Ambitious Politicians, Radical Voters and the Threat of Secessionism: On the Conditions for Two-Party Dominance

Nicolas Motz*

Universidad Carlos III de Madrid

December 2020

Abstract

It is widely accepted that first-past-the-post elections induce a tendency towards only two parties competing in a given election. Less clear are the conditions required for two parties to dominate all elections across separate districts and at different levels of government, as is the case in the US for example. In this paper, I propose a novel model of party formation in which politicians use parties as a vehicle to establish a public profile in the early stages of their career, but may later on abandon the party and run as independents. I then use the model to derive three necessary conditions for the existence of equilibria in which only two parties form and no independent candidates run: *i*) there cannot be parties that span the entire ideological spectrum, *ii*) there cannot be districts with radical median voters relative to the distribution of voters at the national level, and *iii*) politicians need to be sufficiently motivated by the desire to win elections at higher levels of government. In addition, I show through an example of a two-party equilibrium that these necessary conditions are also sufficient for existence. An extension that introduces regionalism shows that high salience of this second dimension of policy is by itself not enough to rule out two-party equilibria.

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

*nicolas.motz@uc3m.es, Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, Getafe, 28903, Spain. An earlier version of this paper formed 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, Zoë Kuehn, Gilat Levy, Aureo de Paula, Michael Ting, and Lukas Wenner. Declarations of interest: none.

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 parties with no realistic chance of winning are either abandoned by voters or decide to drop out of the race until only two parties remain. However, this line of reasoning applies to a single election, but not to elections held across separate districts or at multiple levels of government (see Cox 1994). 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. Similarly, two parties should be contesting each individual seat in Congress, but more than two parties may be represented in Congress overall. In fact, one does not have to look far to find political systems relying on FPTP elections where more than two parties attract significant vote shares: The Houses of Commons of the UK and Canada each currently feature four parties represented by more than ten MPs. National politics in India, on the other hand, are dominated by two parties, but many more parties enjoy success at the regional level. What are the conditions required to enable two parties to dominate all elections across a country? When do more than two parties emerge?

In this paper, I propose a novel model of party formation and derive necessary conditions for the existence of equilibria in which two parties form. The model features elections at different levels of government and imposes no ad-hoc restrictions on the number of parties that compete across these elections. Politicians standing at the beginning of their career join 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 regional election then has a chance to become their party's candidate at the national level. In addition, politicians also have the option of contesting elections as independents at any point in their career.

I start by considering a one-dimensional policy space. A crucial feature of the model is that there is a minimum amount of heterogeneity in voter preferences across regions. In a two-party equilibrium, this heterogeneity forces parties to adopt a sufficiently broad ideological profile if they want to cater to voters' diverse tastes and prevent entry of additional parties. But if parties allow a range of politicians to join, this creates internal competition for party nominations in the run-up to regional elections. As a consequence, members potentially have an incentive to join smaller parties with a narrower ideological profile. The necessary conditions for the existence of two-party equilibria all follow from the requirement that parties provide sufficient opportunities for members while limiting the success of defectors.

The first result that I present is that no two-party equilibrium exists in which one party allows all types of politicians to join. Such a catch-all party is unstable for two reasons, depending on the nature of the second party. If a catch-all party competes with a centrist party, the latter dominates elections at the national level. More extreme politicians, who are thus limited to winning regional elections, prefer to do so as members of smaller parties more targeted at the preferences of the voters they face there. If the second party instead also allows at least some types of partisans to join, this raises the possibility that both parties nominate candidates who fall on the same side of the national median voter, creating an opening for a moderate independent or third-party candidate.

The second necessary condition for the existence of two-party equilibria requires that there are no regions where the electorate favours radical policies relative to the distributions of voters overall. The presence of such regions would force parties to extend their memberships to the extremes as third parties could otherwise successfully contest elections there. This shift towards the extremes may make parties vulnerable to entry of a centrist candidate in the national election. The combination of a few radical regions and a fairly moderate national electorate, in particular, leaves parties stretched too thin and entry of a third party must occur. As I argue in the text, failure of this condition provides a plausible explanation for the existence of the Liberal Democrat Party in the UK. A polarisation of the electorate, in contrast, is not necessarily a threat to a two-party equilibrium. Even though polarisation similarly leads to both parties becoming more likely to nominate extreme candidates, centrist third-party candidates may lack the moderate voter base required to propel them to success in the national election.

The first two necessary conditions together imply that the equilibrium parties dominate the national election even when a third candidate enters the race. These parties can thus compensate their members for the intense internal competition that they generate at the regional level with the prospect of success on the national stage. A necessary prerequisite for politicians to be willing to make this trade-off, however, is that they value opportunities at the national level sufficiently strongly relative to winning at the state level. The third necessary condition for the existence of two-party equilibria is therefore that politicians have strong career concerns.

In a subsequent step, I establish the existence of an equilibrium with a centre-left and a centre-right party. This equilibrium exists if and only if the three previously established conditions are satisfied, demonstrating that these conditions are also jointly sufficient for the existence of two-party equilibria. In addition, the equilibrium illustrates that parties have an incentive to commit to giving all of their members a shot at winning nominations instead of always favouring the candidates who are most likely to win. Otherwise defections by politicians who are popular in their region but less well-positioned to win nationally are hard to prevent.

I then extend the model by introducing a second dimension of policy intended to capture a concern specific to a subset of regions, like the presence of an ethnic minority or an independence movement. Interestingly, regionalism in itself is not a threat to the existence of two-party equilibria, even if regionalism is the dominant issue for voters. This is the case because the presence of a regionalist party can stabilise a party representing all remaining voters, which would otherwise suffer from defections. However, the possibility of a regionalist candidate running for the national election relaxes the conditions required for a moderate independent to win. In particular, two-party equilibria are unlikely to exist if regionalism is salient and regionalist voters make up a small share of the national electorate. Given that Scottish voters account for less than ten percent of the electorate in UK general elections, the model thus provides an explanation of the existence of the Scottish National Party.

Aldrich & Lee (2016) also highlight the importance of political ambitions in explaining why only two parties exist in the US. To make this point, the 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. My paper extends the analysis of Aldrich & Lee in several ways. In particular, I provide an explanation for why the chances of winning a state election should be lower as a member of the national party in the first place. In the model presented here, national parties are less attractive due to a higher level of internal competition for nominations, which arises endogenously. In addition, I present a number of additional requirements for two-party dominance that are not discussed in Aldrich & Lee (2016).

This paper is also 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. Closest to the current paper is Callander (2005), who studies competition between two parties in multiple single-member districts with threat of entry of independent candidates at the district level. Parties, whose formation is not part of the model, 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, as are the existence conditions for two-party equilibria. 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 nor allows for regionalism.

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 consisting of multiple politicians while endogenising the number of parties existing in equilibrium (Jackson & Moselle 2002, Levy 2004, Morelli 2004, Osborne & Tourky 2008, Eguia 2011). To the best of my knowledge, I am the first to do so employing the concept of parties as informative labels.[1] Given the need to include multiple elections with separate electorates, affiliation choices of politicians, as well as assumptions about candidate selection, the model is necessarily relatively complex. Nevertheless, the model is tractable and naturally lends itself to the purpose of investigating other questions, such as the interplay between social diversity and electoral rules in determining the number of political parties (Dickson & Scheve 2010, Milazzo et al. 2018). Indeed, my results suggest that a theoretical analysis of the number of parties competing

[1] 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.

in a particular district may be misleading if linkages across levels of government are not taken into account. The extension introducing regionalism indicates that the model can accommodate multi-dimensional policy spaces.

The rest of the paper is organized as follows: Section 2 explains the details of the model, while Section 3 presents the theoretical results. Section 4 extends the model to allow for regionalism. Robustness of the results to relaxing some of the assumptions is discussed in Section 5. Section 6 concludes.

2 The Model

A federal country consisting of $S \geq 3$ states selects federal and state governments through FPTP elections. Political parties nominate candidates for these elections, but politicians who are not nominated can decide to run as independents. 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. In any of these elections, any politician who has not been nominated by a party can run as an independent candidate. The following sections describe the elements of the model in detail.

2.1 Players

The strategic players of the game are voters and politicians.

2.1.1 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 $i \in \mathbb{R}$. The set of voters in state s is described by a measure Λ_s that assigns to any subset of $\mathbb{R}$ the share of voters whose ideal policies lie in this subset. Let m_s denote the ideal policy of the median voter of state s. All voters in any state vote in the federal election and their distribution is described by the measure Λ_f with median $m_f = 0$.

As the analysis below shows, a crucial aspect of the model is how spread out the positions of state median voters are relative to the variation in the distribution of voters in the federal election. To pin down the latter, I assume

$\Lambda_f([-1,1]) > 0.5$ and $\Lambda_f((-1,1)) \leq 0.5$, such that a majority of voters in the federal election is located in the interval $[-1,1]$. As long as no restriction is placed on the distribution of state median voters, the preceding assumption is essentially without loss of generality and merely amounts to a normalisation.[2]

The only more substantial assumptions placed on voter distributions specify that there is some minimum amount of heterogeneity in voter preferences across states: for some positive number d, let there be exactly one state s such that $m_s < -d$, at least one state s' such that $\Lambda_{s'}((-d,d)) > 0.5$, and exactly one state s'' such that $m_{s''} > d$. As the labels of states are arbitrary, it is without loss of generality to denote these states as states 1, 2, and 3, respectively. The variable d determines how big the variation in voter preferences across states is compared to the distribution of voters overall. For small values of d, state median voters are much less spread out than voters in general. For large values of d, on the other hand, the median voters of at least some states are radicals in the sense that most voters prefer more moderate policies. Note that the assumptions on Λ_f place no restrictions on the value of d, at least as long as states 1 and 3 contain less than half of all voters.

2.1.2 Politicians

Each state s has a finite set of politicians. Every politician is endowed with a platform, which is an element of the set $\mathcal{T} = \{-2d, 0, 2d\}$. There are three politicians in each state, none of which share the same platform. Put differently, there is one politician located at each possible platform in each state. Whenever a politicians wins an election, they are committed to implementing their platform. I show in Section 5.1 that the results extend to a setting where politicians care strongly about the policy that they choose once elected, but are otherwise free to implement any platform.

Politicians with platform 0 are referred to as moderates, while all remaining politicians are labelled as partisans. The latter politicians can be thought of as the most radical candidates who can win a state election, since the median voter of state 1 (state 3) prefers the policy $-2d$ (the policy $2d$) over any other platform. Nevertheless, for small values of d the positions of partisans are themselves fairly moderate relative to the positions of the most extreme voters.

[2]For any discrete distribution Λ_f of voters in the federal election with median $m_f = 0$ it is possible to find an ideal point $i > 0$ such that $\Lambda_f([-i,i]) > 0.5$ and $\Lambda_f((-i,i)) \leq 0.5$, unless more than half of all voters are located at zero. Since distances between ideal points have no meaning without a reference point, the distribution can then be rescaled by dividing all ideal points by i to satisfy the assumptions made above.

2.2 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 platform falls within this subset can join. One way to think of this assumption 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 are denoted by capital letters and for any such party P the interval representing the party is given by I_P. Let $\mathcal{I}$ be the collection of all possible shapes parties can have. Since there are three platforms,

$$\mathcal{I} = \{[-2d, 2d], [-2d, 0], [0, 2d], [-2d], [0], [2d]\} \ .$$

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 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 is described in more detail below.

In addition to the set of politicians who are allowed to join a particular party, it also needs to be specified how parties nominate candidates for elections. Denote by $\mathcal{M}_{P,l}$ the set of politicians who are eligible to be nominated for the election in region $l \in \{1, ..., S, f\}$ by party P. In a state election, this set consists of all politicians who have joined the party in the state. At the federal level, all politicians who have won a state elections as a member of party P can be nominated. The probability that a politician with platform p who belongs to $\mathcal{M}_{P,l}$ is nominated for the election in region l is given by a function $\eta_{P,l}(p|\mathcal{M}_{P,l})$. The only restrictions placed on this function are that $\eta_{P,l}(p|\mathcal{M}_{P,l}) > 0$ for all $p \in \mathcal{M}_{P,l}$ as well as those needed to ensure that $\eta_{P,l}$ yields well-defined probabilities. More specifically, the codomain of $\eta_{P,l}$ is equal to $[0, 1]$ and for any $\mathcal{M}_{P,l}$ the nomination probabilities across all members of $\mathcal{M}_{P,l}$ sum to one.

2.3 Timing

The game proceeds as follows: In a first step, all politicians decide whether to join a party. 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. After observing the candidates put forward by parties, all politicians who have not been nominated decide simultaneously whether to run as independents for the election in their state. Once the lists of candidates have been determined, each voter casts a vote at the election in their state and the winner in each state is the candidate who receives the highest number of votes. Ties are 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. Again, any politician who has not been put forward as a candidate can decide to contest the election as an independent, after observing the candidates nominated by parties. All voters vote in the federal election. The winner is once more the candidate who receives the highest number of voters, who then implements their platform as the federal policy.[3]

2.4 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 the game and only knows how their platforms are distributed. As there is one politician for each of the three possible platforms in each state, the prior belief of voters over the platform of a randomly selected politician thus assigns probability one-third 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

[3]Even in a presidential system, policymaking often requires passing legislation, which in turn 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 are 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.

candidate prior to casting their vote at the state-level election. Suppose, for example, that a candidate in a certain state s is a member of a party P of shape $[0, 2d]$ and voters believe that only a politician with platform 0 has joined this party. Voters consequently believe that the candidate of the party must have platform 0. If the electorate instead believes that two politicians with platforms 0 and $2d$ have joined party P with certainty, then the beliefs over the platform of the nominee are equal to the probabilities that either politician is nominated, given by the nomination technology $\eta_{P,s}$. For example, if both members are nominated with equal probability, voters accordingly attach equal probability to the two possible platforms of the candidate of party P.

The winner of a state election implements her platform at the state level, thus revealing it to voters. Given that parties nominate state winners for the federal election, voters accordingly have full information about party candidates at the federal level.

Finally, all agents are fully informed about the distribution of voters in all states and at the federal level.

2.5 Strategies

Voters are assumed to vote 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 candidates in each respective election. Politicians, on the other hand, choose their party affiliation and whether to run as an independent in their state election and in the federal election. The decision to run as an independent is a simple binary choice, which is made conditional on the candidates previously nominated by parties. It is assumed that politicians always accept the nomination of their party. Accordingly, only politicians who have not been nominated for a particular election by a party can become independents.

This leaves the choice of party affiliation to be described in more detail. 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 platform 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 at the point when party affiliations are chosen is given by $\mathcal{P}(p) \cup \{\emptyset\}$, where $\emptyset$ represents the choice not to join a party.

2.6 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 candidates and the behaviour of other voters, the objective of a voter with ideal policy $i \in \mathbb{R}$ in an election in region $l \in \{1, ..., S, f\}$ is to maximize

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

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

The payoffs of politicians are determined by their individual electoral success. The utility of a politician who does not win any elections is normalised to zero. In contrast, 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.

In order to clearly define the utility of a politician, let π_s be the probability that a politician wins the election in her state, independent of whether this happens after being nominated by a party or running as an independent. Similarly, let π_f give the probability that a politician wins the federal election. Both these probabilities are determined in equilibrium. The expected utility of a politician is then given by

$$\pi_s y_s + \pi_f y_f \ .$$

2.7 Equilibrium

The party-formation game described in the previous sections features incomplete information and requires a corresponding equilibrium concept. I focus on

[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.

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 wins. Nevertheless, it seems plausible that voters are able to coordinate on electing a candidate who is favoured by a strict majority.[5]

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 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, $\mathcal{P}^*$ denotes the set of active parties in equilibrium, while $N^* \equiv |\mathcal{P}^*|$.

2.8 Discussion

In the following I discuss some of the features of the model. First of all, I would like to highlight how general the assumptions regarding candidate selection are. While they imply that parties commit in advance to nominating specific candidates with fixed probabilities, these probabilities can be arbitrarily close to the case of strategic nomination of candidates without commitment. For example, the aim of maximising the number of elections won is in most settings with two parties achieved by nominating a moderate whenever possible.[6] The model

[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.

[6]At the state level, nominating a moderate even when other politicians have joined the party is consistent with maximising the chance of winning the state election either because *i*) all politicians of the state have joined and there is no competition, *ii*) the moderate is located closest to the state median voter among the members, or *iii*) the fact that another member is located closer to the state median voter implies that the moderate is located closer to the median voter than the member of the competing party.

allows parties to implement this strategy with only a very small probability of mistakes. The results below show that parties do in fact have an incentive to commit to giving all types of politicians a chance at winning nominations in order to avoid defections. Such a commitment can be achieved, for example, by selecting candidates through primary elections or caucuses, which limit the control of the national party committee over the nomination process.

The assumption that there are only three platforms primarily serves the purpose of ease of exposition. Allowing for an arbitrary number of platforms, as a previous version of this paper did, requires the introduction of additional notation and substantially lengthens the proofs while leaving the results largely unchanged.

An assumption that may appear strong is that there can be parties that perfectly reveal the platforms of their members. However, the essential feature of the model is not that there can be such "singleton parties", but that there can be parties that reveal different amounts of information about their members. The Tea Party movement in the US (Arceneaux & Nicholson 2012), while not a party of its own, illustrates that it is possible to send a more fine-grained signal to voters than, for instance, the Republican party label does. A separate issue is whether a single politician should be allowed to form a new party. I discuss this question in Section 5.2, where I consider the possibility of allowing for joint deviations by groups of politicians.

Some readers may wonder about the difference between deviating to joining a singleton party and running as an independent. Conceptually, forming a new party should be thought of as a process that requires more time and preparation, while the decision to run as an independent can be made even late in a race.[7] This is highlighted by the fact that independent candidates in the model can make their entry decision conditional on the candidates put forward by established parties. It is true, however, that at the state level no politician would ever strictly prefer to run as an independent over forming a singleton party. The choice of allowing independent candidates to enter state elections is thus made merely for completeness. At the federal level, on the other hand, politicians who have previously been elected in a state are much less dependent on their party.

[7] In exceptional cases, an already established and popular politician can create a party organisation at surprising speed, as the example of Emmanuel Macron's En Marche movement in France illustrates (Mény 2017). Given that En Marche was constructed entirely around the persona of Macron, this case is, nevertheless, better described as entry of an independent candidate and contrasts sharply with the establishment of parties out of grassroots movements, such as in the case of European green parties (Kaelberer 1998).

Given that such candidates are already known to voters, running a successful campaign as an independent can be a viable option.

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.[8]

3 Results

The aim of this paper is to investigate conditions under which equilibria with two active parties exist. It is useful to introduce a formal definition of the type of equilibrium that is the main focus.

Definition 2 (Two-Party Equilibrium)**.** *A two-party equilibrium is a party-formation equilibrium such that $N^* = 2$ and no independent candidates run in any elections along the equilibrium path.*

A two-party equilibrium thus requires not only that the number of active parties is equal to two, but also that these parties face no competition from independent candidates. In most democratic countries, independent candidates are a common occurrence, but independent candidates with a serious chance of winning are much rarer.

In the following subsections, I establish three conditions that are necessary for the existence of any two-party equilibrium. Subsequently, I demonstrate that these conditions are also sufficient for existence through of a particular example of a two-party equilibrium. To start with, however, I introduce two preliminary results, which are at the core of what is to follow. The first of these states that in any two-party equilibrium, all types of politicians must be able to join one of the active parties.

Lemma 1. *Consider any two-party equilibrium and denote the parties that are active in equilibrium by A and B. Then it has to be the case that $\{-2d, 0, 2d\} \subset I_A \cup I_B$.*

Proof. See Appendix. □

[8]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.

The preceding result is a consequence of the heterogeneity in voter preferences across states. If a particular type of politician is excluded from the existing parties, there is at least one state where this type can successfully contest the state election by forming a new party. Two-party equilibria therefore require that the two parties existing in equilibrium span a sufficiently "big tent" to accommodate all politicians who appeal to some state median voter.

A second feature of any two-party equilibrium is that any politician who has a chance to join a party that can win the election in the politician's state must do so. A central ingredient of this result is the assumption that voters cannot observe the membership of a party at the state level. A politician joining a party out of equilibrium therefore does not affect voters' expectations regarding the platform of the party's candidate and the outcome of the elections remains unchanged. Even politicians who are not well aligned with the median voter of a state can thus join a party without jeopardising the electoral success of the party.

Lemma 2. *Consider any two-party equilibrium and suppose there is a politician in some state s who is eligible to join a party that wins the election in the state with positive probability. Then the equilibrium strategy of this politician must place zero probability on not joining a party or on joining a party that loses the state election with certainty.*

Proof. See Appendix. □

Broad ideological profiles in combination with the fact that successful parties attract members imply that a necessary part of any two-party equilibrium is internal competition for nominations at the state level by different factions within a party. In isolation, this factor gives party members an incentive to defect to smaller parties with a narrower ideological profile in order to increase their chances of winning state elections. As a consequence, two parties can only form an equilibrium if they offer their members sufficient success at the federal level while simultaneously limiting that of defectors. The necessary conditions for the existence of two-party equilibria provided in the next three sections all follow from this logic.

3.1 No Catch-All Parties

The first condition that any two-party equilibrium must satisfy is that none of the parties that are active in equilibrium allow all types of politicians to join.

Proposition 1. *Consider any constellation of two parties A and B. Then a two-party equilibrium such that $\mathcal{P}^* = \{A, B\}$ only exists if $[-2d, 2d] \notin \{I_A, I_B\}$.*

Proof. See Appendix. $\square$

While Lemma 1 requires that parties in a two-party equilibrium are not "too narrow", they also cannot be too broad. The reason is that broad and therefore overlapping parties generate the possibility that both parties nominate partisans for the federal election whose platforms lie on the same side relative to the median voter. This enables a moderate independent to enter and win. The exception would be a constellation with a catch-all party and a party that allows only moderates to join, but then partisans would have no chance of winning the federal election. Consequently, these politicians would have no reason to put up with the internal competition they are facing at the state level and prefer to join smaller parties.

3.2 No Outlier States

The second necessary condition for the existence of two-party equilibria imposes a limit on how far the median voters of states can be shifted towards the tails of the distribution of voters overall.

Proposition 2. *Any two-party equilibrium exists only if $d \leq 1$.*

Proof. See Appendix. $\square$

As was already established by Lemma 1, the active parties have to allow partisans to join in any two-party equilibrium. As a consequence, the situation may arise that both parties nominate partisans for the federal election. Since no party can allow all types of politicians to join, these candidates will necessarily be drawn from opposite ends of the political spectrum. Voters with an ideal policy close to zero would then prefer a moderate candidate. However, if d is small and partisans are located close to zero, the number of voters in favour of a moderate would be small and a campaign by such a candidate would not be successful. If d exceeds 1, in contrast, the support for a moderate independent or third-party candidate is strong enough to allow them to win the federal election. Proposition 2 can thus be interpreted as requiring that parties are not pulled too much to the extremes by the presence of some states whose electorates are radical compared to the distribution of voters in general.

The specific level of the threshold on d follows from the assumption that more than half of all voters in the federal election are located in the interval $[-1, 1]$. All of these voters prefer a moderate over any partisan when $d > 1$. Another way of expressing the condition in Proposition 2 is therefore that there cannot be states whose median voters are “outliers” in the distribution of voters overall in the sense that their position is more radical than the ideal policies of more than half of all voters.

Proposition 2 is interesting due to the situations that it does not rule out. In particular, a polarisation of the electorate per se poses no threat to a two-party equilibrium as long as the number of remaining moderate voters is sufficiently small. To see this, consider an example of a country in which voters are extremely polarised in the sense that the median voters of half of all states are located at $-d$, while the remaining half is located at d. Then the normalisation $\Lambda_f([-1, 1]) > 0.5$ implies that the largest possible value of d is one, since otherwise less then half of all voters would be located in the interval $[-1, 1]$. Accordingly, Proposition 2 does not rule out the existence of two party equilibria in this case. This logic may help explain why the two-party system is alive and well in the US despite increasing polarization both of political elites and of the electorate (Iyengar et al. 2019). In 2016, there was a chance that both the Democratic and the Republican Party were going to nominate radical candidates for the presidential election, namely Bernie Sanders and Donald Trump. Anticipating this scenario, it was rumoured that Michael Bloomberg was considering a run as an independent (Burns & Haberman 2016). However, it is far from clear that Bloomberg’s campaign would have attracted sufficient support from moderate voters to gain any kind of momentum. It is perhaps telling that Bloomberg chose to pursue the Democratic ticket for the 2020 election instead.

Two examples from UK politics lend further support to the relevance of the logic underlying Proposition 2. The formation of the Social Democratic Party in 1982 was an attempt to seize the centre ground vacated by other parties that, at least for a while, looked likely to succeed (Crewe & King 1995). The party was founded by four senior members of the Labour party at a time when the leaders of both Labour and the Conservatives were pursuing radical policies. The electoral alliance that the Social Democratic Party entered with the previously marginalised Liberal Party soon found itself topping opinion polls, indicating that the radicalisation of parties was in this case not driven by a radicalisation of the electorate as a whole. It is widely held that it was only the Falklands War that turned the tables in favour of Margaret Thatcher. The entry of the Social Democratic Party could thus be explained by failure of the condition presented in Proposition 2. The recent polarisation in British politics induced or uncovered by Brexit also lead to the formation of a centrist group of breakaway MPs. However, this new party never managed to attract significant support from voters and was dissolved less than a year after its inception. It thus appears that, in this instance, defections were more driven by a marginalisation of moderate members within their parties rather than by the prospect of electoral success as members of a new party.

3.3 Career Concerns

The conditions introduced in the preceding two sections follow from the necessity of ensuring that the success of defectors does not extend beyond the state level. If these conditions are satisfied, the relatively broad parties that are necessarily part of any two-party equilibrium can compensate their members for the internal competition that they generate with opportunities at the federal level. Politicians need to value these opportunities sufficiently strongly, otherwise defections cannot be prevented.

Proposition 3. *For any constellation of two parties $\{A, B\} \subset \mathcal{P}$, there exists a constant $\bar{y} > 0$ such that a two-party equilibrium in which $\mathcal{P}^* = \{A, B\}$ only exists if $y_f / y_s \geq \bar{y}$.*

Proof. See Appendix. □

In words, Proposition 3 requires that the ratio between the payoff from winning the federal election and the payoff of winning a state election must exceed a certain threshold for any two-party equilibrium to exist. The level of this threshold depends on the specific equilibrium under consideration, but is generally driven by how likely different politicians are to win elections in equilibrium. These probabilities in turn depend on the candidates nominated by the opposing party and the intensity of internal competition. At the federal level, internal competition increases with the number of state elections won by a party: the greater the electoral success of a party at the sub-national level, the greater the number of party members with a public profile who would make viable candidates for the federal election. Furthermore, the chances of any given politician of securing the federal nomination depend on how their party selects candidates. A party that gives priority to moderates due to their electability is likely to suffer defections by partisans. If the party heavily favours the latter type of candidate, in contrast, moderates may be reconsidering their options. It should be kept in mind, however, that not all politicians are equally able to achieve success in state elections as members of a third party. When deciding the federal nomination, parties thus have an incentive to give priority to members whose platform is especially popular in their state. This group includes, in particular, partisans in "safe seats", that is, states where the median voter strongly favours one party. Finally, the probability that the *opposing* party nominates a partisan candidate for the federal election unequivocally reduces the chances that a politician deviates and joins a third party.

While the payoffs y_s and y_f depend on the intrinsic motivations of politicians, other factors such as financial rewards, public visibility, or the competencies and powers associated with an office equally play a role. The strength of career concerns accordingly does not depend only on the internal organisation of parties, but also on the setup of the political system more generally.

I conclude this section with a brief comparison between the US and Canada. The former country features two stable parties that dominate elections at all levels of government, while Canada is home to a multitude of parties whose federal and provincial branches are only loosely connected (Johnston 2017). The two countries also differ strongly in the typical career paths of politicians. US politicians start their careers at the local or state level and follow a relatively linear path towards offices at the federal level (Diermeier et al. 2005). Political life in Canada, in contrast, has historically been characterised by parallel career tracks at the federal and the provincial level (Barrie & Gibbins 1989). While politicians in the US thus face strong incentives to stay loyal to established parties, the absence of a clear path for provincial politicians towards federal offices in Canada has possibly contributed to the fracturing of the Canadian party system over time. Alternatively, the separation between provincial and federal careers in Canada may itself be a consequence of a relatively equal appeal of federal offices and opportunities at the provincial level.

3.4 An Example of a Two-Party Equilibrium

While the previous sections have introduced a number of necessary conditions for the existence of any two-party equilibrium,the following proposition shows that these necessary conditions are also jointly sufficient.

Proposition 4. *Consider a constellation of two parties, L and R, with* $I_L = [-2d, 0]$ *and* $I_R = [0, 2d]$ *and suppose that parties use some combination of* *nomination technologies satisfying the assumptions made in Section 2.2. Then an equilibrium such that $\mathcal{P}^* = \{L, R\}$ exists if*

i) $d \leq 1$ *and*

ii) $\frac{y_f}{y_s}$ *is larger or equal to some threshold* $\bar{y} > 0$.

Proof. See Appendix. □

The equilibrium presented in Proposition 4 features a centre-left and a centre-right party that overlap in the centre of the policy space. This constellation is essentially the one that has been dominating US politics ever since the end of the civil war and also captures the party landscape in Britain in the decades after World War II well. Naturally, the equilibrium features no catch-all parties, as required by Proposition 1. Proposition 2, on the other hand, reappears as Part *i)* of Proposition 4. Finally, Proposition 3, which requires sufficiently strong career concerns, is reflected in Part *ii)* of Proposition 4.

To illustrate the factors determining the level of the threshold $\bar{y}$ in Proposition 4, I now consider a specific example of nomination technologies. I assume that any party P in any state s uses the technology

$$\eta_{P,s}(p|\mathcal{M}_{P,s}) = \eta^u(p|\mathcal{M}_{P,s}) \equiv |\mathcal{M}_{P,s}|^{-1} \ .$$

Any politician who has joined a particular party in a particular state is therefore equally likely to be nominated for the state election. At the federal level, on the other hand, all parties are assumed to nominate candidates according to a function η^ε. The variable $0 < \varepsilon < 1$ denotes the probability that a party nominates a partisan for the federal election when the candidate pool contains both moderates and partisans. Within the groups of moderates and partisans, on the other hand, any politician is equally likely to be selected.[9] For ε equal to one-half, neither type of politician is favoured by the nomination process. For values of ε close to zero, on the other hand, the nomination of a partisan happens merely by accident or if no other type of politician is available. This

[9] To define the nomination technology η^ε formally, let $\mathcal{M}_{P,l}(p)$ be the set of members of $\mathcal{M}_{P,l}$ with platform p. Then, for any p and $\mathcal{M}_{P,f}$ such that $\mathcal{M}_{P,f}(p) \neq \emptyset$

$$\eta^\varepsilon(p|\mathcal{M}_{P,f}) = \begin{cases} (1-\varepsilon)/|\mathcal{M}_{P,f}(0)| & \text{if } p = 0 \text{ and } \mathcal{M}_{P,f}(0) \neq \mathcal{M}_{P,f}, \\ \varepsilon/|\mathcal{M}_{P,f} \setminus \mathcal{M}_{P,f}(0)| & \text{if } p \neq 0 \text{ and } \mathcal{M}_{P,f}(0) \neq \emptyset, \\ 1/|\mathcal{M}_{P,f}| & \text{otherwise.} \end{cases}$$

latter case approximates the situation of strategic candidate nomination at the federal level when parties aim to maximise the chance of winning.

Proposition 5. *Suppose* $\eta_{P,f} = \eta^{\varepsilon}$ *for all* $P \in \mathcal{P}$ *with* $\varepsilon \leq 0.5$ *and* $\eta_{P,s} = \eta^{u}$ *for all* $P \in \mathcal{P}$ *and* $s \in \{1, ..., S\}$*. Then an equilibrium such that* $\mathcal{P}^* = \{L, R\}$ *with* $I_L = [-2d, 0]$ *and* $I_R = [0, 2d]$ *exists if*

i) $d \leq 1$ *and*

ii) $\frac{y_f}{y_s} \geq (S-1)/\left(\varepsilon + (0.5-\varepsilon)/2^{S-2}\right)^2$.

Proof. See Appendix. □

Given that Proposition 5 assumes particular functions determining nomination probabilities, Condition *ii* of the proposition provides a specific threshold for the strength of career concerns sufficient for existence, rather than just stating that such a threshold exists. This condition derives from the fact that any partisan politician who has a chance of winning the election in their state as member of a smaller party must be compensated by a sufficiently high chance of winning the federal election. This probability of winning the federal election depends not least on the likelihood of being nominated, which is driven by two main factors: First of all, competition for the nomination is higher the more states a party wins. As a consequence, the right-hand side of Condition *ii* is increasing in the number of states S. Second, the parameter ε captures how likely each party is to nominate a partisan. Accordingly, Condition *ii* is less likely to be satisfied the smaller ε is. The focus on partisan candidates derives from the fact that, at least for $\varepsilon \leq 0.5$, moderate politicians are both more likely to be nominated for and to win the federal election in equilibrium. Out of equilibrium, on the other hand, moderates find it just as impossible as partisans to win the federal election as long as condition *i* is satisfied. Moderate politicians are thus less likely to deviate.

Under the constellation of parties in Proposition 4, moderate politicians are able to choose between either of the equilibrium parties. By Lemma 2, they must join a party that wins the state election with positive probability in equilibrium. In states with a moderate median voter, either party can win the state election depending on the affiliation choice of the moderate politician. When the median voter is more radical, only the party located close to the median voter can win independent of the choice of the moderate politician. A comparative static generated by the equilibrium is therefore that a shift in the position of the median voter of a state can lead to a change in the affiliation of the moderate of the state. If one interprets the realignment of voters in the US South since the 1960s from the Democratic to the Republican Party as such a shift in voter preferences, then the model would predict the occurrence of state politicians changing their affiliation from Democrat to Republican. This is exactly what could be observed in the 1990s and early 2000s, as documented by McKee & Yoshinaka (2015). In addition, the equilibrium generates a distribution of vote margins also reminiscent of US politics. Elections in states with centrist median voters generate relatively equal vote shares while states where the mass of voters is shifted away from the centre can be characterised as "safe seats" where election results heavily favour one party.

While the equilibrium established in Proposition 4 is appealing due to its similarity to two-party systems observed in reality, it is in fact also one of only three possible two-party equilibria. Lemma 1 and Proposition 1 in combination imply that the two remaining possibilities are a centre-left party competing with a right-wing party, and the symmetric situation of a left-wing party facing a centre-right party. Equilibria featuring these constellations of parties can be shown to exist under similar conditions as the equilibrium above.

4 Regionalism

In this section, I introduce a second dimension of policy and derive an additional necessary condition for the existence of two-party equilibria in this extended version of the model. The second dimension of the policy space represents an issue or characteristic specific to some states, such as an independence movement or the presence of an ethnic minority that is concentrated in a subset of states. In line with these examples, the secondary issue is modelled as binary. Accordingly, the policy space is now given by $\mathbb{R} \times \{0, r\}$ with $r > 0$ and the set of possible platforms is $\mathcal{T} = \{-2d, 0, 2d\} \times \{0, r\}$. Voters and politicians located at r along the second dimension of policy will be referred to as regionalists and all others as non-regionalists. The terms partisan and moderate continue to describe the position of a politician along the first, ideological dimension of the policy space.

In line with the basic version of the model, I assume that the median voter along the ideological dimension in the federal election is located at zero:

$$\min\{\Lambda_f((-\infty, 0] \times \{0, r\}), \Lambda_f([0, \infty) \times \{0, r\})\} > 0.5 \ .$$

Again building on the basic model, I assume $\Lambda_f([-1,1] \times \{0\}) > 0.5$. Beyond serving as a normalisation of the policy space, this assumption now has the additional implication that regionalism is a minority issue. In most states, there are only three politicians with platforms $(-2d, 0)$, $(0, 0)$, and $(2d, 0)$ and there are no regionalist voters, that is, $\Lambda_s(\mathbb{R} \times \{0\}) = 1$. In a subset $\mathcal{S}^r$ of states, however, there are six politicians, one for each possible platform in $\mathcal{T}$, and the distribution of voters in the state is not restricted to $\mathbb{R} \times \{0\}$. To ensure that the results from the previous sections also apply to the extended version of the model, I assume that there are no regionalist voters in states 1, 2, and 3, that is, $\{1, 2, 3\} \cap \mathcal{S}^r = \emptyset$. In addition, the assumptions about the distributions of voters in these states are maintained. While regionalism is a minority issue in the federal election, there exists at least one state $s \in \mathcal{S}^r$ such that $\Lambda_s([-1,1] \times \{r\}) > 0.5$ and regionalist voters form the majority.

Given that the shapes of parties were given by intervals in the basic model, a natural generalisation is to require party shapes to be convex subsets of $\mathbb{R}^2$ in the extended model.

The aim of a voter with ideal policy (i, j) in an election in some region $l \in \{1, ..., S, f\}$ is now to maximise

$$\mathbb{E}[u(|p_{l,1} - i|) + u(|p_{l,2} - j|)] \ ,$$

where $(p_{l,1}, p_{l,2})$ is the policy implemented in region l.

The necessary conditions for the existence of two-party equilibria discussed in the previous section carry over to the extended model. In particular, career concerns need to be sufficiently strong, the value of d cannot exceed 1, and parties that allow all types of politicians to join are hard to sustain.[10] Intuitively, one might expect that the extended model yields the additional requirement that the salience of regionalism is low enough to prevent entry of a regionalist party. Salience of the regionalist issue is determined by the relationship between the parameters d and r. If r is small relative to d, the position of a politician along the regionalist dimension has a negligible impact on voters' utility relative to the ideological position. In this case there may even be equilibria in which no active party allows any regionalists to join.[11] If r is large relative to d,

[10]Strictly speaking, there are two-party equilibria of the extended model where one party allows all types of politicians to join. However, all of these cases require that the catch-all party does not win any elections in regionalist states. For any such equilibrium there therefore exists an equivalent equilibrium where the party in question does not allow any regionalists to join.

[11]Since there are states in which no regionalist politicians are present, the reverse case is

on the other hand, regionalism becomes the decisive issue for regionalist and non-regionalist voters alike. While the latter case does indeed imply that any equilibrium must feature a regionalist party, the following example illustrates that two-party equilibria can exist no matter how strongly voters care about regionalism.

Proposition 6. *Let $d = 0.5$. Suppose there are four states with equal-sized populations, $\mathcal{S}^r = \{4\}$, and*

$$\begin{aligned} &\Lambda_1([-1,-0.5] \times \{0\}) \\ =\ &\Lambda_2((-0.5,0.5) \times \{0\}) \\ =\ &\Lambda_3([0.5,1] \times \{0\}) \\ =\ &1 \end{aligned}$$

while

$$\Lambda_4([-1,-0.5] \times \{r\}) \ = \ \Lambda_4([0.5,1] \times \{r\}) \ = \ 0.5 \ .$$

Then there exists a two-party equilibrium such that $\mathcal{P}^* = \{N, R\}$ with $I_N = [-1,1] \times \{0\}$ and $I_R = \{(-1,r)\}$ if $r > 2$ and y_f/y_s exceeds some threshold $\bar{y} > 0$.

Proof. See Appendix. □

A specific case satisfying the assumptions made in Proposition 6 is illustrated in Figure 1, where grey discs indicate that voters are located at the centre of the disc. The size of each disc is proportional to the share of voters in the federal election in the specified location and grey numbers indicate which state the voters belong to. Party N is a non-regionalist party with a broad ideological profile, which wins the elections in states 1, 2, and 3. Party R, in contrast, is a regionalist party with a very narrow profile, which wins the election in the sole regionalist state and never wins the federal election. The equilibrium requires that the salience of the regionalist issue is sufficiently high to ensure that regionalist voters vote en bloc for the regionalist candidate at the federal election. For lower values of r, a coalition of regionalist and non-regionalist voters would enable entry of a moderate independent in the federal election when both parties nominate candidates located at -1 along the ideological dimension.

not possible.

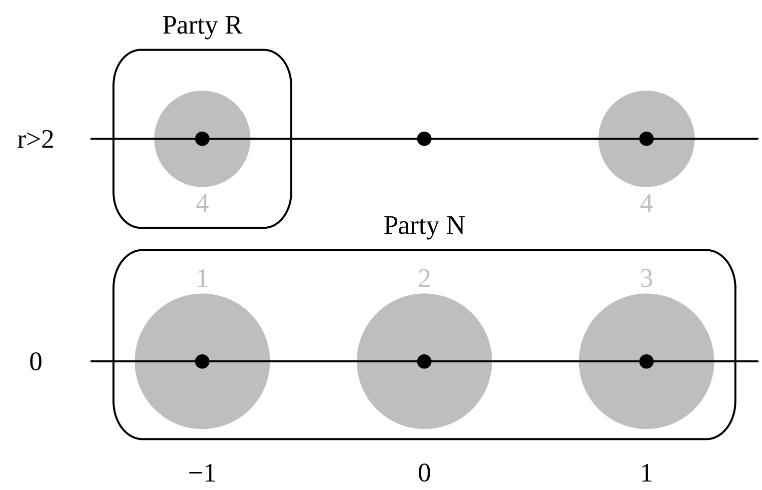

Figure 1: Example of a Two-Party Equilibrium that Exists for Arbitrarily Large Values of r/d

A grey disc indicates that voters are located at the centre of the disc. The size of each disc is proportional to the share of voters in the federal election at the indicated location. Grey numbers give the state that the voters in question belong to. Rounded rectangles indicate the shapes of the parties that are active in equilibrium, that is, which of the politicians located at black dots are allowed to join each party.

While two-party equilibria can exist for any value of r/d, salience of regionalism in combination with a second condition can rule out all such equilibria.

Proposition 7. *For any $d > 0$ there exists a threshold $\bar{r}_d$ such that no two-party equilibrium exists if $r > \bar{r}_d$ and*

$$\min\{\Lambda_f((-\infty, d) \times \{0\}), \Lambda_f((-d, \infty) \times \{0\})\} > 0.5 \ . \tag{1}$$

Proof. See Appendix. $\square$

Inequality (1) requires that there is both a majority of voters with ideal points belonging to the region $(-\infty, d) \times \{0\}$ and to the region $(-d, \infty) \times \{0\}$. If this condition is satisfied, a non-regionalist moderate can successfully run for the federal election as an independent whenever parties nominate a regionalist

a)	$\Lambda_f([-2,-d] \times z)$	$\Lambda_f((-d,d) \times z)$	$\Lambda_f([d,2] \times z)$
$z = \{r\}$	0	0.01	0.02
$z = \{0\}$	0.48	0.44	0.05

b)	$\Lambda_f([-2,-d] \times z)$	$\Lambda_f((-d,d) \times z)$	$\Lambda_f([d,2] \times z)$
$z = \{r\}$	0	0.4	0
$z = \{0\}$	0.11	0.38	0.11

c)	$\Lambda_f([-2,-d] \times z)$	$\Lambda_f((-d,d) \times z)$	$\Lambda_f([d,2] \times z)$
$z = \{r\}$	0.1	0.1	0.1
$z = \{0\}$	0.15	0.4	0.15

Table 1: Examples of Distributions of Voters

Each entry in the tables gives the share of voters participating in the federal election who are located in the specified region.

and a non-regionalist partisan—a situation that must arise in any two-party equilibrium as soon as regionalism is the salient dimension. If there were no regionalist voters, Inequality (1) would always be satisfied by the assumption that the median of the distribution of voter ideal points along the ideological dimension lies at 0. The higher the share of regionalist voters, on the other hand, the easier it is to violate Inequality (1) and the likelier two-party equilibria are therefore to exist. The two situations in which independent candidates cannot successfully enter the federal election are either that there are relatively many regionalist voters or that the distribution of non-regionalist voters is heavily skewed in favour of either the platform $-2d$ or the platform $2d$. To illustrate this, Table 1 provides examples of distributions of voters. In panels a) and b) of the table, Inequality (1) is not satisfied and two-party equilibria may exist depending on the values of other parameters. In particular, it must be the case that $d \leq 1$, as was already discussed above. The subsequent discussion of the examples given in Table 1 assumes that $d \leq 1$ is satisfied.

In Panel a) of Table 1, Inequality (1) is violated as $\Lambda_f((-d,\infty] \times \{0\}) = 0.49$. While the share of regionalist voters is small, the high number of voters in favour of the policy $(-2d, 0)$ prevents that a moderate independent would have enough support to win the federal election when a regionalist politician and a politician with platform $(-2d, 0)$ are running for the federal election. On the other hand, the simultaneous nomination of a regionalist candidate and candidate with platform $(2d, 0)$ would create an opening for a moderate non-regionalist to run as an independent. To prevent that the latter constellation can arise, a two-party equilibrium cannot feature a regionalist party and a party allowing all non-regionalist politicians to join as in Proposition 6. A constellation of parties that can form a two-party equilibrium depending on other parameters is that, for example, one party allows politicians with platforms $(-2d, 0)$, $(0, 0)$, and $(0, r)$ to join, while the second party admits members with platforms $(2d, 0)$ and $(2d, r)$.[12]

In Panel b) of Table 1, non-regionalist voters are highly concentrated around the policy $(0, 0)$, but the number of regionalist voters is so large that Inequality (1) is nevertheless violated. If two parties nominate candidates for the federal election, one of them a regionalist and one of them a non-regionalist partisan, less than half of all voters would be strictly in favour of a non-regionalist moderate running as an independent. Accordingly, the assumed threshold required for voters to coordinate on electing the latter candidate is not met.

Note that $d > 1$ would imply that Inequality (1) is satisfied since $\Lambda_f([-1, 1] \times \{0\}) > 0.5$. However, the reverse is not true: Panel c) of Table 1 illustrates that Inequality (1) can also hold for any value of $d \leq 1$. In addition, Inequality (1) only has bite when r is sufficiently large. The necessary conditions established by Proposition 2 and Proposition 7 can therefore be fulfilled or violated independently.

To what extent do the results presented in this section provide an explanation for the existence of, for example, the Scottish National Party (SNP)? The consequence of the condition on the value of r in Proposition 7 is that non-regionalist voters prefer a moderate non-regionalist over any regionalist, which is plausibly the case in UK politics. In addition, Scottish voters account for less than 10 percent of the electorate in UK general elections, which makes it likely that Inequality (1) is satisfied as was argued above. The necessary conditions for entry of a third party provided by the model thus appear to be met in the

[12]To sustain the pair of parties given in the text as an equilibrium, it is crucial that the first party does not win any elections in regionalist states. As a consequence, the party always nominates a politician with platform $(-1, 0)$ or $(0, 0)$ for the federal election, which prevents entry of independent candidates. Nevertheless, it is also crucial that the same party allows regionalists to join. These politicians then serve as spoiler candidates that keep members of the second party with platform $(1, r)$ from defecting successfully.

case of the UK. Note that this section assumed throughout that there is at least one state/constituency where a majority of voters favours regionalism. That the SNP emerged as a political force only in the 1970s may thus be due to a lack of sufficient voter support for the issue of independence in the previous decades. As Proposition 7 makes clear, however, an up-tick in voter support is by itself not enough to predict entry of a regionalist party.

5 Robustness

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

5.1 Policy Choices

The assumption that politicians are committed to implementing their platform is not satisfying. While 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), it would be more appealing to see this behaviour emerge as part of an equilibrium rather than imposing it from the outset. Endowing politicians with policy preferences alongside their office motivations introduces two additional difficulties: First, partisan politicians may want to pretend to be a moderate when choosing state policy in order to increase their chances of winning the federal election. Second, politicians take into account how their choices affect the policies chosen by other politicians. Particularly the latter issue creates difficulties, since it is hard to track how the decision of a politician to join or not to join a party affects events in the federal election. However, it is possible to incorporate the role of state policies as signals of policy preferences without having to deal with the second type of complication. To do so, I follow Snyder & Ting (2002) and Ashworth & Bueno de Mesquita (2008) and assume that politicians only care about policy once elected. The utility of a politician with ideal policy $p \in \{-2d, 0, 2d\}$ is now be given by

$$\pi_s[y_s + \alpha\ v(|p_s - p|)] \ + \ \pi_f[y_f + \alpha\ v(|p_f - p|)]\ ,$$

where α measures the relative weight that politicians attach to policy, v is a continuous decreasing function with $v(0) = 0$, and the notation otherwise follows Section 2. Parties then allow only politicians with certain ideal policies to join. In addition, politicians can freely choose the policy they implement at any stage from the set $\{-2d, 0, 2d\}$. All other elements of the game remain unchanged.

The following result shows that politicians are always willing to forgo a higher chance of winning the federal election in favour of implementing their own ideal policy at the state level if α is sufficiently large.

Proposition 8. *Suppose $\alpha > -y_f/v(2d)$. Then any politician must implement their own ideal policy at any point of the game in any equilibrium.*

Proof. See Appendix. □

When all politicians always select their ideal point when choosing policy, the utility function considered here simplifies to the one assumed in the benchmark model and all results go through unchanged. In the case that α does not satisfy the condition of Proposition 8, separating equilibria may nevertheless exist depending on the technologies used to select candidates. However, when α is sufficiently small and candidates are mainly motivated by the spoils of office, partisans cannot be prevented from imitating moderates in order to increase their electoral prospects. Given this change in behaviour, it is not clear to what extent the results of the benchmark model carry over. The following proposition shows that at least the existence conditions of the equilibrium of Proposition 5 remain qualitatively unchanged even when all candidates pool on implementing the policy zero at the state level.

Proposition 9. *Assume that*

$$\alpha \leq -\frac{\left(\varepsilon + \frac{0.5-\varepsilon}{2(S-2)}\right)\ y_f}{2\ (S-1)\ v(2d)}\ .$$

Furthermore, suppose $\eta_{P,f} = \eta^{\varepsilon}$ for all $P \in \mathcal{P}$ with $\varepsilon \leq 0.5$ and $\eta_{P,s} = \eta^{u}$ for all $P \in \mathcal{P}$ and $s \in \{1, ..., S\}$. Then an equilibrium such that

i) $\mathcal{P}^* = \{L, R\}$ with $I_L = [-2d, 0]$ and $I_R = [0, 2d]$ and

ii) all politicians choose the policy 0 at the state level

exists if

i) $d \leq 1$ *and*

ii) $y_f/y_s \geq 2(S-1)/(\varepsilon + (0.5-\varepsilon)/2^{S-2})$.

Proof. See Appendix. □

The assumption on the value of α in the first sentence of Proposition 9 ensures that pooling on the policy 0 at the state level is consistent with equilibrium. Otherwise the existence conditions are essentially the same as those of Proposition 5. The only difference is in the level of career concerns required to support the equilibrium. The reason is that partisans increase their chances of winning the federal election by imitating moderates, which raises their equilibrium payoff and thus relaxes the requirements on the value of y_f/y_s.

5.2 Coalitional Deviations

Even though the formation of parties has previously been modelled as a non-cooperative game (Morelli 2004, Osborne & Tourky 2008), this is a setting where deviations by coalitions of players are a natural consideration. Members of the same party, in particular, are well placed to coordinate their actions. A prominent contribution to the literature on party formation allows for parties to be broken up by subsets of their membership (Levy 2004). In the model presented in the current paper, the incentives for joint deviations are limited. One example where such considerations matter would be a situation with three parties competing in a state election, where coordination failure among voters gives two politicians an incentive to cooperate and join the same party.[13] Since coordination failure does not occur when there are only two candidates—at least under the assumed restrictions on voting behaviour—the results on two-party equilibria are not subject to this issue. In the context of two-party equilibria, there are essentially two ways in which coalitional deviations could matter: The first is that politicians might jointly deviate to forming a single party. Unless at least one of the existing parties allows all types of politicians to join—which cannot be the case in equilibrium—all politicians would have to participate in this deviation. However, it is impossible that all politicians are made better off

[13] Considering the basic version of the model, suppose that there is a state where a third of voters is located at each of the policies $-2d$, 0, and $2d$. If all politicians have joined a singleton party, neither candidate is favoured by a strict majority of voters and each of them may win the election. Suppose the politician with platform $2d$ wins. Then the remaining politicians would have an incentive to jointly deviate and join a party of shape $[-2d, 0]$, since two thirds of voters would strictly prefer the candidate nominated by this party over a candidate with platform $2d$.

at once, given the constant-sum nature of the game. The second possibility is that a coalition of politicians could have an incentive to break up a party where a defection by a single politician would not be profitable. Cooperation of politicians across state borders on forming a third party would need to have the aim of affecting the outcome of the federal election. However, given that two-party equilibria require that no independent candidates can successfully run for the federal election, the same must be true of third party candidates. An exception would be that the coalitional deviation wipes out one of the equilibrium parties, but then it is not clear how this move would increase the payoffs of defectors as the deviation would have to include all previous members. Within a state, on the other hand, joint deviations by a strict subset of members would not be of relevance since there are no two-party equilibria where any party has more than two members in any state anyway.

While it thus seems unlikely that allowing for coalitional deviations poses a threat to the results presented above, coordination among politicians would be relevant in a slightly modified version of the model. In particular, the assumption that a single politician can form a new party can be dropped if one allows for coalitional deviations while assuming that there are multiple politicians with the same platform in each state. In many situations, all politicians with the same platform in a given state would join the same party in equilibrium[14] and these politicians then form a homogeneous faction within a possibly broader party. The same incentives that drive defections by individual candidates in the basic version of the model would then apply to such a faction. Suppose, for example, that all politicians with platform $-2d$ in state 1 have joined a party that also has as least some politicians with other platforms as members in the same state. If all politicians with platform $-2d$ in the state jointly deviate and join a party of shape $[-2d]$, any member of the group would benefit from a reduction in internal competition, while any payoffs received after the state nomination has been decided do not depend on the number of party members in the state anyway. The perhaps unrealistic assumption that a single politician can form a new party could thus be replaced with the assumption that a minimum number of members is required for a new party to compete, at least as long as this number is no larger then the number of politicians with a specific platform in each state. In this sense, the politicians in the original model can be thought of as representing a faction rather than an individual person.

[14]The only possible exception may occur when there is a tie in a state election.

6 Conclusion

Why are the same two parties competing in elections in the US across all levels of government, while more than two parties attract significant vote shares in other countries relying on FPTP such as the UK or Canada? The model of party formation presented in this paper provides a number of potential explanations. Specifically, any two-party equilibrium of the model requires that a number of conditions are satisfied: First, there cannot be any parties that allow all politicians to join as such parties are too vulnerable to defections. Second, the most radical median voters in regional elections cannot lie in the tails of the distribution of voters overall, since this situation would force parties to extend their memberships too far to the extremes, creating the risk of entry of a centrist party or independent candidate. Third, politicians need to be sufficiently motivated by career concerns to prevent them from joining parties more targeted at the preferences of voters in specific regions. Finally, an issue that splits regions into two camps, such as ethnic cleavages or an independence movement, must either be less salient than the classical left-right divide or it must be the case that regionalist voters make up a sufficiently large share of the electorate. If any of these conditions fail, only equilibria with three or more parties exist. In this sense, the necessary conditions for the existence of two-party equilibria suggest explanations for the emergence of new parties in different settings. An interesting prediction of the model is that only centrist parties should be expected to be formed top-down by politicians in advanced stages of their career,[15] while parties located in the political wings emerge as grassroots movements at the regional or local level.

In the absence of a regionalist movement and subject to the conditions given above, the model has an equilibrium featuring a centre-left and a centre-right party very much in line with the party system of the US. More generally, any two-party equilibrium must feature a left-leaning and a right-leaning party. When regionalism is salient, on the other hand, two-party equilibria may feature a division of parties into a regionalist and a non-regionalist camp.

Finally, it is worth highlighting that the model introduced in this paper provides a flexible tool for investigating other questions surrounding political

[15] Strictly speaking, the model does not allow for the formation of parties by politicians who have already moved beyond the state level. However, the model does allow the possibility of independent candidates in the national election. In parliamentary systems, such as the UK, independent candidates would essentially be forced to form a new party since becoming prime minister requires support of a majority in Parliament.

parties, such as the strategic choice of the rules that govern the nomination of candidates, for example. While I have argued that parties have an incentive to commit to giving different types of candidates a shot at winning nominations, I assumed the mechanisms used for candidate selection to be exogenously determined. However, endogenising this feature of parties would be feasible. The section on regionalism, though relatively simple, indicates that the model can generate insights even under multidimensional policy spaces.

Appendix: Proofs

The following three lemmas are not presented in the text:

Lemma 3. *The following politicians win the election in their state with certainty whenever they join a party that allows no other type of politician to join: the politician with platform* $-2d$ *in state 1, the politician with platform 0 in state 2, and the politician with platform* $2d$ *in state 3.*

Proof. Suppose that the politician with platform $-2d$ in state 1 joins a party D of shape $I_D = [-2d]$. Since no other politicians in the state are able to join this party, voters believe that the candidate of party D has platform $-2d$ with certainty, in or out of equilibrium. The expected utility of a voter with ideal policy i if the candidate of party D is elected is then given by $u(|-2d-i|)$. As u is strictly decreasing, it holds for any $i < -d$ and $0 \leq \pi \leq 1$ that

$$\begin{aligned} u(|-2d-i|) &> u(|0-i|) \\ &\geq \pi\ u(|0-i|) + (1-\pi)\ u(|2d-i|)\ . \end{aligned}$$

The right-hand side of the preceding inequality subsumes the expected utility of a voter with ideal point i if any candidate other than the candidate of party D is elected. Since $m_1 < -d$, this implies that a strict majority of voters strictly prefers the candidate of party D over any other candidate and the politician with platform $-2d$ thus wins the state election with certainty. An analogous argument shows that the politician with platform $2d$ in state 3 wins the state election with certainty whenever this politician joins a party with shape $[2d]$.

Finally, suppose that the politician with platform 0 in state 2 joins a party D of shape $I_D = [0]$. Voters then believe that the candidate of party D has platform 0 with certainty. The expected utility of a voter with platform i if the candidate of party D is elected is then given by $u(|0-i|)$. For any $-d < i < d$ and $0 \leq \pi \leq 1$ it holds that

$$u(|0-i|) > \pi\ u(|-2d-i|) + (1-\pi)\ u(|2d-i|)\ .$$

The right-hand side of the preceding inequality subsumes the expected utility of a voter with ideal point i if any candidate other than the candidate of party D is elected. Since $\Lambda_2((-d,d)) > 0.5$, this implies that a strict majority of voters strictly prefers the candidate of party D over any other candidate and the politician with platform 0 thus wins the state election with certainty. □

Lemma 4. *In any two-party equilibrium, each party must nominate a partisan for the federal election with positive probability.*

Proof. Suppose there was a two-party equilibrium such that $\mathcal{P}^* = \{A, B\}$ and party A never nominates a candidate other than a moderate for the federal election. This implies that no partisan joins party A in any state where party A wins the state election with positive probability, since any member of the federal candidate pool is nominated with positive probability. It follows that party B allows both types of partisans to join, as otherwise either the politician with platform $-2d$ in state 1 or the politician with platform $2d$ in state 3 would have no chance of winning any election and would prefer to join a singleton party by Lemma 3. Furthermore, in any state where party B wins with positive probability both partisans must join the party by Lemma 2.

Two cases will be considered separately. First, suppose that party A always nominates a moderate for the federal election along the equilibrium path. This implies that the politician with platform $-2d$ in state 1 has no chance of winning the federal election, since there are at most two candidates in any federal election in equilibrium and $m_f = 0$. The politician must therefore join a party that maximises their chance of winning the state election. By Lemma 3, this politician can win the state election with certainty by joining a singleton party. In contrast, as a member of party B the politician competes with the other partisan of the state for the nomination. The politician with platform $-2d$ in state 1 therefore has an incentive to deviate and join a party of shape $[-2d]$.

Now suppose that party A never nominates a partisan for the federal election, but sometimes does not nominate a candidate at all, which requires that party A wins any state election with less than certainty. Suppose that party A wins the election in state 2 with positive probability. As was argued above, this implies that both partisans of the state join party B with certainty. In equilibrium voters therefore believe that the candidate nominated by party A has the platform 0. Consequently, party A must win the state election with certainty since $\Lambda_2((-d, d)) > 0.5$. Therefore, party B can only win the election in state 2 if the moderate candidate of the state has joined the party and party A does not nominate a candidate. Conditional on having won all state elections, the situation arises with positive probability that party B nominates a partisan for the federal election while the moderate of state 2 has won the state election there. Since the state elections resolved any uncertainty about the platforms of these politicians, the moderate can run as an independent and defeat the partisan given that $m_f = 0$, upsetting the equilibrium. $\square$

Lemma 5. *Let P be a party of shape $[-2d, 0]$ or $[0, 2d]$ that wins $w > 0$ state elections in a pure-strategy equilibrium. If $\eta_{P,f} = \eta^{\varepsilon}$ and $\eta_{P,s} = \eta^{u}$ for all $s \in \{1, ..., S\}$, the probability that party P nominates a partisan for the federal election is given by*

$$\varepsilon + \frac{0.5 - \varepsilon}{2^{(w-1)}} \ , \tag{2}$$

while the probability that a particular partisan is nominated is given by

$$\frac{2}{w}\left(\varepsilon+\frac{0.5-\varepsilon}{2^{(w-1)}}\right)\ . \tag{3}$$

Proof. In a pure-strategy equilibrium, party P must be joined by both eligible politicians in any state where the party wins the state election by Lemma 2. Given that the party uses the nomination technology η^u in all states, any politician who wins a state election as a member of party P is therefore equally likely to be a moderate or a partisan. Any combination of platforms of the w state winners thus occurs with probability $1/2^w$. There are 2^w such combinations, all but two of which contain both moderates and partisans. If the pool of candidates contains only moderates, a partisan is nominated with probability zero, while a partisan is nominated with probability one in the opposite case. In any other case a partisan is nominated with probability ε. Overall, the probability that party P nominates a partisan is given by

$$\frac{1}{2^w}\left[(2^w-2)\varepsilon\ +1\right]\ ,$$

which can be rewritten in the form given in the statement of the lemma. The probability that any given partisan is nominated is just the probability that a partisan is nominated divided by the average number of partisans in the candidate pool, which is equal to $w/2$. □

Proof of Lemma 1. Consider a two-party equilibrium such that $\mathcal{P}^*=\{A,B\}$ and suppose there is a platform $p\in\{-2d,0,2d\}$ such that $p\notin I_A\cup I_B$. This implies that either the politician with platform $-2d$ in state 1, the politician with platform 0 in state 2, or the politician with platform $2d$ in state 3 neither joins a party in equilibrium nor runs as an independent and thus receives a payoff of zero. By Lemma 3, however, either of these politicians can guarantee themselves a payoff of at least $y_s>0$ by joining a party that allows no other type of politician to join, contradicting equilibrium. □

Proof of Lemma 2. Recall that voters observe whether a party nominates a candidate for a state election, but not how many politicians have joined a party. Voting behaviour can therefore only be conditional on which parties have nominated candidates. Let party A be a party that wins the election in some state s with positive probability. If party A wins with positive probability due to the other party not nominating a candidate, then this event does not become any less likely due to additional politicians joining party A. Suppose the other party nominates a candidate with certainty. Then party A must have a chance of winning in this situation as well, given that party A was assumed to win with positive probability. But then party A continues to win the state election if additional members join, either because voters cannot detect that this has occurred, or because the second party has no members left. Party A therefore always continues to win the state election if any politician joins the party with higher probability. Joining party A in state s therefore yields a positive payoff and any politician who has this option can therefore never chose the strategies of remaining passive or joining a party that loses the state election with certainty, given that the latter two strategies lead to a payoff of zero. $\square$

Proof of Proposition 1. Suppose there was a two-party equilibrium such that $\mathcal{P}^* = \{A, B\}$ with $[-2d, 2d] \in \{I_A, I_B\}$. Without loss of generality, let $I_A = [-2d, 2d]$. It will be shown that there must be one type of partisan such that both parties are joined by at least one such politician with positive probability. Towards a contradiction, assume that all politicians with platform $-2d$ join party A with certainty while all politicians with platform $2d$ join party B. This assumption is without loss of generality as party B must allow at least one type of partisan to join by Lemma 4. This situation would have the consequence that party A wins the election in state 1 with certainty independent of the choice of party affiliation of the moderate in the state since $m_1 < -d$. But by Lemma 2 this would contradict that the politician with platform $2d$ in the same state joins party B given eligibility to join party A.

Without loss of generality, let politicians with platform $2d$ be the type that joins both parties with positive probability. Since every state election is won by one of the two equilibrium parties, all politicians with platform $2d$ must join one of them or mix over joining either by Lemma 2. Accordingly, at least one party must be joined by more than one politician with platform $2d$ with positive probability. Given that every member of a party is nominated for the election in their state with positive probability, there is a point on the equilibrium path where both parties have a member with platform $2d$ who has won a state election. As each winner of a state election is nominated for the federal election with probability greater than zero, it is possible that both parties nominate a politician with platform $2d$. But then a moderate winner of a state election would have an incentive to run as an independent for the federal election. Given that $m_f = 0$ and that the state elections have resolved any uncertainty about the platforms of these politicians, the independent candidate would be favoured by a strict majority over any candidate with platform $2d$. Accordingly, the only way to rule out that an independent candidate runs for the federal election and upsets the equilibrium would be if the nomination of two candidates with platform $2d$ implies that no moderate has won a state election. If both parties allow moderates to join, then any moderate must win the election in their state with positive probability by Lemma 2. Since the identities of state winners are independent of each other, it then cannot be ruled out that a moderate has won a state election while both parties nominate candidates with platform $2d$. Accordingly, assume that $I_B = [2d]$. But then party A must win the election in states 1 and 2 with certainty. In the case of state 1, this is true as soon as any politician joins party A, who therefore also has an incentive to do so. In the case of state 2, party A not winning would imply that the moderate politician of the state has an incentive to join a singleton party by Lemma 3. All politicians must then join party A in states 1 and 2 by Lemma 2. Accordingly, it can occur that party A nominates a politician with platform $2d$ for the federal election while a moderate has also won a state election for party A, given that any member of the candidate pool is nominated with positive probability. Said moderate politician then has an incentive to run as an independent for the federal election. □

Proof of Proposition 2. Suppose three candidates with different and known platforms are running for the federal election. Since voters prefer candidates with platforms close to their ideal points, a voter with ideal point i strictly prefers the moderate candidate over any other candidate if and only if $i \in (-d, d)$. If $\Lambda_f((-d, d)) > 0.5$, the moderate candidate must then win the federal election by the definition of party-formation equilibrium.

Now consider the following situation: in a two-party equilibrium, one party has nominated a candidate with platform $-2d$ for the federal election while the other has nominated a candidate with platform $2d$. At the same time, there is a moderate politician who has won a state election. Suppose this moderate deviates and runs for the federal election as an independent candidate. Then the independent candidate wins by the previous paragraph, given that the state election has resolved any uncertainty about the identity of the politician in question. The deviation is therefore profitable.

To complete the proof, it will be shown that the situation described in the previous paragraph must arise with positive probability in any two-party equilibrium. Proposition 1 and Lemma 1 in combination imply that in any two-party equilibrium one party must allow politicians with platform $-2d$ to join but not politicians with platform $2d$, while the reverse is true for the other party. Refer to these parties as A and B, respectively. Party A (party B) must then win the election in state 1 (state 3) with positive probability with the politician with platform $-2d$ (platform $2d$) as a member. Otherwise one of these politicians would have a profitable deviation by Lemma 3. As any winner of a state election is nominated by their party with positive probability for the federal election, it thus happens with probability greater than zero that party A nominates a politician with platform $-2d$ while party B nominates a politician with platform $2d$. At the same time, the politician with platform 0 in state 2 must win the state election with positive probability by Lemma 3. As the identities of state winners are independent of each other conditional on the set of existing parties, it thus occurs with positive probability that a candidate with platform $-2d$ and a candidate with platform $2d$ are nominated for the federal election while a moderate has won the election in state 2. □

Proof of proposition 3. Consider a two-party equilibrium such that $\mathcal{P}^* = \{A, B\}$. Any politician can only win a state election with certainty in a two-party equilibrium if they are member of a singleton party with certainty. Otherwise at least one of the parties that the politician joins with positive probability would attract other members with positive probability by Lemma 2. As a consequence, there would be a non-zero chance that the politician misses out on the nomination for the state election. Given that any politician can only win a state election with certainty by joining singleton parties, it is impossible that the politician with platform $-2d$ in state 1, the politician with platform 0 in state 2 and the politician with platform $2d$ in state 3 all win the election in their state with certainty, since this would require that three singleton parties are active. Given that no independent candidates run in a two-party equilibrium, at least one of these politicians therefore achieves a payoff in equilibrium of at most

$$\pi(y_s + y_f) \ ,$$

where $0 < \pi < 1$ denotes the probability that the politician wins the state election.

By Lemma 3, the same politician can achieve a payoff of at least y_s by deviating and joining a singleton party. This deviation is profitable if $y_s > \pi(y_s + y_f)$, which can be rewritten as $y_f/y_s < (1-\pi)/\pi$ where $(1-\pi)/\pi > 0$. A necessary condition for the existence of the equilibrium is therefore $y_f/y_s \geq (1-\pi)/\pi > 0$. □

Proof of Proposition 4. Suppose $\mathcal{P}^* = \{L, R\}$, all politicians use pure strategies, every politician joins one of the two active parties, and no politician ever runs as an independent. Since moderates can join either party, assume that moderates in a state such that $m_s \leq 0$ join party L and all other moderates join party R. This implies that in each state the party joined by the moderate wins, independent of nomination technologies. Party L thus wins the election in state 1 and party R the election in state 3, so both parties compete in the federal election and one of the candidates closest to the federal median voter located at zero wins. Assume indifferent voters mix such that either candidate wins with equal probability when two partisans or two moderates have been nominated.

It needs to be verified that no politician has an incentive to deviate. First, it will be shown that no politician has an incentive to run as an independent in the federal election if $\Lambda_f((-d, d)) \leq 0.5$. An independent candidate running out of equilibrium would have to be strictly preferred by a strict majority of voters over any other candidate, which is only possible if the platform of this candidate is equal to 0 with positive probability and both of the party candidates are partisans. The only situation in which an independent candidate could win is therefore that party L has nominated a politician with platform $-2d$ and party R a candidate with platform $2d$. Voters in the interval $(-\infty, -d)$ (in the interval (d, ∞)) strictly prefer a candidate with known platform $-2d$ (platform $2d$) over any candidate who has platform 0 with positive probability, while voters located at -d (at d) would at least weakly prefer the latter candidate. This implies that there is no strict majority strictly in favour of an independent candidate if $\Lambda_f((-d,d)) \leq 0.5$.

Next consider a politician who decides to run as an independent in a state election. As politicians always accept their party's nomination, only a politician whose party has two members in the state can consider becoming an independent. After this out-of-equilibrium event, it is possible to assign voters either the belief that the partisan runs as an independent and the moderate is running for the party or the reverse case. Since it is impossible that a strict majority of voters strictly prefers the independent in both cases, the equilibrium can be constructed such that an independent in a state election does not win.

Finally, it needs to be shown that no politician wants to change their party affiliation. Only moderates can switch between equilibrium parties. If a moderate does so, voters do not observe this deviation. Accordingly, the party abandoned by the moderate continues to win and the deviation is not profitable. Furthermore, no politician prefers not to join a party since this option gives a payoff of zero. In order to demonstrate that no politician wants to join a previously passive party, it is sufficient to show that no politician wants to join a singleton party. This is true as it is always possible to assign voters the same belief after a politician has deviated and joined a party that allows more than one type to join as in the case of a deviation to a singleton party. Suppose thus that a politician deviates and joins a singleton party instead of party $P \in \{L, R\}$. A partisan who has no chance of winning a state election already is the sole member of their party in their state and essentially does not change their situation by joining a singleton party. Any other politician can achieve a payoff no higher than y_s by deviating, as the same argument that was used to show that independent candidates cannot win the federal election also applies to third-party candidates. In equilibrium, on the other hand, such a politician wins the federal election with positive probability. This is true by the assumption that each member of the candidate pool for each election is nominated with non-zero probability and because each candidate wins the federal election with equal probability when two partisans or two moderates compete. For a particular politician, who is nominated for and wins their state elections with probability $0 < \pi_s < 1$ and, conditional on doing so, is nominated for and wins the federal election with probability $0 < \pi_f < 1$, deviating to joining a singleton party is therefore not profitable if

$$\pi_s(y_s \ + \ \pi_f \ y_f) \geq y_s \Leftrightarrow \frac{y_f}{y_s} \geq \frac{1-\pi_s}{\pi_s \pi_f} > 0 \ .$$

There thus exists a threshold $\bar{y} > 0$ such that no politician has a profitable deviation if $y_f/y_s \geq \bar{y}$. □

Proof of Proposition 5. Building on the proof of Proposition 4, it only needs to be shown which level of y_f/y_s is required to prevent that politicians deviate and join singleton parties. Suppose thus that a partisan deviates and joins a singleton party instead of party $P \in \{L, R\}$. A partisan who has no chance of winning a state election already is the sole member of their party in their state and essentially does not change their situation by joining a singleton party. A partisan who can win in equilibrium, on the other hand, can achieve a payoff no higher than y_s by deviating, as established by the proof of Proposition 4. In equilibrium, on the other hand, such a politician achieves a payoff of

$$0.5 \left[y_s + \frac{2}{w^*} \left(\varepsilon + \frac{0.5 - \varepsilon}{2^{w^*-1}} \right) \left(\varepsilon + \frac{0.5 - \varepsilon}{2^{S-w^*-1}} \right) \frac{1}{2} \, y_f \right] ,$$

where w^* denotes the number of state elections won by party P. The equilibrium payoff takes this shape since the politician faces one competitor for the state nomination and is thus nominated with probability one-half; is nominated for the federal election with a probability given by Equation (3); and can only win the federal election if tying against another partisan. The probability that the competing party nominates a partisan is given by Equation (2), keeping in mind that this party wins $S - w^*$ state elections. The condition that ensures that a partisan does not profit from the deviation is thus

$$\frac{y_f}{y_s} \geq \frac{w^*}{\left(\varepsilon + \frac{0.5-\varepsilon}{2^{w^*-1}} \right) \left(\varepsilon + \frac{0.5-\varepsilon}{2^{S-w^*-1}} \right)} .$$

The right-hand side of this inequality is increasing in the number of states won by party P and in the number of states won by the competing party. As neither party can win more than $S - 1$ state elections, a sufficient condition for the deviation by the partisan not to be profitable is given by

$$\frac{y_f}{y_s} \geq \frac{S - 1}{\left(\varepsilon + \frac{0.5-\varepsilon}{2^{S-2}} \right)^2} .$$

In the case of a moderate joining a singleton party, the highest possible deviation payoff is the same as for a partisan since third party candidates cannot win the federal election under the same conditions as independent candidates. The equilibrium payoff, on the other hand, is higher as a moderate is (weakly) more likely to be nominated for and win the federal election. The condition that ensures that partisans do not profit from any deviation is therefore sufficient to ensure the same is true for moderates. □

Subsequent proofs refer to the extended model introduced in Section 4.

Proof of Proposition 6. To complete the description of the equilibrium, assume that all eligible politicians have joined party N or party R, while all remaining politicians remain passive. Since all voters are located in the interval $[-1, 1]$ along the ideological dimension and $r > 2$, any non-regionalist voter prefers any non-regionalist candidate with platform $(p, 0)$ over any regionalist candidate with platform (p', r) and vice versa.

This is true as

$$u(|p-i|)+u(0)\geq u(2)+u(0)>u(r)+u(0)\geq u(|p'-i|)+u(r)\ .$$

Accordingly, party N wins the elections in states 1, 2, and 3, party R wins the election in state 4, and any candidate of party N wins the federal election. The politician with platform $(-1,r)$ in state 4 thus achieves a payoff of y_s and cannot improve on this through any deviation as long as party N nominates a candidate for the federal election.

Consider the possibility of any independent or third-party candidates other than the politician with platform $(-1,r)$ from state 4 running in the federal election. If a non-regionalist enters the race, a fourth of all voters strictly prefers the candidate of party R over the entrant while another fourth of voters at least weakly prefers the candidate of party N. If a regionalist enters, all voters in states 1, 2, and 3 strictly prefer the candidate of party N. Therefore no strict majority strictly prefers any third candidate and a voting equilibrium can be constructed such that these candidates do not win the federal election. In state elections, on the other hand, beliefs over the platform of any independent candidates can be assigned such that these candidates do not win.

It remains to be verified that no politicians can gain from changing their party affiliation. Deviating to not joining a party is never profitable. In state 4, at least half of all voters strictly prefer the candidate of party R over any other candidate and no additional party can enter successfully. Suppose a politician in any other state could win the state election by joining a third party. This deviation achieves a payoff of at most y_s, since third party candidates cannot win the federal election by the argument above. In equilibrium, in contrast, each of the politicians in states 1, 2, and 3 wins the federal election with positive probability, receiving a payoff of $\pi_s(y_s+\pi_f y_f)$ where π_s and π_f are positive probabilities. The deviation is therefore not profitable if $y_s\leq\pi_s(y_s+\pi_f y_f)$, which can be rewritten as

$$\frac{y_f}{y_s}\geq\frac{1-\pi_s}{\pi_s\pi_f}\ . \tag{4}$$

The right-hand side of the preceding inequality is strictly greater than zero since none of the politicians in states 1, 2, or 3 are nominated for the state election with certainty. □

Proof of Proposition 7. As a first step, it will be shown that there exists a threshold $\bar{r}_d$ such that all voters with ideal point $(i,j)\in[-1,1]\times\{0,r\}$ strictly prefer a candidate with platform $(0,j)$ over any candidate with platform $(p,r-j)$ if $r>\bar{r}_d$. The statement is true since when $r > 1$ it holds that

$$u(|0 - i|) + u(0) \geq u(1) + u(0) > u(r) + u(0) \geq u(|p' - i|) + u(r)$$

for any of the ideal points under consideration. The threshold $\bar{r}_d$ therefore exists and is no larger than 1.

Next, it will be established that there must be at least one active party in equilibrium that allows regionalists to join when $r > \bar{r}_d$. Suppose there was an equilibrium such that no regionalist politician joins a party. Accordingly, all of these politicians receive a payoff of zero. By assumption, there exists at least one state $s \in \mathcal{S}^r$ such that $\Lambda_s([-1, 1] \times \{r\}) > 0.5$. Suppose the politician with platform $(0, r)$ in this state joins a singleton party. If $r > \bar{r}_d$, all voters in the region $[-1, 1] \times \{r\}$ and thus a strict majority prefer the candidate of the newly-formed party by the previous paragraph. By the definition of two-party equilibrium, the entrant must win the state election and the deviation is profitable.

Now assume $r > \bar{r}_d$ and consider a two-party equilibrium such that $\mathcal{P}^* = \{A, B\}$. Let s be a state such that $\Lambda_s([-1, 1] \times \{r\}) > 0.5$, which exists by assumption. By the previous paragraph there must exist at least one party that allows at least one regionalist to join and this party must win the election in state s with positive probability. Let this party be party A.

To demonstrate that Inequality (1) implies that no two-party equilibrium exists, assume that the condition is satisfied. Then a strict majority of voters in the federal election prefers a candidate with platform $(0, 0)$ over any partisan or regionalist candidate. Accordingly, a winner of a state election with platform $(0, 0)$ who has not been nominated for the federal election would have an incentive to run as an independent whenever one party nominates a partisan and one party a regionalist or whenever both parties nominate a regionalist. Equilibrium therefore requires that the preceding constellations of candidates are either impossible or imply that no politician with platform $(0, 0)$ has won a state election. Two cases need to be distinguished.

First, suppose that a politician with a platform other than $(0, 0)$ can win a state election as a member of party B in equilibrium. Then it can occur that party B nominates this politician for the federal election while party A nominates a regionalist. To prevent entry of an independent candidate who is a known non-regionalist moderate, this constellation of candidates must imply that no politician with platform $(0, 0)$ has simultaneously won a state election. By Lemma 3, such a politician must win the election in state 2 with positive probability. If she does so as a member of party A, it cannot be ruled out that she has an opportunity to run for the federal election as an independent. If she does so as a member of party B, on the other hand, ruling out such an opportunity requires that the politician with a platform other than $(0, 0)$ who can win a state election as member of party B also belongs to state 2 and thus has platform $(-2d, 0)$ or $(2d, 0)$. Since the latter politician may be nominated for the federal election, it follows that no politician with platform $(0, 0)$ can win in any state other than states s or 2 for any party. Since non-regionalist moderates are allowed to join party B, this party cannot win the elections in state 1 or 3. But by Lemma 3, party A winning both state 1 and state 3 is only consistent with equilibrium if this party allows politicians with platforms $(-2d, 0)$ and $(2d, 0)$ to join. Convexity of party shapes then implies $(0, 0) \in I_A$ and a politician with this platform must thus win the election in either state with positive probability. Subsequently, these politician find themselves in a situation where they have an incentive to run as an independent with positive probability.

The second case to consider is that only politicians with platform $(0, 0)$ can win state elections as a member of party B. This is impossible as shown by the proof of Lemma 4, which applies to the extended model since it focuses on the behaviour of politicians in states 1, 2, and 3. $\square$

Proof of Proposition 8. Consider a politician who has won the election in their state. If they implement their own ideal policy, this could imply at worst that they win the federal election with probability 0. The corresponding continuation payoff would be 0. In contrast, choosing any other policy at best yields a payoff of $\alpha v(2d) + y_f$, which would be the case if the politician subsequently wins the federal election with certainty and implements their own ideal policy federally. Choosing the own ideal policy at the state level leads to a strictly greater payoff if $\alpha > -y_f/v(2d)$, taking into account that $v(2d) < v(0) = 0$. $\square$

Proof of Proposition 9. Suppose $\mathcal{P}^* = \{L, R\}$, all politicians use pure strategies, every politician joins one of the two active parties, and no politician ever runs as an independent. Furthermore, let all politicians implement their ideal policy when elected at the federal election, while all politicians implement policy 0 when elected at the state level. Since moderates can join either party, assume that moderates in a state such that $m_s \leq 0$ join party L and all other moderates join party R. Voters in any state election are indifferent between both parties, since all candidate implement the same policy at the state level and it was assumed that voters do not look forward to the federal election when voting in a state election. The equilibrium can therefore be constructed such that in each state the party joined by the moderate wins. Party L thus wins the election in state 1 and party R the election in state 3 and both parties compete in the federal election. Given that parties use the nomination technology η^u at the state level, any candidate for the federal election is a moderate with probability one-half. The federal median voter is thus indifferent and both parties tie at the federal election.

It needs to be verified that no politician has an incentive to deviate. Since the game ends after the choice of the federal policy, the winner of the federal election always implements her own ideal policy. Continuing to proceed by backwards induction, consider entry of an independent candidate in the federal election. Given that any politician is a partisan with positive probability at this point of the game, the belief can be assigned to voters that any independent running out of equilibrium is a partisan with certainty. Accordingly, such a candidate is not strictly preferred by a strict majority of voters over any other candidate and a voting equilibrium can be selected such that the candidate loses the federal election. Deviations may also occur at the point of choosing policy at the state level. Let out-of-equilibrium beliefs be such that voters believe that a politician who implements a policy other than 0 is a partisan with certainty. As a consequence, any politician who does so is sure to lose the federal election. Moderates therefore certainly have no incentive to choose the policy $-2d$ or $2d$ at the state level, since this would make them both less likely to win and suffer a loss from implementing a policy different from their most preferred one. Partisans, on the other hand, may benefit from choosing their ideal policy instead of the policy 0. In equilibrium, the continuation payoff of such a politician at the point when state policy is chosen is given by

$$\alpha \ v(2d) + \frac{2}{w^*} \left(\varepsilon + \frac{0.5 - \varepsilon}{2^{(w^*-1)}} \right) \frac{1}{2} \ y_f, \tag{5}$$

where w^* denotes the number of state elections won by the party of the politician. The probability that the politician is nominated for and wins the federal election is derived from Equation (3) in combination with the feature of the equilibrium that all candidates tie in the federal election along the equilibrium path. If the politician deviates and implements their ideal policy, the continuation payoff is equal to zero, since there is no loss from policy and the politician loses the federal election. The deviation is therefore not profitable if

$$\alpha \ \leq -\frac{\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(w^*-1)}} \right) \ y_f}{w^* \ v(2d)} \ .$$

As the right-hand side of the preceding inequality is decreasing in w^* and $w^* \leq S-1$, this condition is satisfied if

$$\alpha \ \leq -\frac{\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}} \right) \ y_f}{(S-1) \ v(2d)} \ .$$

Since

$$-\frac{\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}} \right) \ y_f}{(S-1) \ v(2d)} \ > -\frac{\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}} \right) \ y_f}{2 \ (S-1) \ v(2d)} \ , \tag{6}$$

the assumption on the value of α in the statement of the proposition is sufficient to ensure that the deviation is not profitable.

Next, consider a politician who decides to run as an independent in a state election or joins a third party. In these cases it is possible to assign the same beliefs to voters as if the politician had joined a singleton party. It is therefore sufficient for existence that the latter type of deviation is not profitable. Joining a singleton party fully reveals a politician's ideal policy to voters. If this politician subsequently runs for the federal election, she can only be strictly preferred by a strict majority of voters over any other candidate if she has ideal policy 0. Otherwise the median voter and all voters either to the left or the right of 0 would prefer either the candidate of party L or party R. In case the politician does have the ideal policy 0, only voters located in the interval $(-d, d)$ strictly prefer this candidate. The condition $\Lambda_f((-d, d)) \leq 0.5$ therefore ensures that the equilibrium can be constructed such that independent or third-party candidates do not win the federal election. If the deviating politician wins the election in their state, she must accordingly implement her ideal policy. This implies that a politician who cannot win a state election in equilibrium cannot win after joining a singleton party either, as the median voter of the state prefers the policy zero by construction of the equilibrium. The only politicians who can achieve a positive payoff after joining a singleton party are those who win a state election with positive probability in equilibrium. Among these, partisans achieve the lowest equilibrium payoff, which, building on Expression (5), is given by

$$\begin{aligned}
&\frac{1}{2}\left[y_s + \alpha\ v(2d) + \frac{1}{w^*}\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(w^*-1)}}\right)\ y_f\right] \\
\geq\ &\frac{1}{2}\left[y_s + \alpha\ v(2d) + \frac{1}{S-1}\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}}\right)\ y_f\right]
\end{aligned}$$

since the payoff is decreasing in the number of states won and neither party wins more than $S-1$ states. Given the arguments above, a deviation to joining a singleton party by such a politician yields at most a payoff of y_s. The condition that

$$\alpha \leq -\frac{\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}}\right)\ y_f}{2\ (S-1)\ v(2d)}$$

is equivalent to the condition

$$\alpha\ v(2d) + \frac{1}{S-1}\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}}\right)\ y_f \geq \frac{1}{2(S-1)}\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}}\right)\ y_f\ .$$

Using this latter formulation, it holds that

$$\frac{1}{2}\left[y_s + \alpha\ v(2d) + \frac{1}{S-1}\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}}\right)\ y_f\right] \geq \frac{1}{2}\left[y_s + \frac{1}{2(S-1)}\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}}\right)\ y_f\right]\ .$$

Given the assumption on the value of α in the statement of the proposition, a sufficient condition ensuring that deviating to joining a singleton party is not profitable is therefore

$$\frac{1}{2}\left[y_s + \frac{1}{2(S-1)}\left(\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}}\right)\ y_f\right] \geq y_s\ ,$$

which can be rewritten as

$$\frac{y_f}{y_s} \geq \frac{2(S-1)}{\varepsilon + \frac{0.5-\varepsilon}{2^{(S-2)}}}\ .$$

Finally, it needs to be shown that no politician wants to remain passive. This deviation would only be profitable if a politician incurs a negative equilibrium payoff, which can only occur when a politician implements a policy different from their ideal policy. However, it was already verified above that all politicians achieve a positive payoff, as otherwise partisans would prefer to implement their ideal policy after winning a state election. □

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.

Arceneaux, K. & Nicholson, S. (2012), 'Who wants to have a tea party? The who, what, and why of the Tea Party movement', *PS: Political Science & Politics* **45**(4), 700–710.

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

Barrie, D. & Gibbins, R. (1989), 'Parliamentary careers in the Canadian federal state', *Canadian Journal of Political Science* **22**(1), 137–145.

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.

Burns, A. & Haberman, M. (2016), 'Bloomberg is considering 3rd-party bid', *New York Times*, January 24 2016, p. 1.

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.

Crewe, I. & King, A. (1995), *SDP: The Birth, Life and Death of the Social Democratic Party*, Oxford University Press.

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.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N. & Westwood, S. (2019), 'The origins and consequences of affective polarization in the united states', *Annual Review of Political Science* **22**(1), 129–146.

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

Johnston, R. (2017), *The Canadian party system: an analytic history*, UBC Press.

Kaelberer, M. (1998), 'Party competition, social movements and postmaterialist values: Exploring the rise of green parties in France and Germany', *Contemporary Politics* **4**(3), 299–315.

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.

McKee, S. & Yoshinaka, A. (2015), 'Late to the parade: Party switchers in contemporary US southern legislatures', *Party Politics* **21**(6), 957–969.

Mény, Y. (2017), 'A tale of party primaries and outsider candidates: the 2017 French presidential election', *French Politics* **15**(3), 265–278.

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.