion at the state level and each individual member is thus less than certain to win the state election. A politician who can contest a state election by joining a previously passive party with less internal competition may thus be tempted to do so. As it turns out, in any two-party equilibrium there are at least some politicians who have this opportunity.

Proposition 1. *In any equilibrium such that $N^* = 2$ there exists at least one state that is contestable.*

Parties do not necessarily have the same set of members across all states, even in equilibrium. As the proof of Proposition 1 shows, the heterogeneity in voter distributions across states nevertheless makes it impossible for two parties to appeal to all state electorates to the same extent. There will thus always be some politicians in any two-party equilibrium who could deviate and be successful at the state level as a member of a third party. If these politicians care strongly about being elected at the state level relative to career opportunities at the federal level this deviation will be profitable and any constellation of two parties is not stable. This logic yields the following proposition, which is the main result of the paper.

Proposition 2. *For any constellation of two parties $\{A, B\} \subset \mathcal{P}$, there exists a constant $\bar{y} > 0$ such that an equilibrium in which $\mathcal{P}^* = \{A, B\}$ exists only if $y_f/y_s \geq \bar{y}$.*

The previous result shows that two-party equilibria can only exist if politicians care sufficiently strongly about career prospects at the federal level. However, it still remains to be shown that a two-party equilibrium exists at all. I will therefore now derive sufficient conditions for the existence of an equilibrium in which two parties, L and R, are formed along the equilibrium path, where party L allows politicians located to the left of the federal median voter to join, while party R admits only politicians to the right of the federal median voter. This constellation of parties closely resembles that observable in the US.

Constructing such an equilibrium requires parties to extend their membership far enough to the extremes such that any politician who can contest a state election is able to join a party. Otherwise an additional party would certainly enter. For any state s such that $m_s \leq 0$, denote by $p_{L,s}$ the smallest $x \in \mathcal{T}$ such that

$$u(|x - m_s|) \;\geq \sum_{p \in (x,0] \cap \mathcal{T}} \frac{1}{|(x,0] \cap \mathcal{T}|}\; u(|p - m_s|)\;.$$

That is, $p_{L,s}$ denotes the platform of the most left-wing politician that is preferred by the median voter of state s to the nominee of party L in case that party L allows all politicians between $p_{L,s}$ and 0 to join. As u is decreasing, it must be true that $p_{L,s}$ falls in between $2m_s$ and the smallest element of $\mathcal{T}$ that is larger than m_s. Define

$$p_L \equiv \min_{s \ s.t. \ m_s \leq 0} p_{L,s} \ .$$

If party L has the shape $[p_L, 0]$, no party located to the left of party L can enter and successfully contest a state election. Similarly, for any state s such that $m_s \geq 0$, let $p_{R,s}$ equal the largest $x \in \mathcal{T}$ such that

$$u(|x - m_s|) \ \geq \sum_{p \in [0,x) \cap \mathcal{T}} \frac{1}{|[0,x) \cap \mathcal{T}|} \ u(|p - m_s|)$$

and define

$$p_R \equiv \max_{s \ s.t. \ m_s \geq 0} p_{R,s} \ .$$

Proposition 3. *An equilibrium of the party formation game where $\mathcal{P}^* = \{L, R\}$, with $I_L = [p_L, 0]$ and $I_R = [0, p_R]$, exists if*

i. for each party $P \in \{L, R\}$ there exist at least two states such that the median voter of the state weakly prefers party P over party $-P \in \{L, R\} \backslash P$ when all eligible politicians have joined party P and among the remaining politicians all those eligible have joined party $-P$; and

ii. for any politician p in some state s who joins party $P \in \{L, R\}$ in equilibrium the following conditions are satisfied:

a) If party P wins in state s and politician p can contest the election then

$$\Lambda_f \left(\left[\frac{p_L + p}{2}, \frac{p + p_R}{2} \right] \right) \leq 0.5 \ .$$

b) If party P wins in state s and politician p can contest the election then

$$\frac{y_f}{y_s} \geq \frac{1 - \frac{1}{|\mathcal{M}_{P,s}|}}{\frac{1}{|\mathcal{M}_{P,s}|} \eta_P^*(p) \ \omega_{P,f}^*(p)} \ . \tag{2}$$

c) If party P does not win the election in state s then politician p cannot contest the election.

Conditions ii.b and ii.c are also necessary for the existence of such an equilibrium.

I will refer to the equilibrium in the preceding proposition as the L-R equilibrium. Condition i guarantees that the equilibrium can be constructed such that each party wins at least two state elections. This has the consequence that whenever a third party enters and wins a state election, no less than three parties compete at the federal election. Given that this is the case, Condition $ii.a$ implies that there would not be a strict majority of federal voters in favour of the candidate of the newly formed party, even if parties L and R both nominate their most extreme member. The only benefit that newly formed parties can offer their members under these conditions is that they may enable politicians to win the state election with higher probability due to lower internal competition for the party nomination. Parties L and R, on the other hand, offer career prospects at the federal level. If politicians value such opportunities sufficiently strongly—as expressed in condition $ii.b$ of Proposition 3—entrant parties are unable to attract members. Condition $ii.b$ is a reflection of Proposition 2, which states that the existence of any two-party equilibrium requires sufficiently strong career concerns. Inequality (2) is easier to interpret if rewritten as

$$\frac{1}{|\mathcal{M}_{P,s}|}\ \eta_P^*(p)\ \omega_{P,f}^*(p)\ y_f \geq \left(1 - \frac{1}{|\mathcal{M}_{P,s}|}\right)\ y_s\ . \tag{3}$$

Suppose that a politician is a member of party L in equilibrium in a state where party L wins. Then the left-hand side of Inequality (3) captures the expected payoff from winning the federal election for this politician. By deviating and joining a third party, the politician misses out on this payoff. The resulting drop in utility must be as least as large as the value of the increase in the probability of winning the state election, which is given by the right-hand side of Inequality (3). Otherwise the deviation would be profitable and the L-R equilibrium fails to exist.

As can be seen from condition $ii.b$ of Proposition 3, the ability of parties to prevent their members from defecting depends on the nomination technology η_P used at the federal level. Parties can increase their chances of winning the federal election by nominating centrist politicians with high probability. The lower the probability that extremist politicians are nominated, however, the more likely they are to join a third party. If they do so, this would be highly problematic for the party they are leaving behind. Suppose, for example, that

some members of party L defect and form a more left-wing party. This can lead to a split in the left-wing vote, handing victory at the federal election to party R. Party L would therefore prefer to grant extremist politicians a somewhat higher probability of nomination if this maintains the unity of the party.

Finally, condition $ii.c$ of Proposition 3 is a necessary condition for any equilibrium: if a politician does not have any chance of winning the election in their state, the same must be the case after a change in party affiliation.

It is noteworthy that the L-R-equilibrium is able to reproduce the pattern in US election results presented in the introduction: States with extreme median voters will display large majorities in favour of one of the parties, while in states with median voters close to zero the margin of victory will be small. In addition, the federal election will be competitive, particularly if neither party extends much further to the extremes than the other and both of them use a similar nomination technology for the federal election, as is the case in the US. The model thus not only provides an explanation for the absence of successful third parties in the US, but is also able to match broad patterns in election results.

I will conclude this section by discussing other types of equilibria. If career concerns are not sufficiently strong and the L-R equilibrium does not exist, other two-party equilibria might. Excluding some centrist politicians from both parties would lower internal competition and make extremist politicians less likely to defect. If the gap between parties becomes too large, however, centrist politicians not affiliated to any party will become able to contest elections in centrist states, upsetting the equilibrium. On the other hand, a two-party equilibrium with stronger overlap than in the L-R equilibrium might also exist if politicians care strongly about success at the federal level. Such overlap can be hard to maintain in equilibrium though: it creates the possibility that both parties nominate politicians for the federal election who are located on the same side of the federal median voter. This would make it relatively easy for voters to coordinate on electing a third-party candidate. For example, a centrist candidate of a third party can attract all voters to left of the centre if the remaining candidates are both located on the right. But if centrist politicians can do well both in state elections and in the federal election after joining a third party, nothing can prevent them from defecting.

For similar reasons an equilibrium with only one party is very unlikely to exist. If only a single party is nominating a candidate for the federal election, there is effectively no coordination problem for voters even if the candidate of a

second party enters the race. Unless the equilibrium party is itself very likely to nominate a candidate with platform 0, centrist politicians stand a good chance of winning the federal election after joining a second party. And this probability need not be large in order to make the deviation attractive: A single party can only win all elections if it is sufficiently broad to allow even extreme politicians to join, who would otherwise be able to contest elections in states with similarly extreme median voters. As a result, there is intense internal competition for nominations. Centrist politicians are therefore likely to be better off as members of a different party. This is captured by the following proposition:

Proposition 4. *Let $\tilde{\eta}_P(0)$ denote the probability that party P nominates a politician with platform 0 for the federal election if the party wins every state election with exception of the election in state 2. An equilibrium such that $N^* = 1$ does not exist if for any party $P \in \mathcal{P}$ such that $[-1,1] \subseteq I_P$ it is true that*

$$\frac{1}{|\mathcal{T} \cap [-1,1]|} \;\leq\; 1 - \tilde{\eta}_P(0) \;. \tag{4}$$

To understand the condition provided in Proposition 4, note that the probability of any politician winning the federal election is limited by their chance of winning a state election in the first place. This latter probability depends on how many members the party has. Suppose there was only a single active party in equilibrium denoted by P. Then at least all politicians in the interval $[-1,1]$ must be allowed to join party P by Lemma 4. The left-hand side of Inequality (4) is therefore an upper bound on the probability of any single politician winning the federal election in equilibrium. Now recall that state 2 was assumed to have a centrist median voter. When only a single broad party is active in equilibrium, the politician with platform 0 in state 2 can contest the state election by joining a second party with a sufficiently narrow ideological profile. In the resulting situation party P wins all state elections with exception of the election in state 2 and $\tilde{\eta}_P(0)$ gives the probability that party P will nominate a politician with platform 0. Whenever party P does *not* nominate a politician with platform 0, then the deviating politician wins the federal election. The right-hand side of Condition (4) is accordingly a lower bound on the probability that the deviating politician wins the federal election, and the condition as a whole is sufficient to ensure that the deviation is profitable. While a tighter (if less intuitive) sufficient condition for the non-existence of a one-party equilibrium can be derived, Condition (4) is easily satisfied. If there are only three platforms in the interval $[-1,1]$, internal competition is at its lowest possible

level. But even then a single party would need to nominate centrists with a probability of at least two thirds out of equilibrium to prevent defections. A one-party equilibrium is therefore highly unlikely to exist.[10] This final result is noteworthy as there are no democratic countries where one party dominates all levels of government.[11]

4 Robustness

The basic model of party formation presented above requires a number of simplifying assumptions for tractability. This section will discuss some of these in more detail.

4.1 Candidate Selection and Mixed Strategies

It was assumed that any member of a party in a state is nominated with equal probability for the state election. Generally, assuming that candidate selection at the state level is a fairly noisy process is realistic. The potential candidates are largely unknown to voters, making it difficult for parties to commit to nominating a particular type of politician. Furthermore, different factions within the party will be in disagreement about the ideal candidate. While the party leadership at the national level has an interest in supporting moderates who will later on make suitable candidates for federal offices, party activists within a state will likely be pushing for more extreme nominees. Furthermore, ignoring parts of the party membership when deciding the nomination will likely lead to defections.

It therefore seems doubtful that parties will be able to target their candidates with great precision at the median voters of different states. And small differences in the expected platforms of candidates of the same party in different states are unlikely to upset the results. While this would make it easier for two parties to fend off entry of a third one, other assumptions already favour incumbents over entrants. In particular, the bar for voters to coordinate on electing a third-party candidate was set relatively high. In addition, I only assume a minimum amount of heterogeneity in voter tastes across states. A small increase

[10] Even if there is no noise in the federal nomination technology, the nomination of a politician with platform 0 is not guaranteed. This is true because no potential candidate with platform 0 is available if no such politician has won a state election.

[11] It should be clear from the discussion above that the existence of a one-party equilibrium depends on the career concerns of politicians just as in the case of equilibria with two parties.

in the ability of parties to differentiate could be countered by a small increase in the assumed amount of heterogeneity and Proposition 1 survives.

Mixed strategy equilibria may also enable parties to have a different ideological profile in different states since politicians may be playing different mixed strategies in different states. The same arguments apply as in the previous paragraph. Furthermore, mixed strategies can only have a substantial impact on the expected platform of candidates in cases where parties overlap strongly. While this is difficult to show in general, it seems unlikely that two broad and largely overlapping parties could preclude entry of a third party even if mixed strategies are taken into consideration.

4.2 Policy Choices

The assumption that politicians are committed to implementing their platform is not satisfying. While the empirical literature quoted above seems to suggest that policy preferences of politicians are the main driver of their choices in office, it would be more appealing to see this behaviour emerge as part of an equilibrium rather than imposing it from the outset. In the model, extremist politicians can often increase their chances of winning the federal election by pretending to be a centrist when choosing state policies. To address this concern I will briefly consider a more general utility function for politicians that includes both career concerns and policy preferences. For a politician with ideal policy i who joins a party that wins at least one election let the utility function now be given by

$$y_w \ + \ \pi_s(y_s + \pi_f y_f) \ - \ \alpha \sum_{l \in \{s,f\}} u(|p_l - i|) \ ,$$

where α measure the the relative weight that politicians attach to policy and the notation is otherwise the same as in Section 2. Parties then allow only politicians with certain ideal policies to join. In addition, assume that politicians can freely choose the policy they implement at any stage. All other elements of the game remain unchanged. This more general version of the model is challenging to solve in its entirety. However, focusing on the subgame reached after state elections have been held, it is clear that a separating equilibrium exists where politicians implement their ideal policy at the state level if α is sufficiently large: Given that politicians behave in this way, some politicians might be able to increase their chances of being elected at the federal election by choosing a different policy after winning a state election. In an equilibrium where everyone behaves

truthfully, voters will then expect this politician to implement the same policy if elected federally and might be more likely to vote for her. The gain in utility associated with this increase in the likelihood of winning the federal election is clearly finite though. If α is sufficiently large, the loss in utility associated with implementing a less-than-ideal policy at the state level will weigh more heavily and the deviation is not profitable.

Even if such a separating equilibrium does not exist, however, behaviour doesn't necessarily change drastically. The party that a politician belongs to puts limits on the ideal policy that a politician can have and therefore also on the beliefs that voters can form about this ideal policy. An equilibrium in this case might see all winners of state elections pool on the most centrist policy that a member of their party is allowed to pursue. In some cases, such as in the L-R equilibrium presented above, this can mean that all politicians chose the same policy at the state level (the policy 0 in the case of the L-R equilibrium). Voters would then nevertheless benefit from voting for parties positioned closely to them as this can pay off if one of their members gets elected at the federal election.

5 Conclusion

Why are the same two parties competing in elections in the US across all levels of government, while parties in other countries relying on FPTP elections are much less integrated between the national and the regional level? In this paper I have highlighted the career concerns of politicians as a possible explanation. In my model, heterogeneity in voter tastes across states forces two parties to adopt a broad ideological profile. This leads to intense internal competition for nominations. Joining a smaller party more targeted at the preferences of voters in a particular state would allow some politicians to win the state election with higher probability. The drawback of this move is that politicians miss out on the career opportunities that federally successful parties offer. As a consequence, two-party equilibria exist only if politicians value such opportunities sufficiently strongly.

In addition, I provide existence conditions for a particular equilibrium with two parties that looks very similar to what we observe in the US. In this equilibrium both parties allow centrist politicians to join, while one party extends its membership far enough to the left to prevent entry of a left-wing party and

the other party does the same on the opposite end of the political spectrum. As a consequence, this equilibrium reproduces the pattern in election results presented in Figure 1 in the introduction.

In applying the model to the US, one possible reservation may be that the office of the President is the only office with a truly national electorate. Indeed it seems unlikely that the chance of becoming President will be a variable that exerts a strong influence over the choices of politicians running for, say, a seat in a state legislature. However, I believe that this interpretation is too narrow. The federal election in the model should be thought of as representing all opportunities that are available to members of a federally successful party, such as becoming member of the federal cabinet or floor leader in Congress. Even being a backbencher in the House of Representatives is more attractive as a member of the majority party than as a member of a small minority party. And while individual seats in the US Congress are decided by the votes of a fraction of all citizens, the majorities in the Senate and the House of Representatives do depend on the electorate as a whole.

While this paper has focused on FPTP elections, a similar logic should also apply to countries using proportional representation, such as Austria or Germany. Both countries feature two main parties that traditionally (if less so recently) receive the vast majority of votes. Importantly, this is true federally as well as at the state level. It thus seems that the major parties allow for and attract a membership that is ideologically broad enough to ensure a strong position across states. Preventing entry of any additional parties, in contrast, would require extending the party membership far to the extremes under proportional representation and could be too costly in terms of votes lost at the federal level and in more moderate states. Career concerns would then again be an important factor in that they prevent fringe parties from luring politicians away from the major parties. The technical difficulties involved in modelling systems of proportional representation make this a challenging subject for further research.

Appendix: Proofs

The following three lemmas are not presented in the text:

Lemma 3. *Consider an equilibrium of the party-formation game and suppose that a politician with platform p in some state s deviates and joins a party P with shape $I_P = [p]$. Then voters' belief at the information set reached must place full probability on politician p in state s having joined party P while the behaviour of all other politicians has not changed.*

Proof. Index politicians by $j \in \mathcal{J} = \{1, ..., S \cdot |\mathcal{T}|\}$ and denote by p_j the platform of politician j. Let $\mathcal{N}$ be a node of some information set $\tilde{\mathcal{N}}$ reached after politicians have made their affiliation choices, that is, $\mathcal{N}$ assigns one element of $\mathcal{P}(p_j) \cup \{\emptyset\}$ to each politician j. $\mathcal{N}(j)$ is the action assigned to politician j at the node $\mathcal{N}$. Let $\sigma_j : \mathcal{P}(p_j) \cup \{\emptyset\} \to [0,1]^{|\mathcal{P}(p_j)|}$ describe a strategy of a politician while σ is a vector of strategies for all politicians.

Let $\mathcal{N}^*$ be the node of information set $\tilde{\mathcal{N}}^*$ reached along the equilibrium path of some pure-strategy equilibrium where each politician j uses strategy σ_j^*. Sequential equilibrium requires that beliefs are consistent with equilibrium play. In addition, there must be a sequence of fully mixed strategies σ^n such that $\sigma^n \to \sigma^*$ and the beliefs of players are a limit of the sequence of beliefs generated by σ^n (see Fudenberg & Tirole 1991, pp. 337). Sequential equilibrium therefore implies

$$\lim_{n\to\infty} \frac{\prod_{j\in\mathcal{J}} \sigma_j^n(\mathcal{N}^*(j))}{\sum_{\mathcal{N}\in\tilde{\mathcal{N}}^*} \prod_{j\in\mathcal{J}} \sigma_j^n(\mathcal{N}(j))} = \frac{\prod_{j\in\mathcal{J}} \sigma_j^*(\mathcal{N}^*(j))}{\sum_{\mathcal{N}\in\tilde{\mathcal{N}}^*} \prod_{j\in\mathcal{J}} \sigma_j^*(\mathcal{N}(j))} = 1 \ , \tag{5}$$

where the second equality holds as σ^* is a vector of pure strategies. More specifically, the numerator and denominator in first expression in equation (5) must both approach one as n gets large. Furthermore,

$$\lim_{n\to\infty} \prod_{j\in\mathcal{J}\setminus k} \sigma_j^n(\mathcal{N}^*(j)) = 1 \quad \forall\ k \in \mathcal{J} \ . \tag{6}$$

Now suppose that some politician k in some state s deviates from σ_k^* and joins a party P' with shape $[p_k]$. Voters then observe that party P' nominates a candidate in state s. Let $\tilde{\mathcal{N}}'$ be the information set reached after the deviation. As politician k is the only politician who can join party P' in state s, $\mathcal{N}'(k) = P'$ for all $\mathcal{N}' \in \tilde{\mathcal{N}}'$. The belief that some node $\mathcal{N}'$ of information set $\tilde{\mathcal{N}}'$ has been reached in the sequential equilibrium under consideration can therefore be written as

$$\lim_{n\to\infty} \frac{\prod_{j\in\mathcal{J}\setminus k} \sigma_j^n(\mathcal{N}'(j))}{\sum_{\mathcal{N}\in\tilde{\mathcal{N}}'} \prod_{j\in\mathcal{J}\setminus k} \sigma_j^n(\mathcal{N}(j))} \ . \tag{7}$$

Now let $\mathcal{N}'$ be the node actually reached after the deviation of politician k. Then it

is true that $\mathcal{N}'(j) = \mathcal{N}^*(j)$ for all $j \neq k$. This, together with equation (6), implies

$$\lim_{n\to\infty} \prod_{j\in\mathcal{J}\setminus k} \sigma_j^n(\mathcal{N}'(j)) = \lim_{n\to\infty} \prod_{j\in\mathcal{J}\setminus k} \sigma_j^n(\mathcal{N}^*(j)) = 1 \ ,$$

which shows that the limit in expression (7) must also be equal to one in this case. In words, the belief of voters at information set $\tilde{\mathcal{N}}'$ must assign probability one to the node where politician k has joined party P' and all other politicians behave as in equilibrium. □

Lemma 4. *In any equilibrium*

i. state 1 is contestable unless there exists at least one party P such that $\bar{p}^*_{P,1} \leq -1$,

ii. state 2 is contestable unless there exists at least one party P such that $\bar{p}^*_{P,2} \in (-1, 1)$, and

iii. state 3 is contestable unless there exists at least one party P such that $\bar{p}^*_{P,3} \geq 1$.

Proof. It will be shown that state 1 is contestable if $\bar{p}^*_{P,1} > -1$ for all $P \in \mathcal{P}^*$. Assume this condition was satisfied and suppose the politician with platform -1 in state 1 joins some party D with shape $[-1]$. According to Lemma 3, voters must then believe that politician -1 has deviated while all other players still follow their equilibrium strategies. But then the expected platform of the candidate of party D is equal to -1 with certainty while the expected platforms of all other parties' candidates must be greater than -1 even after the deviation. Accordingly the median voter and all voters located further to the left must strictly prefer the candidate of party D and party D wins the state election. Similarly, state 2 is contestable by a politician with platform 0 if $\bar{p}^*_{P,2} \notin (-1, 1)$ for all $P \in \mathcal{P}^*$ as $\Lambda_2((-0.5, 0.5)) > 0.5$, while state 3 is contestable by a politician with platform 1 if $\bar{p}^*_{P,3} < 1$ for all $P \in \mathcal{P}^*$. □

Lemma 5. *Consider any equilibrium such that $N^* = 2$ and denote the parties active in equilibrium by A and B. If the sets $I_A \setminus I_B$ and $I_B \setminus I_A$ are both non-empty and convex, then in any state election along the equilibrium path*

i. the party strictly preferred by the median voter of the state must win if such a party exists, and

ii. either party winning is consistent with equilibrium if the median voter is indifferent between both parties.

Proof. Consider the election in some state s. If the conditions in the statement of the lemma are satisfied, Lemma 1 implies that along the equilibrium path any politician with some platform p in state s has joined party $P \in \{A, B\}$ if and only if p belongs to some interval $\tilde{I}_P \subseteq I_P$ with $\tilde{I}_A \cap \tilde{I}_B = \emptyset$. Denote by μ_P the midpoint between the

two most extreme platforms belonging to $\tilde{I}_P$. Let

$$U_P(i) = \sum_{p \in \mathcal{T} \cap \tilde{I}_P} \frac{1}{|\mathcal{T} \cap \tilde{I}_P|} u(|p - i|) \ .$$

Since voters hold correct beliefs about which politician has joined which party along the equilibrium path, $U_P(i)$ is the expected utility of a voter with ideal policy i if party P wins the state election. Given that $u(|p - i|)$ is a concave function of i, the same is true for $U_P(i)$. In addition, $U_P(i)$ is symmetric around the midpoint μ_P of the set of platforms of the party's members in the state, given that this set itself is symmetric. Combined with concavity this implies that the unique maximum of $U_P(i)$ is attained for $i = \mu_P$. Assume without loss of generality that $\mu_A < \mu_B$. For any $i \leq \mu_A$, any member of party B is further away from i than any member of party A. It follows that $U_A(i) > U_B(i)$ for any such i. Equivalently, $U_B(i) > U_A(i)$ for any $i \geq \mu_B$. Combined with the fact that both functions are continuous and monotone in the interval $[\mu_A, \mu_B]$, it follows that there exists a unique i such that $U_A(i) = U_B(i)$. As a consequence, a strict majority of voters must strictly prefer the party preferred by the median voter whenever such a party exists. According to the definition of equilibrium, this party must therefore win the state election. If the median voter is indifferent, in contrast, there is no strict majority in favour of either party. Either party winning is then consistent with equilibrium. For example, all voters voting in favour of party A or party B are both possibilities. □

Proof of Lemma 1. Suppose there was an equilibrium such that some party P wins at least one election and that there is a politician with platform p in state s such that $p \in I_P$ who does not join a party in equilibrium. This politician therefore receives a payoff of zero. If the politician joins party P instead, she receives a positive payoff unless party P wins no elections other than in state s in equilibrium and loses even this election after the deviation. But if party P wins the election in state s in equilibrium then it must also have some members in state s in equilibrium. As voters only observe whether a party is campaigning at the state level, but not how many politicians have joined, voters therefore do not detect that a deviation has taken place and party P continues to win the state election. Hence, the deviation is profitable. On the other hand, any politician who joins a party that does not win any elections receives a negative payoff by assumption and would prefer not to join a party.

To show the final part of the statement, assume that some party P wins the election in some state s and there is a politician with platform p in state s such that $p \in I_P$ who does not join party P in equilibrium. If the politician joins party P, voters can only observe that a deviation has taken place if the politician was a member of a different party before and this party has no other members in the state after the deviation. But in any two-party equilibrium this would imply that party P is the only remaining

party in the state after the deviation, as it was argued above that parties that do not win elections have no members in equilibrium. Party P will therefore always continue to win the sate election after the deviation and as each member of a party at the state level is nominated with positive probability, the deviation is profitable, contradicting equilibrium. □

Proof of Lemma 2. Suppose $I_P \cap (-1,1) = \emptyset$ for all $P \in \mathcal{P}^*$. The expected platform of any nominee therefore lies outside of $(-1,1)$ and Lemma 4 accordingly implies that politician 0 in state 2 is able to contest the state election. But $I_P \cap (-1,1) = \emptyset$ for all $P \in \mathcal{P}^*$ requires that politician 0 in state 2 has not joined a party in equilibrium and this politician therefore has a profitable deviation. The remaining cases can be shown analogously. □

Proof of Proposition 1. Consider an equilibrium such that $\mathcal{P}^* = \{A, B\}$. Denote by $\mathcal{M}_P$ the set of politicians in any given state that is allowed to join party $P \in \{A, B\}$ and let $\mathcal{M}_{P\setminus -P} = \mathcal{M}_P \setminus \mathcal{M}_{-P}$ with $P \in \{A, B\}$ and $-P \in \{A, B\} \setminus P$. Let $\bar{p}_P(\mathcal{M})$ be the expected candidate of party P in any state where the set of politicians $\mathcal{M}$ has joined party P, that is

$$\bar{p}_P(\mathcal{M}) \equiv \sum_{p \in \mathcal{M}} \frac{1}{|\mathcal{M}|}\ p\ .$$

Now assume that in equilibrium no state was contestable. By Lemma 1, it must be true in equilibrium that

$$\bar{p}^*_{P,s} \in \{\bar{p}_P(\mathcal{M}_P), \bar{p}_P(\mathcal{M}_{P\setminus -P})\} \quad \forall\ s \in \{1, ..., S\}, P \in \{A, B\}\ . \tag{8}$$

The two parties combined can thus generate up to four different expected platforms. According to Lemma 4, one of these must lie weakly below -1, one in the interval $(-1,1)$, and one weakly above 1. Without loss of generality, assume that $\bar{p}_A(\mathcal{M}_A) \leq \bar{p}_B(\mathcal{M}_B)$. Together with the fact that candidates are nominated uniformly at random this implies

$$\bar{p}_A(\mathcal{M}_{A\setminus B})\ \leq\ \bar{p}_A(\mathcal{M}_A)\ \leq\ \bar{p}_B(\mathcal{M}_B)\ \leq\ \bar{p}_B(\mathcal{M}_{B\setminus A})\ , \tag{9}$$

as long as all of these quantities are defined.[12] The restrictions in (9) combined with the requirements of Lemma 4 immediately rule out a number of possible cases. For example, if party A wins in state 1 and $\bar{p}_A(\mathcal{M}_A) \leq -1$ and party B wins in state 3 and $\bar{p}_B(\mathcal{M}_B) \geq 1$, then condition (9) implies that there can be no candidate in state 2 whose expected platform lies in the interval $(-1,1)$, as required by Lemma 4.

[12] As the expected platform of party A with full membership is weakly smaller than the expected platform of party B with full membership, removing anyone who is able to join party B from party A must remove more members above the expected platform of party A (if any). Thus it must hold that $\bar{p}_A(\mathcal{M}_{A\setminus B}) \leq \bar{p}_A(\mathcal{M}_A)$.

Three possible cases are less straightforward to rule out: *i*) party A wins in state 1 and state 3 and $\bar{p}_A(\mathcal{M}_A) < -1$ and $\bar{p}_B(\mathcal{M}_{B\setminus A}) > 1$, *ii*) party B wins in state 1 and $\bar{p}_A(\mathcal{M}_{A\setminus B}) < -1$ while party A wins in state 3 and $\bar{p}_B(\mathcal{M}_{B\setminus A}) > 1$, or *iii*) party B wins in state 1 and state 3 and $\bar{p}_A(\mathcal{M}_{A\setminus B}) < -1$ and $\bar{p}_B(\mathcal{M}_B) > 1$. It will be shown that all of these constellations are impossible. Consider case *i*. As candidates are selected uniformly at random, $\bar{p}_A(\mathcal{M}_A) < -1$ requires $I_A \subseteq [-2, 0]$. Similarly, it must be true that $\mathcal{M}_{B\setminus A} \subseteq [0, 2]$. But then any possible candidate of party B in state 3 is located closer to the median voter of state 3 than any possible candidate of party A, contradicting that party A wins the state election in equilibrium. Next consider case *ii*. $\bar{p}_A(\mathcal{M}_{A\setminus B})$ and $\bar{p}_B(\mathcal{M}_{B\setminus A})$ both exist only if neither party is a subset of the other. $\bar{p}_A(\mathcal{M}_{A\setminus B}) < -1$ requires $\mathcal{M}_{A\setminus B} \subseteq [-2, 0]$. The utility of the median voter of state 1 if party A wins the state election must therefore be greater than $u(|0 - m_1|)$ as $m_1 \leq -1$ and the worst case from the perspective of the median voter would therefore be that the candidate of party A has platform 0. The median voter can then only prefer party B with full membership if $\bar{p}_B(\mathcal{M}_B) < 0$ as voters are risk averse. The assumed situation in state 3 implies analogously that it must be true that $\bar{p}_A(\mathcal{M}_A) > 0$. But this contradicts the assumption that $\bar{p}_A(\mathcal{M}_A) \leq \bar{p}_B(\mathcal{M}_B)$. Finally, case *iii* can be shown to yield a contradiction analogously to case *i*. □

Proof of Proposition 2. Consider an equilibrium such that $\mathcal{P}^* = \mathcal{P}_2$ and suppose there is a politician in a state s with platform p who is able to contest a state election. Suppose that in equilibrium politician p is a member of party A. If party A allows more than one type of politician to join then the equilibrium payoff of politician p is bounded from above by

$$y_w + \frac{1}{2}(y_s + y_f) \ .$$

If politician p deviates and joins a party D that allows only politicians with platform p to join, then she achieves at least a payoff of $y_w + y_s$ as it was assumed that the politician can contest the state election. If $y_s > y_f$, then politician p in state s strictly prefers joining party D, contradicting equilibrium.

As at least one state is contestable in any two-party equilibrium by Proposition 1, the only possibility of a two-party equilibrium that exists for any value of $y_f/y_s > 0$ would then be that one of the parties allows only one type of politician to join. Suppose such an equilibrium existed and denote the parties that attract members in equilibrium as A and B. Let A be the party that allows only one type of politician to join. By Lemma 2, party B must then allow more than one type of politician to join. It will be shown that at least one politician who is not a member of party A can contest a state election. Suppose $I_A \subset [-2, -1]$. Then state 3 is contestable unless $\bar{p}^*_{B,3} > 1$ by Lemma 4. This requires $I_B \subseteq [0, 2]$. But in this case parties A and B do not overlap, so $\bar{p}^*_{B,s} > 1$ must hold across all states, and accordingly the politician with platform zero in state 2 must be able to contest the state election, again using Lemma

4. The case $I_A \subset [1,2]$ can be ruled out analogously. The remaining case to consider is $I_A \subset (-1,1)$. In this case Lemma 4 requires that either $\bar{p}_B(\mathcal{M}_B) < -1$ and $\bar{p}_B(\mathcal{M}_{B\setminus A}) > 1$ or the reverse inequalities hold (using the notation defined in the proof of Proposition 1). But as $\mathcal{M}_B$ and $\mathcal{M}_{B\setminus A}$ differ by only one politician with a platform $p \in (-1,1)$, this is impossible. □

Proof of Proposition 3. If $\mathcal{P}* = \{L, R\}$, equilibrium behaviour can be characterised as follows: By Lemma 1 all eligible politicians in each state must have joined either party L or party R, politicians with platform zero must have joined the party that wins the election in their state, and all remaining politicians must remain passive. Given this behaviour of politicians, denote by $\mathcal{S}_P$ the set of states such that the median voter weakly prefers party P if the politician with platform 0 of the state has joined party P. Each state must belong to at least one of the sets $\mathcal{S}_L$ and $\mathcal{S}_R$. To see this, note that the expected utility of the median voter when party P gets elected and has been joined by the politician with platform 0 can be written as

$$\alpha\ U_{P-0}(m_s)\ +\ (1-\alpha)\ u(|m_s|)$$

where α is some probability and $U_{P-0}(m_s)$ is the expected utility of the median voter when party P gets elected and has not been joined by the politician with platform zero. That

$$\alpha\ U_{L-0}(m_s)\ +\ (1-\alpha)\ u(|m_s|)\ <\ U_{R-0}(m_s)$$

and

$$\alpha\ U_{R-0}(m_s)\ +\ (1-\alpha)\ u(|m_s|)\ <\ U_{L-0}(m_s)$$

hold simultaneously is only possible if $u(|m_s|) < U_{L-0}(m_s), U_{R-0}(m_s)$. But this is impossible since $u(|m_s|) > U_{R-0}(m_s)\ \forall m_s \leq 0$ and $u(|m_s|) > U_{L-0}(m_s)\ \forall m_s \geq 0$. Given that every state belongs to $\mathcal{S}_L \cup \mathcal{S}_R$, Lemma 5 implies that party P winning the election in state s is consistent with equilibrium if and only if $s \in \mathcal{S}_P$.

It will be established that the equilibrium can be constructed such that each party wins at least two state elections if Condition i in the statement of the proposition is satisfied. Select two states $s \in \mathcal{S}_L$, giving priority to states that do not belong to $\mathcal{S}_R$. Under condition i this is possible. Since $S \geq 4$, this leaves at least two states unassigned. Suppose fewer than two of these remaining states belong to $\mathcal{S}_R$. Since each state must belong to $\mathcal{S}_L \cup \mathcal{S}_R$, this implies that there is one unassigned state that only belongs to $\mathcal{S}_L$. Given that priority was given to states that only belong to $\mathcal{S}_L$ in the initial step, the two states selected then do not belong to $\mathcal{S}_R$ either. But this contradicts that $|\mathcal{S}_R| \geq 2$. It is thus possible to construct the equilibrium such that either party wins at least two state elections.

Given these preliminaries, it remains to be established that no politician has an incentive to deviate. Politicians who are unable to join either of the equilibrium parties

cannot win the election in their state by joining an additional party. This is true by construction as parties extend from p_L to p_R. Politicians who can join party L or party R but are unable to win the state election if nominated would be better off if they could join some party that enables them to contest the state election. Consequently, this must be impossible, making condition *ii.c* a necessary condition. Finally, consider politicians whose party wins the state election in their state and who can successfully contest the state election by joining some third party D. As such a politician is the only member of party D at the point of deviating, she will be nominated for the federal election with certainty and face competing candidates from parties L and R, given that both those parties win at least two state elections in equilibrium. If condition *ii.a* is satisfied, there is never a strict majority of voters in favour of the candidate of party D and accordingly there exists a voting equilibrium such that this candidate loses the federal election. The highest payoff a politician can achieve by deviating to joining party D is therefore $y_w + y_s$. The equilibrium payoff of such a politician who is a member of party $P \in \{A, B\}$ on the other hand is

$$y_w + \frac{1}{|\mathcal{M}^*_{P,s}|}\left[y_s + \eta^*_P(p)\ \omega^*_{P,f}(p)\ y_f\right] \ .$$

The inequality that states that the equilibrium payoff must be at least as large as the deviation payoff is equivalent to condition *ii.b*. If a politician can contest a state election, $y_w + y_s$ is also the lowest possible payoff that this politician can achieve by deviating, implying that condition *ii.b* is a necessary condition for equilibrium existence. □

Proof of Proposition 4. Suppose one party wins all elections and call this party A. By Lemma 2, $[-1, 1] \subset I_A$ and combined with Lemma 1 this implies that party A has at least $|\mathcal{T} \cap [-1, 1]|$ members in any state while all other parties have no members. A politician with platform 0 therefore achieves a payoff in equilibrium that is no greater than

$$y_w + \frac{1}{|\mathcal{T} \cap [-1, 1]|}(y_s + y_f) \ . \tag{10}$$

Denote by $D \in \mathcal{P}$ a party such that $I_D = [0]$. If the politician with platform 0 joins party D in state 2, voters observe that party D starts campaigning in the state and therefore know that the only eligible politician must have joined while all other politicians have not changed their behaviour by Lemma 3. As the expected platform of the candidate of party D is then equal to the ideal policy of the state median voter with certainty while the platform of the candidate of party A has positive variance, the median voter strictly prefers the candidate of party D and so do all voters either to the left or the right of the median voter. Party D must accordingly win the state election. Politician 0 from state 2 is then nominated with certainty as party D's candidate for the federal election. As only two parties are competing in the federal election, party

D must win with certainty whenever party A nominates a candidate with platform unequal to 0, which is the federal median voter's ideal policy. Denote by $\tilde{\eta}_A(0)$ the probability that party A nominates a politician with platform 0 in this situation. The payoff of politician 0 in state 2 from the deviation is thus at least equal to

$$y_w + y_s + (1 - \tilde{\eta}_A(0))\ y_f\ ,$$

which is greater than expression (10) for any possible values of y_s and y_f if

$$1 - \tilde{\eta}_A(0)\ \geq\ \frac{1}{|\mathcal{T} \cap [-1, 1]|}\ .$$

If this last inequality is satisfied for any party $P \in \mathcal{P}$ such that $[-1, 1] \subseteq I_P$, no equilibrium such that $N* = 1$ exists. □

References

Acemoglu, D., Egorov, G. & Sonin, K. (2009), Equilibrium refinement in dynamic voting games, Mimeo.

Aldrich, J. & Lee, D. (2016), 'Why two parties? Ambition, policy, and the presidency', *Political Science Research and Methods* **4**(2), 275–292.

Ashworth, S. & Bueno de Mesquita, E. (2008), 'Informative party labels with institutional and electoral variation', *Journal of Theoretical Politics* **20**(3), 251–273.

Bernhardt, D., Campuzano, L., Squintani, F. & Câmara, O. (2009), 'On the benefits of party competition', *Games and Economic Behavior* **66**(2), 685–707.

Besley, T. & Coate, S. (1997), 'An economic model of representative democracy', *The Quarterly Journal of Economics* **112**(1), 85–114.

Bhalotra, S. & Clots-Figueras, I. (2014), 'Health and the political agency of women', *American Economic Journal: Economic Policy* **6**(2), 164–97.

Callander, S. (2005), 'Electoral competition in heterogeneous districts.', *Journal of Political Economy* **113**(5), 1116–1145.

Chattopadhyay, R. & Duflo, E. (2004), 'Women as policy makers: Evidence from a randomized policy experiment in India', *Econometrica* **72**(5), 1409–1443.

Cox, G. (1994), 'Strategic voting equilibria under the single nontransferable vote', *The American Political Science Review* **88**(3), 608–621.

Dickson, E. & Scheve, K. (2010), 'Social identity, electoral institutions and the number of candidates', *British Journal of Political Science* **40**(2), 349–375.

Diermeier, D., Keane, M. & Merlo, A. (2005), 'A political economy model of congressional careers', *American Economic Review* **95**(1), 347–373.

Eguia, J. (2011), 'Voting blocs, party discipline and party formation', *Games and Economic Behavior* **73**(1), 111–135.

Eyster, E. & Kittsteiner, T. (2007), 'Party platforms in electoral competition with heterogeneous constituencies', *Theoretical Economics* **2**(1), 41–70.

Fudenberg, D. & Tirole, J. (1991), *Game Theory*, MIT Press, Cambridge, MA.

ICPSR (1995), Candidate name and constituency totals, 1788-1990, Dataset ICPSR00002-v5, Inter-University Consortium for Political and Social Research.

Jackson, M. & Moselle, B. (2002), 'Coalition and party formation in a legislative voting game', *Journal of Economic Theory* **103**(1), 49–87.

Lee, D., Moretti, E. & Butler, M. (2004), 'Do voters affect or elect policies? Evidence from the U.S. House', *The Quarterly Journal of Economics* **119**(3), 807–859.

Levitt, S. (1996), 'How do senators vote? Disentangling the role of voter preferences, party affiliation, and senator ideology', *American Economic Review* **86**(3), 425–41.

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

Milazzo, C., Moser, R. & Scheiner, E. (2018), 'Social diversity affects the number of parties even under first-past-the-post rules', *Comparative Political Studies* **51**(7), 938–974.

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

Myatt, D. (2007), 'On the theory of strategic voting', *The Review of Economic Studies* **74**(1), 255–281.

Osborne, M. (1993), 'Candidate positioning and entry in a political competition', *Games and Economic Behavior* **5**(1), 133–151.

Osborne, M. (2000), 'Entry-deterring policy differentiation by electoral candidates', *Mathematical Social Sciences* **40**(1), 41–62.

Osborne, M. & Slivinski, A. (1996), 'A model of political competition with citizen-candidates', *The Quarterly Journal of Economics* **111**(1), 65–96.

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

Palfrey, T. (1984), 'Spatial equilibrium with entry', *The Review of Economic Studies* **51**(1), 139–56.

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

Yoshinaka, A. (2015), *Crossing the Aisle: Party Switching by U.S. Legislators in the Postwar Era*, Cambridge University Press, New York.