DOCTORAL THESIS

# Parties, Interest Groups, and Political Outcomes

*Author:*
Nicolas MOTZ

*Supervisor:*
Professor Ian PRESTON

*A thesis submitted in fulfilment of the requirements*
*for the degree of Doctor of Philosophy*

*in the*

Department of Economics

June 2015

# Declaration of Authorship

I, Nicolas MOTZ, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Signed:
_____

Date:
_____

# *Abstract*

This thesis is concerned with the role of parties and interest groups in politics. The reason for the presence of these organisations is seen in the limited information of voters. Two chapters explore the idea that the role of parties is to provide information about politicians to voters. It is asked whether this concept of parties can be used to explain systematic differences in the degree of political competition between national and regional elections, as well as across regions. To do so, a novel model of party formation is put forward. The second chapter about parties analyses in more detail how the choice of candidate for an election by the leadership of a party can reveal information about this candidate. It is shown that the act of nomination can change how voters perceive a candidate, while the nature of this change depends on how much competition the candidate will face in the subsequent election. A separate chapter looks at the role of interest groups. When voters are not fully informed political advertising becomes an important element of election campaigns. This, in turn, gives interest groups political clout, as they can provide the required campaign funding. The chapter entertains the possibility that interest groups can exploit competition between candidates and use threats of donations to achieve policy goals.

# *Acknowledgements*

# Contents

# List of Figures

# Chapter 1

# Introduction

In an ideal representative democracy, one might be tempted to think, the only actors of relevance are voters and those seeking to represent them. This ideal is not fulfilled in reality: Parties and other forms of special interest groups are ubiquitous in politics, seeking to shift policy outcomes in their favour. This thesis adds to our understanding of the role that these organisations play. The view taken is that their existence is due to informational frictions. It has long been argued that voters are not in a position to acquire all the knowledge relevant to their voting decisions, or, at the very least, not willing to get fully informed in the face of substantial constraints on their time and resources.[1] Parties and interest groups then emerge in order to mitigate or exploit the informational deficits of voters.

For example, professional politicians can be plausibly expected to possess superior information about the state of the economy as they constantly engage in policy debates and rely on expert advisers. Voters can benefit from this knowledge if politicians use it to design adequate policies. However, politicians may also be tempted to exploit the ignorance of voters to their own benefit. A number of papers have explored which outcome seems more likely (Heidhues and Lagerlöf, 2003, Kartik et al., 2015, Martinelli, 2001, Schultz, 1996). What these contributions have in common with most of the literature on political economy is that they do not differentiate between individual candidates and political parties and therefore do not provide a justification for the emergence of the latter. In this thesis parties will be treated as organisations consisting of multiple politicians.

Chapters 2 and 3 both explore the idea that the role of political parties is to provide information to voters who are otherwise unable to observe crucial characteristics of politicians, such as their ideological views or their competence. Chapter 2 follows Snyder

---

[1]This was already discussed in Downs (1957) seminal work.

and Ting (2002) and treats parties as "informative labels" that reveal information about the ideology of their members. It is then asked whether this concept of parties can be usefully integrated into a model of party formation aimed at explaining a particular empirical pattern: Across a number of federal democracies, elections at the federal level tend to be highly competitive, while elections at the state level are often monopolized by individual parties. This raises the question of why there is no entry of additional parties trying to contest such monopolies. A crucial element of the model presented in this chapter is that politicians use state elections as a stepping stone towards running for elections at the federal level. This proves to be a disadvantage for parties targeted at winning elections in particular states, as their members have no chance of advancing their career to the federal level. Consequently, entry of regional parties does not occur and monopolies at the state level persist if politicians value career opportunities at the federal level sufficiently strongly. On the other hand, the model predicts federal elections to be competitive as parties are located symmetrically around the federal median voter.

Chapter 3 looks in more detail at a mechanism through which parties can reveal information about their candidates. The setting is one were the choice of candidates rests with the party leadership, which is better informed about the competence of individual politicians than voters are. The question is then under what circumstances the party leadership can be expected to nominate the candidate that voters prefer and whether voters gain any information from observing the parties decision. While the answer to the second question is generally positive, the answer to the first question depends on the degree of political competition. Low competition implies that the party leader can get away with selecting candidates according to her own preferences, which are unlikely to coincide with those of voters. However, as competition increases, the party leader is forced to take into account which candidate voters prefer. Interestingly, this does not simply mean that the party leader more often chooses the candidate that voters prefer based on their own information. Instead, the party leader increasingly frequently nominates the candidate that voters would choose if they had the same information as the leader does. This chapter also demonstrates that the gains to the party of introducing primary elections as a means of selecting candidates are much less clear than previously argued in the literature.

Chapter 4 turns to interest groups. The presence of these organisations has been argued to be beneficial or detrimental to voters by different authors, but is always seen as related to informational issues. For example, Grossman and Helpman (2001) discuss how interest groups can convey policy relevant information to politicians or voters in a variety of settings, despite their possible bias in favour of particular policies. In a world where voters are incompletely informed, however, politicians will naturally engage in political advertising in order to gain an advantage over their competitors. The need to fund such

activities may make politicians susceptible to offers from interest groups, who want to trade policy favours for campaign donations (Besley and Coate, 2001, Grossman and Helpman, 1994, 1996). If it is accepted as true that interest groups seek influence in this way, however, one is confronted with the so called Tullock Paradox: When considering the value of government subsidies and public procurement to particular industries, the amount of campaign donations made by these industries seems surprisingly small in comparison. Chapter 4 demonstrates that it may not be necessary to make actual contributions in order to influence policy choices; just the threat of contributions can be enough. To illustrate this, suppose an incumbent is aiming for re-election against a challenger. An interest group may then be able to influence the policy choices of the incumbent simply through the threat of making donations to the campaign of the challenger. It is shown that this logic remains valid even if there is competition among interest groups with opposing interests.

Some concluding remarks are offered in chapter 5, which also discusses directions for future research.

# Chapter 2

# How Political Parties Shape Electoral Competition

## 2.1 Introduction

Across federal democracies a common pattern can be observed. At the federal level, elections are competitive: Multiple parties participate and more than one of them stands a chance of emerging as the winner. Accordingly, a single party rarely manages to hold on to power for more than two or three electoral cycles. In contrast, it is not uncommon to observe that elections in a particular state are dominated by one party. In the United States, it is well known that many states in the south have become strongholds of the Republican Party. In the German state of Bremen, the Social Democratic Party has been in control of the state legislature for more than 60 years. Similarly, the Austrian People's Party has ruled the states of Tyrol and Vorarlberg ever since the end of the Second World War.

To demonstrate this pattern more systematically, I collect election results for a number of federal countries. My sample consists of Australia, Austria, Canada, Germany, and the U.S. Using this data, I construct a measure of how competitive elections in a particular region are, as described in detail in section 2.2.

The results are presented in figure 2.1, which clearly shows that federal elections are typically about as competitive as the most competitive states in the respective country, while in each country there are states where competition is substantially lower than at the federal level.[1] Particularly puzzling is the existence of states that are practically

---

[1]A more in-depth discussion of the figure is provided in section 2.2, where I also address alternative explanations of the patterns in the data.

FIGURE 2.1: Political Competitiveness
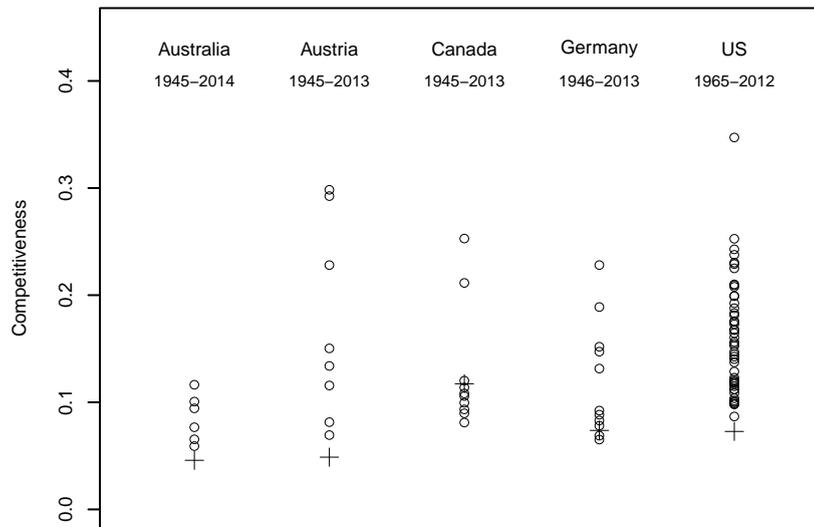


Notes: Each circle represents elections in a given state, while crosses stand for federal-level elections. A higher number implies a lower degree of competition.

monopolized by one party. This is generally the case when my measure of competitiveness takes values of about 20 or higher. The examples mentioned above all fall into this category. The reason why the existence of such regional monopolies is puzzling is that they are typically held by national parties with a relatively broad ideological profile. Why is there no entry of regional parties better able to cater to the views of voters in such states? And if something allows parties to dominate a region, why is this force not at play at the national level as well? Providing a satisfactory answer to these questions requires a concept of what exactly the role of political parties is.

In the model that I construct in this chapter—and in line with a growing body of empirical research to be discussed below—the policy choices of a politician are largely determined by this politician's preferences. Voters therefore care about the preferences of politicians, but are initially poorly informed about these. Following Snyder and Ting (2002), parties provide some of this information by not allowing politicians of all political shades to join. Seeing that a politician is a member of a particular party thus tells voters that this politician must fall into a specific region of the political spectrum. In contrast, there would be no role for parties in the model if voters were fully informed.

In order to compete, parties thus need to attract the right kind of politicians. Importantly, politicians also care about their chances of getting elected. This concern is a main driver of the choice of party affiliation. A key insight that emerges from the model is

that political parties that are successful in national elections can maintain regional monopolies because they offer career prospects at the federal level. As an example, consider a state like Nebraska where the Republican Party currently controls all major elected offices. Suppose strongly conservative members of the Republican Party in Nebraska could form a separate party and do equally well in state elections. This would have the benefit of eliminating internal competition for nominations from politicians belonging to the more moderate wing of the party. It would, however, also deprive members of the newly formed party of any chance of advancing to the federal level. If these career prospects are valuable enough, conservative politicians in Nebraska prefer to remain a part of the Republican Party, which can then maintain its hold on the state.

While the moderate wing of the Republican Party is a burden to conservatives in a right-leaning state, the conservative wing is a detriment to the electoral chances of the party at the national level. It makes the party more extreme and thus less attractive to voters in the political centre. The national party itself might therefore have an incentive to try to exclude its most conservative members. But this would result in the establishment of a more extreme party and the risk of a split in the conservative vote. This threat of entry is the force that prevents the national party from moderating itself in the model and may explain why the establishment of the Republican Party has been relatively accommodating towards the radical Tea Party movement.

The main result of the chapter is that there exists an equilibrium where two parties are formed, one centre-left and one centre-right. Both parties win with equal probability at the federal level while dominating some state elections. This equilibrium is maintained by the forces described above: Neither party can shift further towards the centre without inducing entry of a third party, while in equilibrium such entry is precluded as politicians have no incentive to deviate towards joining a new party.

State monopolies exist in this equilibrium because parties have strongly differentiated ideological profiles. This enables each party to capture a large share of votes in particular states. For example, in a state with a median voter located far to the left, the centre-left party dominates state elections, while states with a more moderate median voter will be more competitive. The model is thus able to recreate the pattern displayed in figure 2.1. As I will argue in section 2.2, other factors may be at play as well, but cannot convincingly explain the data by themselves.

To the best of my knowledge, the contrasting patterns of political competition at the state and at the federal level have previously not been demonstrated as clearly as in figure 2.1. Besley et al. (2010) discuss the wide variation in the degree of competition observable across U.S. states, but do not refer to the federal level. Their empirical results are nevertheless closely related to the analysis here, in that they show that in

states where competition is lower policies tend to be less favourable to growth and actual growth is reduced as well. This indicates that the dominance of one party has negative consequences and highlights the need for a better understanding of how such political monopolies emerge.

A related theoretical paper is provided by Callander (2005), who studies competition between two parties in multiple single-member districts with threat of entry at the district level. Parties, which are not explicitly modelled, are free to choose any platform. Callander finds that the threat of entry leads to the divergence of party platforms. The mechanism through which entry is deterred is different though. In addition, the equilibrium presented by Callander requires specific assumptions on the distribution of voters across districts, while the restrictions imposed on voter distributions in this chapter are mild. This is because entry in the model of Callander implies the loss of one district, while entry has much wider consequences in the current model as explained above. Previous contributions to the literature on political competition with entry consider only a single district (Osborne, 1993, 2000, Palfrey, 1984).

Political parties clearly form a central element of the political system of democratic countries, yet they have received surprisingly little attention, at least in terms of formal modelling. Few papers have attempted to fully endogenize the number parties existing in equilibrium as I do here (Eguia, 2011, Jackson and Moselle, 2002, Levy, 2004, Morelli, 2004, Osborne and Tourky, 2008). As mentioned above, the concept of political parties that I employ is taken from Snyder and Ting (2002). These authors, as well as other contributions building on their approach (Ashworth and Bueno de Mesquita, 2008, Bernhardt et al., 2009), consider the behaviour of a given number of parties. I show how the concept of parties as "informative labels" can yield an equilibrium with two parties that looks very similar to what we observe in a number of countries. Furthermore, I demonstrate that career concerns of politicians can be a driving force behind the number and shapes of parties that form in equilibrium. Previously, attention has mainly focused on variations in the electoral system as a determinant of the number of parties (see Morelli, 2004). Overall, I feel that the success of the model presented here in reproducing and explaining empirical regularities indicates that thinking of parties as informative labels is a fruitful approach.

The rest of this chapter is organized as follows: Section 2.2 details the construction of the measure of political competitiveness I use and addresses some alternative explanations of the pattern displayed in figure 2.1. In section 2.3 I discuss a number of empirical results that lend support to some of the assumptions made in the model, which is laid out in section 2.4. Section 2.5 gives the theoretical results. Robustness of the results to

relaxing some of the assumptions made in the basic version of the model is discussed in section 2.6. Section 2.7 concludes.

## 2.2 Measuring Competitiveness

I want to illustrate how political competitiveness varies across regions. In selecting countries to include in my sample I focus on federal states for three main reasons: First of all, the result of Besley et al. (2010) that limitations on political competition are harmful was established at the state level for the U.S.. Secondly, federal states have stable regional boundaries that are less subject to manipulation by politicians than is the case for other kinds of administrative units. This rules out gerrymandering as an explanation of regional monopolies. Finally, state elections carry some weight, making it harder to argue that the formation of parties is entirely driven by considerations regarding the national level.

My sample consists of state and federal elections for the countries Austria, Australia, Canada, Germany and the U.S.[2] I focus on elections that directly or indirectly determine the selection of federal or state executives. Accordingly, the elections I consider are for state and federal parliaments. The only exception is given by the U.S., where I compare popular voter shares for presidential elections with results of gubernatorial elections. My data generally includes all such elections between 1945 and June 2014. For the U.S. I restrict the sample to elections held after the passage of the voting rights act of August 1965. Prior to this event the Democratic Party dominated the U.S. South, partially through limiting the ability of African-Americans to vote. In Germany I include only the 11 states belonging to the Federal Republic of Germany prior to 1990.

My measure of the competitiveness of an election is the vote margin between the highest and the second-highest vote getter. Denote this vote margin for an election at time $t$ in administrative unit $r$ by $d_r^t$, where $r$ stands either for a particular state or the federal level of a country. I then measure the competitiveness of elections in region $r$ by computing average vote margins over time:

$$\frac{1}{T} \sum_t d_r^t \ ,$$

---

[2]Election results were retrieved from the following sources:

Austria and Germany: www.parties-and-elections.eu/

Australia: elections.uwa.edu.au/

Canada: www.electionalmanac.com

U.S. presidential elections: www.ropercenter.uconn.edu/elections/common/pop_vote.html

U.S. gubernatorial elections up to 1990: ICPSR (1995)

U.S. gubernatorial elections after 1990: library.cqpress.com/elections/ .

where $T$ is the total number of elections in region $r$ included in the sample. These are the values displayed in figure 2.1.[3]

Austria, Germany, and the U.S. show the same pattern of highly contested federal elections and wide vote margins in at least some states. The picture for Australia is similar, but less extreme. In fact, no Australian state is dominated by one of the two main parties of the country.[4] It would seem that this is a consequence of relatively homogeneous distributions of voters across states. In Canada, on the other hand, competition at the federal level is relatively low. This reflects the success of the Liberal Party in the nineteenth century, but also the landslide victories of the Progressive Conservative Party in 1958 and 1984. The more important difference between Canada and the other countries, however, is not visible in the picture: Canadian federal parties are only loosely connected to state parties and successful regional parties that play no role at the federal level exist. Such regional parties can be observed mostly in countries with strong regional identities such as Canada, Belgium, or Spain. Their presence highlights the question of why such parties fail to exist in other countries. I will return to this issue at the end of section 2.5.2.

Of course, my model is not able to explain why voter preferences are distributed in a certain way, but takes this as a given input. However, the model suggests that voter preferences may play a surprisingly small role in determining the number and shapes of parties. The assumptions on voter heterogeneity I impose below are mild. Once this minimal amount of heterogeneity is given, further increases in heterogeneity do not induce existing parties to change their positions nor do they lead to the entry of additional parties. A different issue would be to allow for the possibility that voters develop preferences for particular parties rather than just over policies. This could allow parties to consolidate their position in a region over time. Apart from the tautology inherent in saying that a party wins because voters want it to win, this would also leave the question unanswered of what drives party formation in the first place.

To conclude this section, I will briefly discuss two other factors beyond heterogeneity in voter preferences that may play a role in shaping the outcomes shown in figure 2.1. The first one is the role of incumbents running for re-election. It is well known that, at least in the U.S., incumbents tend to enjoy an electoral advantage. Some states in the U.S. also have less strict term limits for governors. This raises the question to what extent the difference between federal and state elections is a consequence of the presence of incumbents at either level. I demonstrate for the U.S. that this factor is

---

[3]The measure of competitiveness employed by Besley et al. (2010) is the absolute value of the distance of the Democratic vote share from one-half. This number is a linear function of the measure employed here in races where no more than two parties participate, but clearly less appropriate in other settings.

[4]I treat the coalition of the Liberal Party and the National Party as a single party. Keeping them separate makes Australian elections look somewhat less competitive.

FIGURE 2.2: Political Competitiveness in the U.S.
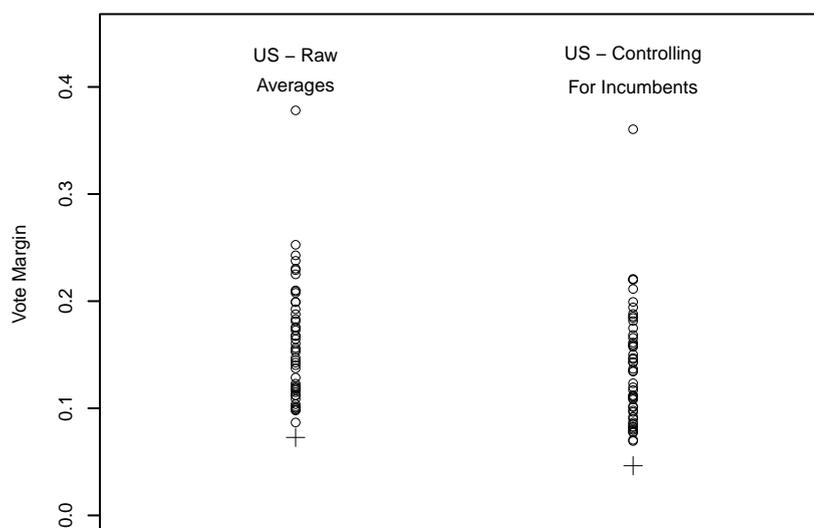


Notes: Each circle represents governors elections in a given state, while crosses stand for presidential elections.

of minor importance. To do so, I run a regression of presidential and gubernatorial vote margins in my sample on a set of state dummies as well as an indicator for the presence of an incumbent. The coefficient on the incumbent dummy indeed turns out to be highly significant with a magnitude of slightly more than five percent. The effect on the remaining coefficients is small, as can be seen in figure 2.2. The first column in this graph reproduces the raw average vote margins for the U.S. as shown before. The second column plots predicted vote margins with the incumbent dummy set to zero. The presence of incumbents does not appear to be a main driver of the low competitiveness of many gubernatorial elections.

A second concern I want to address is that differences in the rate of turnout between state and federal elections might be of importance. Lower turnout could potentially make election results more volatile and thus result in higher average vote margins. This would not explain the persistently high vote margins in favour of one party though, as they are observable across a range of states in different countries. It is harder to rule out that differential rates of turnout among different groups of voters could result in larger vote margins. However, the results of Levine and Palfrey (2007) point in the opposite direction: In an experiment on voter turnout, these authors find that those favouring a disadvantaged candidate are more likely to vote. In addition, turnout is lower in the experiment when the expected closeness of the election is reduced. This suggests that

low turnout in state elections might be a consequence rather than a cause of high vote margins.

## 2.3 Related Empirical Evidence

This section will discuss empirical evidence supportive of some of the assumptions featured in the model or the general ideas behind it. First of all, a growing literature investigates the determinants of the policy choices of elected officials. Chattopadhyay and Duflo (2004) and Bhalotra and Clots-Figueras (2014) find that policy preferences of politicians matter. Both papers establish that an exogenous increase in female representation in India leads to a greater provision of public services typically utilised by women. The results of Lee et al. (2004) go further: According to their estimates the voting behaviour of individual members of the U.S. House of Representatives is independent of their electoral odds. This seems to indicates that the preferences of politicians not only influence but largely determine policy choices. Otherwise one would expect legislators in close election to alter their voting behaviour in an attempt to cater to voter tastes. Similar results are obtained by Levitt (1996) for the U.S. Senate, who additionally controls for a potential role of party discipline imposed on legislators.

A second strand of evidence relates to the nature of political parties. Casual observation suggests that in many countries the vast majority of votes is cast for two parties, one located left and one located right of the political centre, which tend to alternate in power. In view of the studies cited above, the interpretation of the statement that a party is centre-left would have to be that the politicians belonging to such a party prefer centre-left policies, at least on average. This is also formally confirmed by research estimating the ideological positions of politicians. Poole and Rosenthal (1997, 2001) do so using voting records from the U.S. Congress. According to their results the membership of both the Republican and the Democratic Party spans a wide range of positions, but with very little overlap between them. Barberá (2015) finds very similar results using data from the social network Twitter. He also applies his method to five European countries, and again a similar picture emerges.[5]

All of this is consistent with the view of parties as collections of similar-minded politicians. In this case voters can learn something about a politician's views from observing which party she is a member of, even if they cannot observe preferences of politicians directly. This is the idea formalised by Snyder and Ting (2002) and applied in this

---

[5]As the author remarks, the reason that he finds somewhat more overlap between parties in Europe may be due to the possibility that variation in political preferences is less well captured by a single dimension there.

chapter. Snyder and Ting also demonstrate empirically that voters' knowledge of an individual politician's position is almost entirely captured by her party affiliation. They use estimates of these positions to predict how voters place candidates on an ideological scale. As it turns out, a simple dummy for party affiliation does just as well in explaining the variation in voter knowledge.

Finally, career concerns of politicians are a driving force behind the results. In the context of the model, it seems natural to assume that politicians progress from the regional to the national level. After all, winning a regional election reveals information about a politician, which is a strong advantage when facing a competitor who is unknown to voters. At least for the U.S., it is well documented that politicians indeed use elections as stepping stones towards higher offices. For example, Diermeier et al. (2005) collect a sample of members of the U.S. Congress in the period 1947 to 1994. They find that 78 percent of these politicians held a different local, state, or federal elected office before joining Congress. About ten percent of representatives in their sample run for a Senate seat. Of those who leave Congress, 35 percent stay in politics. In Germany, candidates for the office of federal prime minister as well as federal ministers are frequently recruited from among state prime ministers (Detterbeck, 2012, p. 193).

## 2.4 The Model

A federal state consisting of $S \geq 4$ states selects federal and state governments through plurality rule elections. Candidates for these elections are nominated by political parties. I divide the game into two main stages: A party formation stage and an election stage. Each stage will be described in more detail below once some of the basic elements of the game have been introduced. Note that the model features no "states of nature" and the word state therefore always refers to a geographical unit.

### 2.4.1 Voters, Politicians, and Parties

Each state $s$ has an infinite set of citizens and each citizen votes in two elections: The election for the government of state $s$ and the election for the federal government. Let $p_s$ and $p_f$ denote the policies that are implemented in state $s$ and at the federal level, respectively. The objective of voters in election $l \in \{s, f\}$ is to maximize

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

where $u : \mathbb{R} \to \mathbb{R}$ is a decreasing function while $i \in \mathbb{R}$ is the ideal policy of the voter. As will become clear later, the outcomes of state elections may affect events at the federal

level, but it is assumed that voters do not take this interdependence into account when voting at the state level. This is done simply to make the proofs presented in subsequent sections more compact, but is otherwise not necessary.[6]

Each state also has a finite set of politicians. As discussed in section 2.3, preferences over policies appear to be the main driver of the choices that politicians make in office. In the basic version of the model, I assume that these preferences are sufficiently strong to make any politician always implement her ideal policy. Every politician is thus associated with a policy that she has to implement if elected to any state or federal office. In order to avoid confusion with preference parameters of voters I refer to the policy of a politician as her platform. I will allow politicians to be more flexible in their policy choices in section 2.6. An additional simplification that is required for tractability is that politicians have only three possible platform, namely -1, 0, and 1. The number of cases to consider increases rapidly in the number of possible platforms and is already large with three platforms. A possible interpretation of this assumption is that voters have a coarse perception of the policies chosen by politicians. Evidence from psychology indicates that people tend to think in simplifying categories (For a discussion of some of this research see Fryer and Jackson, 2008). Each state has three politicians and none of them share the same platform. Put differently, there is one politician located at each of the possible policies -1, 0, and 1. I will apply the labels centrist, extremist, rightist, and leftist to politicians in the obvious manner.

For an election at level $l \in \{s, f\}$ the winning candidate receives a payoff of $y_l > 0$. A politician who has won a state election but does not win the nomination of her party for the federal election nevertheless receives a payoff of $y_P > 0$ if her party wins the federal election. This payoff has a number of interpretations. It may represent opportunities to move upwards in the party hierarchy that arise when a party wins the federal election or the chance of becoming a member of the federal government. A second interpretation of this payoff is that career opportunities in the private sector become more valuable if a politician is well connected within the party in power. It is also possible to think of $y_P$ as representing "pork": The leader of a state government will be likely to receive more federal money if her party is in control of the federal government.

In order to clearly define the utility of a politician, let $\pi_s$ be the probability that a politician is nominated for and wins the state election in her state. Conditional on doing so, let $\pi_n$ give the probability that a politician is nominated for the federal election. $\pi_f$ is then the likelihood of winning conditional on receiving the nomination, while $\pi_P$ is the probability that the party wins conditional on some other candidate having received

---

[6]If voters were forward-looking, this would only change their behaviour where voting for their less preferred candidate at the state level would somehow yield a sufficiently large benefit at the federal level. In all of the specific situations dealt with below it can be shown that this is impossible.

the nomination. All of these probabilities will later be determined in equilibrium. The expected utility of a politician who has joined a party is given by

$$\pi_s(y_s + \pi_n \pi_f y_f + (1 - \pi_n)\pi_P y_P) \ ,$$

while a politician who is not part of a party receives a payoff of zero. It is assumed that $y_f > 2y_P$.

A political party is basically a subset of the policy space and only politicians whose platforms fall within this subset can join. This idea is based on Snyder and Ting (2002), where the leadership of a party chooses a platform and politicians pay a cost for joining the party that depends on the distance between this platform and their own ideal policy. Two interpretations of this cost are given: First, politicians could find it costly to be members of an organisation that pursues goals that differ from their personal views. Second, parties could be actively screening their members and only promote those who agree with the party line. As a result, only politicians with an ideal policy belonging to an interval centred around the party platform join. The size of the interval depends on the membership cost in the first interpretation, or the effectiveness of screening in the second interpretation. I simplify things by giving parties full control over the size of the interval that represents the party. Given that the space of platforms consists of integers, parties will be given by "integer intervals": For $a, b \in \{-1, 0, 1\}$ define

$$[a..b] \equiv \{p \in \{-1, 0, 1\} : a \leq p \leq b\} \ .$$

If $a$ equals $b$ I simply write $[a]$. The set of all possible shapes a particular party can have is

$$\mathcal{I} = \{[-1], [0], [1], [-1..0], [0..1]\} \ .$$

Note that parties that allow all types of politicians to join are not allowed for. Including them would give rise to two additional equilibria, which have some implausible features and can easily be eliminated with an additional refinement. In order to not distract from the main arguments, I exclude this type of party for now. Section 2.6 will discuss the consequences of also admitting parties of shape $[-1..1]$ in more detail.

Parties are organized nationally, meaning that the interval that represents the party is the same in all states. The set of politicians that joins a party does not have to be the same across states, however, as politicians in different states might face different incentives. Individual parties will be denoted by capital letters. For any such party $P$ the shape of the party is given by $I_P \in \mathcal{I}$. Multiple parties are allowed to have the same shape.

### 2.4.2 The Election Stage

In order to describe the election stage, let there be $N$ existing parties, collected in the set $\mathcal{P}$. Let $\mathcal{P}(p)$ denote the possibly empty set of existing parties that include the policy $p$. The strategy set of a politician with platform $p$ in this subgame is then given by $\mathcal{P}(p)$. Note that this means that a politician can join at most one party and that a politician who has the ability to join at least one party must do so. The latter assumption is made for convenience and could easily be replaced with a small payoff that a politician receives once she joins a party.

The election stage starts with politicians making their affiliation decisions, followed by simultaneous state elections, which in turn are followed by the federal election. Immediately prior to each election every party nominates a candidate, who is drawn uniformly at random from the candidate pool of the party for the election in question. For a particular state the candidate pool of a party consists of all politicians of that state who have joined this party. Each winner of a state election then becomes a member of the candidate pool of their party for the federal election. Of course, candidate selection is generally an important strategic decision. As it turns out, however, parties have an incentive to commit to a candidate selection mechanism that gives extremists a sufficiently high chance of being nominated. I explain this in more detail in section 2.6.

The policy that is implemented in a state is equal to the platform of the politician elected in the state election, just like the policy at the federal level is equal to the platform of the politician elected in the federal election. The winner of each election is the candidate that achieves the highest number of votes with ties resolved randomly.

### 2.4.3 The Party-Formation Stage

Parties are formed by "founders". Founders can choose to propose a party or remain passive and are divided into two groups, both of which have a countably infinite number of members. The first group consists of early movers. The action space of these founders is given by $\mathcal{I} \cup \emptyset$, where $\emptyset$ stands for the decision to not propose a party. Once a founder has proposed a party, I will also refer to this founder as a party leader. Each member of the second group of founders, so called late movers, randomly draws a shape $I$ from a distribution with full support on the set $\mathcal{I}$. The action space of such a founder is then given by $\{I, \emptyset\}$.

The party formation stage then proceeds as follows: Initially, all early movers simultaneously decide to propose a party or to stay passive. I will refer to all parties formed at this stage as incumbent parties. Subsequently, all late movers simultaneously choose

whether they want to field a party. I will refer to all such parties as entrant parties. This timing has the effect that founders who move first have to take into account that their actions my induce entry of additional parties, similar to a standard model of entry with competing firms. Reducing the flexibility of late movers in the shape of party that they can offer is necessary, as the complexity of determining equilibria of subgame otherwise makes the model intractable.

Each founder pays a cost $c > 0$ for proposing a party, while she receives a payoff of $x_w > 0$ for every state election that her party wins, as well as a payoff of $x_f > 0$ if her party wins the federal election. Denoting by $\rho_r$ with $r \in \{1, ..., S, f\}$ the equilibrium probability that the party of a founder wins election $r$, the expected utility of a founder who proposes a party is given by

$$\sum_{s=1}^{S} \rho_s x_w + \rho_f x_f - c \ .$$

Passive founders always achieve a utility of zero. I assume $c < x_w$, so that a founder whose party wins at least one state election does not want to deviate to remaining passive.

### 2.4.4   Information

A crucial feature of the concept of political parties employed here is that voters have limited information about politicians. At the beginning of the game, the electorate cannot distinguish between different politicians, but knows how their platforms are distributed. In contrast, politicians and founders observe platforms. Everything apart from platforms is common knowledge. In particular, voters know which parties have been proposed and how many politicians have joined each one of them in each state. Knowing that a candidate belongs to a certain party therefore allows voters to update their beliefs about this politician's platform prior to casting their vote for the state-level election. The winner of the election then implements her platform at the state level, thus revealing it to voters. Voters accordingly have full information about candidates at the federal level. All agents are also fully informed about the distribution of voters in all states and at the federal level.

### 2.4.5   Equilibrium

The timing of the game is summarised in figure 2.3. Any set of parties $\mathcal{P}$ that gets proposed in the first two stages of the model leads to a proper subgame comprised

FIGURE 2.3: Timing
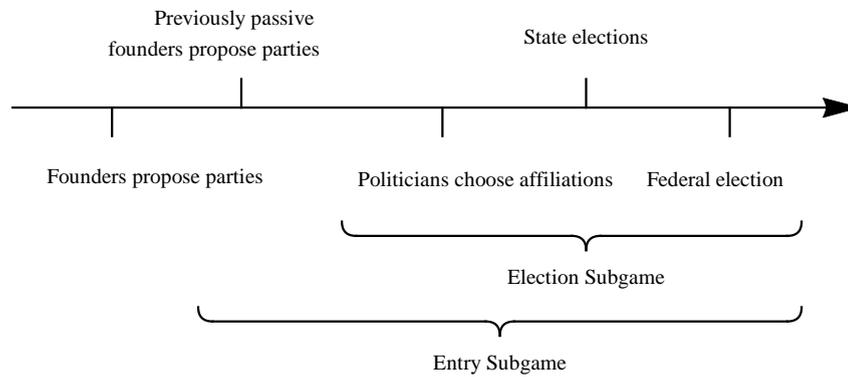


of the steps previously referred to as the election stage. I will refer to this as the election subgame under the set of parties $\mathcal{P}$. Similarly, any set of incumbent parties $\mathcal{J}$ proposed at the first stage of the game leads to proper subgame starting with the possible formation of entrant parties. I will refer to this as the entry subgame under the set of parties $\mathcal{J}$.

Given that the game features incomplete information, the appropriate equilibrium concept is perfect Bayesian equilibrium. By itself, this would entail the possibility of a huge number of equilibria that exist when voters are allowed to vote strategically. Other papers in the literature on party formation assume sincere voting to avoid this problem. I generally allow for strategic voting, but impose three plausible restrictions: First of all, I consider only equilibria in weakly undominated strategies. The exclusion of weakly dominated strategies is a common way of refining voting equilibria and excludes the possibility that voters vote for their least preferred candidate. The second restriction reads as follows: If a candidate is the unique most preferred option of a strict majority of voters, then this candidate wins the election. In general, there may exist voting equilibria where a different candidate gets elected in this situation, but it nevertheless seems likely that voters will be able to solve the coordination problem in this case. The third restriction is akin to a tie breaking rule: I assume that all candidates receive an equal number of votes if all voters are indifferent between all candidates, but only if the election takes place along the equilibrium path. Imposing this restriction along the equilibrium path only can be interpreted as "party loyalty": If an additional party enters, indifferent voters may continue to vote for one of the previously existing parties out of habit.

The following definition summarises the equilibrium concept:

**Definition 2.1.** A party-formation equilibrium is a perfect Bayesian equilibrium of the party-formation game that satisfies the following conditions:

i) No player uses a weakly dominated strategy.

ii) If a candidate in some election is the unique most preferred option of a strict majority of voters, then this candidate wins the election.

iii) Along the equilibrium path all candidates receive an equal share of votes if all voters are indifferent between all candidates.

I will restrict attention to equilibria in which only incumbent parties form along the equilibrium path. This is without loss of generality regarding the shapes and numbers of parties that can be supported in equilibrium. I will uses stars to denote equilibrium objects. In particular, $\mathcal{P}^*$ will denote the set of parties formed in equilibrium, while $N^* \equiv |\mathcal{P}^*|$. In addition, $w_P^*$ will denote the equilibrium number of state elections won by party $P$.

### 2.4.6 Voter Distributions

A crucial input of the model is the set of voters. I will only make relatively weak assumptions in this regard. More specifically, the results in subsequent sections require a minimum amount of heterogeneity in voter tastes. Figure 2.1 seems to indicate that actual heterogeneity is often substantial.

Before stating my assumptions, I need to take a step towards analysing behaviour in the model. Suppose there are two parties, $A$ and $B$, contesting a state election. Party $A$ has the politician with platform -1 as the unique member, while the remaining two politicians have joined party $B$. Assume voters are aware of this. Each voter then knows that the candidate nominated by party $A$ has platform -1. The candidate of party $B$, on the other hand, is equally likely to have either platform 0 or 1 due to the assumption of random candidate selection. Let $p^-$ be the unique real number such that a voter with ideal policy equal to $p^-$ is indifferent between the candidate of either party, that is $p^-$ solves

$$u(|-1-p^-|) = \frac{1}{2}[u(|-p^-|) + u(|1-p^-|)] \ .$$

As the utility of voters is symmetric around their ideal policy and $u$ was assumed to be decreasing, it must be the case that $p^- \in (-0.5, 0)$.[7] Next, consider the situation that

---

[7]These assumptions imply that a voter with ideal policy equal to -0.5 is indifferent between the policies -1 and 0, but strictly prefers -1 over 1. She must therefore strictly prefer the candidate of party $A$ over the candidate of party $B$. A voter with ideal policy zero, on the other hand, is indifferent between the policies -1 and 1 and therefore strictly prefers party $B$ over party $A$ due to the possibility that party $B$ nominates a centrist candidate.

would result if the politician with platform 0 were to switch from party $B$ to party $A$. In this case a voter with ideal policy $|p^-|$ would be indifferent between voting for either party. Denote this policy by $p^+$.

I assume that the set of voters in any state $s$ can be described by a measure $V_s$ over possible ideal policies. Let $m_s$ denote the median associated with this measure. Similarly, let $V_f$ be the measure of voters at the federal election with median $m_f$. It is assumed that $m_f$ is equal to zero. It will often be important to know what share of voters in some region $r \in 1, .., S, f$ is located in the interval $[-0.5, 0.5]$. I will therefore define

$$\Lambda_r([-0.5, 0.5]) \equiv \frac{V_r([-0.5, 0.5])}{V_r[\mathbb{R}]} \ .$$

The first more substantial assumption regarding voter preferences specifies that there is some minimum amount of heterogeneity in voter distributions across states: Let there be at least one state $s$ such that $m_s < -0.5$, at least one state $s'$ such that $m_{s'} \in (p^-, p^+)$ and $\Lambda_{s'}([-0.5, 0.5]) > 0.5$, and at least one state $s''$ such that $m_{s''} > 0.5$. Note that -0.5 (0.5) is the ideal policy at which a voter is indifferent between the platforms -1 and 0 (0 and 1). I will refer to states with median voter below -0.5 or above 0.5 as extremist states, while states with median voter between $p^-$ and $p^+$ are called centrist states. Purely for convenience, I will also assume that there is no state with median voter located at $p^-$ or $p^+$.

The second assumption on voter distributions says that voters at the federal level are not too concentrated in the centre of the policy space: $\Lambda_f([-0.5, 0.5]) \leq 0.5$. This requirement would be satisfied, for example, if $V_f$ was the probability measure associated with a uniform distribution with support on an interval of length at least equal to two, or a normal distribution with variance slightly above one-half.

## 2.5 Results

The model described in the previous section has many equilibria. This should come as no surprise: After all, it features two coordination problems—one between politicians and one between voters—as well as unrestricted out-of-equilibrium beliefs that can be freely chosen to support a specific equilibrium. In particular, voters may believe that a

politician who deviates has a platform that the median voter of the state dislikes, which makes it unlikely that the deviation is successful.[8]

Given the multiplicity of equilibria, I will proceed as follows: Given their empirical relevance, my main interest is in equilibria with two parties. It turns out that this class can be fully characterized, as I discuss in the next section. This section contains the main results of the chapter. In general, however, the number of equilibria is large. Section 2.5.2 will discuss this in more detail and suggest a refinement.

### 2.5.1 Equilibria with Two Parties

Following the discussion in section 2.3, a natural starting point is a situation with a centre-left and a centre-right party. The most obvious formalisation of this would be an equilibrium where the set of proposed parties is equal to $\{L, R\}$ with $I_L = [-1..0]$ and $I_R = [0..1]$.

Given that parties $L$ and $R$ are the only existing parties, how will politicians behave? Those with platform -1 and 1 will become members of the unique party available to them by assumption. In a state with median voter below $p^-$ it then does not matter which party the politician with platform 0 joins: The median voter always prefers party $L$.[9] As this is her only chance of getting elected, politician 0 will therefore always join party $L$ in such states. Analogously, a politician with platform 0 will join party $R$ in a state where the ideal policy of the median voter is greater than $p^+$.

In centrist states, in contrast, politician 0 can make either party the winner of the state election by joining. The probability with which politician 0 is nominated for and wins the state election in such states is thus the same independent of which party politician 0 becomes a member of. Conditional on receiving the nomination at the federal level, the probability of winning is also independent of the choice of party. This is because each party has a moderate and an extremist member in states where it wins and also nominates these with equal probability for the state election. Accordingly, both parties have an equal number of politicians of either type in their federal candidate pool in expectation. This means that the "expected opponent" at the federal election is equally strong no matter which party a politician joins. The only factor affecting the utility of

---

[8]Such beliefs are not entirely unrestricted though. Consider for example the case where all politicians in some state have joined a party that allows any politician to join. If one politician deviates and joins a party with shape $[-1..0]$, voters find themselves at an information set that has two nodes; one for the case in which politician -1 has deviated and one for the case in which politician 0 has deviated. Voters may assign arbitrary weights to either node. Naturally, they may not attach any weight to the possibility that politician 1 has deviated.

[9]Note that voters always know who has joined which party in this setting as there are no information sets that contain more than one node.

a politician with platform 0 that may differ between parties is the probability of being nominated at the federal level. This probability is decreasing in the number of states elections won by the party. It follows that politicians with platform 0 in centrist states will join the party that wins fewer elections, a force that works towards equalizing the number of states won by each party. Equilibrium is reached if centrist politicians in centrist states have either joined the party that wins fewer elections or are distributed across parties such that the number of state elections won differs by at most one between them. For example, suppose there are $2n$ states, in $n$ of which the median voter's most preferred policy is -1 and party $L$ accordingly wins the state election. Then party $R$ must win all other states, whether they are centrist or rightist.

The observations made in the preceding paragraphs are collected in the following lemma.

**Lemma 2.1.** *Suppose $\mathcal{P} = \{L, R\}$ with $I_L = [-1..0]$ and $I_R = [0..1]$. Then*

i) *politician -1 (politician 1) joins party L (party R) in every state,*

ii) *politician 0 joins party L (party R) in any state s such that $m_s < p^-$ ($m_s > p^+$),*

iii) *in any state s such that $m_s \in (p^-, p^+)$ politician 0 joins party L (party R) if $w_L^* < w_R^*$ ($w_R^* < w_L^*$) and may be a member of either party if $w_L^* - 1 \leq w_R^* \leq w_L^* + 1$,*

iv) *and in each state the party that politician 0 joins wins.*

The equilibrium of the election subgame given in lemma 2.1 has an interesting feature: Parties are more extreme in states where the distribution of voters does not favour them. For example, members of the centre-left party are more left-leaning on average in states where the distribution of voters is strongly skewed towards the right. This is because centrist politicians join party $R$ in this case as this is the only party that gives them a chance of winning elections. Only politicians with a left-wing platform remain in party $L$. Translated to the context of U.S. politics, this would imply that members of the Democratic Party are more moderate on average in a very liberal state like Massachusetts, and more extreme in a conservative state like Texas. It should be noted that this prediction does not necessarily carry over to nominated candidates, if the assumption that candidates are selected randomly is relaxed. There would be an additional selection issue when trying to test this using observations on elected politicians. Regarding the membership itself, this seems like a robust prediction of the model though.

It was explained above that each party nominates a moderate or an extremist politician with equal probability at the federal level, which is true independent of the number of states won by either party. Given the behaviour imposed on completely indifferent

voters, it follows that each party wins the federal election with equal probability ex-ante. The reason for this is that both of them are able to recruit moderate politicians that are attractive to the federal median voter. In contrast, one party would gain a substantial advantage if it were able to attract all centrists. As described above, electoral concerns of politicians make this impossible. Due to the importance of this result in the context of the chapter, I restate it as a proposition.

**Proposition 2.1.** *Suppose* $\mathcal{P} = \{L, R\}$ *with* $I_L = [-1..0]$ *and* $I_R = [0..1]$. *Then each party wins the federal election with equal probability.*

It remains to establish that there actually exists an equilibrium of the game as a whole where parties $L$ and $R$ are formed and no other parties enter. This is confirmed in the following proposition, subject to a condition on payoffs being satisfied.

**Proposition 2.2.** *An equilibrium of the party formation game where* $\mathcal{P}^* = \{L, R\}$, *with* $I_L = [-1..0]$ *and* $I_R = [0..1]$, *exists if*

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

*Proof.* First, consider deviations by passive founders, may they be early or late movers. It is sufficient to show that conditional on the affiliation behaviour given in lemma 2.1, no politician wants to deviate to joining some entering party that admits only politicians with a particular platform. Politicians with platform -1 do not gain by joining a party with shape $[-1]$ if they are in a state with median greater than -0.5 as they would subsequently lose the state election. In a state with median voter below -0.5, on the other hand, a majority of voters would strictly prefer the new party, which would then win the state election by assumption. In equilibrium a politician with platform -1 in such a state achieves

$$\frac{1}{2}\left[y_s + \frac{1}{4w_L^*} y_f + \left(1 - \frac{1}{w_L^*}\right)\frac{1}{2}y_P\right] ,$$

as each member of party $L$ gets nominated with equal probability and there are two members in the state; each of the $w_L^*$ state-winners of party $L$ are nominated for the federal election with equal probability; extremists can only win the federal election if party $R$ also nominates an extremist (which happens with probability one-half) and a tie results; and both parties win the federal election ex-ante with equal probability. If the same politician were to join the entering party her payoff is $y_s$ as she loses the federal election with certainty. This is because there will then be three parties competing at the federal level and a politician with platform -1 can never be strictly preferred over the other two candidates by a strict majority. Accordingly, there always exists a voting equilibrium where some other party wins. As it was assumed that $y_f > 2y_P$, the equilibrium utility decreases as $w_L^*$ increases. As no party can win more state elections than there are states, a sufficient condition for deviations of this type not being profitable is

$$\frac{1}{2}\left[y_s + \frac{1}{4S} y_f + \left(1 - \frac{1}{S}\right)\frac{1}{2}y_P\right] \geq y_s ,$$

which can be rewritten to give the condition in the statement of the proposition.

For politicians with platform 0 the equilibrium payoff is given by

$$\frac{1}{2}\left[y_s + \frac{3}{4w_j^*}y_f + \left(1 - \frac{1}{w_j^*}\right)\frac{1}{2}y_P\right],$$

with $j \in \{L, R\}$. In case such a politician deviates to a party with shape [0] she can at best hope to win the state election. This is because it was assumed that $\Lambda_f([-0.5, 0.5]) \leq 0.5$ and there accordingly exists a voting equilibrium of the federal election where one of the other two parties wins. This is true even if all parties nominate a candidate with platform 0, as the restriction that all candidates receive an equal share of all votes when all voters are indifferent was only imposed on the equilibrium path. The payoff from the deviation is then $y_s$, which is smaller than the equilibrium payoff. This follows because the payoff of centrist politicians is greater then the payoff of extremist politicians for a given number of state election won by the party. Accordingly, centrists do not deviate as long as extremists do not deviate.

It remains to check whether any founder has an incentive to reposition their party. After any such deviation, there exists an equilibrium of the entry subgame reached where two late movers propose additional parties of shape $[-1..0]$ and $[0..1]$ and no politicians join any of the incumbent parties. It was already established in the first part of the proof that this behaviour of politicians is an equilibrium and that no additional parties can successfully enter. As both late movers then win at least one state election, neither of them wants to deviate to remaining passive. This ensures that any deviation by an early mover is not profitable. □

I will refer to the equilibrium in the preceding proposition as the *L-R* equilibrium. Under this constellation of parties no third party can successfully enter for reasons easily illustrated in an example: Suppose a party were to enter that admits only politicians with platform -1 as members. Given that such a party attracts members, it may do well at the state level, but would not be able to win the federal election due to the presence of the two already established parties. Politicians with platform -1 in states with very left-leaning median voters may nevertheless be tempted to join the entering party, as this eliminates competition from centrist politicians for the party nomination at the state level. The incumbent party prevents them from defecting by offering career prospects at the federal level. These include the possibility of becoming the party's nominee for the federal election as well as the payoff $y_P$. In the context of many countries this payoff represents positions in the federal government or high up in the party hierarchy. In the U.S., parties also control assignments to posts on congressional committees. Such positions give individual legislators greater influence. This influence may then also translate into higher earnings outside of politics.

FIGURE 2.4: Possible Two-Party Equilibria



I)                                                II)
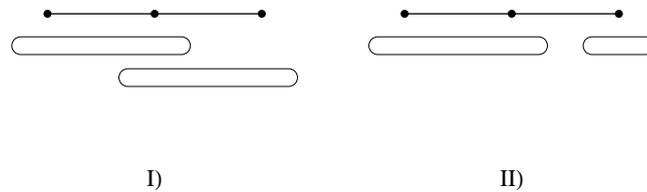
While the set of parties $\{L, R\}$ is robust to entry in many cases, the threat of entry is nevertheless required to support the equilibrium. Either party leader would otherwise have an incentive to exclude extremist politicians from the party, winning at least as many state elections as before and increasing the probability of winning the federal election. With entry this move is not profitable: That a part of the political spectrum is not covered by any party means that entry must occur. A new party can easily recruit the politicians otherwise unable to join a party and win at least some state elections. Consequently, this new party will also nominate a candidate for the federal election. With three candidates competing at the federal level there exists a voting equilibrium where the candidate of party $L$ loses. My interpretation of this is that entry of even an extremist party poses a serious threat because of the danger that the established party loses its core support.[10] Real-life examples of this abound: The UK Conservative Party currently fears the rise of the UK Independence Party (UKIP), not so much because UKIP itself is expected to win many elections, but because the votes lost to UKIP may hand victory in the upcoming general election to the Labour Party. In Germany, the market-oriented reforms unexpectedly pushed through by chancellor Gerhard Schroeder between 2002 and 2003 alienated the socialist wing of his party and fostered the formation of left-wing party The Left. The Social Democrats have been trailing the Christian Democrats in federal elections ever since.

Defeat at the general election due to a split in the left-wing vote implies that party $L$ becomes less attractive and all eligible politicians prefer to join party $R$. Even entry of an extremist party would consequently deter the leader of party $L$ from deviating. There thus exists a mutual dependence: Extremist politicians prefer being members of a more moderate party due to the opportunities that come with being a member of federally powerful party. The party, on the other hand, is happy to offer these opportunities as it benefits from having extremist politicians as members rather than as competitors.

---

[10]For simplicity, the proof uses the fact that a harsher punishment can be assigned after this deviation by the leader of party $L$. However, even entry of a party of shape $[-1]$ as it is discussed here is sufficient to make this deviation not profitable.

FIGURE 2.5: Existence of Two-Party Equilibria
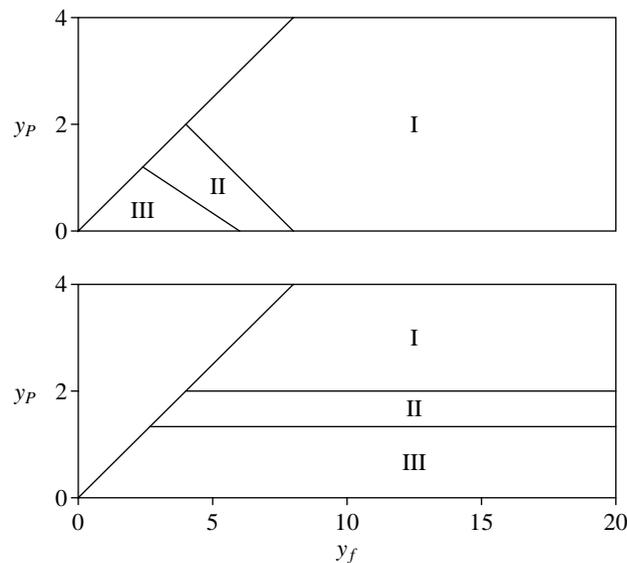


Notes: Both panels show equilibrium existence in $y_f$-$y_P$ space for $y_s = 1$ and a fixed number of states. In the upper panel the number of states is equal to four. The lower panel presents the limit case as the number of states increases towards infinity.

The set of all equilibria with two parties is characterized in appendix A.1 and illustrated in figure 2.4. The constellation in panel I) is the $L$-$R$ equilibrium discussed previously. If career opportunities at the federal level are not sufficiently valuable to prevent extremists from joining smaller parties, this equilibrium fails to exist. There then nevertheless exists a two-party equilibrium, which is exemplified in panel II) of figure 2.4. Here an extremist left-wing party (right-wing party) faces a centre-right (centre-left) party. I will refer to this equilibrium as the $M$-$E$ equilibrium. The moderate party wins the federal election with higher probability than parties in the $L$-$R$ equilibrium do, increasing the payoff of its members. This can prevent them from joining a third party, even when the $L$-$R$ equilibrium does not exist.

If even the $M$-$E$ equilibrium is not robust to members of the larger party defecting to a smaller party, no equilibrium with two parties exists. This is illustrated in figure 2.5 for $y_s$ equal to one. In the upper panel the number of states is equal to four, while the lower panel corresponds to the limit case as the number of states goes towards infinity. Under combinations of values for the payoffs $y_f$ and $y_P$ that fall into region I the $L$-$M$ and the $M$-$E$ equilibrium both exist, while in region II the $M$-$E$ equilibrium is the unique equilibrium with two parties. While the $M$-$E$ equilibrium exists more broadly, the $L$-$R$ equilibrium seems more empirically relevant, as it generates competitive federal elections and also allows for party switching of moderate politicians, as is observable

in the U.S. for example. Comparing both panels illustrates how internal competition increases with the number of states. When the number of states is low, each state winner has a high chance of being nominated for the federal election. This prospect alone can be enough to keep politicians from deviating to joining smaller parties, who would offer less competition at the state level. As the number of states increases, the probability with which each individual politician wins the federal nomination of her party goes to zero. A sufficiently high value for the payoff $y_P$ is then required in order to prevent defections. In region III of the figure, no equilibrium with two parties exists. The space above the diagonal line, on the other hand, is not part of the parameter space due to the assumption $y_f > 2y_P$.

## 2.5.2 Equilibria with Any Number of Parties

This section considers the possibility of equilibria where the number of parties formed is not equal to two. A result that can easily be established is that equilibria with only one party do not exist.

**Proposition 2.3.** *There is no equilibrium such that $N^* = 1$.*

*Proof.* Suppose only one party exists and call this party $A$. As party $A$ cannot cover all platforms, there must be some policy $p$ such that politicians with this platform are not allowed to join party $A$. Then a passive founder could successfully form a party $D$ that accepts only politicians with platform $p$: By assumption, there exists a state with a median voter who strictly prefers platform $p$ over any other platform. In this state party $D$ would accordingly win the state election, as there are only two parties nominating candidates. This contradicts that only one party exists. □

If there was only one party, this party would not be able to allow all politicians to join by assumption. But if some politicians are unable to join any party, a new party that allows just these politicians to join could easily recruit them. Due to the assumption on heterogeneity in voter preferences across states, this party would then also win at least one state election.

It should be stressed that the result that no single party can fend off entry would continue to hold even if parties were allowed to cover all policies. The reason is that such a party would create intense competition for nominations within the party, making at least some politicians willing to join a second party.

No such clear-cut results can be established once equilibria with three or more parties are considered, as this set is large. What is more, the equilibrium number of parties can also be large as demonstrated by the following example.

**Lemma 2.2.** *Suppose there are $S$ states, with two and only two states $s$ such that $m_s \in (p^-, p^+)$. Also, let it be true for any state $s'$ such that $m_{s'} \notin (p^-, p^+)$ that $\Lambda_{s'}([-0.5, 0.5]) \leq 0.5$. Then there exists an equilibrium such that $N^* = S$.*

*Proof.* Consider the following strategy profile: $S$ founders propose parties. Two of them propose a party of shape $[0]$. Call these centrist parties $C_1$ and $C_2$. All other parties either have the shape $[-1]$ or $[1]$. Call these parties leftist and rightist. The number of parties with shape $[-1]$ is equal to the number of states $s$ such that $m_s < p^-$. Collect these states in the set $\mathcal{S}_l$. The number of parties with shape $[1]$ equals the number of states $s$ such that $m_s > p^+$. Collect these states in the set $\mathcal{S}_r$. Each of the centrist parties wins the election in one of the two states that do not belong to the set $\mathcal{S}_l \cup \mathcal{S}_r$ and accordingly has a member in that state. Each leftist (rightist) party has a member in one and only one of the states belonging to $\mathcal{S}_l$ ($\mathcal{S}_r$) and wins the state election in that state. At the federal election the candidates of the centrist parties tie and no other parties receive any votes.

Given that each election features at least three candidates and the assumptions on voter distributions, there always exists a voting equilibrium such that the specified candidate wins. Centrist politicians do not gain from changing their party affiliation as they either continue to lose the state or federal election or simply substitute external competition for internal competition at the federal level. Similarly, extremist politicians do not benefit from switching parties either. Any additional parties would not win at the federal level due to strategic voting and all politicians who win state elections are already the sole member of their party in their state, implying that they would not increase their utility by joining a newly formed party. No founder wants to deviate to remaining passive, as each founder wins one state election. Finally, no founder wants to change the shape of their party, as any such deviation can be punished by entry of an additional party of the same shape that the party of the deviating founder had prior to the deviation. By virtue of the proceeding steps of the proof, there then exists an equilibrium of the election subgame reached where the party of the deviating founder does not gain any members. □

The proof of the preceding proposition illustrates the possibility that a large number of parties forms, only a few of which play a role at the federal level. What may be the reason that such an equilibrium is not typically observed in reality? One possibility is that constellations with many parties do not persist because there are incentives for parties to merge. In fact, mergers are commonly observed in reality. To name just a few examples: The Liberal Democrats of the UK were formed as a fusion of the Liberal Party and the Social Democratic Party. The Conservative Party of Canada came into

existence as a merger of the Progressive Conservative Party and the Canadian Alliance. In Australia, the coalition between the Liberal Party and the National Party has existed for so long that it is often treated as a single party.

One gain from a merger may be that an alliance of parties wins more elections than a number of small parties taken together would. A second potential reason for a merger of parties lies in synergies such as reduced administrative costs, more effective fund-raising and advertising campaigns, and greater recognition by voters in states where the party does not have a strong presence. In the model, such an incentive for a merger comes from the presence of the fixed cost $c$ that founders pay when forming a party. Avoiding this cost can make a merger profitable even if the newly formed party does not enjoy greater success.

Note that any merger that happens for reasons of increased electoral success also leads to a reduction in fixed costs, while the converse is not true. In trying to integrate mergers into the model I will therefore focus on fixed costs as the driving force, which is also simpler to implement than the case where electoral motivations play a role. Doing so requires that leaders are somehow able to share the benefits of their joint ventures. I allow for this by giving founders the ability of committing to transfers to other founders in the beginning of the game that are conditional on whether or not the founder receiving the transfer proposes a party. In reality these transfers will often take the form of powerful positions in the party hierarchy or in government, which are pledged to party leaders who agree to enter into an alliance. The ability to credibly commit to these transfers, on the other hand, can be justified by reputational concerns. Formally, such a promise takes the shape of a map $\{0,1\} \to \mathcal{R}_+$, which assigns an amount of utility to be transferred to a founder to an indicator for whether this founder has proposed a party or not. Accordingly, the strategy space of each founder is extended to allow for a choice of a vector of infinite length consisting of such maps. Let $\bar{t}$ be the sum of all transfers a founder receives net of any transfers the founder carries out. The utility of this founder is then given by

$$\sum_{s=1}^{S} \pi_s x_w + \pi_f x_f - c + \bar{t} \, ,$$

using the same notation as in section 2.4. Call this version of the game the extended party-formation game.

Beyond transfers, mergers also require the possibility of joint deviations by groups of founders. One way to allow for this would be to look for equilibria that are coalition-proof as defined by Bernheim et al. (1987). However, a much simpler approach, which turns out to be equivalent in this particular case, is to look for equilibria that are Pareto

efficient among founders in the extended party-formation game. I say that such an equilibrium is robust to party mergers.

As is shown in appendix A.2, transfers cannot be used to maintain formations of parties that are not stable in the original game. This would require sufficiently large transfers to all passive founders in order to prevent them from taking the opportunity of forming a successful party. In reality, one would expect it to be impossible to pay off all individuals in a position to initiate the formation of a party. The same is true here due to the assumption that the number of founders is infinite. The result would also hold for a finite but sufficiently larger number of founders.

Furthermore, it is necessary and sufficient for an equilibrium to be robust to party mergers that there exists no other equilibrium with a lower number of parties (Lemma A.2 in appendix A.2). Only an equilibrium with a lower number of parties generates a greater sum of utilities among founders, which can then be redistributed through transfers such that a Pareto improvement results. The reason that the sum of utilities is greater when there are fewer parties is due to the fixed cost $c$ that any party leader has to pay.

Combined these results make it possible to fully characterize the number of parties formed in party-formation equilibria that are robust to party mergers.

**Proposition 2.4.** *The number of parties in any party-formation equilibrium that is robust to party mergers is*

  i) *no lower than two and no greater than three,*

 ii) *equal to two whenever a party-formation equilibrium exists in which two parties are formed.*

*Proof.* See appendix A.2. □

A second look at figure 2.5 illustrates the set of equilibria that are robust to party mergers. In regions I and II, two-party equilibria exist and all other types of equilibria are therefore ruled out. In region III the only equilibria that are robust to party mergers feature three parties. The equilibrium used to establish this result is one where all existing parties allow only one type of politician to join, there is one such party for each possible platform, and the candidates of the two extremist parties tie at the federal election. Other three-party equilibria exist as well. An interesting possibility are equilibria which are "almost identical" to one of the two-party equilibria. As an example, consider the $M$-$E$ equilibrium with $I_M = [-1..0]$ and $I_E = [1]$. A comparison of the upper and the

lower panel of figure 2.5 shows that this equilibrium sometimes exists only if the number of states is sufficiently low. Otherwise too much internal competition makes extremist members of party $M$ willing to defect to a newly formed party of shape $[-1]$. This suggests the existence of an equilibrium where this smaller party forms in addition to the parties of the $M$-$E$ equilibrium and voters at the federal level behave as if the smaller party did not exist (this is possible under strategic voting). While not successful at the federal level, the additional party wins some state elections and thereby reduces internal competition among members of party $M$ sufficiently to prevent further deviations.

The model thus allows for the formation of parties that win state elections and play no role federally, but only if the expected value of career opportunities at the federal level is sufficiently low. Thinking more broadly, the model suggests an additional reason for the existence of parties whose success is confined to specific regions. Take the Scottish National Party (SNP) as an example. A politician who firmly believes that Scotland should be an independent country cannot expect to have a career at the national level of the UK. Naturally, such politicians will then try to establish a party of their own, explaining the existence of the SNP. The same reasoning applies to other regions with separatist movements, such as Catalonia in Spain, Quebec in Canada, or South Tyrol in Italy.

## 2.6 Robustness

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

### 2.6.1 Comprehensive Parties

Parties that allow all politicians to join were ruled out in the basic version of the model. This section will discuss what additional equilibria exist if parties of shape $[-1..1]$ were included in the action space of founders. As was already argued above, even a party that allows all politicians to join would not be able to deter entry of additional parties due to intense internal competition. However, focusing on the class of two-party equilibria, two additional stable constellations of parties emerge. One features either the equilibrium set of parties $\{[-1..1], [-1..0]\}$ or $\{[-1..1], [0..1]\}$. This constellation can actually be essentially equivalent to the $M$-$E$ equilibrium. This is the case when all eligible politicians join the smaller party in all states. Out-of-equilibrium beliefs of voters then prevent any deviations by politicians towards joining the larger party.

The second case is that an equilibrium with two parties of shape $[-1..1]$. This is possible because out-of-equilibrium beliefs can be used to make both parties have only one member in any state, as in the previous paragraph. It seems unlikely though that a party could maintain such widely varying ideological profiles across different states. If both parties had three members in most states, on the other hand, entry of a party of shape $[0]$ would be possible. This is because the larger parties then create too much internal competition. Furthermore, a centrist party could also do well federally, as it is relatively likely that both of the larger parties nominate a candidate with platform -1 or that both of them nominate a candidate with platform 1. In both cases a strict majority of voters would prefer a centrist candidate, who would accordingly win by assumption.

A way of eliminating the equilibria just discussed would be to allow politicians within a state to make joint-deviations. This would make it impossible that a party of shape $[-1..1]$ wins a state election with a single member in that state, as the remaining two politicians would be better off by coordinating on joining the winning party. By the same logic as in the preceding paragraph, entry of of a party of shape $[0]$ would then occur. Other equilibria discussed above, on the other hand, are robust to allowing for such joint deviations. On the equilibrium path, the constellations of parties considered actually give no scope for joint deviations, as at most one politician has a choice between different parties. Off the equilibrium path, there is no point in making joint deviations. The reason why politicians may find a deviation to an additional party attractive is that this can reduce the degree of internal competition they face at the state level. A joint-deviation would eliminate this benefit.

### 2.6.2   A Greater Number of Politicians

Allowing for a greater number of politicians is difficult, as this leads to a greater number of cases to consider. It also increases the scope for coordination failure at the stage at which party affiliations are chosen and thus the number of equilibria. Nevertheless, one effect of a higher number of politicians populating each state is clear: More politicians imply greater internal competition for nominations at the state level. This would make existence of the two-party equilibria discussed above less likely, as the magnitude of the payoffs $y_s$ and $y_P$ required in order to keep politicians from joining smaller parties increases proportionally with the number of party members. This may not be too much of a concern: The politicians in the model should be thought of as those who have already achieved some prominence within state parties and are therefore in a position to be considered for nominations. At any given point in time the number of such individuals will be limited. In addition, an emerging party focused on issues already covered by an existing party would find it hard to achieve credibility if it fails to attract any of the more

prominent members of that party. In fact, the formation of new parties is typically the product of a whole faction of an existing party defecting jointly. It is possible to interpret each politician in the model as representing factions who coordinate their actions.

### 2.6.3 Policy Choices

The assumption that politicians are committed to implementing their platform is not satisfying. While the empirical literature presented in section 2.3 seems to suggest that policy preferences of politicians are the main driver of their choices in office, it would be more appealing to see this behaviour emerge as part of an equilibrium rather than imposing it from the outset. In the model, extremist politicians can often increase their chances of winning the federal election by pretending to be a centrist when choosing state policies. To address this concern I will consider a more general utility function for politicians that includes both career concerns and policy preferences. For a politician with ideal policy $i$ let the utility function now be given by

$$\pi_s(y_s + \pi_n \pi_f y_f + (1 - \pi_n)\pi_P y_P) \; - \alpha \sum_{l \in \{s,f\}} (p_l - i)^2 \; ,$$

using the same notation as in section 2.4. In addition, assume that politicians can freely choose the policy they implement at any stage. All other elements of the game remain unchanged. This more general version of the model is challenging to solve in its entirety. I will present results for the election subgame reached after parties $L$ and $R$ have been proposed.

Proceeding by backwards induction, it is clear that any politician elected at the federal level will implement her ideal policy. All other stages are less straightforward. I will start by asking under what conditions a separating equilibrium exists where politicians implement their ideal policy at the state level and otherwise behave as in lemma 2.1. In this case a politician with ideal policy -1 achieves a continuation utility of

$$\frac{1}{w_L^*} \left[ \frac{1}{4} y_f + \frac{1}{4}(-\alpha 4) + \frac{1}{2}(-\alpha) \right] + \left( 1 - \frac{1}{w_L^*} \right) U_n$$

after winning a state election, where the term $U_n$ captures the utility in case the politician is not nominated for the federal election. Choosing the policy 0 after the state election results in a utility of

$$-\alpha + \frac{1}{w_L^*} \left[ \frac{3}{4} y_f + \frac{1}{4}(-\alpha) \right] + \left( 1 - \frac{1}{w_L^*} \right) U_n \; .$$

The politician now incurs a cost for a suboptimal policy choice but increases her chance of winning the federal election, simultaneously reducing the probability that an extremist from party $R$ gets elected. Note that voters do not observe that a deviation has taken place as the politician elected in the state has ideal policy 0 with positive probability ex-ante. It can then be shown that the separating equilibrium exists as long as the ratio $\alpha/y_f$ is no smaller than $\max_{j \in \{L,R\}} 2/(4w_j^* - 5)$.[11] The greater the number of elections won by a party, the lower the probability that any given politician will get nominated for the federal election, which in turn makes extremist politicians less likely to benefit from pretending to be a centrist.

Proceeding as above it can be shown that if the ratio $\alpha/y_f$ is smaller than

$$\min_{j \in \{L,R\}} 2/(4w_j^* - 5)$$

then there only exists a pooling equilibrium where all politicians implement the policy 0 at the state level.[12] Behaviour nevertheless remains very close to the one given in lemma 2.1. Dropping the assumptions that voters vote myopically (which was introduced above purely as a simplification), voters in state elections will base their choice on considerations regarding federal policies as all politicians implement the same policy at the state level. Federal policies are determined by the winner of the federal election, who is effectively picked at random from among state winners. State voters thus benefit from adding politicians to this pool that have similar preferences to them.[13] For example, a voter with ideal policy -1 in a state where politician 0 has joined party $L$ would not vote for party $R$ as a victory by party $R$ increases the probability that a politician with platform 1 wins the federal election. As long as politicians in both parties pool, the only difference between them is the degree of internal competition for the federal nomination, at least from the perspective of a centrist politician in a centrist state. As before, these politicians will therefore tend to sort into the party that wins fewer state elections.

The empirical evidence is in favour of the separating equilibrium, where all politicians implement their own ideal policies. Nevertheless, the discussion above suggests that the overall results do not necessarily change much even if extremist politicians try to pass off as centrists.

---

[11]This condition applies as long as each party wins at least two state elections, which implies that $4w_P^* - 5$ is positive.

[12]Cases where politicians in one party separate and politicians in the other party pool are more complicated and will not be considered here.

[13]Voters at the federal election are split at zero between both parties as they cannot distinguish between candidates exactly, but know that the candidate of either party may be a centrist or an extremist with equal probability.

### 2.6.4 Candidate Selection

Parties are assumed to nominate candidates randomly in the basic version of the model. This hurts the chances of either party of winning the federal election. However, if parties were to nominate a centrist politician for the federal election whenever one is available, this would reduce the expected utility of extremist party members and potentially induce them to join a third party. In order to prevent this, the party leader may have an incentive to commit ex-ante to nominating extremists with a sufficiently high probability. In practice this could be achieved through choosing a particular mechanism for candidate selection, such as primaries or nomination through voting by delegates at the party convention. As it turns out, the requirement to keep extremist party members satisfied may lead to extremists being nominated with even greater probability than in the basic model. This possibility is demonstrated in appendix A.3, which analyses an extension of the model where the party leader commits to a probability of nominating either type of politician at the federal election.

A second and closely related possibility of keeping the extremist wing of the party satisfied would be to increase the probability with which extremist politicians get nominated in extremist states. To some extent I feel that the assumption that the nomination process is fairly noisy from the perspective of party-outsiders is more appropriate at the state level than at the federal level. After all, state-level candidates are often little known to the public and different factions within the party will be pulling in different directions, with the party leader (the party establishment) certainly favouring moderate candidates due to their electability at the federal election. Nevertheless, nominating candidates through primary elections may be one way to ensure that candidates fit the preferences of the state median voter well. If this is the case, then is becomes even easier to see how national parties maintain regional monopolies: Extremist politicians in extremist states would basically face no internal competition from moderates and the existence conditions for the two-party equilibrium of proposition 2.2 would be greatly relaxed. It is noteworthy though that even in this case no single party would be able to monopolize all elections. Intense internal competition for the federal nomination would still lead to the successful entry of a second party and proposition 2.3 continuous to hold.

## 2.7 Conclusion

The aim of this chapter was to provide an explanation for why federal elections are typically strongly competitive, while state elections are often dominated by one party.

This pattern is a product of the interaction between political demand (voter preferences) and supply (political parties). The model presented here has focused on political parties while taking heterogeneity in voter preferences as given. It was shown that an equilibrium with two parties, one centre-left and one centre-right, can explain the main features of the data. In this equilibrium, states with centrist median voters will display small vote margins while state elections with more extreme median voters are dominated by one party. Such regional monopolies are possible for two reasons: First of all, parties cannot promise to implement policies other then the ones preferred by its members. This makes it impossible for, say, a left-wing party to appeal to right-wing voters. Secondly, and relatedly, a third party would have to attract some of the members of an existing parties in order to successfully contest a regional monopoly. The existing party, however, prevents its members from defecting by offering its members career opportunities (inside the party and at the federal level) that they would not have as members of a smaller regional party.

The federal level, on the other hand, turns out to be competitive as both parties are able to attract a symmetric candidate pool. If, in contrast, all moderate politicians were to sort into one party, this would skew the electoral odds in favour of that party. This is impossible. For example, if all centrists were to join the centre-right party, there would be some states with a very left-leaning electorate where the left-wing party nevertheless wins elections. In such states moderate politicians simply have to join the centre-left party if they want to win elections. An extremist state thus becomes a source of centrist candidates for the party that can win elections there.

It is, however, not necessarily by choice that each party maintains an ideological profile that makes it particularly attractive towards voters in some states. Moderating the party could potentially increase success at the federal level while also raising the number of votes received in some centrist states. If both parties did so, there would be no clear difference between them and no reason why one of them should do much better than the other in any state election. The force that prevents this outcome is the threat of entry of more extreme parties. It was shown that a single state with a voter located relatively far to the left and a single state with median voter located relatively far to the right is sufficient to make this threat of entry effective. The prospect of winning a single state election is enough for a third party to form, which would then also be able to participate in the federal election. The established parties do not shift towards the political centre due to the risk that the third party that would subsequently enter could attract enough votes to make the deviating party lose the federal election.

The data presented in this chapter was taken from a sample of federal states and the model was also framed in this context. I believe that the mechanisms discussed here

should nevertheless be useful for thinking about other types of institutional arrangements as well. For example, consider a politician who is planning to contest a seat in the House of Commons in the UK. If this person stands politically between the Labour Party and the Conservative Party, it seems likely that she will base her choice of affiliation on which party is more popular in her district. Having such moderate members will affect how either party is perceived by voters and accordingly also influence election results. Furthermore, all democratic countries hold regional elections. Winning these can serve to raise the profile of a newly formed party, even if this election does not receive quite as much attention as one for a state parliament would. The threat of entry would then operate in a similar way as it was described above.

A more serious shortcoming is that the analysis has focused on plurality rule, while other electoral systems were given no consideration. In particular, it would be interesting to know how well the results extend to proportional representation, where the number of parties achieving a significant share of votes is typically larger. Due to the already complicated nature of the model it would have been a challenge to incorporate the additional complexity inherent in a system of proportional representation. I hope to address this question in future work.

# Chapter 3

# Who Emerges from Smoke-Filled Rooms? Political Parties and Candidate Selection

## 3.1 Introduction

Before the emergence of primary contests, U.S. presidential candidates were selected by the leadership of their respective parties. The popular cliché of the nominee being chosen in "smoke-filled rooms" by men in dark suits with big cigars captures the sentiment that this process was undemocratic, intransparent, and ultimately to the disadvantage of voters. In the face of expensive and drawn-out primary elections, other observers have held that party establishments consist of professional politicians who know their potential candidates well and can judge which politician has the best chances of getting into office. Which of these competing views should be considered more reasonable?

This chapter tries to shed some light on this question through constructing a theoretical model of candidate selection through party elites. A key feature of this model is that the leadership of the party is better informed about potential candidates than voters are. In general, this enables parties to use their superior information to make informed decisions on behalf of voters. Whether they will do so, however, is not immediately clear. Parties often have interests that differ from those of voters and this is the second central assumption of the model. In this setting, can it ever be expected that parties will select the candidate that voters prefer?

The answer, it turns out, depends crucially on the degree of political competition. When competition is low, the party will win the election no matter which candidate it puts

forward and will consequently decide the nomination based on its own preferences. As competition increases, the party is forced to take into account which candidate voters prefer. Interestingly, this does not simply mean that the party leader more often chooses the candidate that voters prefer based on their own information. Instead, the party leader increasingly frequently nominates the candidate that voters would choose if they had the same information as the leader does.

Providing an intuition for this result requires a closer look at the model. The party leader chooses among two potential candidates and these politicians differ along two dimensions: Their ideological position and their quality[1]. Voters are not fully informed about these characteristics of politicians, while the party leader is. The conflict of interest between the party leadership and voters is assumed to be particularly strong along the ideological dimension, such that the median voter and the party leader would choose different candidates if the choice was purely based on ideology. On the other hand, everyone agrees that candidates of higher quality are more desirable, even though the weight that the party leader places on quality may be arbitrarily small.

Now suppose that the election is competitive, meaning that there exists a second party whose candidate is sufficiently attractive to voters in a sense to be made precise below. This enables voters to play a strategy such that the politician whose ideological position is further away from the one favoured by the median voter is less likely to be elected. In equilibrium, this lower electability of more extreme candidates neutralises the ideological bias of the party leader and as a result the nomination is decided based on quality. This also implies that the choice of candidate serves as a signal about the quality of the nominated candidate.

A party with polarized interests can thus be induced to select candidates in the interest of voters as long as competition is sufficiently strong. In fact, it may even be the case that the ideological bias of the party leader works to the benefit of voters. This result requires that the weight that the party leader attaches to quality is small. In this case competition is always required to induce the party leader to select candidates of high quality. Eliminating the ideological bias of the party leader has the effect that she more frequently nominates the politician that the median voter prefers based on ideology. But this effectively reduces competition, resulting in the selection of candidates of lower quality.

---

[1]Quality here describes a characteristic of politicians that is valued by voters independently of the implemented policy, such as honesty or competence. The political economics literature often uses the term "valence" instead of quality.

The choice of the party leader also serves as a signal about candidate quality to voters. The exact nature of the information revealed, however, depends on the level of competition. For example, the nomination of the moderate politician is a sign that she is of hight quality when competition is low. This is due to the preference of the party leader for the more extreme politician, which implies that the moderate can only win the nomination if she is of sufficiently high quality to make up for this disadvantage. When competition is high, on the other hand, the electoral advantage of the moderate politician reverses this logic.

The elite centred approach to candidate selection is of more than historical interest, as it is still used today even in well-established democracies. For example, in France the responsibility for the nomination of candidates for the National Assembly rests largely with central parties. In Germany, candidates for the office of prime minister are formally elected at national party conventions. However, there is typically only one candidate, namely the candidate previously announced by the party leadership. This example also fits well with another assumption of the model: Voters know who the potential nominees are and possess at least some partial information about them. In Germany, public discussions about who could become the next candidate start well in advance of the actual nomination. Here, the assumption that will be maintained throughout most of this chapter is that voters are well informed about the policies a politician stands for while they know little about quality. It could be argued that the careers of politicians prior to being considered for a nomination are more informative about policy than quality. After all, politicians make political decisions along similar ideological fault lines throughout their career. On the other hand, higher offices may require skills that a politician was not able to demonstrate before. This argument notwithstanding, a later section suggests that the results are robust to some uncertainty along the policy dimension as well.

This chapter is not the first attempt to analyse how parties generate candidates for elections. In a paper by Snyder and Ting (2002) voters initially have no information about individual politicians. By joining a party politicians can reveal their policy preferences to some extent, as parties impose costs on politicians who are located too far from the party platform. This model seems most appropriate for politicians in early stages of their career who voters know little about. Many of the papers where parties play a more active role in nominating candidates have considered how different methods of selecting candidates induce homogeneous candidates to supply effort (Caillaud and Tirole, 2002, Castanheira et al., 2010) or have focused exclusively on either the quality/valence dimension or the policy dimension. Quality is the centre of attention in Mattozzi and Merlo (2007, 2010), and Snyder and Ting (2011), while Cadigan and Janeba (2002) and

Jackson et al. (2007) are concerned with policy.[2] Contributions that feature both quality and policy are Adams and Merrill (2008), Serra (2011), and Boleslavsky and Cotton (2015). None of the papers mentioned so far feature a party leadership with superior information about the characteristics of politicians, while Snyder and Ting (2011) is the only paper where the degree of competition that the party faces plays an important role.

There are other papers that do not deal with candidate selection directly, but are nevertheless related. Callander (2008) and Carrillo and Castanheira (2008) show how more extreme platforms can be used to signal high quality under certain circumstances. The same may be true here, but the relationship between quality and ideology is more subtle: When competition is weak, nominating a more ideologically extreme candidate can actually be a signal of low quality. Caillaud and Tirole (1999) argue that ideological conflict within a party is required for platform choice to reveal information about quality. This chapter shows that all that is required for voter learning is superior information on the side of the party leadership.[3]

Among the papers given above, Adams and Merrill (2008), Serra (2011), and Snyder and Ting (2011) investigate the question of why parties may choose to adopt primaries to select their candidates. They take the benefit from primaries to be that they reveal information about the quality of politicians, with the most competent one going on to win the nomination. This can give the party a competitive edge. The benchmark that this is compared to, however, is that the party has only one potential candidate or chooses randomly. As Snyder and Ting (2011) point out (p. 783, footnote 8), "Naturally, introducing a primary would benefit a party less electorally if it had an alternative selection mechanism that more frequently generated the voter's preferred candidate."

The answer that this chapter provides to the point raised by Snyder and Ting is that the revelation of information during a primary campaign may not benefit the party at all when compared to candidate selection through an informed leadership, as demonstrated in section 3.3.4.2. This is true even when the interests of the primary electorate are perfectly aligned with the party leadership and primaries are very effective at revealing information to the general electorate. The reason for this is that having more information than the electorate can work in favour of the party leadership. As long as voters are uncertain about the quality of a candidate, even an incompetent one can get elected.

The basic model will be presented in the next section. Section 3.3 describes the different shapes that equilibrium takes depending on the degree of competition. In addition,

---

[2]These last two papers are quite similar to the current chapter in that they extend a citizen-candidate model by candidate nomination through parties. Compared to those contributions, the results here show that there is less policy convergence when candidates also differ in quality.

[3]There is voter learning even when the party leader has the same ideal policy as the median voter, as discussed in section 3.3.4.3.

results on welfare and some comparative statics are presented. Subsequently, section 3.4 relaxes some of the assumptions made in the basic version of the model. Section 3.5 concludes.

## 3.2 The Model

$N$ voters ($N$ odd) care about two characteristics of politicians. The first is their policy preference: Each politician has an ideal policy $i \in \mathbb{R}$. The second characteristic is quality. A politician can either be of low or high quality $q \in \{0, 1\}$.[4] While the quality of policy maker enters the utility function of voters directly, they care about policy preferences because it is assumed that elected politicians implement their ideal policy. This assumption is supported by a number of empirical studies (Bhalotra and Clots-Figueras, 2014, Chattopadhyay and Duflo, 2004, Lee et al., 2004, Levitt, 1996). The utility of a voter with ideal policy $i$ from a policy $x$ implemented by a policy maker with quality $q$ is

$$u_i(x, q) = -(i - x)^2 + q .$$

In this setting the outcome of the election is determined by the median voter, whose ideal policy is assumed to equal zero. More general utility functions could easily be accommodated. The utility of voters over policies could be given by any concave function that is uniquely maximized at $i$. It would also be possible to introduce a weight on quality. These changes would merely shift the boundaries where different equilibria occur in the parameter space but not the nature of the equilibria themselves. The additive separability between policy and quality is discussed in section 3.4.

Politicians belong to either one of two parties. The current incumbent belongs to party $I$ and through acting as policy maker has already revealed her quality $q_I$ and ideal policy, which is also denoted by $I$ and assumed to be smaller than zero. An incumbent is introduced purely to simplify the exposition. It would also be possible to let two parties compete by choosing candidates, which would yield qualitatively very similar results. Denote by

$$\mathcal{I} \equiv -I^2 + q_I$$

the utility that the median voter would receive from re-electing the incumbent.

The second party, party $C$, has a party leader whose role it is to nominate one of two politicians as the party's candidate for the election. The ideal policies of these two politicians lie in the interval $[0, 1]$. The politician located further away from zero is

---

[4]It would also be possible to let quality be a continuous variable. The binary representation of quality is chosen for simplicity.

referred to as the extremist and her most preferred policy is given by $E \in (0, 1]$. Her competitor for the party nomination is called the moderate, with ideal policy given by $M$ with $0 \leq M < E$. Politicians are identified by their ideal policies. Voters know that their respective qualities, $q_M$ and $q_E$, independently take the value one with probability $\pi$, which is also the unconditional expectation of quality. The party leader, on the other hand, observes qualities directly. All other variables are common knowledge.

The party leader can be thought of as representing the group at the top of the party hierarchy, which controls the nomination process. The ideal policy of the party leader is different from the one preferred by the median voter and assumed to be equal to one. The utility function of the leader is given by

$$u_C(x, q) = -(1 - x)^2 + w \cdot q + \mathbb{1}_{\omega \in \{M, E\}} Y,$$

where $\omega$ indicates the winner of the election, $Y \geq 1$ is a payoff that the leader receives if the winner belongs to her party, and $w \in (0, 1]$ is the weight that the party leader attaches to quality. $Y$ is introduced to make sure that the party leader never prefers the re-election of the incumbent over the election of one of the politicians belonging to party $C$. Allowing $w$ to be smaller than one implies that the party leader may put less weight on policy than voters do, which yields the most interesting results.[5]

It is worth pausing here for a moment to further discuss some of the features of the model. The assumption that the ideal policy of the party leader is one is made for simplicity. What is actually crucial for the results is that the party leader is located closer to the extremist than to the moderate. Regarding the politicians of party $C$, a noteworthy assumption is that the moderate and the extremist are never at a distance greater than one. This implies that competition takes place in a range where quality trumps policy: Voters always prefer any high quality politician over any low quality politician. Allowing politicians to be further away from each other would not introduce any additional types of equilibria. Finally, restricting attention to two potential candidates is necessary to keep the model tractable. It would seem though that the qualitatively important feature is that the number of politicians competing for the candidacy is "small". As the number of competing politicians grows the trade-off between policy and quality that the party leader faces disappears as high quality candidates become more and more abundant. The assumed scarcity of potential nominees seems, however, to also be a realistic choice. Parties rarely recruit outsiders and in order to be considered for nomination for a higher office party members typically need to have gained some experience as well as a public profile through serving in regional or local offices. Another restriction is that regional

---

[5]One reason why the party leader may put less weight on quality is that she faces pressure to nominate the extremist from the more radical members of the party, who may withdraw their support if they feel that their interests are not sufficiently taken into account. See chapter 2.

offices seem to require regional candidates.[6] Only a limited number of politicians will satisfy these criteria at any point in time. Finally, a viable candidate also has to stand a chance to win the election once nominated. As the following sections will show, even when the party generally has two politicians available there may be situations when only one of them is able to compete, making the presence of a second potential candidate practically irrelevant.

The strategic players in this game are the party leader and the median voter. After observing the quality of her politicians the party leader nominates one of them as the party's candidate for the election. The party leaders strategy is given by the function $\eta_M(q_M, q_E)$, which gives the probability that the leader will nominate the moderate given any realization of the qualities of both politicians. While this is sufficient to fully describe the strategy of the party leader, it will be convenient to directly refer to the probability of nomination of the extremist as well, which is given by $\eta_E(q_M, q_E) = 1 - \eta_M(q_M, q_E)$. After the nomination decision has been made, voters update their priors and vote for the incumbent or the challenger nominated by party $C$. The outcome of the election is driven by the median voter and it is therefore sufficient to focus on her behaviour. Let $r(p)$ be the probability that the median voter elects the candidate of party $C$ given that politician $p$ has been nominated.

The structure of the game is that of a signalling game, where the party leader is the sender and the median voter is the receiver. In the language of signalling games, the type $q_C \equiv (q_M, q_E)$ of the party leader is the combination of qualities she observes and the type-space is $Q \equiv \{0, 1\}^2$. The posterior probability that the nominated politician is of high quality is denoted by $\bar{\pi}_p$.

Signalling games typically have many perfect Bayesian equilibria, as it is possible to assign any belief that supports an equilibrium at information sets that are off the equilibrium path. The same is true here: For example, if voters believe that the extremist has quality zero, always nominating the moderate independent of actual qualities is an equilibrium. To be able to make sharper predictions it is therefore imposed that beliefs off the equilibrium path satisfy the refinement of Universal Divinity due to Banks and Sobel (1987), which has previously been applied in the literature (Banks, 1990, Callander, 2008). To give an informal description of the requirements of Universal Divinity, suppose that voters observe that the party leader unexpectedly nominates a certain politician. Voters then believe with certainty that the quality of the unexpectedly nominated politician must be such that it makes the leader most likely to gain from this move. The notion of "most likely to gain" is formalized as the type of leader that gains

---

[6]Members of the U.S. senate, for example, are almost always native to the state where they were elected. Furthermore, they also tend to highlight this fact in the biographical section of their website or even directly on the homepage.

in utility for the greatest set of voter responses: Let $\Lambda(p|q_C)$ be the set of election probabilities such that the party leader of type $q_C$ receives a greater expected utility from nominating politician $p$ rather than her competitor. If politician $p$ never gets nominated then $\bar{\pi}_p$ is restricted to be consistent with the belief that $q_C \in Q^*$, where $Q^*$ contains all $q^*$ that satisfy $\Lambda(p|q^*) \supseteq \Lambda(p|q') \; \forall q' \in Q$.

An additional issue more specific to this particular model is that the party leader is indifferent between all possible strategies once neither politician belonging to party $C$ can get elected. As a consequence the party leader could be playing the strategy "always nominate the politician with the lowest quality", which in turn could make it a best response for the median voter to re-elect the incumbent with certainty. However, it seems implausible that voters would expect the party leader to behave in this way. In order to circumvent this issue all equilibria that feature weakly dominated strategies are excluded. As intended this requirement only affects equilibria where both the extremist and the moderate get defeated by the incumbent with certainty.

## 3.3   Results

Whether or not a candidate nominated by the leader of Party $C$ stands a chance of getting elected depends on her political position as well as the expectation of voters regarding the quality of that candidate. Candidates that are very close to the median voter's most preferred policy can get elected even if they are perceived as being of low quality. Conversely, even a candidate far from the centre can be appealing to the median voter if her expected quality is high enough. However, this expectation of high quality is difficult to maintain. Suppose that the extremist gets elected with certainty once nominated because voters believe that the party leader nominates the moderate if the extremist turns out to be of low quality. Given this high probability of winning, the leader then actually prefers to nominate the extremist even when she is of low quality, since the extremist is politically closer to the leader. This undermines the initial expectation that the extremist is of high quality.

The exact shape of equilibrium therefore depends on the positions of both potential candidates of party $C$. If both are located close enough to the median the incumbent never gets re-elected. This case is referred to as "No Competition". The case labelled "Limited Competition" describes the situation where only the moderate can get elected. This requires that the moderate is close to the centre while the extremist is indeed too extreme and the median voter can never be persuaded to elect her. The most interesting case, called "Full Competition", features a positive probability of election for either politician belonging to party $C$ as well as the incumbent. The next three sections

explore each case in more detail. Finally, it is also possible that neither the moderate nor the extremist stands a chance of being elected. Obviously, this requires that both politicians are relatively far from the centre. Determining the exact conditions under which this is an equilibrium, however, is a rather technical exercise, which is therefore relegated to appendix B.1.

### 3.3.1 No Competition

Characterizing this equilibrium is straightforward: If both politicians of party $C$ are located close enough to the median voter the incumbent never gets re-elected: $r(M) = r(E) = 1$. Depending on the distance between the moderate and the extremist, the party leader may then behave in two different ways. In the first case, the two potential candidates are located so far from each other that the party leader always nominates the extremist independent of qualities. Accordingly, voters expect that the extremist has average quality: $\bar{\pi}_E = \pi$.[7] The second case applies if the two politicians are so close to each other ideologically that the party leader prefers a moderate of high quality over an extremist of low quality, but nominates the extremist in all other cases.[8] This implies that voters expect the moderate to have high quality if nominated ($\bar{\pi}_M = 1$) while the posterior quality of a nominated extremist is given by

$$\bar{\pi}_E = \frac{\pi}{\pi + (1 - \pi)^2} \ .$$

Both versions of this equilibrium exist as long as the median voter at least weakly prefers the extremist over the incumbent, which is equivalent to the condition $E \leq \sqrt{\bar{\pi}_E - \mathcal{I}}$.

In this equilibrium the median voter has no means to discipline the party leader who chooses her preferred politician without having to worry about electability. Consequently, the median voter would be better off if the ideal policy of the party leader was closer to her own ideal policy. The threshold on the position of the party leader at which her nomination strategy changes is the point at which she is equidistant from both politicians: A party leader who is located closer to the moderate than to the extremist would nominate the moderate whenever both politicians have the same quality.

---

[7]To give a complete description of this equilibrium the belief of voters over the quality of the moderate would have to be specified as well. According to lemma 3.1 below, Universal Divinity implies $\bar{\pi}_M = 1$, which makes $r(M) = 1$ the best response of the median voter to the nomination of the moderate: As the median voter prefers an extremist of average quality over the incumbent, she must also prefer a moderate of high quality over the incumbent.

[8]This case applies if and only if

$$w > -(E - 1)^2 + (M - 1)^2 \ .$$

## 3.3.2 Limited Competition

When only one politician in Party $C$ can successfully challenge the incumbent this is also the only politician that can get nominated. Nominating the candidate that loses for sure could only be optimal for the party leader if the utility from the other candidate getting elected was lower than the utility from the incumbent being re-elected. Due to the assumption that the payoff $Y$ from winning the election is at least one this is impossible. It follows that the party leader must always be nominating the politician that wins with positive probability.

In this situation voters cannot use Bayes' rule to update their belief over the quality of the politician that never gets nominated. The restrictions imposed on this off-equilibrium path belief by Universal Divinity are given by the following lemma.

**Lemma 3.1.** *Fix some $p \in \{M, E\}$. An equilibrium in which $\eta_p(q_C) = 0$ for all $q_C \in Q$ satisfies Universal Divinity if and only if $\bar{\pi}_p = 1$.*

*Proof.* Let $p'$ denote the competitor for the party nomination of politician $p \in \{M, E\}$. The interim utility of the party leader under a strategy profile $\sigma = (\eta_p, r)$ where $\eta_p(q) = 0$ for all $q \in Q$ (politician $p$ is nominated only off the equilibrium path) is given by

$$r(p')[-(p'-1)^2 + w \cdot q_{p'} + Y] + (1 - r(p'))[-(I-1)^2 + w \cdot q_I] .$$

Suppose politician $p$ would be elected with probability $\lambda$ if nominated. The utility of the party leader from nominating $p$ would then be

$$\lambda[-(p-1)^2 + w \cdot q_p + Y] + (1 - \lambda)[-(I-1)^2 + w \cdot q_I] .$$

Equating the two utilities and solving for $\lambda$ yields the probability of electing politician $p$ that makes the party leader indifferent between nominating either politician:

$$\lambda_p(q_c) = \frac{r(p')[-(p'-1)^2 + (I-1)^2 + w(q_{p'} - q_I) + Y]}{[-(p-1)^2 + (I-1)^2 + w(q_p - q_I) + Y]}$$

As $q_p$ only shows up in the denominator of this expression, the minimum of $\lambda_p(q_c)$ can only be attained for $q_p$ equal to one. Universal Divinity therefore implies $\bar{\pi}_p = 1$ as it holds that $\Lambda_p(q_c) = (\lambda_p(q_c), 1]$. $\square$

Intuitively, as the party leader puts a positive weight on quality, she is most likely to gain from nominating a candidate if that candidate has high quality. Universal Divinity accordingly requires that voters believe that unexpectedly nominated politicians have high quality. The politician that never gets elected must consequently be the extremist. Otherwise the median voter would strictly prefer an unexpectedly nominated moderate

over the incumbent and $r(M) = 0$ would not be a best response to the nomination of the moderate.

As the moderate is always nominated she is expected to be of average quality: $\bar{\pi}_M = \pi$. The median voter has to at least weakly prefer her over the incumbent in order to elect her with positive probability, which is equivalent to the condition $M \leq \sqrt{\pi - \mathcal{I}}$. Whenever this holds as a strict inequality the moderate is elected with certainty. In addition, not electing the extremist must be a best response. This requires that the median voter at least weakly prefers the incumbent over an extremist of high quality, which is the posterior implied by Universal Divinity. This implies the condition $E \geq \sqrt{1 - \mathcal{I}}$.

Limited Competition is the exact opposite of No Competition in the sense that in the former case the party leader is completely constrained in her choice of which politician to nominate. Accordingly, the preferences of the party leader over policies are of no consequence for the outcome of the nomination process.

### 3.3.3 Full Competition

In the two previously discussed cases electoral incentives were either too weak to discipline the party leader or too strong to enable her to choose candidates based on quality. The type of equilibrium discussed in this section falls in between those extremes. Here, the extremist is not clearly better or worse than the incumbent from the perspective of the median voter, who prefers the extremist only if she believes her to be of sufficiently high quality. The choice of the party leader to nominate the extremist must then be a credible signal that this is indeed the case. This becomes possible because in equilibrium the extremist is less likely to be elected than the moderate. This lower electability offsets the ideological bias of the party leader, with the consequence that the extremist is chosen only if her quality is high enough. The following proposition states the conditions under which this equilibrium exists. In particular, the extremist must be located such that the median voter would prefer her over the incumbent if she had high quality with certainty but not if she had average quality. In addition, the moderate cannot be located too far from the median either.

**Proposition 3.1.** *An equilibrium where both politicians belonging to party $C$ and the incumbent get elected (i.e. $r(M) > 0$, $r(E) > 0$, and $r(M) + r(E) < 2$) exists whenever $\sqrt{\pi - \mathcal{I}} \leq E \leq \sqrt{1 - \mathcal{I}}$ and*

$$M \leq \sqrt{\frac{\pi(\mathcal{I} + E^2)}{\mathcal{I} + E^2 - \pi(1 - \pi)} - \mathcal{I}} \ .$$

*Furthermore, $r(M) = 1$ in any such equilibrium.*

*Proof.* First of all, it is stated without formal proof that it is impossible that $0 < r(M) < 1$ and $0 < r(E) < 1$ simultaneously. This would require that the median voter is indifferent between all candidates, which in turn would require that the party leader plays a mixed strategy under more than one combination of politician qualities. Otherwise it is impossible to generate the posterior beliefs that make the median voter indifferent. As should become clear below, however, indifference of the party leader between her pure strategies can only hold for one pair of politician qualities at a time.

Next, assume that the politician getting elected with certainty was the extremist. This would imply that the moderate either never gets nominated or is chosen only in the case $q_C = (1, 0)$, depending on the value of $w$. Both cases lead to the posterior belief $\bar{\pi}_M = 1$. But if the median voter is willing to elect the extremist then she must certainly prefer a moderate of high quality over the incumbent as well, contradicting that $r(M) + r(E) < 2$.

It must therefore be true that $r(M) = 1$ and $r(E) < 1$. This can only hold if the median voter is indifferent between the incumbent and the extremist, which requires

$$\bar{\pi}_E = \mathcal{I} + E^2 \ . \tag{3.1}$$

To generate this posterior expected quality of the extremist the party leader must be playing a mixed strategy. In equilibrium mixing is only possible for one particular realization of qualities, as different combinations of qualities require different election probabilities to achieve indifference of the party leader. As the moderate gets elected with certainty the expected utility of the party leader from nominating the moderate is

$$-(M-1)^2 + w \cdot q_M + Y$$

while nominating the extremist gives

$$r(E)[-(E-1)^2 + w \cdot q_E + Y] + (1 - r(E))[-(I-1)^2 + w \cdot q_I] \ .$$

Equating the two utilities it is possible to derive the following identity:

$$r(E) = \frac{[-(M-1)^2 + w \cdot q_M + Y] - [-(I-1)^2 + w \cdot q_I]}{[-(E-1)^2 + w \cdot q_E + Y] - [-(I-1)^2 + w \cdot q_I]} \ . \tag{3.2}$$

Given the restrictions on parameters the expression on the right-hand side is always positive. In the case of $q_M = q_E = 0$ the numerator is smaller than the denominator and accordingly there exists an election probability $r(E)$ that leaves the party leader indifferent between nominating either a moderate or an extremist of low quality.

Indifference between politicians of low quality implies that under the quality combinations $(1, 0)$ and $(0, 1)$ the party leader nominates the politician of high quality, while in the case of both having high quality the party leader strictly prefers to nominate the moderate. The last point can be seen by recognizing that in this case the utility from nominating the moderate is equal

to the utility of nominating a moderate of low quality plus $w$ and the utility from nominating the extremist equal to the utility of nominating an extremist of low quality plus $r(E)w$. Hence, indifference in the $(0,0)$-case implies that the difference in utilities from nominating the moderate and the extremist is equal to $w(1-r(E))$ in the $(1,1)$-case, which is positive. Given this strategy of the party leader, posterior expectations are given by

$$\bar{\pi}_M = \frac{\pi}{\pi \; + \; (1-\pi)^2(1-\eta_E(0,0))} \tag{3.3}$$

and

$$\bar{\pi}_E = \frac{\pi}{\pi \; + \; (1-\pi)\eta_E(0,0)} \; .$$

Solving this last equality for $\eta_E(0,0)$ and using equation (3.1) to substitute for $\bar{\pi}_E$ gives

$$\eta_E(0,0) = \frac{\pi(1-\mathcal{I}-E^2)}{(1-\pi)(\mathcal{I}+E^2)} \; . \tag{3.4}$$

For this expression to be no greater than 1, it must be true that $\mathcal{I} \geq -E^2 + \pi$. This first necessary condition for the existence of this equilibrium implies that the denominator is positive. The second condition, which ensures that the numerator is non-negative, is $\mathcal{I} \leq -E^2 + 1$. Finally, it has to be true that the median voter weakly prefers the moderate over the incumbent: $\mathcal{I} \leq -M^2 + \bar{\pi}_M$. After substituting equation (3.4) into equation (3.3) this condition can be written as

$$\mathcal{I} \leq -M^2 + \frac{\pi(\mathcal{I}+E^2)}{\mathcal{I}+E^2-\pi(1-\pi)} \; .$$

If the election strategy of the median voter was such that the party leader was indifferent if $q_M = 0$ and $q_E = 1$, then the party leader would strictly prefer to nominate the moderate whenever the quality of the extremist is zero. This implies $\bar{\pi}_E = 1$ and contradicts that the median voter could be indifferent between the incumbent and the extremist.

Indifference under $q_M = 1$ and $q_E = 0$, on the other hand, is possible only if $w$ is sufficiently small. As a consequence the extremist would be nominated whenever she has high quality and when both politicians have low quality. The posterior beliefs are then

$$\bar{\pi}_M = 1$$

and

$$\bar{\pi}_E = \frac{\pi}{\pi + (1-\pi)^2 + (1-\pi)\pi\eta_E(1,0)} \; .$$
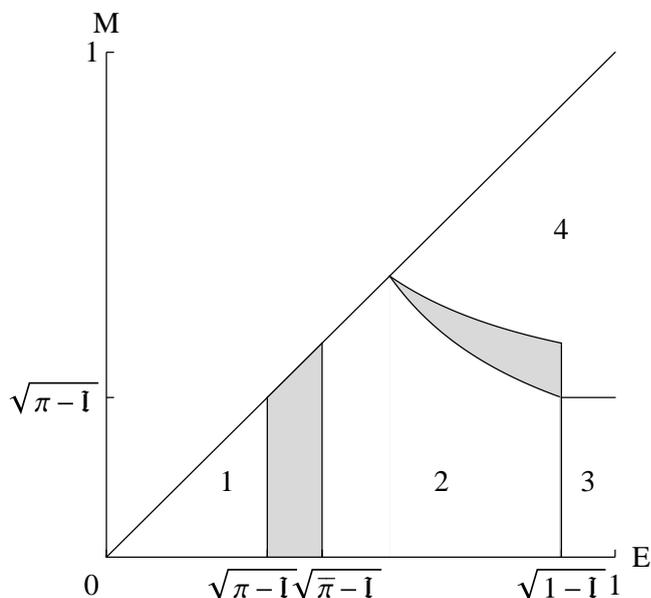
Solving this last equality for $\eta_E(1,0)$ and using equation (3.1) to substitute for $\bar{\pi}_E$ gives

$$\eta_E(1,0) = \frac{\pi - [\pi + (1-\pi^2)](\mathcal{I}+E^2)}{(1-\pi)\pi(\mathcal{I}+E^2)} \; . \tag{3.5}$$

The necessary and sufficient conditions for this expression to be positive and no greater than one are

$$-E^2 + \pi \leq \mathcal{I} \leq -E^2 + \frac{\pi}{\pi + (1-\pi^2)}.$$

FIGURE 3.1: Equilibrium for Different Positions of Politicians



The requirement that the median voter at least weakly prefers the moderate over the incumbent in this case is equivalent to the condition $\mathcal{I} \leq -M^2 + 1$.

Finally, suppose the party leader is indifferent between nominating either politician if both are of high quality. Proceeding as before, an equilibrium with this feature can be shown to exists under the same conditions as in the previous paragraph □

As the preceeding proof shows, there are up to three equilibria that satisfy the definition of Full Competition. All of them require that the median voter does not always vote for the extremist. In other words, the median voter must be playing a mixed strategy in case the extremist is nominated. In order to achieve the required indifference between the incumbent and the extremist the party leader herself must be playing a mixed strategy under a particular combination of qualities. In one possible equilibrium the party leader is indifferent between nominating either politician if both are of low quality ("low quality indifference") while in the other two cases indifference holds for other combinations of qualities. The equilibrium with low quality indifference exists more widely, as can be seen in figure 3.1, which exemplifies the existence conditions for the different types of equilibria for given values of $\mathcal{I}$, $\pi$, and $w$.[9] The possible combinations of $M$ and $E$ lie below the 45-degree line, as it holds that $M < E$. All areas where more than one equilibrium exists are shaded. The No Competition equilibrium exists in area 1, when

---

[9]For simplicity, the figure shows the limit case as $w$ approaches zero. Otherwise region 1 would have to be subdivided according to the two different types of No Competition-equilibria.

both politicians are relatively close to zero. Full Competition occurs in area 2 and both shaded areas. Within this area, the equilibrium with low quality indifference exists everywhere while the other two cases are confined to the shaded region bordering on area 1. In area 3, where the extremist is located far from the median, the Limited Competition equilibrium is the unique equilibrium. Finally, in area 4 and the bordering shaded region equilibria exist where no politician of party $C$ can get elected as both of them are too far from zero.[10]

Not all equilibria that exist under Full Competition discipline the party leader to act in the interest of the median voter to the same extent. However, as the degree of competition increases eventually the equilibrium with low quality indifference becomes the unique equilibrium.[11] In this equilibrium electoral incentives work very well in disciplining the party leader. This is in line with general message of this chapter that electoral competition needs to be sufficiently strong in order to induce the part leader to select "good" candidates. For the remainder of the chapter it will therefore implicitly be the equilibrium with low quality indifference that is referred to when Full Competition is mentioned. Under this equilibrium, the only case where the party leader does not always nominate the politician preferred by the median voter is the case of both politicians having low quality. Importantly, the fact that the party leader otherwise follows the preference of the median voter in her nomination choice is not driven by the party leader's own preference for politicians of high quality. In fact, $w$ can be arbitrarily small as long as it remains positive. This is because the ideological appeal of the extremist is neutralised by her lower electability in equilibrium. The party leader's decision is therefore driven by quality even when she attaches little value to quality in general.[12]

While the expected quality of the extremist is always such that the median voter is indifferent between the extremist and the incumbent, the relationship between the expected quality of the moderate and of the extremist depends on the utility that the median voter gets from re-electing the incumbent, $\mathcal{I}$, which can be interpreted as the degree of competition. If $\mathcal{I}$ is relatively low and there is only moderate competition, expected quality is higher for the moderate than for the extremist as in the No Competition case described above. As $\mathcal{I}$ becomes larger and competition intensifies this relationship reverses. In short, it is electability that determines which choice of nominee signals higher quality. The ideological preferences of the party have a tendency to make the extremist

---

[10]The boundaries on this region are derived in appendix B.1.

[11]The degree of competition is captured by the strength of the incumbent $\mathcal{I}$, as will be discussed in more detail in section 3.3.4.1.

[12]Even though the re-election of the incumbent is certainly the worst outcome for the party leader, she does not always nominate the politician who is most likely to defeat the incumbent. This is noteworthy as observers sometimes chide primary voters for not voting for the candidate with the highest chance of winning the general election.

look like the weaker candidate, but this is not true if this nomination choice implies a significant drop in the chance of winning the election.

The mixed strategy that the median voter plays when the extremist is nominated reflects the difficulty in maintaining the expectation that the extremist has high quality. Electing her any more frequently would make the extremist too attractive from the perspective of the party leader, which in turn would lower her expected quality and render this candidate a sure loser. A second interpretation of the mixed strategy is that the party leader is uncertain over the exact position of the median voter, which shows that the assumption of full information about the distribution of voters can be relaxed.[13]

### 3.3.4 Comparative Statics and Welfare Analysis

As has been pointed out in the previous section, parties do a pretty good job at selecting high quality candidates under Full Competition once electoral incentives are strong enough. In fact, parties maximize the average quality of their candidates in this equilibrium.[14] This is simply a consequence of the fact that the party leader never nominates a politician of low quality when a politician of high quality is available.[15] The maximal average quality is equal to the probability that at least one politician has high quality, which is $1 - (1 - \pi)^2$.

Using this result, it is possible to derive a simplified expression of the utility of the median voter under Full Competition, which is generally given by

$$\tilde{\eta}_M(-M^2 + \bar{\pi}_M) + \tilde{\eta}_E[r(E)(-E^2 + \bar{\pi}_E) + (1 - r(E))\mathcal{I}] \ ,$$

where $\tilde{\eta}_p$ denotes the ex-ante probability that politician $p$ gets nominated. In the Full Competition case the median voter is indifferent between the extremist and the incumbent: $-E^2 + \bar{\pi}_E = \mathcal{I}$. The previous expression can therefore be written as

$$\tilde{\eta}_M(-M^2 + \bar{\pi}_M) + \tilde{\eta}_E(-E^2 + \bar{\pi}_E) \ .$$

---

[13]This possibility will be discussed in more detail in section 3.4

[14]This is also true for the equilibrium in which the party leader is indifferent between politicians of high quality, but not for the equilibrium in which indifference applies to a moderate of high quality and an extremist of low quality. See the proof of proposition 3.1 for details.

[15]Proof: The average posterior quality of the moderate and the extremist is equal to the sum of their posterior qualities weighted by their respective nomination probabilities. The posterior quality of each candidate is given by the probability of being nominated conditional on having high quality divided by the unconditional nomination probability. The nomination probability therefore cancels out and the average quality is given simply by the sum of the nomination probabilities conditional on having high quality. This is maximized when no low quality candidate is nominated whenever a high quality candidate is available.

Using the fact that average quality is equal to $1 - (1-\pi)^2$ yields

$$\tilde{\eta}_M(-M^2) + \tilde{\eta}_E(-E^2) + 1 - (1-\pi)^2 \;. \tag{3.6}$$

This expression will be useful in the welfare comparisons below.

### 3.3.4.1 Increasing Competition

A crucial determinant of the shape that equilibrium takes is the strength of the incumbent, as given by the utility $\mathcal{I}$ that the median voter experiences in the case of re-election of the incumbent. A natural interpretation of $\mathcal{I}$ is that it represents the degree of electoral competition that the party of the challenger faces. From this perspective the model generates a prediction about the relationship between electoral competition and the expected quality of politicians. This can be seen by fixing a combination of political positions for the moderate and the extremist, i.e. a point in figure 3.1. For low enough values of $\mathcal{I}$ any such point will lie in region 1, where party $C$ faces No Competition. Increasing $\mathcal{I}$ (increasing competition) shifts the boundaries that separate the different types of equilibria towards the origin. Therefore, eventually the "Full Competition"-case applies. This is connected with an increase in the quality of nominated politicians, as higher competition forces the incumbent to select candidates of higher quality.[16] Increasing competition even further can have either one of two effects. If the moderate and the extremist are located close to each other (their position generate a point close to the 45-degree line in figure 3.1) increasing $\mathcal{I}$ will make the incumbent the certain winner of the election. If, on the other hand, there is a clear political difference between the moderate and the extremist, there exists an interval for values of $\mathcal{I}$ in which equilibrium takes the shape of what was labelled Limited Competition, where only the moderate has a chance of defeating the incumbent. The step from Full Competition to Limited Competition leads to a reduction in the quality of nominated politicians, as the party leader effectively has a smaller set of politicians to choose from.

Taken as a whole, the model predicts a nonlinear relationship between competition and quality. Starting from a low level, increasing competition is beneficial as it disciplines parties to select high quality candidates. In contrast, more intense competition can force parties to select candidates more on policy and less on quality when the election is already fiercely contested.

---

[16]This statement is true irrespective of which of the three equilibria that satisfy the definition of Full Competition is selected.

It should be noted that competition has been framed from the perspective of one party in this discussion. More generally speaking, one would actually consider the most competitive situation to be the one in which both parties face equal chances. In this perspective the model indeed predicts higher competition to lead to the nomination of candidates of higher quality. The unexpected result here is that reduced competition leads to worse outcomes both on the advantaged and the disadvantaged side.

### 3.3.4.2 Comparison to Primaries

A number of papers have argued that primaries reveal information about participating politicians and thus allow parties to select candidates of higher quality (Adams and Merrill, 2008, Serra, 2011, Snyder and Ting, 2011). The way that candidates are generated in the absence of primaries in these papers, however, is that either there is only one candidate or that the nomination occurs at random, while the quality of the nominee remains unknown in either case. If parties were instead selecting candidates as described here, the advantage of primaries would be much less clear. To demonstrate this point, this section will compare the results presented so far to the outcomes under a simple version of primaries where the nomination is decided by a vote among the party's rank and file. It will be assumed that the median voter among primary voters is decisive and thus effectively chooses between the extremist and the moderate. Two additional assumptions skew the odds heavily in favour of primaries. First of all, the disadvantage of holding primaries from the perspective of the party leadership in Adams and Merrill (2008) and Serra (2011) is that primary voters may have differing ideological preferences. This disadvantage will be eliminated here by positing that the median voter in the primary election has the same utility function as the party leader. Secondly, it will be assumed that campaigning in the run-up to the primary perfectly reveals quality. Despite these assumptions the party leadership may still prefer to retain control over the nomination, as will be shown below. The timing of the game under primaries is as follows: First, nature draws qualities and these then become perfectly observable to all players during the campaign leading up to the primary election. Subsequently the primary election is held, followed by the general election between the incumbent and the winner of the primary.

The outcome of the primary election can be easily summarised: Under any realisation of qualities, the median voter in the primary selects the politician she most prefers among those politicians who are able to defeat the incumbent at the general election. As the median party member has the same preferences as the party leader, this is also the politician that the party leader would select if there was no asymmetric information. Primaries in this simple setting are therefore essentially as if the party leader was giving

up her informational advantage. Without primaries there is always at least some pooling going on. That is, there is always a chance from the perspective of voters in the general election that the nominee of party $C$ is of high quality. This enables the party leader to get even politicians of low quality elected. If the quality of candidates becomes observable, this may no longer be the case. If the incumbent is particularly strong, on the other hand, pooling may be a disadvantage if it makes all potential candidates of party $C$ unelectable. Primaries may then enable the party to get at least politicians of high quality into office. A second potential advantage is that primaries increase the electability of an extremist of high quality relative to the case of Full Competition, which makes it worthwhile to nominate such an extremist more frequently as well. If $\pi$ is large and candidates are likely to be of high quality, primaries will therefore tend to work in the party leader's favour. But if average quality is low, the disadvantages of primaries will outweigh the benefits and this is true despite the assumption that there is no gap in terms of ideological preferences between the party leadership and primary voters.[17]
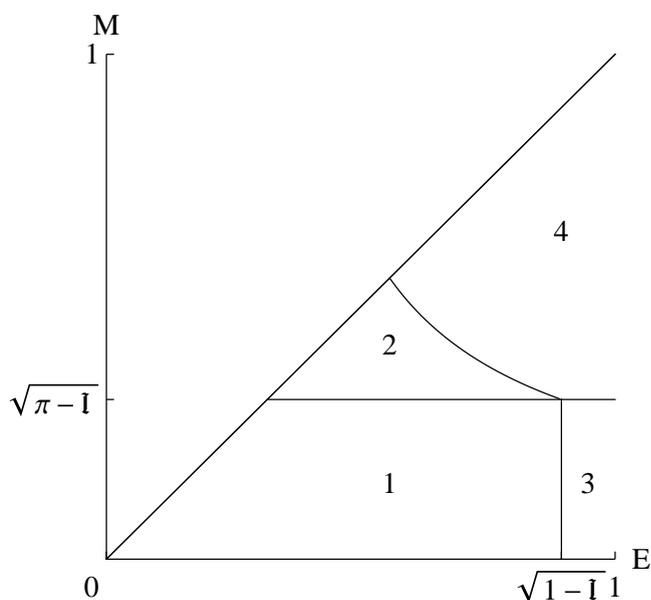
### 3.3.4.3  Common Interests

A central question raised in the introduction was whether the special interests of the party imply that it will select "bad" candidates. As was pointed out in previous sections, in the case of No Competition the median voter would indeed be better off if the party leader shared her political interests. In the case of Limited Competition, on the other hand, the preferences of the party leader over policies were of no consequence. What has not been taken into account so far though is that the existence conditions for the different types of equilibria also depend on the preferences of the party leader. These boundaries are shown in figure 3.2 for a party leader located at zero.[18] The boundaries on the equilibrium where the incumbent always gets re-elected (area 4) and the Limited Competition equilibrium (area 3) are unchanged. In contrast, the equilibrium where both the moderate and the extremist get elected with certainty exists much more widely, namely in area 1 in figure 3.2. Previously, the binding constraint on the existence of the No Competition equilibrium was that the median voter had to prefer an extremist of average quality over the incumbent. A party leader with the same preferences as the median voter, in contrast, always selects the moderate while the extremist is believed to be of high quality according to lemma 3.1. This shifts the boundary on the existence of this equilibrium outwards. Full Competition occurs in area 2.

---

[17]An example of parameter values under which holding a primary makes the party leader better of is $M = 1/2$, $E = 3/4$, $\pi = 4/5$, $w = 1/2$, and $Y = 2$. Reducing $\pi$ to $1/4$ while holding all other parameters constant implies that the party leader is better off selecting candidates herself.

[18]The derivation of the equilibria in the $C = 0$-case will not be given here as it proceeds exactly as in the case of $C = 1$. The figure again displays the limit case as $w$ approaches zero.

FIGURE 3.2: Equilibrium under Common Interests



When the party leader puts a relatively small weight on quality,[19] the change of the welfare of the median voter in the area where under diverging interests Full Competition applies while under common interests the moderate is always nominated is not immediately clear. Diverging interests lead to the nomination of higher quality candidates, while common interests result in the nomination of the preferred politician in terms of policy. In the latter case, the utility of the median voter is $-M^2 + \pi$. Under diverging interests, expression (3.6) shows that the expected utility of the median voter is given by

$$[\pi + (1 - \pi)^2(1 - \eta_E(0,0))](-M^2)$$
$$+ [\pi(1 - \pi) + (1 - \pi)^2\eta_E(0,0)](-E^2) + 1 - (1 - \pi)^2 .$$

Using equation (3.4) to substitute for $\eta_E(0,0)$, some tedious but straightforward algebra shows that the difference in the utilities can be written as $-\mathcal{I} - M^2$. Therefore, the median voter is better off under common interests in this particular case if and only if $-M^2 > \mathcal{I}$. In words, the median voter has to prefer a moderate of low quality over the incumbent - a rather strong condition.

Similarly, common interests can work to the advantage or the disadvantage of the median voter in the case where Full Competition applies both under common and diverging

---

[19]That is, $w$ is below the threshold at which the party leader nominates a moderate of high quality under No Competition when the extremist has low quality.

interests, which will not be shown formally here. Without knowing the distributions that the characteristics of politicians are drawn from, it is therefore not clear whether the special interests of the party make the median voter better or worse off. However, there seems to be substantial possibility of the former at least in the case where the party leader attaches little weight to quality. The reason for this is that a party leader is more likely to nominate a candidate who is ex-ante attractive to the median voter if they share the same political interests. This, however, reduces competition and therefore leads to the selection of candidates of worse quality. The effect on all voters depends additionally on the distribution of ideal points, but the possibility that the utility of the median voter can be lower under common interests certainly implies that the same can be true for the sum of all voter utilities as well.

## 3.4 Robustness

The model features a number of assumptions that can be relaxed. First of all, the results are robust to adding some uncertainty over the position of the median voter. As was already mentioned in the discussion of the case of Full Competition, it is possible to interpret the mixed strategy that the median voter is playing in this vein. The belief of the party leader over the position of the median voter would have to be given by a smooth density, which would make the election probability of the extremist a smooth function of her posterior quality. In contrast, all other equilibria do not feature mixing by the median voter but are nevertheless robust in a similar way. Here the differences between the possible candidates are so large that uncertainty over the position of the median voter would not translate into uncertainty over the outcome of the election.

Two further assumptions that will be discussed in more detail in the following two subsections are the additive separability of quality in the utility function of voters and the discrepancy between full information over politicians' positions and uncertainty over their quality.

### 3.4.1 Non-Additive Quality

Specifying quality as additively separable from policy has received criticism in the past. The main argument is that it seems implausible that, for example, a left-wing voter would want a right-wing candidate to be very effective at implementing policy. Put differently, quality should become a bad for a sufficiently high political distance. It

would be possible to allow for this effect by giving voters the following utility function:

$$-(i-x)^2 + h(|x-i|) \cdot q$$

where the function $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ is decreasing and positive at zero. The difficulty that arises with this specification is that the median voter may no longer be decisive, which would at the very least complicate the analysis of the model. However, additional assumptions would ensure the applicability of the median voter theorem (a proof can be found in appendix B.2) while still allowing for an interaction between ideology and quality as described above. These assumptions are that the function $h$ is concave and all voters are located in an interval $[-d, d]$ with $d > 0$ such that $h(d) \geq 0$.

If it is assumed in addition that $d \geq 1$, all the results remain qualitatively the same. A recent paper by Gouret et al. (2011) lends empirical support to the latter assumption. Using data from the French presidential election of 2007 the authors find that a utility function that allows for an interaction between quality and policy fits the data well while the simple additive utility function is rejected. However, the parameter estimates indicate that the main candidates are well within the range in which higher quality is beneficial to the median voter.

### 3.4.2 Uncertainty about Politicians' Policy Preferences

The distribution of information imposed in the model may seem to lack a strong justification. While voters know much about the policies a candidate stands for they know little about quality. Furthermore, many of the findings seem to rest on this skewed information structure: Voters observe policy preferences and are able to make inferences about the quality of candidates based on this observation. This section will argue that it is possible to introduce uncertainty about the policy positions of politicians while leaving the main results intact.

To this end, suppose that the policy positions of the candidates of party $C$, $M$ and $E$, are drawn from the distributions functions $F_M$ and $F_E$ respectively. For the moment these will not be specified any further. A party leader confronted with a particular draw of positions and qualities will decide whom to nominate based on a comparison of the expected utility resulting from either choice. This utility depends on the chance of each politician winning the election. To keep things reasonably simple, the disutility from policy will now be given by the absolute value, rather than the square, of the difference between policy and ideal position of an agent. Furthermore, assume that the party leader expects that the moderate would get elected with certainty while the extremist would get elected with probability $r(E)$, as in the Full Competition case above. The

decision rule of the party leader is then to nominate the moderate if and only if

$$-|M - 1| + w \cdot q_M + Y \geq r(E)[-|E - 1| + w \cdot q_E + Y] +$$
$$(1 - r(E))[-|I - 1| + w \cdot q_I]$$

or equivalently

$$M - r(E)E \geq r(E)[w \cdot q_E + Y] + (1 - r(E))(I + w \cdot q_I) - w \cdot q_M - Y$$
$$\equiv K(q_C) .$$

This choice rule implies that under different quality combinations politicians will be nominated with different probabilities and the nomination choice can therefore still be a signal of quality. The expected quality of a moderate nominated according to this rule is

$$\bar{\pi}_M = \frac{\sum_{q \in \{0,1\}} \pi \, Pr[q_E = q] \, Pr[M - r(E)E \geq K(q_C)|q_C = (q, 1)]}{\sum_{q \in Q} Pr[q_C = q] \, Pr[M - r(E)E \geq K(q_C)|q_C = q]} ,$$

which is simply the probability that the moderate gets nominated conditional on being of high quality divided by the unconditional nomination probability. One way to find an expression for $Pr[M - r(E)E \geq K(q_C)]$ is to first derive the density of the random variable $M - r(E)E$ at some point $t$. This is given by

$$\int_{\text{supp}(F_E)} f_E(e) f_M(t + r(E)e) \, de .$$

Appropriately integrating over this density one obtains the desired probability. The expression for the posterior quality of the extremist can be derived analogously.

Beyond quality the nomination choice can now also be a signal of the policy position of a candidate. Considering the decision rule of the party leader, one observation is immediate: If all possible candidates are closer to the median than the party leader, then it is impossible that the expectation of the posterior distribution of the policy position of a nominated politician is below the expectation of the prior distribution. If the party leader prefers to nominate the moderate for a given $M$ then she must ceteris paribus prefer to nominate the moderate for any higher $M$ as well, implying that the posterior distribution first order stochastically dominates the prior distribution. The same holds for the extremist. Therefore, if a nomination tells voters anything about the policies a candidate stands for then that these are more extreme than previously thought. In other words, politically extreme parties are bad for the median voter in terms of the political views of the candidates they select.

To find an expression for the expected policy position of a moderate nominated according to the decision rule above, first note that according to Bayes' rule the posterior

probability density over $M$ conditional on a certain quality combination $q$ is given by

$$f_{M|q}(m) \equiv f_M(m) \frac{Pr[M - r(E)E \geq K(q_C)|M = m, q_C = q]}{Pr[M - r(E)E \geq K(q_C)|q_C = q]}$$

with

$$Pr[M - r(E)E \geq K(q_C)|M = m, q_C = q] = F_E\left([m - K(q_C)]/r(E)\right) .$$

The unconditional expected policy position of a nominated moderate is then given by the weighted sum of the conditional expectations:

$$\frac{\sum_{q \in Q} Pr[q_C = q] \; Pr[M - r(E)E \geq K(q_C)|q_C = q] \int_{\text{supp}(F_M)} m \; f_{M|q}(m) \; dm}{\sum_{q \in Q} Pr[q_C = q] \; Pr[M - r(E)E \geq K(q_C)|q_C = q]} .$$

Again, the expected policy position of the extremist follows analogously.

Giving a general description of equilibrium is beyond the scope of this chapter. Instead, a specific example will be given to illustrate that the characteristics of the Full Competition equilibrium emphasized above remain unchanged in the extended model. It is assumed that both $M$ and $E$ are uniformly distributed with support $[0.2, 0.5]$ and $[0.4, 0.7]$, respectively, while incumbent is located at -0.8 and has high quality. Note that the moderate is expected to be closer to the median than the extremist, but the opposite might be the case in actuality. In addition, $\pi = w = 0.5$ and $Y = 1$ will be used.

Figure 3.3 plots the expected utility of the median voter from electing either politician of party $C$, which can be calculated using the expressions above, as a function of the probability $r(E)$ that the extremist will get elected. The dashed line represents the utility that the incumbent receives in case the incumbent is re-elected. For low values of $r(E)$ the party leader always selects the moderate and both expected utilities are flat in this region.[20] As $r(E)$ increases the party leader finds it worthwhile to nominate the extremist for high values of $E$ in the case where the extremist has high quality and the moderate has low quality, and eventually also for lower values of $E$. This makes the extremist less extreme in expectation and explains the initial increase in the expected utility from electing her. For even higher values of $r(E)$ the extremist gets nominated under other quality combination as well, which lowers her expected quality and results in a decrease in utility for the median voter. The increase in the expected utility from electing the moderate, on the other hand, stems from the fact that her expected quality increases as it becomes more attractive to nominate the extremist.

---

[20]In the extended model Universal Divinity implies that an unexpectedly nominated politician $p$ is of high quality and located as close to the party leader as possible given the distribution $F_p$.

Figure 3.3: Expected Utilities with Uncertain Policy Positions
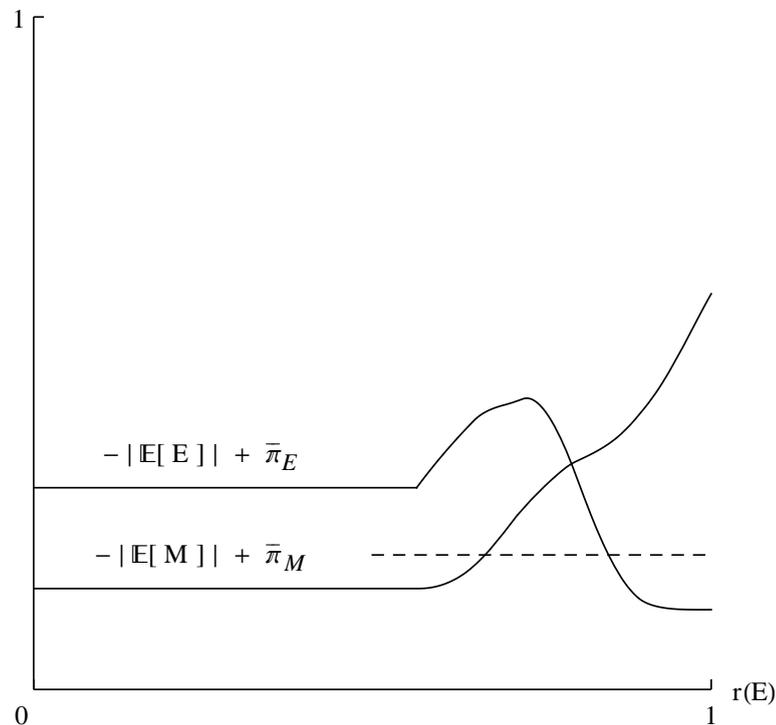


The figure shows that there is an election probability of the extremist such that the median voter is indifferent between the extremist and the incumbent while strictly preferring the moderate. This is equivalent to the Full Competition equilibrium described above.

## 3.5    Conclusion

This chapter has presented a model of candidate selection through party elites where the central premise was that the party leadership has more information about the characteristics of potential candidates than voters do. Given that the party leadership itself has preferences over these characteristics, the nomination choice often reveals information about the chosen candidate to voters. What exactly voters learn depends on the degree of competition a party faces. When competition is low, the nomination of an extreme candidate serves as a signal of low quality, while the opposite can be true when competition is more intense. In the latter case, electoral incentives strongly discipline the party leadership to select candidates in the interest of the median voter. Voters can therefore benefit when parties are polarized as this tends to increase competition compared to a situation where one party is located in the political centre.

An important implication of these results is that parties do not necessarily need to introduce primaries in order to generate candidates of high quality. In the model presented here the party leadership is often better off retaining control over the nomination of candidates even when many of the disadvantages of primaries discussed in the literature are absent. This raises the question whether alternative explanations for the introduction of primaries should be given closer consideration. For example, Hortala-Vallve and Mueller (2015) argue that primaries could help heterogeneous parties to prevent factions from defecting. This argument is also supported by the analysis in chapter 2.

From the perspective of voters, the potential downside of candidate selection through party elites is that parties prioritize ideology and select low quality candidates when competition fails. As was shown in chapter 2, such failure is common at sub-national levels of government where often only one party stands a realistic chance of holding office. This chapter thus points out one potential reason why such political monopolies are problematic.

# Chapter 4

# Competing Candidates, Competing Interest Groups and the Efficacy of Political Threats

## 4.1 Introduction

Political advertising is a fundamental element of any modern election campaign. Candidates use a multitude of means such as television or newspaper adverts, bill boards, or door-to-door canvassing in order to convince members of the public to vote for them. The importance that candidates attach to these activities is probably best illustrated by the share of their time devoted to raising the required funds. In the U.S., for example, a survey of former candidates found that more than 50 percent of those running for state-wide office spent more than a quarter of their time eliciting campaign money. 23 percent of candidates even reported that such activities took up more than half of their time (Herrnson and Faucheux, 2000). U.S. presidents attend fund-raisers not only during their own campaigns, but even during their second and final term in office. Former president Bill Clinton attended 471 such events during his second term, nearly three times more than during his first four years in office (Doherty, 2013). Such observations have raised concerns that the need to finance campaigns has significant opportunity costs in terms of less time spent fulfilling official duties. Even more worrying to some is the possibility that their high demand for campaign funds makes politicians willing to trade policy favours in return for donations.

It has been questioned, however, whether such quid pro quo is actually occurring. The argument is based on a simple observation first made by Tullock (1989) and later re-emphasised by Ansolabehere et al. (2003), now commonly referred to as the Tullock

paradox: When compared to the value of government regulations and subsidies, the amount spent on lobbying efforts and campaign contributions seems small. If these expenditures are viewed as political investments, simple back-of-the-envelope calculations reveal exorbitant rates of return. Ansolabehere et al. (2003) list a number of U.S. industries whose sum of campaign contributions is dwarfed by the gains that government policies imply for them. For example, in 2000 the U.S. government spent \$ 134 billion on defence procurement contracts while the defence industry gave \$ 23.8 million in campaign contributions over the 1999-2000 election cycle. Even assuming modest profit margins this amounts to a rate of return of several hundred percent on political donations. One would expect competition for these contracts among firms in the defence sector to eliminate such excessive rates of return. Why does this not seem to happen?[1]

In this chapter I propose a possible explanation of the Tullock paradox. I show that it might not be necessary to make contributions in order to have influence; the mere threat of contributions may be enough. In the model presented in this chapter an incumbent is facing re-election under competition from a challenger. The incumbent is willing to trade policy favours in return for donations from an interest group. Now it might be possible for the interest group to secure the same favours, not by contributing, but simply by threatening the incumbent with a donation to the challenger. Due to the zero-sum nature of the situation, what really matters to the incumbent is not the absolute amount she spends on advertising, but by how much she out-spends the challenger. This makes threats of donations just as effective as actual contributions.

Interest groups may nevertheless give money in equilibrium. Making a contribution to the campaign of the incumbent allows for the combined threat of withdrawing this donation and simultaneously giving money to the challenger. This gives the interest group even more influence over policy choices. As these policies then reflect the contribution the incumbent receives as well as the threat she is subject to, the model generates the appearance of very high returns on actually carried out donations.

Crucially—and in contrast to the existing literature—I show that this logic remains valid even when there is more than one interest groups. As explained above, the question of why competition among interest groups does not dissipate excessive rents is at the heart of the puzzle.

In order to explain how I achieve these results I will first describe the modelling approach generally used in the literature. Since the seminal work by Bernheim and Whinston (1986), and in particular through the contributions of Grossman and Helpman (1994,

---

[1]It is commonly argued that collective action problems may prevent certain groups from launching effective lobbying efforts. Such barriers to entry could help explains high rates of return in some cases, but it is not clear why this should apply to the example of firms competing for contracts.

1996), it has become customary in models of campaign contributions to allow interest groups to offer schedules of donations to candidates. These schedules make the money an interest group gives to a politician conditional on the politician's policy choices or campaign promises. Importantly, these schedules are viewed as being representative of the commitment power that interest groups would have in a game of repeated elections. In other words, contracts are seen as relational rather than legal. The motivation behind introducing these contracts is that they enable interest groups to make donations with the explicit aim of influencing policy, rather than just increasing the chances of a particular candidate once campaign platforms have been announced.

While the argument just given justifies the use of contribution schedules it says little about the exact nature of these contracts. As it turns out, this is crucial. In Grossman and Helpman (1996) donations are offered to a candidate as a function of this candidate's campaign platform only. This means that it is impossible to threaten candidates as their choice of platform cannot have any influence on how much money their competitor receives. In contrast, threats are possible in this chapter because the policy choice of the incumbent determines how much money both she and the challenger will receive.

When two interest groups are present, the additional issue arises that each one of them might want to change its own donations in response to donations made by the other group. In fact, if contribution schedules are viewed as representing informal commitments made by lobbies in the beginning of the game, it is hard to argue why they should not have the ability of making these commitments conditional on their opponents actions as well. I therefore allow interest groups to make their contributions a function of policy as well as of donations made by the other interest group. Similar contracts arise in other contexts, for example when retailers promise to match the prices of competitors, making prices a function of other prices. Peters and Szentes (2012) discuss other examples.

Independently of the particular choice of what donations can be conditioned on, the number of equilibria is large as soon as more than one interest group is present. This stems partially from the fact that interest groups are almost unconstrained in their commitments to contributions at policies that are not chosen in equilibrium. Threatening to make sufficiently high donations can ensure that these donations never actually have to be carried out. It seems desirable to introduce at least some chance that interest groups will be held to their word. I therefore require that equilibria are robust to small perturbations of the game where there is a small probability that the incumbent turns out to be an "ideologue" who sticks to some platform irrespective of how likely this choice is to lead to electoral success. Ex-ante interest groups thus perceive a small chance that they will have to carry out promises that would otherwise never be tested. This greatly

reduces the complexity of possible schedules and allows me to characterize the set of equilibria more fully.

I find that the set of possible equilibrium outcomes is largely determined by the maximum amount of contributions that each interest group is able to pledge. In particular, policies are always skewed in favour of the group with deeper pockets. This lobby is also the only one that makes contributions in equilibrium and only to the incumbent. In fact, I obtain the striking result that the weaker interest group remains almost completely passive, in the sense that it does not promise any donations for any policies that the incumbent might choose. The presence of the weaker group matters for outcomes nevertheless, as the general nature of contribution schedules allows this lobby to become active if the group with deeper pockets should try to gain even more influence. Intuitively, the stronger lobby does not exert as much pressure as it potentially could, because it knows that doing so would provoke a reaction from the so-far passive group.

While the equilibrium policy is moderated by the existence of a second interest group, the stronger group still uses a combination of threats and actual contributions in order to influence policies. This generates high rates of return on donations in the same way as outlined above. Bidding wars are possible out of equilibrium, but are not initiated by the weaker group in the knowledge that they would not bring any advantage.

This chapter is not the first contribution to allow interest groups to commit to more general contribution schedules. Chamon and Kaplan (2013) present a model with two candidates who compete by announcing campaign platforms. However, they only allow for one interest group and rely on much stronger, parametric assumptions in deriving their results. This is partially due to the more applied nature of their paper, in which they also provide empirical evidence in favour of the theoretical results. As the present chapter, their model predicts that split contributions—where one interest group contributes to both candidates competing in a race—should not occur, which stands in contrast to the theory of Grossman and Helpman (1996). Using data from contributions to candidates for the U.S. House of Representatives Chamon and Kaplan show that split contributions are indeed rarely observed. A second feature of their theoretical results is that candidates who win with a higher vote margin should be receiving higher contributions. Again, this is confirmed by the data. The same pattern can be generated by the model presented here. I view their work as complementary to mine.

This chapter is organized as follows: In the following section I briefly discuss the related literature. Section 4.3 gives the details of the model. In section 4.4 the case with one interest group is analysed. The main results of the chapter are contained in section 4.5, which presents the characterisation of equilibrium in the presence of two interest groups. Section 4.6 summarises the findings.

## 4.2   Related Literature

The idea of giving interest groups the ability to commit to contribution schedules has featured in the literature on campaign contributions ever since the introduction of the concept of a "menu auction" by Bernheim and Whinston (1986), where a number of bidders submit schedules to a seller. These schedules specify the transfers the participants in the auction will make to the seller depending on the allocation that the seller decides to implement. The main application of this theory that Bernheim and Whinston had in mind was influence seeking.

Grossman and Helpman (1994) develop such an application in the context of the design of trade policy by a single policy maker who maximizes a weighted sum of voter welfare and campaign contributions. Their model has since been widely used in both the theoretical and empirical literature on trade barriers. To justify their choice for the objective function of the policy maker in the previous paper, Grossman and Helpman (1996) analyse an election in which two candidates compete by announcing campaign platforms and spending money on advertising. Contributions to one candidate are nevertheless a function of this candidates platform only. While the authors give arguments supporting why interest groups should have the ability to make binding commitments, they do not justify the restrictions they impose on these "contracts". According to their results, interest groups make contributions to both candidates in order to influence the policies that each one of them proposes and sometimes give more to one candidate in order increase that candidate's electoral chances in particular. I show that more general contracts allow interest groups to influences campaign platforms through the mere threat of contributions.

The insight that externalities among agents (candidates in this context) can enable a principal to extract a large share of the surplus has been present in the literature on contract theory and mechanism design (Aghion and Bolton, 1987, Crémer and McLean, 1985, Jehiel et al., 1996, Segal, 1999, Spiegler, 2000). Most of these papers feature only one principal. As fare as I am aware, Spiegler (2000) is the only paper in this literature that simultaneously features both multiple agents and multiple principals.

Within the context of political economy, the literature on vote buying (Dal Bó, 2007, Helpman and Persson, 2001, Morgan and Várdy, 2011) has identified the possibility of influencing policies without having to actually carry out any transfers. This can be achieved by promising committee members bribes in case their vote should be pivotal. As voters do not care about their vote as long as they are not pivotal, voting in favour of the interest group becomes weakly dominant for large enough promised bribes. The consequence is that all votes are cast in favour of the interest group. Consequently no

single voter is pivotal and no transfers have to be carried out. While certainly related, these papers are quite different both in their context and the structure of the models they develop. In particular, legislators are not subject to threats in these papers.

A final related paper is Polborn (2006), which analyses a game between two players who are both trying to capture the status quo in a dynamic setting. It is shown that the expenditures made in order to change the status quo are low relative to the value of achieving this in equilibrium. This result can be applied to a lobbying situation. Due to the more abstract setting that it is derived in, the paper does not yield empirical predictions about the donations made by interest groups beyond their relatively low level.

## 4.3 The Model

The model features two types of actors: Politicians and interest groups. Both will be introduced in turn below.

### 4.3.1 Politicians

A politician, called the incumbent, chooses a policy $p$ from the set $\mathcal{P} \equiv [-1, 1]$. She knows that this policy choice will have an impact on her probability of getting re-elected. She can also influence this probability by spending an amount of money $a_I \in \mathbb{R}_+$ on political advertising. Similarly, her challenger at the upcoming election is going to spend an amount $a_C \in \mathbb{R}_+$ on campaign activities. The probability that the incumbent wins the election is given by the continuous and differentiable function $\varphi : \mathcal{P} \times \mathbb{R}_+^2 \to [0, 1]$ that maps the policy choice and amounts spent on advertising into probabilities. The incumbent cares only about winning the election and simply maximizes the probability of doing so.

I assume that the function $\varphi$ is increasing in policy on $[-1, 0)$ and decreasing in policy on $(0, 1]$ in any point where no candidate wins with certainty. The policy zero can thus be thought of as the policy most preferred by voters or at least the median voter. Implementing the policy zero is not enough for the incumbent to win the election with certainty, but at least guarantees a positive chance in the absence of campaign expenditures. That is, $0 < \varphi(0, 0, 0) < 1$. Furthermore, I require that $\varphi$ is increasing in the campaign expenditure $a_I$ of the incumbent and decreasing in the campaign expenditure $a_C$ of the challenger, again in any point where no candidate wins with certainty.

Beyond these basic assumptions I need one restriction on the relative productivity of money spent by either candidate and on how the policy choice of the incumbent affects this productivity. Namely, I assume that campaign advertising is at least as effective for the incumbent as it is for the challenger: $\varphi(p, a, a) \geq \varphi(p, 0, 0) \ \forall p \in \mathcal{P}$. This assumption can be relaxed. In particular, one may want to allow for the possibility that the relative productivity of money spent by the incumbent decreases as policy becomes more extreme. This is possible as long as this decrease does not occur too quickly.

In general, the function $\varphi$ can be thought of as representing some model of probabilistic voting where voters use the policies implemented by the incumbent to update their expectations about the utility they would receive in case the incumbent got re-elected. The influence of political advertising on voting behaviour is often interpreted as a purely psychological effect in the literature. It is also possible (if less easily so) to think of advertising as conveying actual information.

### 4.3.2 Interest Groups

Candidates do not have any funds of their own to spend on advertising. Instead they have to rely on interest groups for campaign contributions. Interest group $i \in \{L, R\}$ chooses to make donations $a_I^i$ and $a_C^i$ and has a utility function given by

$$U_i(p, a_I^i, a_C^i) = u_i(p) - a_A^i - a_B^i$$

where the function $u_i : \mathbb{R} \to \mathbb{R}$ is strictly decreasing (increasing) on $\mathcal{P}$ for $i = L$ ($i = R$).

Interest groups make contributions according to contribution schedules communicated to candidates at the beginning of the game. I allow interest groups to condition their contributions on the policy choice of the incumbent and the contribution received by each candidate from the other interest group. Allowing for such general contracts can lead to problems of infinite regress. Suppose, for example, that conditional on a certain policy the schedules of the two interest groups take the following form: Interest group $L$ commits to a contribution of $x$ to the incumbent if the challenger receives no contributions and does not make any donations otherwise. Interest group $R$, on the other hand, contributes the same amount that the incumbent has received to the challenger. The final payments made under these contracts are then indeterminate. To prevent cases like this, I impose that the contributions made to any candidate by interest group $i$ have to be a weakly increasing function of the payments received by candidates from other interest groups. This assumption is admittedly ad hoc, but required in order to ensure that the game is well-defined. I would like to stress that all of the equilibria of

the model are robust to deviations to contribution schedules that are not weakly increasing in donations by other interest groups, as long as it is possible to determine the contributions that result from this deviation. An additional restriction on contribution schedules is that no interest group can credibly promise donations greater than a group-specific amount $A^i$, which can be thought of as the total budget the group has available. To simplify notation I let $A^i$ be the upper bound on promises made to each candidate separately instead of the upper bound on the sum of all promises.

I can now define the action spaces of interest groups formally. Let $a_c^{-i}$ be the contribution received by candidate $c$ from the interest group other than group $i$ and let $\tilde{S}^i$ be the set of all maps $s_i : \mathcal{P} \times \mathbb{R}_+^2 \to [0, A^i]^2$, with $s_{i,c}$ giving the contributions to candidate $c$ specified by the map $s_i$. The action space of interest group $i$ is then given by the set

$$\mathcal{S}^i \equiv \{s_i \in \tilde{\mathcal{S}}^i : \ s_{i,c}(p, a_I^{-i}, a_C^{-i}) \text{ weakly}$$
$$\text{increasing in } a_I^{-i}, \ a_C^{-i} \text{ for } c \in \{I, C\}\} \ .$$

The restrictions on schedules ensure that final transfers are always well defined: For a given policy choice $p$ any vector of contracts defines an increasing self-map on the space $\times_{i \in \Theta}[0, D^i]^2$ equipped with the product order. Tarski's theorem therefore guarantees the existence of at least one fixed point. If more than one fixed point exists I pick the one for which all contributions are lowest (the infimum of the set of fixed points under the product order). This assumption is not restrictive: All equilibria derived in this chapter rely on contribution schedules that have a unique fixed point for any given policy. For any vector of contribution schedules $\mathbf{s}$ let $a_{i,c}(p|\mathbf{s})$ be the contribution by interest group $i$ to candidate $c$ corresponding to this lowest fixed point. Also, let

$$a_i(p|\mathbf{s}) = a_{i,I}(p|\mathbf{s}) + a_{i,C}(p|\mathbf{s}) \ .$$

As advertising expenditures increase the probability of winning, candidates will always spend all of the donations they receive. I therefore denote by $a_c$ the sum of contributions to candidate $c$ as well as candidate $c$'s expenditure.

### 4.3.3 Timing and Equilibrium

The timing of the game is simple: In the first stage all interest groups commit to a contribution schedule. Subsequently, the incumbent chooses a policy. Contributions are then made according to the previously announced schedules. Finally, the winner of the election is determined according to the function $\varphi$. I look for subgame perfect equilibria of this game, given by a vector of contribution schedules $\mathbf{s}^*$ and a function

$P^*$ that returns the policy choice of the incumbent for any possible pair of contribution schedules. I focus on pure strategy equilibria, as is standard in this literature.

The set of such equilibria is large as soon as more than one interest group is present. This stems from the fact that interest groups are almost unconstrained in their commitments to contributions at policies that occur out of equilibrium. Threatening to make sufficiently high donations can ensure that these donations never actually have to be carried out. It seems desirable to introduce at least some chance that interest groups will be held to their word. This is possible by requiring that equilibria are robust to an arbitrarily small chance that the incumbent turns out to be an irrational type, who announces a random policy without caring about her chance of being re-elected. This by itself, however, does not impose the desired degree of discipline on contribution schedules. Any particular out-of-equilibrium policy still occurs with zero probability and interest groups thus remain free to make "unreasonable" promises on a small set of platforms that have a huge impact on the incentives that the incumbent faces. This feature, in turn, rests entirely on the infinity of the policy space. I therefore introduce a perturbed version $G_\varepsilon^{\mathcal{P}_\Delta}$ of the contribution game $G$, which differs from the original game in two ways: First of all, the policy space is replaced by some finite subset $\mathcal{P}_\Delta$ of $\mathcal{P}$ and all functions (and function spaces) are appropriately restricted to $\mathcal{P}_\Delta$. Secondly, the incumbent is either a rational player with probability $1 - \varepsilon$ or irrational with probability $\varepsilon$. An irrational incumbent does not care about her chances of winning the election and chooses a policy that—from the perspective of interest groups—is equally likely to be any point of the policy space $\mathcal{P}_\Delta$. I then consider only equilibria of the original game that are robust in the following sense:

**Definition 4.1** (Robust contribution equilibrium)**.** Consider an equilibrium $E = (\mathbf{s}^*, P^*)$ of the contribution game $G$. Let $\mathcal{P}_E$ be a finite subset of the policy space $\mathcal{P}$ that contains the points $P^*(\mathbf{s}^*)$ and zero. Let $\tilde{P}^*$ be such that $\tilde{P}^*(\mathbf{s}) \in \arg\max_{p \in \mathcal{P}_E} \varphi(p|\mathbf{s})$ with $\tilde{P}^*(\mathbf{s}^*) = P^*(\mathbf{s}^*)$. Denote by $\mathbf{s}^*|_{\mathcal{P}_E}$ the restriction of the equilibrium contribution schedules to the set $\mathcal{P}_E$.

Then $E$ is said to be a robust contribution equilibrium if, for any $\mathcal{P}_E$ and some corresponding $\tilde{P}^*$, there exists a positive probability $\bar{\varepsilon}$ such that for any $\varepsilon < \bar{\varepsilon}$ the strategy profile $(\mathbf{s}^*|_{\mathcal{P}_E}, \tilde{P}^*)$ is an equilibrium of the perturbed game $G_\varepsilon^{\mathcal{P}_E}$.

The inclusion of the equilibrium platforms in the discretised policy space greatly simplifies the definition and the application of the concept but is otherwise not essential.

Before proceeding to the description of the results, I need to introduce one more bit of notation. I define $\varphi(p|\mathbf{s})$ as the election probability of the incumbent under the policy choice $p$ and given the contributions made at $p$ under the vector of schedules $\mathbf{s}$.

## 4.4 One Interest Group

I will now describe the solution to the model in the case where interest group $R$ is the only active lobby. This serves mainly as an introduction to the logic underlying the more general results in the following section. The equilibrium presented here is a special case of proposition 4.3 in the next section with $A^L = 0$. Therefore, no proofs will be given here.

In the absence of any contributions, the incumbent would maximise his probability of getting re-elected by choosing the policy zero. Interest group $R$ would like to shift the chosen policy to the right. One way of achieving this would be to promise contributions to the campaign of the incumbent that can then be spent on political advertising. If the amount given is sufficiently high this could compensate for the votes lost due to the less voter-friendly policy. That is, the promised amount $a$ would have to satisfy $\varphi(0, 0, 0) \leq \varphi(p, a, 0)$ in order to make the incumbent implement the policy $p$. However, it would also be possible for interest group $R$ to threaten the incumbent to give the same amount $a$ to the challenger if the incumbent chooses any policy below some policy $p'$. The incumbent then chooses the policy $p'$ as long as $\varphi(0, 0, a) \leq \varphi(p', 0, 0)$, but the lobby does not have to make any actual contributions. Whether the policy $p'$ is greater or smaller than the policy $p$ depends on the shape of the function $\varphi$. For example, if campaign money is much more effective in the hands of the incumbent than when spent by the challenger, it will be true that $p > p'$.

It is not the case though that the interest group has to decide exclusively between making promises or threats. It may give money to the challenger at a certain policy, but threaten to withdraw this money and give it to the challenger at any policy closer to zero. But while the use of promises depends on their effectiveness, threats give influence at no cost and will therefore always be employed. Accordingly, the equilibrium policy and contribution is given by the solution to the maximisation problem

$$\max_{p,a} \ u_R(p) - a$$
$$\text{s.t.} \ \ \varphi(p, a, 0) \geq \varphi(0, 0, A^R) \ ,$$
$$0 \leq a \leq A^R \ .$$

The first constraint is a participation constraint that ensures that the incumbent is willing to locate at the targeted policy. The second constraint ensures that the lobby does not exceed its budget. The first order conditions for this problem can be rewritten to yield

$$u'_R(p) = -\frac{\partial \varphi / \partial p}{\partial \varphi / \partial a} \ .$$

The right-hand side of this condition is the increase in campaign contributions required to satisfy the participation constraint due to an increase in the targeted policy, as can be seen from the implicit function theorem. The condition therefore simply says that the marginal utility of policy must be equal to the marginal cost of achieving this policy at the optimum.

The interest group may rely entirely on threats, or may fully exploit its ability of making both threats and actual donations, or the equilibrium may lie anywhere in between these two extremes. If an observer was to attribute the shift in the policy of the incumbent away from zero solely to contributions received by the incumbent, this would potentially give the impression of very high rates of return: There is no upper bound on the ratio between the utility gain of the interest group relative to the policy zero and equilibrium donations, as the latter may be arbitrarily small. The model with one interest group thus provides an explanation for the Tullock paradox. As will be shown in the next section, this explanation is robust to entry of an additional interest group.

## 4.5   Two Interest Groups

I begin by deriving conditions that contribution schedules need to satisfy in an equilibrium that is robust in the sense of definition 4.1. In essence, these say that interest groups only commit to contributions where these are required to support the equilibrium policy choice of the incumbent. No interest group will promise contributions at policies that it prefers over the equilibrium policy, in particular. These would take the form of donations intended to make the incumbent deviate. If this fails there is no need to maintain these promises.

**Lemma 4.1.** *Consider an equilibrium* $(\mathbf{s}^*, P^*)$ *of the contribution game and let* $p^* = P^*(\mathbf{s}^*)$. *This equilibrium is robust only if*

i) $a_R(p|\mathbf{s}^*) = 0$ *for any* $p > p^*$,

ii) $a_L(p|\mathbf{s}^*) = 0$ *for any* $p < p^*$,

iii) $\varphi(p, 0, 0) \leq \varphi(p^*|\mathbf{s}^*)$ *implies* $a_i(p|\mathbf{s}^*) = 0$ *for* $i \in \{L, R\}$.

iv) $\varphi(p, 0, 0) > \varphi(p^*|\mathbf{s}^*)$ *implies* $\varphi(p|\mathbf{s}^*) = \varphi(p^*|\mathbf{s}^*)$.

*Proof.* To show part *i)*, consider any $p > P^*(\mathbf{s}^*)$ and assume $a_R(p|\mathbf{s}^*) > 0$. Suppose lobby $R$ reduces all of its contributions at $p$ to zero. If the payoff of the incumbent at $p$ is now higher than her equilibrium payoff she would change her policy to $p$, making

group $R$ better off. If, on the other hand, the payoff of the incumbent at $p$ remains at or below her equilibrium payoff then the equilibrium is not robust. To see this, note that there exists some finite subset $\mathcal{P}_p$ of the policy space that contains the policy $p$ besides the policies $P^*(\mathbf{s}^*)$ and 0. For any $\varepsilon > 0$ the policy $p$ is chosen by the irrational type of the incumbent with positive probability in the perturbed game $G_\varepsilon^{\mathcal{P}_p}$. As lobby $R$ does not change the behaviour of the rational type of the incumbent by reducing its contributions at $p$ to zero, but lowers its expected donations, it would prefer to do so. The necessity of the second part of the statement can be shown analogously.

Thus, for any policy other than $p^*$ at most one interest group makes a contribution. Suppose $\varphi(p, 0, 0) < \varphi(p^*|\mathbf{s}^*)$ for some $p$ and some interest group $i$ makes a contribution at $p$. As in the previous paragraph, there exists some perturbed version of the game where the policy $p$ is chosen with positive probability by the irrational type of the incumbent. Group $i$ would therefore like to lower its contributions at $p$ as long as this does not change the behaviour of the rational type of the incumbent. As $i$ is the only lobby making a contribution at $p$ and donations are increasing in the contributions of other groups, the donations of the second lobby have to remain at zero if group $i$ reduces its contributions at $p$. The condition $\varphi(p, 0, 0) < \varphi(p^*|\mathbf{s}^*)$ is therefore sufficient to guarantees that $i$ can lower its contributions at $p$ to zero without inducing a deviation by the incumbent. This implies that part *iii*) is required for robustness.

Finally, the condition $\varphi(p, 0, 0) > \varphi(p^*|\mathbf{s}^*)$ entails that the incumbent must receive a donation at any such $p$, otherwise equilibrium would be violated. However, if it was the case that $\varphi(p|\mathbf{s}^*) < \varphi(p^*|\mathbf{s}^*)$, the continuity of the function $\varphi$ enables the interest group making donations at $p$ to lower these without affecting the behaviour of the incumbent. As in the previous paragraph, there exists a perturbed version of the game where the interest group also has the incentive to do so. This completes the proof. □

I now introduce the concept of a net contribution: A candidate is said to receive a net contribution if her payoff is higher than it would be if neither candidate received any contributions for a given policy choice of the incumbent. Formally, the incumbent is in receipt of a net contribution at $p$ under a vector of schedules $\mathbf{s}$ if $\varphi(p|\mathbf{s}) > \varphi(p, 0, 0)$. Equivalently, the challenger receives a net contribution when $\varphi(p|\mathbf{s}) < \varphi(p, 0, 0)$.

**Lemma 4.2.** *In any robust contribution equilibrium and for any policy $p$, $a_L(p|\mathbf{s}^*) > 0$ only if the challenger receives a net contribution at the equilibrium policy.*

*Proof.* Let $p^*$ be the policy choice and $\varphi^*$ the payoff of the incumbent in equilibrium. Suppose the challenger does not receive a net contribution at $p^*$, that is $\varphi^* \geq \varphi(p^*, 0, 0)$. For any $p > p^*$ this implies $\varphi^* > \varphi(p, 0, 0)$. By lemma 4.1 it is therefore the case that

neither group promises any contributions at any $p > p^*$ and consequently the payoff of the incumbent at any such $p$ is strictly lower than the payoff $\varphi(p^*, 0, 0)$.

Now consider any policy $0 \leq p < p^*$. Due to lemma 4.1 it must be true that

$$\varphi(p|\mathbf{s}^*) = min\{\varphi^*, \varphi(p, 0, 0)\} \ .$$

It is therefore true that the payoff of the incumbent at any such $p$ is strictly greater than $\varphi(p^*, 0, 0)$.

The last two paragraphs together show that the incumbent would deviate to a policy smaller than $p^*$ if her payoff at $p^*$ was lowered sufficiently. Consequently, interest group $L$ must be making no contributions at $p^*$. Otherwise it could reduce these to zero, with one of two possible consequences: The equilibrium policy remains unchanged but group $L$ saves on contributions, or the incumbent deviates to a preferable policy outcome. In the latter case, lemma 4.1 implies that lobby $L$ makes no contributions at the new policy choice of the incumbent. This shows that dropping all contributions at $p^*$ must be a profitable deviation for interest group $L$. $\qquad \square$

The previous lemma says that any equilibrium where the challenger does not receive a net contribution must have the following feature: The weaker interest group does not promise any donations at any policy unless the stronger interest group deviates and donates more money than promised in equilibrium. If the election probability of the incumbent is not weighed down by donations to the challenger, then there is no risk that she will deviate to policies that are even further away from zero. These policies are bad for the prospects of the incumbent unless they bring contributions from interest group $R$. As was shown before, lobby $R$ will not make such promises. Group $L$ can therefore remain almost completely passive.

The following proposition is the first main result of this section.

**Proposition 4.1.** *In any robust contribution equilibrium, only interest group $R$ may give money and only to the incumbent.*

*Proof.* First, suppose that both interest groups make a contribution at the equilibrium policy $p^*$. By lemma 4.2 this requires that the challenger must be receiving a net contribution at $p^*$. This implies that the challenger must also be receiving a net contribution on some interval $(p^*, p]$ as the incumbent would deviate to one of these policies otherwise. By lemma 4.1 lobby $L$ must be making these contributions. But then group $L$ could reduce the donation to the challenger to zero at some policy $p^* + \varepsilon > p^*$, inducing the incumbent to deviate to this policy. For $\varepsilon$ small enough this move must be profitable,

as it implies a fixed reduction in contributions from $a_L(p^*|\mathbf{s}^*)$ to zero but a negligible loss in utility from policy. It is therefore impossible that both interest groups make a contribution in equilibrium.

Now suppose an interest group gives money to the challenger in equilibrium. It must then be the only group making positive donations. As contributions are increasing functions of contributions by other groups, lowering the donation to the challenger cannot change the amount of money donated by the second interest group. This move must consequently increase the payoff of the incumbent and therefore leaves her policy choice unchanged. This shows that no lobby gives money to the challenger.

To complete the proof, assume that group $L$ gives money to the incumbent. As the challenger receives no contributions, this means that the incumbent would be receiving a net contribution. By lemma 4.2 this contradicts that lobby $L$ would be making any donations. $\qquad\square$

Donations to the challenger have the sole purpose of preventing the incumbent from choosing a particular policy. At the equilibrium policy such donations obviously fail to achieve their purpose and might as well be withdrawn. Giving money to both candidates simultaneously is equally futile. Nevertheless, the complicated nature of contribution schedules makes it less than obvious that such things never occur. The preceding proposition shows that a small amount of uncertainty about the policy choice of the incumbent restricts contribution schedules enough for interest groups to avoid inefficient contributions.

There are two policies that provide bounds on the possible equilibrium choices of the incumbent. The first one is the policy furthest to the right of zero that interest group $R$ can achieve purely by making threats while group $L$ counters these threats by the highest possible promise of donations. I denote this policy by $\hat{p}$ and it is formally defined as the policy $p > 0$ that satisfies $\varphi(p, 0, 0) = \varphi(0, A^L, A^R)$. As will be shown below, $\hat{p}$ provides a lower bound on policy outcomes. An upper bound is given by $\check{p}$, defined as the policy $p > 0$ such that $\varphi(p, A^R, A^L) = \varphi(0, 0, A^R)$. For any policy to the right of $\check{p}$ interest group $L$ can make the incumbent deviate to zero by giving $A^L$ to the challenger, even if lobby $R$ fully exploits its ability of using actual contributions and threats in order to influence the policy choice of the incumbent.

**Proposition 4.2.** *The equilibrium policy must be an element of the interval $[\hat{p}, \bar{p}]$ in any robust contribution equilibrium.*

*Proof.* Suppose there was an equilibrium where the incumbent chooses a policy $p^* < \hat{p}$. According to proposition 4.1 the challenger does not receive any contributions in

equilibrium. Lemma 4.2 thus implies that $a_L(p|\mathbf{s}^*) = 0$ for any policy $p$ and in particular that $a_L(p, 0, 0) = 0$ for any $p > p^*$. Now let interest group $R$ deviate to the contribution schedule $s'$ defined as follows: For some policy $p'$ such that $p^* < p' < \hat{p}$ no candidate receives any donations from group $R$. For any policy other than $p'$ group $R$ gives $A^R$ to the challenger. Under this schedule and the equilibrium schedule of lobby $L$ the payoff of the incumbent from locating at $p'$ is $\varphi(p', 0, 0)$. For the payoff from any other policy $p$ it holds that

$$
\begin{aligned}
\varphi(p|s', s_L^*) &\leq \varphi(p, A^L, A^R) \\
&\leq \varphi(0, A^L, A^R) \\
&= \varphi(\hat{p}, 0, 0) \\
&< \varphi(p', 0, 0) \, ,
\end{aligned}
$$

where the first line holds because the incumbent can at most receive $A^L$ from interest group $L$ at $p$, while the remaining lines use the assumption that the payoff of the incumbent is increasing towards zero and/or the definitions of $p'$ and $\hat{p}$. This shows that interest group $R$ can induce the incumbent to choose the policy $p'$ without actually carrying out any donations, which must increase the utility of group $R$.

Now suppose there was an equilibrium where the incumbent chooses a policy $p^* > \check{p}$. As above, $a_L(p|\mathbf{s}^*) = 0$ for any policy $p$. If lobby $L$ can lower the payoff of the incumbent at $p^*$ sufficiently, the proof of lemma 4.2 together with the implication of proposition 4.1 that the challenger never receives a net contribution show that this must lead to a deviation of the incumbent to a policy smaller than $p^*$. To induce such a deviation group $L$ can deviate to a schedule $s'$ where $L$ gives $A^L$ to the challenger at $p^*$ and otherwise commits to the same contributions as under the schedule $s_L^*$. In this case

$$
\begin{aligned}
\varphi(p^*|s_R^*, s') &\leq \varphi(p^*, A^R, A^L) \\
&< \varphi(\check{p}, A^R, A^L) \\
&= \varphi(0, 0, A^R) \\
&\leq \varphi(0|\mathbf{s}^*) \, ,
\end{aligned}
$$

where the first line holds because the incumbent can at most receive $A^R$ from interest group $R$ at $p^*$, the second line holds as the payoff of the incumbent is increasing towards zero, the third line uses the definition of $\check{p}$, while the final line uses the fact that $\varphi(0, 0, A^R)$ is the worst possible payoff for the incumbent at zero when group $L$ makes no contributions. Lobby $L$ can therefore achieve a policy smaller than $p^*$ without carrying out any contributions, which must me profitable. $\qquad\square$

So far it has been shown that the equilibrium policy must fall within a certain range that favours lobby $R$ and if any contributions occur in equilibrium they will be given to the incumbent by interest group $R$. Interest group $L$, on the other hand, remains almost completely passive; it does not even offer contributions at out-of-equilibrium policies. Lobby $R$ nevertheless has less influence than if group $L$ was not present. This is because group $L$ has the ability to react when group $R$ increases its contributions in order to gain more influence. For example, consider the situation where $L$ does not give any contributions to the incumbent at policies below the one chosen in equilibrium, while group $R$ threatens to give some money to the challenger at said policies. $R$ can then be unable to intensify these threats beyond the equilibrium level, because any additional money given would trigger contributions to the challenger from lobby $L$.

All of the above statements are conditional on equilibrium existence. I now conclude this section by constructing an equilibrium that always exists. To do so, define the policy $\bar{p}$ as the policy $p > 0$ such that $\varphi(p, A^R, A^L) = \varphi(0, A^L, A^R)$. Next, let the policy $p^E$ be defined as the smallest policy $p$ that is part of a solution to the maximisation problem

$$\max_{p,a} \ u_R(p) - a$$

$$\text{s.t.} \ \ \varphi(p, a, 0) = \varphi(0, A^L, A^R) \ ,$$

$$\hat{p} \le p \le \bar{p} \ ,$$

$$0 \le a \le A^R \ .$$

Accordingly, define the contribution schedule $s_R^E$ such that at $p^E$ the incumbent receives the donation $a$ that solves $\varphi(p^E, a, 0) = \varphi(0, A^L, A^R)$, while at any other policy $p$ the challenger receives the smallest possible contribution $a$ that ensures that $\varphi(p, 0, a) \le \varphi(0, A^L, A^R)$. Furthermore, group $R$ commits to increasing its donation to the incumbent at $p^E$ to $A^R$ if lobby $L$ should make any contributions. Similarly, it commits to contributing $A^R$ to the campaign of the challenger at any $p \ne p^E$ if lobby $L$ should make any contributions.

The contribution schedule $s_L^E$ can then be defined by

$$s_{L,C}^E(p, a_I^R, a_C^R) = \begin{cases} A^L & \text{if } p^E \ne p > \hat{p} \text{ and } a_I^R > 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$s_{L,I}^E(p, a_I^R, a_C^R) = \begin{cases} A^L & \text{if } p = 0 \text{ and } a_C^R > s_{R,C}^E(0,0,0) \\ 0 & \text{otherwise} \ . \end{cases}$$

**Proposition 4.3.** *The strategy profile* $(\mathbf{s}^E = (s_L^E, s_R^E), P^E)$ *is an equilibrium for some* $P^E$ *such that* $P^E(\mathbf{s}^E) = p^E$.

*Proof.* The policy $p^E$ is an optimal choice for the incumbent under the schedules $s_L^E$ and $s_R^E$ by construction. It remains to be shown that no interest group wants to deviate.

To see that interest group $R$ cannot achieve a better outcome, first note that it is impossible for $R$ to lower the payoff of the incumbent at 0 below $\varphi(0|\mathbf{s}^E)$. As $\varphi(0,0,0) > \varphi(0, A^L, A^R)$, the definition of the schedule $s_R^E$ implies that the challenger receives a contribution from $R$ at zero such that $\varphi(0|\mathbf{s}^E) = \varphi(0, A^L, A^R)$. Lowering the payoff of the incumbent at zero would require increasing the donation to the challenger, but this would immediately cause lobby $L$ to give $A^L$ to the incumbent. The lowest payoff $R$ can thus achieve is $\varphi(0, A^L, A^R)$, which is the equilibrium level.

There are then three possible cases to consider: Interest group $R$ could try to move the policy of the incumbent to a point below or at $\hat{p}$, to a policy in the set $(\hat{p}, \bar{p}]$, or to an even greater policy. As the definition of $p^E$ implies that group $R$ is either already making the incumbent locate at $\hat{p}$ for free or prefers to pay to make her choose a greater policy, any policy smaller than or at $\hat{p}$ cannot be better for $R$ even if it can be achieved for free. Next, consider any policy in the set $(\hat{p}, \bar{p}]/p^E$. For any such policy $p$ it is true that $s_R^E(p, 0, 0) = 0$ by the definition of $s_R^E$ as

$$\varphi(p, 0, 0) < \varphi(\hat{p}, 0, 0) = \varphi(0, A^L, A^R) . \tag{4.1}$$

By the definition of $s_L^E$ this means that any contribution at $p$ to the incumbent by group $R$ will cause lobby $L$ to give a donation of $A^L$ to the challenger. However, in order to make the incumbent choose the policy $p$ she would have to receive a contribution, as can be seen from condition (4.1). Suppose then that interest group $R$ could make a contribution $a'$ such that $\varphi(p, a', A^L) > \varphi(0, A^L, A^R)$, as would be required to make the incumbent locate at $p$. This donation must clearly be greater than the donation $a''$ necessary to achieve $\varphi(p, a'', 0) = \varphi(0, A^L, A^R)$. It therefore holds that

$$u_R(p) - a' < u_R(p) - a''$$
$$\leq u_R(p^E) - s_{R,I}^E(p^E, 0, 0) ,$$

where the second line holds due to the definition of $p^E$. This shows that such a deviation cannot be profitable.

For the third case, it is clear that inducing the incumbent to locate at some $p > \bar{p}$ is impossible. As in the previous paragraph, this would require a donation to the incumbent, which would provoke a reaction from group $L$. Accordingly, the highest

payoff that would be possible for the incumbent is $\varphi(p, A^R, A^L)$ for which it holds that $\varphi(p, A^R, A^L) < \varphi(\bar{p}, A^R, A^L) = \varphi(0, A^L, A^R)$.

Finally, it needs to be shown the interest group $L$ cannot improve on the equilibrium outcome. To do so, it would have to change the policy choice of the incumbent, which in turn would require $L$ to raise the payoff of the incumbent at some $p \neq p^E$ above $\varphi(0, A^L, A^R)$ or to lower the payoff at $p^E$. The former could only be achieved through giving money to the incumbent, which would be countered by group $R$ with a donation of $A^R$ to the challenger. If follows that the highest possible payoff lobby $L$ could achieve through a donation to the incumbent at $p$ is $\varphi(p, A^L, A^R) < \varphi(0, A^L, A^R)$. Lowering the payoff of the incumbent at $p^E$ is equally impossible as group $R$ reacts to any donations to the challenger and thus $\varphi(p^E, A^R, A^L) \geq \varphi(\bar{p}, A^R, A^L) = \varphi(0, A^L, A^R)$. $\qquad\square$

As the final step, I verify that the equilibrium established above is also a robust contribution equilibrium.

**Proposition 4.4.** *If the strategy profile $(\mathbf{s}^E, P^E)$ is an equilibrium, then it is a robust contribution equilibrium.*

*Proof.* Consider some finite subset $\mathcal{P}_E$ of the policy space that contains the policies zero and $p^E$ and a strategy profile $(\mathbf{s}^E|_{\mathcal{P}_E}, \tilde{P}^E)$, as required in the definition of a robust contribution equilibrium. The arguments of the proof of proposition 4.3 can also be applied to the perturbed game to show that the incumbent must locate at $p^E$ under any schedule that interest group $L$ can propose. $L$ consequently has no profitable deviations. Group $R$, on the other hand, cannot reduce any contributions without causing the incumbent to deviate by the definition of the schedule $s_R^E$ and because the policy zero is included in $\mathcal{P}_E$. It therefore remains to check that $R$ would not want to deviate to some schedule that induces a different policy choice.

The proof of proposition 4.3 implies that

$$u_R(p') - a' < u_R(p^E) - s^E_{R,I}(p^E, 0, 0) \tag{4.2}$$

for any policy $p' > \hat{p}$ that interest group $R$ may be able to achieve and the contribution $a'$ that would be required to do so. Now, $p^E$ has been defined such that either $p^E = \hat{p}$, in which case $s^E_{R,I}(p^E, 0, 0) = 0$ as this policy can be achieved for free, or it must be the case that

$$u_R(\hat{p}) < u_R(p^E) - s^E_{R,I}(p^E, 0, 0) \ .$$

As $\hat{p}$ is also the largest policy that can be achieved for free, this shows that condition (4.2) actually applies to any policy that interest group $R$ can induce against the schedule $s_L^E$.

For any schedule $s$ of interest group $R$ that induces a deviation of the rational incumbent to some policy $p'$, the utility of interest group $R$ in the perturbed game $G_\varepsilon^{P_E}$ can be written as

$$\varepsilon \, \frac{1}{|P_E|} \sum_{p \in P_E} \left[ u_R(p) - a_R(p|s, s_L^E) \right] + (1 - \varepsilon) \left[ u_R(p') - a' \right] \ ,$$

using the same notation as in the previous paragraph. This utility can be no greater than

$$\varepsilon \, \frac{1}{|P_E|} \sum_{p \in P_E} u_R(p) + (1 - \varepsilon) \left[ u_R(p') - a' \right] \ .$$

The difference between the equilibrium utility and this last expression is

$$\varepsilon \, \frac{1}{|P_E|} \sum_{p \in P_E} \left[ -s_R^E(p, 0, 0) \right] + (1 - \varepsilon) \left[ (u_R(p^E) - s_{R,I}^E(p^E, 0, 0)) - (u_R(p') - a') \right] \ .$$

This difference converges to

$$(u_R(p^E) - s_{R,I}^E(p^E, 0, 0)) - (u_R(p') - a')$$

as $\varepsilon$ approaches zero, which is positive by condition (4.2). This shows that there is no profitable deviation for the incumbent from the schedule $s_R^E|_{\mathcal{P}_E}$ for $\varepsilon$ sufficiently small, establishing the robustness of the equilibrium $(\mathbf{s}^E, P^E)$. $\qquad\square$

## 4.6 Conclusion

In this chapter I have shown how interest groups can use threats of contributions to achieve policy favours from politicians. As interest groups may combine threats with actual contributions, this can generate the appearance of large favours for small sums of money. Importantly, these results apply in a setting with competition among interest groups with opposing aims. The chapter therefore provides a robust explanation of the Tullock paradox.

An additional result is that the interest group with less spending power remains entirely passive even at unexpectedly chosen policies. The only case in which this interest group becomes active is if there are out-of-equilibrium donations by another interest group. Together with the possibility of achieving influence in the absence of any donations this helps to explain why a large number of firms seems to not engage in political

contributions at all, which remains true even among large companies (Ansolabehere et al., 2003, p. 108). The prediction that challengers do not receive money also sits well with data from the U.S. (Bombardini and Trebbi, 2011, Stratmann, 2005).

It would be interesting to see to what extent the results derived here carry over to a more general setting. In particular, it would be worthwhile to allow interest groups to have broader motives. If an interest group, for whatever reason, has a strong motive for seeing the challenger elected, this might destroy the result that the challenger never receives any donations. At least in the case with one interest group, however, it is easy to see that the result survives. This is because the presence of interest groups is actually not helping the incumbent. The participation constraint of the incumbent says that his probability of re-election must be at least as high as in the case where she is located at zero and the interest group gives all available money to the challenger. In equilibrium this constraint will be binding, meaning that the interest group lowers the re-election probability as much as if she was giving the whole budget to the challenger—even in a case where the incumbent receives a donation in equilibrium. An interest group that wants to increase the chances of the challenger is thus able to benefit simultaneously from a shift in the incumbent's policy as well as an increase in the election probability of the challenger. Showing whether it is possible to derive this result in the presence of two interest groups is beyond the scope of this chapter.

# Chapter 5

# Conclusion

This thesis has explored how informational limitations of voters can serve as an explanation for the presence of organisations such as political parties and interest groups in representative democracies. A second theme that emerged across all chapters is political competition. Chapter 2 demonstrated that the formation of parties, which reveal information about the ideology of their members, can result in systematic differences in the extent of competition observed in elections across regions. The crucial factor that determines which parties form was shown to be the career concerns of politicians. If these are sufficiently strong, politicians are unwilling to join parties targeted at particular regions as they fear loosing access to a career at the federal level.

Chapters 3 and 4 were concerned with the consequences of competition. Chapter 3 considered a setting where parties are better informed about the characteristics of potential candidates than voters are. The central question was whether parties can be expected to select the same candidates that voters themselves would choose if they were equally well informed. Given that parties have interests that differ from those of voters it was found that the answer is positive only if the degree of political competition is sufficiently strong. Competition thus has a positive effect as it disciplines parties to act on behalf of voters.

While parties were viewed as organisations that potentially help overcome the informational deficits of voters, a second way to do so is through political advertising. This opens the door to special interest groups who are willing to provide campaign funds in return for political favours. Chapter 4 entertained the possibility that interest groups achieve influence without making any actual donations. This is possible by exploiting competition among candidates and threatening one candidate with a contribution to a competitor. Competition among interest groups, on the other hand, can moderate the

influence that any individual group has but does not lead to a substantial increase in the amount of donations made.

As discussed in the individual chapters, some interesting directions for future research remain. It would be desirable to extend the results of chapter 2 to a setting of proportional representation, which is generally held to provide different incentives for party formation. While the results of chapter 4 are in general agreement with the stylized facts on campaign contributions, it is not clear to what extent these predictions survive if interest groups are given more general motives. Finally, the analysis has yielded some testable predictions. According to the model presented in chapter 2, politicians belonging to the Democratic Party should be more ideologically moderate (on average) in a state where the Democratic Party dominates elections compared to a state where is doesn't (and vice versa for the Republican Party). As this prediction does not necessarily carry over to elected politicians, verifying it would require measuring the political views of a broader set of politicians. One of the results of chapter 3 was that there should be a non-linear relationship between the degree of competition a party faces and the quality of its candidates. The challenge in taking this to the data would be to find a suitable proxy for quality.[1]

---

[1]Existing studies have used educational attainment, which is perhaps too easily observable for the current purpose. See for example Fisman et al. (2012).

# Appendix A

# Appendix to Chapter 2

## A.1  Additional Results for the Basic Model

This appendix provides a full characterization of equilibria with two parties beyond the $L$-$R$ equilibrium analysed in the main text and demonstrates the existence of an equilibrium with three parties.

**Proposition A.1.** *An equilibrium of the party-formation game where $\mathcal{P}^* = \{M, E\}$ with either $I_M = [-1..0]$ and $I_E = [1]$, or $I_M = [0..1]$ and $I_E = [-1]$ exists whenever*

$$\frac{1}{S}y_f + \left(1 - \frac{1}{S}\right)\frac{3}{2}y_P \geq 2y_s \ .$$

*Proof.* As the equilibria in the statement of the proposition are symmetric to each other, the proof will focus on the case $I_M = [-1..0]$ and $I_E = [1]$. In this case the affiliation behaviour of politicians is trivial. Party $M$ wins all states $s$ such that $m_s < p^+$ and wins the federal election with probability three-fourth. It will first be shown that entry of additional parties is impossible.

The equilibrium utility of extremist members of party $M$ is

$$\frac{1}{2}\left[y_s + \frac{1}{2w_M^*}y_f + \left(1 - \frac{1}{w_M^*}\right)\frac{3}{4}y_P\right] \ .$$

This expression is decreasing in $w_M^*$ under the assumption that $y_f > 2y_P$. If this politician joins an entering party of shape $[-1]$, she may win the state election, but can be made the certain loser of the federal election. As party $M$ cannot win more state election than there are states, it follows that a sufficient condition for this deviation not being profitable is

$$\frac{1}{S}y_f + \left(1 - \frac{1}{S}\right)\frac{3}{2}y_P \geq 2y_s \ .$$

Centrist politicians achieve a higher utility in equilibrium than extremist members of party $M$ as they win the federal election with higher probability. Due to the assumption that $\Lambda_f([-0.5, 0.5]) \leq 0.5$ and the presence of three candidates at the federal level, there exists a voting equilibrium such that centrist candidates of a third party do not win at the federal level. Their deviation payoff is accordingly also $y_s$. This shows that centrist members of party $M$ do not deviate to joining a third party whenever politicians with platform -1 refrain from doing so.

Finally, members of party $E$ do not gain from joining a party of shape [1] as they at best win all elections with the same probability as before.

It remains to be checked whether any active founders want to deviate. Any such deviation can be punished by entry of an additional party of the same shape that the party of the deviating founder had prior to the deviation. By virtue of the proceeding steps of the proof, there then exists an equilibrium of the election subgame reached where the party of the deviating founder does not gain any members. $\qquad\square$

The equilibrium in the preceding proposition will be referred to as the *M-E* equilibrium.

Any other constellation of two parties not considered so far is never part of an equilibrium.

**Proposition A.2.** *No constellation of two parties other than $\{[-1..0], [0..1]\}$, $\{[-1..0], [1]\}$, and $\{[-1], [0..1]\}$ is part of an equilibrium.*

*Proof.* It is straightforward to verify that under any constellation of two parties not listed in the statement of the proposition there must be some platform $p$ such that politicians with this platform cannot join any party. Suppose a party of shape $p$ enters. Due to the assumptions on voter distributions there exists at least one state where a strict majority of voters strictly prefers the platform $p$ over the expected platform of a candidate of any other party. It then follows from the restrictions on voting behaviour that the newly formed party wins at least one state election. $\qquad\square$

Finally, it can be shown that an equilibrium with three parties exists whenever no equilibrium with two parties does.

**Proposition A.3.** *An equilibrium such that $\mathcal{P}^* = \{A, B, C\}$ with $I_A = [-1]$, $I_B = [0]$, and $I_C = [1]$ exists whenever no equilibrium with two parties exists.*

*Proof.* Suppose the set of parties $\{A, B, C\}$ is formed in equilibrium with $I_A = [-1]$, $I_B = [0]$, and $I_C = [1]$. Then the affiliation behaviour of politicians is trivial. Assume

the candidates of parties $A$ and $C$ tie at the national level, party $B$ wins all centrist states and states $s$ such that $\Lambda_s([-0.5, 0.5]) > 0.5$, and party $A$ (party $C$) wins all state elections in states $s$ such that $m_s < p^-$ $(m_s > p^+)$ and $\Lambda_s([-0.5, 0.5]) \leq 0.5$.

Passive founders are unable to enter with a new party. There would be at least three parties competing at the federal level, implying that there exists a voting equilibrium such that the newly formed party loses. Politicians therefore have no incentive to join this new party, as they either already win the election in their state with certainty or would not win even after the deviation.

It remains to check whether any founder has an incentive to reposition their party. Any such deviation can be punished by entry of an additional party of the same shape that the party of the deviating founder had prior to the deviation. By virtue of the proceeding steps of the proof, there then exists an equilibrium of the election subgame reached where the party of the deviating founder does not gain any members. □

## A.2 Robustness to Party Mergers

This appendix derives some properties of equilibria of the extended party formation game that are robust to party mergers, as defined in section 2.5.2. The first result regarding the extended party-formation game says that transfers cannot be used to maintain formations of parties that are not stable in the original game.

**Lemma A.1.** *Any constellation of parties $\mathcal{P}^*$ that is part of a party-formation equilibrium of the extended game must have the feature that for any additional party $D$ there exists an equilibrium of the election subgame under the set of parties $\mathcal{P}^* \cup \{D\}$ such that party $D$ does not win any state elections.*

*Proof.* Suppose there was an equilibrium of the extended game such that the condition in the statement of the lemma was not satisfied. This implies that without transfers some passive founders would have an incentive to propose additional parties. That this constellation is nevertheless an equilibrium of the extended game would require that any passive founder receives a net transfer at least as large as the utility she could achieve by forming a party. This utility would be at least as great as $x_w - c > 0$. The total sum of transfers from party leaders to passive founders must therefore be infinite, while the total utility of all party leaders is no larger than $S x_w + x_f$ and therefore finite. Consequently, at least one party leader would be achieving negative utility and prefer to remain passive. □

The second result shows that the requirement of robustness to party mergers selects those party-formation equilibria with the lowest number of parties, but does not discriminate between equilibria within that class.

**Lemma A.2.** *A party-formation equilibrium is robust to party mergers if and only if there exists no other equilibrium in which a smaller number of parties is formed.*

*Proof.* Assume there is a party-formation equilibrium $E$ such that $N^* = k$ and there is a potentially empty set $\mathcal{F}_T$ of founders who receive transfers from other founders. Let $\mathcal{F}_E$ be the set of founders who propose parties in this equilibrium. All founders who make transfers must belong to the set $\mathcal{F}_E$. Any founder not belonging to this set could lower the transfers she makes to zero. This would not affect the transfers she receives as these can depend only on whether or not a founder proposes a party. The total utility of all members of the set $\mathcal{F}_E \cup \mathcal{F}_T$ is therefore equal to $Sx_w + x_f - kc$.

To show necessity, suppose that there exists a second party-formation equilibrium $E'$ in which a number of parties $k' < k$ is formed. Now consider the following equilibrium: In the first step, $k'$ members of the set $\mathcal{F}_E$ promise transfers such that any founder belonging to $\mathcal{F}_E \cup \mathcal{F}_T$ achieves a strictly greater utility than in the equilibrium $E$. This is possible because the total utility of all founders is higher in the equilibrium $E'$ than in the equilibrium $E$ by the amount $(k - k')c$. The same $k'$ founders subsequently propose the set of parties that exists along the equilibrium path of $E'$. If any deviation occurs at the first stage, the same set of parties as in the equilibrium $E$ is formed. This ensures that no founder can gain from such deviations. Deviations by passive founders to proposing a party are not profitable as the set of parties proposed in any subgame deters entry by lemma A.1. This shows that the equilibrium $E$ is not robust to party mergers.

For sufficiency, note that any equilibrium $E'$ in which the number of parties is equal to or greater than $k$ generates a total utility that is no greater than the total utility achieved in the equilibrium $E$. It follows immediately that it is impossible that the equilibrium $E'$ Pareto dominates the equilibrium $E$. □

It is then possible to fully characterize the number of parties formed in party-formation equilibria that are robust to party mergers.

**Proposition A.4.** *The number of parties in any party-formation equilibrium that is robust to party mergers is*

*i) no lower than two and no greater than three,*

*ii) equal to two whenever a party-formation equilibrium exists in which two parties are formed.*

*Proof.* Whenever only a single party exists, passive founders must have an incentive to form additional parties by proposition 2.3. Lemma A.1 then implies that there must

be at least two parties in any party-formation equilibrium that is robust to party mergers. Combined with lemma A.2 this establishes claim *ii*). In order to show that there cannot be more than three parties it needs to be demonstrated that there exists a party-formation equilibrium of the extended game in which three parties are formed whenever no equilibrium with two parties exists. This result has been established in proposition A.3 in appendix A.1. □

## A.3   Candidate Selection

Consider a version of the basic model described in section 2.4, where the founder of a party $P$ makes an additional strategic choice in committing to a probability $q_P$. In any situation where party $P$ has both extremist and moderate politicians in its candidate pool for the federal election, the candidate for this election will be randomly drawn from among centrists with probability $q_P$ and from among extremists with probability $1 - q_P$. The choice of $q_P$ is made simultaneous to the proposal of the party. This appendix will provide a proof for the claim made in section 2.6.4 that this extended version of the model may have an equilibrium where extremist politicians are nominated with probability greater than one-half. To do so I will derive an equilibrium of this more general model where parties $L$ and $R$ as defined in the main text get proposed, no other parties can successfully enter, and $q_L = q_R = q$ for some probability $q$. Throughout this section it will be assumed that there are four states, one with a leftist median voter, one with a rightist median voter, and two with centrist median voters. It will also be assumed that $y_f > \frac{29}{9} y_P$.

Start by considering the affiliation behaviour of politicians. Let $\pi_f(q|p)$ be the probability that a candidate for the federal election with platform $p$ wins, given that the other party uses the nomination probability $q$. In the case where party $L$ wins the state election in one other state, the utility of a member of party $L$ with platform -1 in a state where $L$ wins is

$$\frac{1}{2} \left\{ y_s + \left[ \frac{1}{2}(1 - q) + \frac{1}{4} \right] \, \pi_f(q| - 1) y_f + \left( \frac{1}{2} q \, \pi_f(q|0) + \frac{1}{4} \pi_f(q| - 1) \right) y_P \right\} ,$$

with $\pi_f(q| - 1) = \frac{1}{8} + \frac{1}{2}(1 - q)\frac{1}{2}$ and $\pi_f(q|0) = \frac{3}{8} + \frac{1}{2} - \frac{1}{4} q$. For $q = 0$ this expression becomes

$$\frac{1}{2} \left( y_s + \frac{9}{32} y_f + \frac{3}{32} y_P \right) .$$

In the state with median voter at or below -0.5 such a candidate can achieve a utility of at most $y_s$ by deviating to joining a new party. The deviation utility is no greater than

the equilibrium utility for $q = 0$ as

$$\frac{1}{2}\left(y_s + \frac{9}{32}y_f + \frac{3}{32}y_P\right) > \frac{1}{2}\left(y_s + \frac{29}{32}y_P + \frac{3}{32}y_P\right) = \frac{1}{2}(y_s + y_P),$$

where the first inequality holds due to the assumption that $y_f > \frac{29}{9}y_P$. Deviating to joining a new party is consequently worse as long as $y_P \geq y_s$. It follows from the continuity of payoffs in $q$ that politicians with platform -1 or 1 either do not want to deviate even if $q = 1$ or that there exists some threshold $q^e \in (0,1)$ such that the deviation is not undertaken for $q = q^e$, but occurs for any $q > q^e$. In the former case set $q^e = 1$.

Politicians with platform 0 in states with centrist median voters have a choice between joining the same or separate parties. In the latter case each achieves a utility of

$$\frac{1}{2}\left\{y_s + \left(\frac{1}{2}q + \frac{1}{4}\right)\pi_f(q|0)y_f + \left[\frac{1}{2}(1-q)\,\pi_f(q|-1) + \frac{1}{4}\pi_f(q|0)\right]y_P\right\}, \qquad (\text{A.1})$$

with $\pi_f(q|-1)$ and $\pi_f(q|0)$ as given above. For $q = 0$ this simplifies to

$$\frac{1}{2}\left(y_s + \frac{7}{32}y_f + \frac{13}{32}y_P\right).$$

If both politicians join the same party their utility becomes

$$\frac{1}{2}\left[y_s + \left(\frac{1}{12} + \frac{1}{2}q\frac{1}{2} + \frac{1}{4}q\right)\frac{3}{4}y_f + \left(\frac{1}{8} + \frac{3}{16}q + \frac{3}{16}(1-q)\right)y_P\right]. \qquad (\text{A.2})$$

Setting $q$ to zero yields

$$\frac{1}{2}\left(y_s + \frac{1}{16}y_f + \frac{5}{16}y_P\right).$$

Both politicians thus prefer being in separate parties for $q = 0$. It follows from the continuity of payoffs in $q$ that politicians with platform 0 in centrist states either both want to be members of the same party even if $q = 1$ or that there exists some threshold $q^c \in (0,1)$ such that they are indifferent at $q = q^c$, but would prefer being members of the same party for any $q > q^c$. In the former case set $q^c = 1$.

It can be shown that $q^c > 0.5$. To do so evaluate expressions (A.1) and (A.2) at $q = 0.5$. This yields $\frac{1}{2}(y_s + \frac{3}{8}y_f + \frac{1}{4}y_P)$ and $\frac{1}{2}(y_s + \frac{1}{4}y_f + \frac{5}{16}y_P)$, respectively. The first utility is greater than the second utility as long as $y_f > y_P$. This shows that centrist politicians in centrist states prefer to be members of the same party for $q = 0$ and $q = 0.5$. It is easy to show that expression (A.1) is concave in $q$ as long as $y_f > y_P$, while expression (A.2) is linear in $q$. It immediately follows that the former utility must be greater then the latter utility for any $q \in [0, 0.5]$. This demonstrates that the threshold $q^c$ must be greater than 0.5.

Now assume $y_s = y_P = 1$ and $y_f = 5$. In this case it can be calculated that $q^e \approx 0.38$. This is below the threshold $q^c$, which must be greater than 0.5. Is there an equilibrium such that $\mathcal{P}^* = \{L, R\}$ and the founder of each party sets $q$ equal to $q^e$? As long as the affiliation behaviour of politicians does not change, the utility of a founder is increasing in his choice of $q$, as centrist politicians win the federal election with higher probability. Accordingly, neither founder would want to deviate to choosing a lower value of $q$ than the equilibrium one. Increasing the level of $q$ beyond $q^e$ would lead to the entry of a new party, as extremist politicians in the party under consideration would then be willing to deviate by the definition of the cut-off $q^e$. As in the proof of proposition 2.2, the entry of such a party reduces the utility of the party leader undertaking the deviation to zero. As $q^e < q^c$, centrist politicians will join different parties. A straightforward calculation based on the expressions derived above shows that their utility is greater than $y_s$, which is the utility they could achieve by joining a third party.

# Appendix B

# Appendix to Chapter 3

## B.1 Existence of Equilibria where the Incumbent is Re-elected with Certainty

This appendix derives bounds on the existence of equilibria where the incumbent is re-elected with certainty. When no politician of party $C$ is elected with positive probability the party leader is indifferent between any of her pure strategies. Given the restrictions on equilibrium strategies, whether this case can be an equilibrium crucially depends on which posterior beliefs can be generated by weakly undominated strategies.

Fix an arbitrary nomination strategy $\eta$ and let $m(\eta)$ be the ex-ante probability that the moderate gets nominated under $\eta$. A second strategy $\eta'$ weakly dominates $\eta$ only if $m(\eta) = m(\eta')$: In the case $m(\eta) > m(\eta')$ the expected utility of the party leader under $\eta$ would be strictly higher under $\eta$ than under $\eta'$ given that $\varepsilon(M) = 1$ and $\varepsilon(E) = 0$, i.e. the median voter elects the moderate for sure and never elects the extremist. Similarly, if $m(\eta) > m(\eta')$ $\eta$ gives a strictly higher utility for $\varepsilon(M) = 0$ and $\varepsilon(E) = 1$.

Given this first result, the intuition for which strategies are weakly dominated can be given as follows: A strategy $\eta$ is weakly dominated if and only if it is possible to find a second strategy $\eta'$ such that $m(\eta) = m(\eta')$ and $\eta'$ nominates politician $p$ more frequently when this politician is of high quality and less frequently when this politician is of low quality, relative to $\eta$. The remainder of the proof formalizes this idea.

It is claimed that any nomination strategy that features $\eta_M(0, 1) > 0$ and $\eta_M(1, 1) < 1$ is weakly dominated. Construct a second strategy $\eta'_M$ by setting $\eta'_M(1, 1) = \eta_M(1, 1) + \varepsilon$ and $\eta'_M(0, 1) = \eta_M(0, 1) - \frac{\pi}{1-\pi}\varepsilon$ with $\varepsilon > 0$ and leaving all other nomination probabilities unchanged relative to $\eta_M$. Choosing $\varepsilon$ sufficiently small ensures that all probabilities in the new strategy $\eta'_M$ are well defined. By construction, both politicians ex-ante get

nominated with the same probability under $\eta_M$ and $\eta'_M$. The only difference between the two strategies is that for the quality combination $(1, 1)$ the moderate is nominated more frequently under $\eta'_M$ than under $\eta_M$, while for the quality combination $(0, 1)$ the moderate is nominated less frequently. The expected utility of the party leader under the strategy $\eta_M$ can be written as

$$\sum_{q \in Q} Pr[q_C = q] \left\{ \eta_M(q) \left[ r(M)(-(M-1)^2 + Y + w(r(M), r(E))q_M) \right. \right.$$

$$\left. + (1 - r(M))(-(I-1)^2) \right]$$

$$+ (1 - \eta_M(q)) \left[ r(E)(-(E-1)^2 + Y + w(r(M), r(E))q_E) \right.$$

$$\left. \left. + (1 - r(E))(-(I-1)^2) \right] \right\}.$$

Define $U_M \equiv -(M-1)^2 + Y$, $U_E \equiv -(E-1)^2 + Y$, and $U_I \equiv -(I-1)^2$. The difference in the expected utilities under $\eta'_M$ and $\eta_M$ is

$$\pi^2 \, \varepsilon \, \left\{ r(M)(U_M + w(r(M), r(E))) + (1 - r(M))U_I \right.$$

$$\left. - r(E)(U_E + w(r(M), r(E))) - (1 - r(E))U_I \right\}$$

$$- \pi(1-\pi)\frac{\pi}{1-\pi} \, \varepsilon \, \left\{ r(M)U_M + (1 - r(M))U_I \right.$$

$$\left. - r(E)(U_E + w(r(M), r(E))) - (1 - r(E))U_I \right\} ,$$

which is equal to $\pi^2 \, \varepsilon \, r(M) \, w(r(M), r(E))$ and non-negative for any election strategy $r$. This shows that $\eta'_M$ weakly dominates $\eta_M$.

By analogous arguments any strategy such that either $\eta_M(0, 0) > 0$ and $\eta_M(1, 0) < 1$, $\eta_M(0, 0) < 1$ and $\eta_M(0, 1) > 0$, or $\eta_M(1, 0) < 1$ and $\eta_M(1, 1) > 0$, is weakly dominated as well. Now consider a strategy such that $\eta_M(1, 0) < 1$. For this strategy not to be weakly dominated it must be true that $\eta_M(0, 0) = 0$ and $\eta_M(1, 1) = 0$ by the second and fourth rule above, which in turn leads to the requirement $\eta_M(0, 1) = 0$ by the third rule. Any resulting strategy is not weakly dominated, as the construction of a weakly dominating strategy would require reducing the probability of nominating a high quality moderate.

Next, consider a strategy such that $\eta_M(1, 0) = 1$ and $\eta_M(0, 1) > 0$. By the first and third rule given above it must hold that $\eta_M(1, 1) = 1$ and $\eta_M(0, 0) = 1$ for this strategy to not be weakly dominated. Similar to before, to find a strategy that could weakly dominate this strategy it would be necessary to reduce the probability of nominating a

high quality extremist, which would reduce utility against most strategies of the party leader.

Finally, let $\eta_M(1,0) = 1$ and $\eta_M(0,1) = 0$. None of the conditions above imposes any restrictions on $\eta_M(0,0)$ and $\eta_M(1,1)$. Furthermore, any strategy of this kind is not weakly dominated. Raising the probability of nominating a high quality politician while keeping the ex-ante nomination probabilities constant necessarily implies reducing the probability of nominating the second politician when she is of high quality by an equivalent amount.

To summarize, there are only three different types of nomination strategies that are not weakly dominated:

- $\eta_M(1,0) = 1$, $\eta_M(0,1) = 0$, $0 \leq \eta_M(0,0) \leq 1$, $0 \leq \eta_M(1,1) \leq 1$

- $\eta_M(1,0) = 1$, $\eta_M(0,1) > 0$, $\eta_M(0,0) = 1$, $\eta_M(1,1) = 1$

- $\eta_M(1,0) < 1$, $\eta_M(0,1) = 0$, $\eta_M(0,0) = 0$, $\eta_M(1,1) = 0$

The second of these strategies nominates the extremist only if she has high quality and consequently $\bar{\pi}_E = 1$ in this case. For the moderate this strategy implies

$$\bar{\pi}_M = \frac{\pi}{\pi + \pi(1-\pi)\eta_M(0,1) + (1-\pi)^2} \ .$$

This expression achieves its minimum of $\pi$ for $\eta_M(0,1) = 1$. The conditions $\mathcal{I} > -M^2 + \pi$ and $\mathcal{I} > -E^2 + 1$ are therefore jointly sufficient for the existence of an equilibrium where $r(M) = r(E) = 0$. Similarly, the third strategy nominates the moderate only if she has high quality and $\bar{\pi}_M = 1$ must hold, while the lowest posterior expectation over the quality of the extremist that this strategy can generate is $\pi$ for $\eta_M(1,0) = 0$. This implies the joint sufficient conditions $\mathcal{I} > -M^2 + 1$ and $\mathcal{I} > -E^2 + \pi$, where the second condition is satisfied whenever the first condition holds.

For the first of the weakly undominated strategies given above the posterior expectations are

$$\bar{\pi}_M = \frac{\pi(1-\pi) + \pi^2 \eta_M(1,1)}{\pi(1-\pi) + \pi^2 \eta_M(1,1) + (1-\pi)^2 \eta_M(0,0)} \tag{B.1}$$

and

$$\bar{\pi}_E = \frac{\pi(1-\pi) + \pi^2(1 - \eta_M(1,1))}{\pi(1-\pi) + \pi^2(1 - \eta_M(1,1)) + (1-\pi)^2(1 - \eta_M(0,0))} \ . \tag{B.2}$$

This strategy generates $\bar{\pi}_E = 1$ if and only if $\eta_M(0,0) = 1$ and the lowest value of the posterior expectation $\bar{\pi}_M$ that can be achieved in this case is $\pi$, which implies the same sufficient conditions as the first set of conditions given in the previous paragraph. On

the other hand, the lowest value that the right-hand side of equation (B.2) can take is $\pi$. Together with the previous results this shows that no undominated strategy can lead to a posterior expected quality below $\pi$ for any politician. It remains to show which sufficient conditions the current strategy yields if $E$ is such that $-E^2 + \pi \leq \mathcal{I} \leq -E^2 + 1$. This requires for any such $E$ to find the lowest $M$ such that the median voter is indifferent between the incumbent and both politicians of party $C$. This $M$ satisfies $\mathcal{I} = -M^2 + \bar{\pi}_M^*$, where $\bar{\pi}_M^*$ is the solution to the minimization problem

$$\min_{0 \leq x,y \leq 1} \frac{\pi(1-\pi) + \pi^2 x}{\pi(1-\pi) + \pi^2 x + (1-\pi)^2 y}$$
$$s.t. \quad -E^2 + \frac{\pi(1-\pi) + \pi^2(1-x)}{\pi(1-\pi) + \pi^2(1-x) + (1-\pi)^2(1-y)} = \mathcal{I}$$

## B.2 Non-Additive Quality

This appendix provides a proof for the claim in section 3.4.1 that the median voter theorem applies if the utility function of voters is given by

$$-(i-x)^2 + h(|x-i|) \cdot q$$

as long as the following assumptions are satisfied: the function $h : \mathbb{R}_+ \to \mathbb{R}$ is decreasing and concave, there exists a positive constant $d$ such that $h(d) \geq 0$, and all voters are located in the interval $[-d, d]$. It needs to be shown that either all voters to the left or all voters to the right of the median voter agree with the preference ordering of the median voter for all possible combinations of candidates. Without loss of generality, assume that the median voter prefers the extremist over the incumbent: For $i$ equal to zero it then holds that

$$-(E-i)^2 + h(|E-i|) \cdot \bar{\pi}_E \; > \; -(I-i)^2 + h(|I-i|) \cdot q_I \; . \tag{B.3}$$

Consider voters such that $i \in [0, E]$. As $I < 0$ and $E > 0$, inequality (B.3) must hold for these voters: The right-hand side of the expression is decreasing in $i$ while the left-hand side is increasing on this interval.

Now consider voters located in the interval $(E, d]$ in the case where $d > E$. These voters clearly prefer the extremist over the incumbent on ideological grounds. As $h(|E-i|) \geq 0$ for any of these voters, the only way that they could prefer the incumbent over the extremist was if the quality $q_I$ of the incumbent was larger than the expected quality $\bar{\pi}_E$ of the extremist. But this, together with the result shown above that a voter located at $i = E$ must prefer the extremist over the incumbent, implies that all voters in the

interval $(E, d]$ must prefer the extremist as well. To see this note that it follows from $h$ being concave and decreasing, $q_I > \bar{\pi}_E$, and $I < E$ that the function $h(|I - i|) \cdot q_I$ decreases at least as fast as the function $h(|E - i|) \cdot \bar{\pi}_E$ in $i$ on the interval $(E, d]$. It is then clear that inequality (B.3) holds for all $i \in (E, I]$.

# Bibliography

Adams, J. and Merrill, S. (2008), 'Candidate and party strategies in two-stage elections beginning with a primary', *American Journal of Political Science* **52**(2), 344–359.

Aghion, P. and Bolton, P. (1987), 'Contracts as a barrier to entry', *American Economic Review* **77**(3), 388–401.

Ansolabehere, S., de Figueiredo, J. and Snyder, J. (2003), 'Why is there so little money in U.S. politics?', *The Journal of Economic Perspectives* **17**(1), 105–130.

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

Banks, J. (1990), 'A model of electoral competition with incomplete information', *Journal of Economic Theory* **50**(2), 309–325.

Banks, J. and Sobel, J. (1987), 'Equilibrium selection in signaling games', *Econometrica* **55**(3), 647–61.

Barberá, P. (2015), 'Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data', *Political Analysis* **23**(1), 76–91.

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

Bernheim, D., Peleg, B. and Whinston, M. (1987), 'Coalition-proof Nash equilibria I. Concepts', *Journal of Economic Theory* **42**(1), 1 – 12.

Bernheim, D. and Whinston, M. (1986), 'Menu auctions, resource allocation, and economic influence', *The Quarterly Journal of Economics* **101**(1), 1–32.

Besley, T. and Coate, S. (2001), 'Lobbying and welfare in a representative democracy', *The Review of Economic Studies* **68**(1), 67–82.

Besley, T., Persson, T. and Sturm, D. (2010), 'Political competition, policy and growth: Theory and evidence from the U.S.', *The Review of Economic Studies* **77**(4), 1329–1352.

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

Boleslavsky, R. and Cotton, C. (2015), 'Information and extremism in elections', *American Economic Journal: Microeconomics* **7**(1), 165–207.

Bombardini, M. and Trebbi, F. (2011), 'Votes or money? Theory and evidence from the U.S. congress', *Journal of Public Economics* **95**(7-8), 587–611.

Cadigan, J. and Janeba, E. (2002), 'A citizen-candidate model with sequential elections', *Journal of Theoretical Politics* **14**(4), 387–407.

Caillaud, B. and Tirole, J. (1999), 'Party governance and ideological bias', *European Economic Review* **43**(4-6), 779–789.

Caillaud, B. and Tirole, J. (2002), 'Parties as political intermediaries', *The Quarterly Journal of Economics* **117**(4), 1453–1489.

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

Callander, S. (2008), 'Political motivations', *The Review of Economic Studies* **75**(3), 671–697.

Carrillo, J. and Castanheira, M. (2008), 'Information and strategic political polarisation', *The Economic Journal* **118**(530), 845–874.

Castanheira, M., Crutzen, B. and Sahuguet, N. (2010), 'Party organization and electoral competition', *Journal of Law, Economics, and Organization* **26**(2), 212–242.

Chamon, M. and Kaplan, E. (2013), 'The iceberg theory of campaign contributions: Political threats and interest group behavior', *American Economic Journal: Economic Policy* **5**(1), 1–31.

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

Crémer, J. and McLean, R. (1985), 'Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent', *Econometrica* **53**(2), 345–361.

Dal Bó, E. (2007), 'Bribing voters', *American Journal of Political Science* **51**(4), 789–803.

Detterbeck, K. (2012), *Multi-Level Party Politics in Western Europe*, Palgrave Macmillan.

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

Doherty, B. (2013), 'A camapign without end', *New York Times* **21 January 2013**, 21.

Downs, A. (1957), *An Economic Theory of Democracy*, Harper.

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

Fisman, R., Harmon, N., Kamenica, E. and Munk, I. (2012), Labor supply of politicians, NBER Working Papers 17726, National Bureau of Economic Research.

Fryer, R. and Jackson, M. (2008), 'A categorical model of cognition and biased decision making', *The B.E. Journal of Theoretical Economics* **8**(1), 1–44.

Gouret, F., Hollard, G. and Rossignol, S. (2011), 'An empirical analysis of valence in electoral competition', *Social Choice and Welfare* **37**(2), 309–340.

Grossman, G. and Helpman, E. (1994), 'Protection for sale', *American Economic Review* **84**(4), 833–850.

Grossman, G. and Helpman, E. (1996), 'Electoral competition and special interest politics', *The Review of Economic Studies* **63**(2), 265–286.

Grossman, G. and Helpman, E. (2001), *Special Interest Politics*, MIT Press.

Heidhues, P. and Lagerlöf, J. (2003), 'Hiding information in electoral competition', *Games and Economic Behavior* **42**(1), 48 – 74.

Helpman, E. and Persson, T. (2001), 'Lobbying and legislative bargaining', *The B.E. Journal of Economic Analysis & Policy* **0**(1), 3.

Herrnson, P. and Faucheux, R. (2000), 'Candidates devote subtantial time and effort to fundraising'.
**URL:** *http://www.bsos.umd.edu/gvpt/herrnson/reporttime.html*

Hortala-Vallve, R. and Mueller, H. (2015), 'Primaries: the unifying force', *Public Choice* **163**(3-4), 289–305.

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

Jackson, M., Mathevet, L. and Mattes, K. (2007), 'Nomination processes and policy outcomes', *Quarterly Journal of Political Science* **2**(1), 67–92.

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

Jehiel, P., Moldovanu, B. and Stacchetti, E. (1996), 'How (not) to sell nuclear weapons', *American Economic Review* **86**(4), 814–829.

Kartik, N., Tinn, K. and Squintani, F. (2015), Information revelation and pandering in elections, Mimeo.

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

Levine, D. and Palfrey, T. (2007), 'The paradox of voter participation? A laboratory study', *American Political Science Review* **101**(1), 143–158.

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.

Martinelli, C. (2001), 'Elections with privately informed parties and voters', *Public Choice* **108**(1-2), 147–167.

Mattozzi, A. and Merlo, A. (2007), 'The transparency of politics and the quality of politicians', *American Economic Review* **97**(2), 311–315.

Mattozzi, A. and Merlo, A. (2010), Mediocracy, PIER Working Paper Archive 11-002, Penn Institute for Economic Research.

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

Morgan, J. and Várdy, F. (2011), 'On the buyability of voting bodies', *Journal of Theoretical Politics* **23**(2), 260–287.

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. and 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.

Peters, M. and Szentes, B. (2012), 'Definable and contractible contracts', *Econometrica* **80**(1), 363–411.

Polborn, M. (2006), 'Investment under uncertainty in dynamic conflicts', *The Review of Economic Studies* **73**(2), 505–529.

Poole, K. and Rosenthal, H. (1997), *Congress : A political-economic history of roll call voting*, Oxford University Press, New York.

Poole, K. and Rosenthal, H. (2001), 'D-nominate after 10 years: A comparative update to Congress: A political-economic history of roll-call voting', *Legislative Studies Quarterly* **26**(1), 5–29.

Schultz, C. (1996), 'Polarization and inefficient policies', *The Review of Economic Studies* **63**(2), 331–343.

Segal, I. (1999), 'Contracting with externalities', *The Quarterly Journal of Economics* **114**(2), 337–388.

Serra, G. (2011), 'Why primaries? The party's tradeoff between policy and valence', *Journal of Theoretical Politics* **23**(1), 21–51.

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

Snyder, J. and Ting, M. (2011), 'Electoral selection with parties and primaries', *American Journal of Political Science* **55**(4), 781–795.

Spiegler, R. (2000), 'Extracting interaction-created surplus', *Games and Economic Behavior* **30**(1), 142–162.

Stratmann, T. (2005), 'Some talk: Money in politics. A (partial) review of the literature', *Public Choice* **124**, 135–156. 10.1007/s11127-005-4750-3.

Tullock, G. (1989), *The economics of special privilege and rent seeking*, Studies in Public Choice, Kluwer Academic, Boston.