TI 2022-074/VII
Tinbergen Institute Discussion Paper

# Political Regimes, Party Ideological Homogeneity and Polarization

*Micael Castanheira*[1]
*Benoit S Y Crutzen*[2]

1 ECARES, Université Libre de Bruxelles and CEPR

2 Erasmus School of Economics Rotterdam and Tinbergen Institute

# Political Regimes, Party Ideological Homogeneity and Polarization

September 2022

**Micael Castanheira**

ECARES, Université Libre de Bruxelles and CEPR

and

**Benoît S Y Crutzen**

Erasmus School of Economics Rotterdam and Tinbergen Institute

**Abstract**

We develop a model of elections in which parties choose their ideological position and the ideology of their candidates. Tighter candidate selection reduces policy uncertainty for voters. We show that weak institutional constraints, as in a Presidential regime, induce parties to allow their candidates to be ideologically heterogeneous. Tighter constraints or reduced voter polarization induces them to choose an ideologically homogeneous set of candidates. This highlights a multiplier effect of intraparty candidate selection: the parties' best responses amplify institutional and socio-economic changes. These effects rationalize why mainstream parties look so different across the two sides of the Atlantic. Around the middle of the nineteenth century, when facing similar organizational challenges, parties made opposite choices that still apply to this day: the introduction of direct primaries in the US, which decentralized candidate selection, versus the tightening and centralization of selection in Victorian England.

**Keywords**: parties as brands, political regime, intraparty candidate selection, ideology, polarization

# 1 Introduction

In the last decades of the nineteenth century, the two main US parties were facing growing organizational and governance problems.[1] Their response was the introduction of direct primaries. These were introduced in basically all non-Southern states between 1899 and 1915. Direct primaries eventually spread to all remaining states after the abolition of slavery and still characterize US politics to this date. These primaries gave center stage to individual politicians by *loosening* the parties' grip on the selection of congressional candidates; see Ware (2002). Given the districts' heterogeneity in terms of political views and socio-economic characteristics, the advent of primaries contributed to making parties ideologically heterogeneous institutions, leading Katz and Kolodny (1994) to describe them as 'empty vessels' and Schlozman and Rosenfeld (2019) to label them as 'hollow parties'.

English parties were facing similar challenges in Victorian England. Yet, in the UK, parties responded by gradually *increasing* internal ideological homogeneity and cohesiveness. They achieved this by gaining and retaining firm control over the selection of candidates; see Cox (1987). To this date, national party organs still have a decisive say in candidate selection; see Rahat and Cross (2018). Thus, across the two sides of the Atlantic, parties' responded to similar problems in an opposite fashion: loosened party control over candidate selection in the US, tightened control in the UK.

In this paper, we argue that such opposite party strategies are driven, among other things, by the fact that party ideological homogeneity is *endogenous* to the political regime and the heterogeneity of ideological preferences among voters. We also argue that these different party choices impact polarization. To spell out our argument, we develop an electoral game under plurality rule in single-member districts in which parties choose both their ideological position and the ideology of their candidates. For our theory to speak to the US-UK differences highlighted above, we solve for the equilibrium of the game under two different political regimes: a US-style presidential one and a British-style parliamentary one.

[1] The evolution of candidate selection followed a different path in Southern States for reasons that probably were largeley linked to the importance of slavery in those States.

The first key ingredient of our model is that parties are competing *organizations* instead of unitary actors as in Downs (1957). The second key ingredient is that voter preferences are district specific, and so are the preferences of local candidates, be they independents or party representatives. After all, voters and candidates in Liverpool or Detroit have different policy preferences than those of voters and candidates in London or San Francisco. The third key ingredient builds on the findings of Huber (1996) and Diermeier and Feddersen (1998): majority and opposition tend to be more cohesive and homogeneous in a parliamentary system than in a presidential regime. Our model suggests the candidate selection process may help explain this difference.

Our setup builds on a well-understood and universal function of parties: they provide voters with informational shortcuts about the preferred policy of their candidates through the –strategic and publicly observed– choice of both their national platform (Downs 1957) and their candidates (Cox and McCubbins 1993, Snyder and Ting 2002): a tight selection process which leaves little to no ideological leeway to candidates informs perfectly voters about the future choices of a candidate: she cannot deviate from the party platform. If parties opt for a looser selection process, these can put forth policies that together form a *cloud* around the announced party platform. This leaves voters partly uncertain about future policy choices. In the words of Grofman (2008), party candidates are “tethered by a rubber band to the ideology espoused by the parties whose label they run on”.

Building on this function, the existing literature explains why parties exist, but cannot explain why parties look the way they do across different institutional regimes. Our model achieves this by allowing parties to choose their selection process. We first show that selection introduces a *certainty-versus-flexibility* trade-off: if selection is tight, the message sent to voters is very precise but party candidates cannot pander to their local electorate – implying that voters could then prefer to vote for a local independent over the party candidate. If selection is loose, legislators may better represent local preferences, but the informational content of the party label is more limited.

This certainty-versus-flexibility trade-off has opposite implications for the districts close to, and those distant from, the party platform. Close districts always value reduced uncertainty: the party platform represents their ideology well. They thus benefit from tight selection. Distant districts instead benefit from flexibility: if selection is tight, legislators who belong to the party must stay close to the party line, which is disliked by distant voters. Taken together, these opposite preferences imply that parties never value selection processes that are neither tight nor loose: these please neither close nor distant districts. Instead, they choose either tight selection, this helps win close districts but alienates distant ones, or loose selection, which nets distant districts but reduces support in close ones.

How does this trade-off interact with institutional constraints? Parliamentary systems produce relatively tight constraints on intraparty ideological heterogeneity. Surprizingly, we find that parties *do not* freeride on these external constraints: even with moderate legislative constraints, parties are induced to choose a tight selection process. They also avoid strong levels of polarization: since they only attract close districts, they must locate sufficiently close to the ideological centre to maximize their seat share. A US-type presidential regime is associated with laxer legislative constraints. In this case parties may decide to target distant districts: if preference heterogeneity is sufficiently high in the electorate, parties (1) choose a loose selection process and (2) choose polarized platforms. Only if district preferences are very homogeneous do parties select platforms close to the centre, and switch to tight selection.

Our analysis thus shows that if the ideological homogeneity of parties adapts to the political regime, it is largely *through the decisions of parties*, and not simply because of external institutional constraints. This identifies a *multiplier effect* of the way parties organize: party leaders may want to switch from loose to tight selection even when institutional changes are marginal. We document in Section 7 that this multiplier effect is crucial in understanding: 1) the reduction in intraparty ideological heterogeneity in Victorian England, as described by Cox (1987); and 2) the introduction of direct primaries in the US at the end of the nineteenth century, as described by Ware (2002). In both countries, these institutional changes still shape politics to this day.

The rest of the paper is organized as follows. Section 2 reviews some of the existing literature. Section 3 lays out the model. Section 4 identifies the effects of intraparty candidate selection on electoral success, while Sections 5 and 6 solve for the equilibrium of the game in terms of the optimal intraparty selecton process and platform positions. Section 7 shows how our findings can be used to rationalize the institutional reforms that were adopted in the US and Great Britain in the late nineteenth and early twentieth centuries. Section 8 discusses some extensions of the model. Finally, the last section concludes. Most proofs are relegated to the Appendix.

## 2 Related Literature

To contribute to the existing comparative politics literature, we focus on the mapping from the political regime to intraparty ideological heterogeneity. There also exists a literature that focuses on the effects of electoral rules. See Beath et al. (2016) and Buisseret and Prato (2019) for the effect of district magnitude, Galasso and Nannicini (2015, 2017) for a comparison between plurality rule and proportional representation, Matakos et al. (2019) for the effect of the proportionality of the electoral rule and Carroll and Nalepa (2020) for the effect of list flexibility under proportional representation.

Two key ingredients in our model are the tightness of the candidate selection process and the institutional constraints on intraparty ideological heterogeneity. Huber (1996) and Diermeier and Feddersen (1998) demonstrate that the vote of confidence procedure that characterizes Parliamentary regimes produces more disciplined legislative assemblies than a Presidential regime. We operationalize their findings by introducing an exogenous limit to how ideologically heterogeneous parties can be before the winning party looses its ability to govern in a parliamentary regime. Our main idea is that, as a party becomes more ideologically heterogeneous, a growing part of its troops becomes ideologically too distant from their party's stance for them to be willing to support it. Such internal disagreement, if not kept sufficiently in check, leads to the fall of government under a Parliamentary regime, but not under a Presidential one. Our contribution is to show that even subtle changes in this limit can produce opposite party structures in equilibrium, through the party multiplier effect.

To model the tightness of candidate selection, we borrow from Snyder and Ting (2002, S&T henceforth). Whereas our goal is to explain why parties look so different across political regimes, the goal of S&T is to explain why parties exist at all. They propose a model in which parties impose a participation cost to candidates. This cost is exogenous and increasing in the ideological distance between a candidate and the party position. As a result, only the candidates sufficiently close to the party platform run under the party banner. Higher costs mean that party candidates must be closer to the party platform, and the party label is thus more informative. Importantly, both parties and voters always prefer tight selection: if parties could, they would always choose tight selection in the S&T setup. This is largely due to the fact that in S&T all candidates have preferences that are drawn from the *same* distribution.

We build on S&T's "screening technology" as such, but allow each party to strategically choose both their platform and their selection process. We show that parties may either prefer minimal or maximal intraparty ideological homogeneity in equilibrium. Technically, this result stems from the *certainty-versus-flexibility* trade-off that is absent from S&T. Central to this trade-off is our assumption that candidate preferences are district-specific instead of being drawn from a common, national, pool.

The forces driving the strategic location of parties are also different from those in S&T, and produce qualitatively different results. In their setup, parties always select median platforms, unless the party label conveys very little information. In the latter case, parties must polarize to improve the informativeness of their label. Indeed, S&T assume that candidates have an ideology located on a bounded set, say $[-a, a]$. If the party locates close to $-a$ or $a$, the breadth of the set of party candidate preferences gets smaller, which reduces uncertainty for the voters. In our setup instead, polarization has no aggregate effect on informativeness: since candidates are district-specific, polarization reduces breadth in some districts, but increases it in other districts. Polarization is thus driven by a very different rationale in our model: parties have an incentive to pander to different electorates. As the polity's socioeconomic characteristics become more heterogeneous, this incentive increases, as does equilibrium polarization.

Like us, Samuels and Shugart (2010) study the organizational choices of parties. Yet, they study the parties' incentives to control their representative(s) *in government*, while we study the incentives for leaders to control their *legislative troops*. Their findings complement ours: we show that leaders want ideologically homogeneous troops in a parliamentary regime, and they show that parties also want to make sure executives do not deviate from the party platform. Thus, in a Parliamentary regime, high intraparty homogeneity is the outcome, whether this means the party's tight control on its legislators (our contribution) or on its leadership (Samuels and Shugart's contribution). For presidential regimes, both Samuels and Shugart and our paper predict that legislative freedom should be high, even though our mechanism is different from theirs. Whereas we build on our certainty-versus-flexibility trade-off, they build on the hypothesis that the key contest for parties is the presidential election.

# 3 The Model

**The policy space, voters and parties.** The policy space is unidimensional and represented by the real line. We consider an election under first past the post in a large number (formally, a continuum) of single member districts. The median voter of any district $i$ is always pivotal, which implies that the district Condorcet winner is always elected. District $i$'s median voter has single-peaked and quadratic preferences around her ideological bliss point $y_i \in \mathbb{R}$: given any policy choice $x$, the median voter's ultility from this choice is given by $-(y_i - x)^2$.

There are two parties, a center-right and a center-left one, labelled $R$ and $L$ respectively. The two parties compete in the election by making two publicly observable choices. First, they choose a national platform, $x_P$. These platforms inform voters about the type of society, that is, the set of values, rights and duties each party stands for. Thus, the main function of the two party platforms is a *framing* one: it informs voters about the broad view of society each party wishes to promote and defend. Platforms can thus be thought of as being George Lakoff's center-right "Strict Father family" platform for $R$ and the "Nurturant Parent family" left-of-center one for $L$; see Lakoff (2014). Because they are general moral frames about how society should look like, both of these platform leave wiggle room for the party leadership. Such room allows parties to decide how far from the party platform the ideology of district-specific candidates can be. Let this distance be $\phi_P$.[2] The two party platforms send voters clues about the type of policies the candidate who wins the election in their district is likely to support. Parties are brands as

[2]Assuming that parties can directly select $\phi_P$ allows us to save on notation. Snyder an Ting (2002) detail how $\phi_P$ can result from selection processes that impose costs on candidates. Our results would be identical if we allowed parties to choose this cost function.

in Snyder and Ting (2002). Through their platform and the tightness of their candidate selection process, they offer voters a frame to anticipate and reduce uncertainty about which policies candidates are more likely to support once in the legislature. This framing is more informative the more ideologically homogeneous is a party. To allow for the party label to play its informational role, we restrict $\phi_P$ to be bounded from above by 1.

**Candidates.** Each candidate's preferred policy position is $x_c \in \mathbb{R}$, and is private information to her. Candidates are either independents, with no party affiliation, or representatives of a party. An independent running for election in district $i$ cannot reveal more information about her preferences than the fact that $x_c$ is uniformly distributed on $\mathcal{Y}_i \equiv [y_i - 1, y_i + 1]$. This distribution is district-specific, to capture the fact that in many democracies candidates originate from the district in which they run.

Belonging to a party allows a candidate to convey to voters more precise information about her preferences, at a cost in terms of reduced freedom in terms of policy choices. Given $\phi_P$, any candidate running under the banner of party $P$ must have preferences $x_c$ in:

$$x_c \in \mathcal{X}_P \equiv [x_P - \phi_P, x_P + \phi_P] . \tag{1}$$

**Timing.** We consider the following timing:[3]

- $t = 1$: party leaders $L$ and $R$ select their national platforms, $x_L$ and $x_R$.
- $t = 2$: party leaders select $\phi_L$ and $\phi_R$, and candidates are assigned to parties.
- $t = 3$: each district median elects his preferred candidate, and payoffs are realized.

**Institutional and economic environment.** We introduce two (exogenous) parameters that define the country's institutional and socioeconomic environment. The country's institutional environment is summarized by its minimal level of required *legislative cohesion* $\lambda$. The socioeconomic environment is captured by the *heterogeneity of voter preferences*, $\sigma$; $\sigma$ proxies *preference and income heterogeneity* across districts. In

[3]Reversing the timing between periods 1 and 2 produces the same results.

our analysis, we wish to distinguish between parliamentary and presidential regimes. To achieve this, we make the following assumption on the minimal level of legislative cohesion the legislature needs to exhibit for government to survive:

**Assumption 1** *The party in government can only operate if its legislator preferences are within distance $\lambda$ of its platform. $\lambda$ is strictly smaller in a Parliamentary than in a Presidential regime.*

Turning to the parameter that characterizes the economic environment, we make the following assumption:

**Assumption 2** *The distribution of district medians $y_i$ is a centered Normal with standard error $\sigma$:*

$$f(y_i) = \frac{\exp\left[-y_i^2/\left(2\sigma^2\right)\right]}{\sqrt{2\sigma^2}}.$$

In the next three sections, we solve for the perfect Bayesian equilibrium of the game in terms of vote, intraparty ideological homogeneity, and party platforms.

# 4 How Does Selection Impact Voting? ($t = 3$)

Three sets of candidates can run in each district: (1) independent candidates, who are not affiliated with any party; (2) candidates affiliated with party $L$ and (3) candidates affiliated with party $R$. Since voters cannot observe candidate preferences directly, all candidates within one of these sets are *ex ante* identical in the eyes of a voter. The district median's expected utility from electing any local independent is:

$$\begin{aligned} \mathsf{E}u\left(y_i, x_c | x_c \in \mathcal{Y}_i\right) = \mathsf{E}_{x_c \in \mathcal{Y}_i}\left[-\left(y_i - x_c\right)^2\right] &= \int_{y_i-1}^{y_i+1} -\left(y_i - x_c\right)^2 \; f\left(x_c\right)\; dx_c \\ &= -\left(y_i - y_i\right)^2 - 1/3 = -1/3. \end{aligned}$$

Voters have more information about party candidates: first, given that she runs in district $i$, the party candidate must have preferences somewhere in $\mathcal{Y}_i \in [y_i - 1, y_i + 1]$. Second, being a party candidate, she must also have preferences somewhere in $\mathcal{X}_P \equiv$

$[x_P - \phi_P,\, x_P + \phi_P]$ (see (1)). Thus, voters know that a candidate of party $P$ who runs in district $i$ has preferences uniformly distributed on the set:[4]

$$\mathcal{P}_i\left(x_P, \phi_P\right) \equiv \mathcal{Y}_i \cap \mathcal{X}_P.$$

It follows that the median voter's expected utility from electing a candidate of party $P$ is:

$$\mathsf{E}_i \; u\left(y_i, x_c | x_c \in \mathcal{P}_i\left(x_P, \phi_P\right)\right) = -\left(y_i - \mu_i\left[x_P, \phi_P\right]\right)^2 - \sigma_i^2\left[x_P, \phi_P\right], \tag{2}$$

where, by the properties of uniform distributions:

$$\begin{cases} \mu_i\left[x_P, \phi_P\right] = \frac{\max[y_i - 1, x_P - \phi_P] + \min[y_i + 1, x_P + \phi_P]}{2} \\ \\ \sigma_i^2\left[x_P, \phi_P\right] = \frac{\left(\max[y_i - 1, x_P - \phi_P] - \min[y_i + 1, x_P + \phi_P]\right)^2}{12} \end{cases} \tag{3}$$

The district median's decision to vote for either candidate depends on (a) the distance between the median's bliss point $y_i$ and (b) the platform $x_P$ of each party. For a given platform $x_P$, we can separate the districts into those that are *close* and those that are *distant* from $x_P$:

- *Close* districts are districts such that $y_i$ is within distance $1 - \phi_P$ of $x_P$: $|y_i - x_P| \leq 1 - \phi_P$. In these districts, the party set $\mathcal{X}_P$ is *within* the district set $\mathcal{Y}_i$.

- *Distant* districts are such that $y_i$ is further than $1 - \phi_P$ from $x_P$. In these districts, the set of party candidates is both a function of the district and of the party set.

In close districts, the expected position of a party candidate is $x_P$, independently of $\phi_P$. As a consequence, voters in a close district have an unambiguous preference for homogeneous parties: this is the *variance-reduction effect* of tight selection. In distant districts instead, looser selection may be preferred, as it allows a party candidate to have a bliss point which is closer to that of the district median. This is the *legislative freedom effect* of loose selection. Substituting for (3) in (2) shows that expected utility in a distant district is maximized at $\phi_P = |y_i - x_P| + 1/2$ and thus is actually hump-shaped in $\phi_P$. Yet, as the relevant comparison for voters in any district is between the utility from voting for the

[4] If the set is empty, no candidate from party $P$ is expected to enter. In the out-of-equilibrium case one such candidate runs, beliefs are such that the candidate's platform is the relevant boundary of $\mathcal{Y}_i$.

party candidate and that from voting for the independent, distant districts unambiguously prefer minimal homogeneity, as this reduces the expected distance between the candidate's ideological bliss point and $y_i$. When choosing how much freedom of action to grant its candidates, the party will thus have to weigh the preferences of these two sets of districts against one another. This is illustrated in Figure 1.

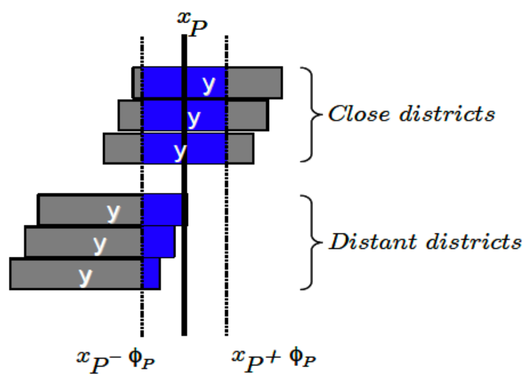


Figure 1: Close and distant districts

Remember that in any district a party $P$ candidate faces two competitors: the independent and the candidate from the other party. This candidate must offer higher expected utility than both competitors to win the electoral seat. Our first step is to identify the set of districts in which a party $P$ candidate beats the independent:

**Definition 1** *The set of districts who prefer a candidate of party $P$ to an independent is party $P$'s* catchment area.

Our first proposition formalizes how this catchment area relates to the party position $x_P$ and to $\phi_P$:[5]

**Proposition 1** *All districts $y_i$ within distance $\kappa(\phi_P)$ of the party platform $x_P$ prefer the party candidate to the local independent. The catchment area of a party is therefore a compact set centered on $x_P$:*

$$\mathsf{E}_i \; u(y_i, x_P) \geq \mathsf{E}_i \; u(y_i, x_I) \Leftrightarrow |y_i - x_P| \leq \kappa(\phi_P),$$

[5]Omitted proofs are in the appendix.

*where* $\kappa(\phi_P) \equiv \max\left[\sqrt{\frac{1-\phi_P^2}{3}}, \phi_P\right]$, *has a global minimum at* $\phi_P \equiv \phi_{\min} = 1/2$, *a local maximum at* $\phi_P = 0$, *and a global maximum at* $\phi_P = 1$.

Figure 2 illustrates this result graphically. The parabolic curve is the outer limit of the set of close districts that vote for the party candidate. The straight lines are the outer limits of the set of distant districts that vote for the party candidate. The catchment area is the outer envelope of these curves.

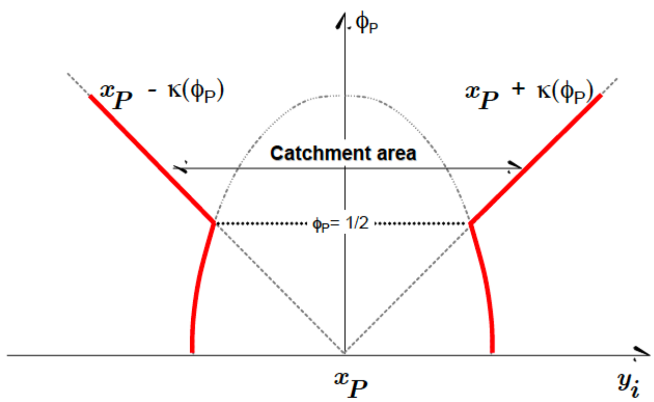


Figure 2: A party's catchment area

Proposition 1 and Figure 2 show how $\phi_P$ maps into electoral support. As we said above, intermediate levels of selection tightness do not maximize expected utility neither in close nor in distant districts. This is why the size of the catchment area $\kappa(\phi_P)$ is minimal in $\phi_P = 1/2$: intermediate levels of $\phi_P$ minimize electoral support. Parties thus prefer "extreme" forms of organization.

Which extreme form do parties choose? The intricacy is that the identity of the marginal district changes with changes in $\phi_P$. If $\phi_P$ is smaller than 1/2, the party catchment area contains close districts only. To expand its catchment area, the party thus benefits from further increasing candidate homogeneity, to cash in on the variance-reduction effect of the party label. A local maximum is found when homogeneity is maximal ($\phi_P = 0$). This is the bottom part of the figure. By contrast, for relatively high levels of intraparty candidate heterogeneity, $\phi_P > 1/2$, the marginal district is distant. In this case, the party has an incentive to further increase heterogeneity: this increases utility in the marginal district and induces the next district to also prefer the party candidate. The global maximum is found when intraparty candidate heterogeneity is maximal ($\phi_P = 1$).

Remark that we can focus on the preference of the marginal district because, from Proposition 1, districts closer to the party keep preferring the party candidate to the independent: the party catchment area is always a compact set. Compactness also implies that electoral support is bounded. This is because the legislative freedom effect of ideological homogeneity implies *party alienation* beyond some distance. Traditional Downsian analyses abstract from party alienation: in the absence of competition from another party, the party catchment area is the whole ideological spectrum. This does not happen in our setup, because of the presence of independent candidates. When voters have the option to vote for independents, they will do so when the party platform is too distant. The boundedness of the catchment area is key to the other findings below.

Finally, the actual shape of the catchment area is only partially due to the specific assumptions we made. For example, the linearity of the catchment area in $\phi_P$ for $\phi_P \geq 1/2$ does not depend on these. We investigate this and other issues in section 8.2 below, in which we discuss how generalizing our assumptions impacts on our results.

# 5 Equilibrium Candidate Homogeneity ($t = 2$)

At time $t = 2$, parties choose intraparty candidate homogeneity to maximize their seat share, taking as given the national platforms chosen in the previous stage. From Proposition 1, we know that a party can win the seat in district $i$ only if this seat is within its catchment area. The two parties' objective function can thus be written as:

$$V_P(\phi_P, x_P) = \int_{x_P-\kappa(\phi_P)}^{x_P+\kappa(\phi_P)} \mathbf{1}\left[u(y_i, x_P|\phi_P) > u(y_i, x_{-P}|\phi_{-P})\right] \, dF(y_i),$$

where $\mathbf{1}[\cdot]$ is the indicator function, taking value 1 when district $i$ prefers the candidate of party $P$ to the candidate of the other party, $-P$.

As this stage, we must distinguish between two cases: the first is when institutional constraints are tight –namely when $\lambda < 1/\sqrt{3}$. The second is when they are loose –namely when $\lambda > 1/\sqrt{3}$.

## 5.1 Case 1: tight institutions

We have:

**Proposition 2** *When institutional constraints are tight ($\lambda < 1/\sqrt{3}$), maximal homogeneity is a dominant strategy for any distribution of districts and any degree of polarization.*

The intuition for this result is a direct consequence of the findings of Section 4. Suppose first that the two party platforms are so distant that their respective catchment areas cannot overlap. When $\phi_P$ cannot exceed $1/\sqrt{3}$, by Proposition 1, maximal homogeneity maximizes the size of the catchment area, and thus $\phi_P = 0$ also maximizes $P$'s seat share.

If the two platforms are close in the sense that the two catchment areas (may) overlap, the two parties are competing directly for some (centrist) districts. By contrast, they only compete against independents in outer districts. Full intraparty homogeneity still maximizes the number of seat won in outer districts. What about centrist districts? Given the institutional constraint $\phi_P \leq \lambda \left(\leq 1/\sqrt{3}\right)$, these districts are at most at distance $1/\sqrt{3}$ from the party platform. The proof of Proposition 2 shows that these districts also prefer maximal homogeneity. Thus, any district that may potentially elect a candidate of party $P$ prefers maximal homogeneity, independently of the distance between party platforms or the distribution of districts.

## 5.2 Case 2: loose institutions

When institutions put less constraint on the parties' choice of internal homogeneity, that is, when $\lambda > 1/\sqrt{3}$, we have:

**Proposition 3** *If $\lambda > 1/\sqrt{3}$, $\phi_P^*$ depends both on party platforms and on the degree of preference heterogeneity $\sigma$:*

*1) If $|x_R - x_L| \geq 2\lambda$, such that the two catchment areas cannot overlap, then $\phi_P^* = \lambda$, that is parties minimize candidate homogeneity in equilibrium.*

*2) If the two catchment areas can overlap, equilibrium homogeneity also depends on voter preference heterogeneity $\sigma$. Set $\lambda = 1$. Then,*

i) *if $-x_L = x_R \equiv x \geq 1/2$, parties set: $\phi_P^* = 1$;*

ii) *if* $-x_L = x_R \equiv x < 1/2$, *there exists a cut-off* $\sigma(x)$ *such that* $\phi_P^* = 0$ *if and only if* $\sigma < \sigma(x)$ *and* $\phi_P^* = 1$ *otherwise.*

Together, Propositions 2 and 3 show how institutional constraints, ideological polarization and the socioeconomic environment interact to determine equilibrium intraparty homogeneity. They reveal a hierarchy of incentives due to, first, institutions, then polarization and then, thirdly, socioeconomic factors:

- First, if institutional constraints are tight, parties choose maximal homogeneity, irrespective of other considerations. If institutional constraints are loose, then parties face a more complex trade-off which depends on the extent of polarization.

- When platforms are highly polarized, parties avoid direct competition. Their primary target is then to maximize the size of their catchment area, which requires maximizing their candidates' ideological heterogeneity. This is the *multiplier effect of selection*: small institutional changes ($\lambda$ being slightly smaller or greater than $1/\sqrt{3}$) produce substantially different levels of intraparty ideological homogeneity (Proposition 2 versus Proposition 3).

- If platforms are close to one another, socioeconomic factors enter into play. In this case, parties face two countervailing incentives. On the one hand, they should minimize homogeneity to increase the size of their catchment area. On the other hand, they should maximize homogeneity to gain seats in the centrist districts for which they compete directly. If preference heterogeneity is high, parties find it more valuable to minimize homogeneity, because there are few centrist districts. By contrast, if preferences are sufficiently homogeneous, many districts are "close". Thus, parties prefer to maximize homogeneity.

# 6 Equilibrium platforms ($t = 1$)

We distinguish again between tight and loose institutional constraints.

## 6.1 Case 1: tight institutions

If $\lambda \leq 1/\sqrt{3}$, we know from Proposition 2 that parties set $\phi_P = 0$ at $t = 2$. The parties' vote shares can then be expressed as:

$$\begin{aligned} V_L\left(\phi_L = 0, x_L; x_R\right) &= \int_{x_L - \frac{1}{\sqrt{3}}}^{\min\left[x_L + \frac{1}{\sqrt{3}},\ \frac{x_L + x_R}{2}\right]} dF\left(y_i\right), \\ V_R\left(\phi_R = 0, x_R; x_L\right) &= \int_{\max\left[x_R - \frac{1}{\sqrt{3}},\ \frac{x_L + x_R}{2}\right]}^{x_R + \frac{1}{\sqrt{3}}} dF\left(y_i\right). \end{aligned} \tag{4}$$

Our next proposition identifies the two equilibrium platforms:

**Proposition 4** *For $\lambda \leq 1/\sqrt{3}$, parties are fully homogeneous ($\phi_P = 0$) and the pair of party platforms is:*

$$\begin{aligned} \left(-x_L = x_R =\right) x &= 0, \ \textit{for } \sigma^2 < 1/\left(6 \log 2\right); \\ &= \sigma\sqrt{2 \log 2} - \sqrt{1/3}, \textit{for } \sigma^2 \in \left[1/\left(6 \log 2\right), 2/(3 \log 2)\right]; \\ &= 1/\sqrt{3}, \ \textit{for } \sigma^2 > 2/(3 \log 2). \end{aligned}$$

Proposition 4 and Figure 3 show that the two parties choose $x_L = x_R = 0$, the preferred platform of the median of the median voters, as in Downs (1957), only when preferences are sufficiently homogeneous across districts. Otherwise, polarization increases in preference heterogeneity. Yet, there is an absolute ceiling to polarization. This stems from the endogenous *alienation effect*: since voters prefer the independent candidate when the party platform is too distant, a party cannot win seats in centrist and outer districts with the same ideological position. The party must choose a sufficiently extreme position to win in outer districts, but then loses in centrist districts. Consider the out-of-equilibrium case in which the two parties are so polarized that their catchment areas are not even tangent. In that case, both parties lose the center to independents. Since there are more centrist than extremist districts, both parties can increase their seat share by moderating their platform. In other words, parties never polarize beyond the point in which they lose the center, which explains the absolute ceiling to polarization in Proposition 4.

This being said, up to which point will the two parties move towards the center? Starting from the point in which the two catchment areas are tangent, any move to the center increases the overlap between the two parties' catchment areas, and thus the extent of direct competition between the two parties. Since both parties choose maximal

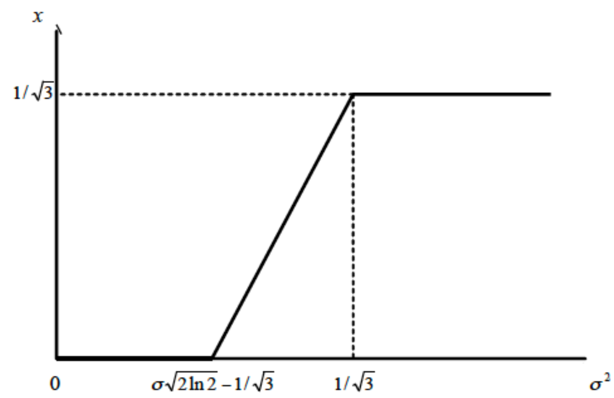


Figure 3: Polarization as function of preference heterogeneity

homogeneity (see Proposition 2), voters prefer the party that is ideologically closest to them. Thus, for $x_L < x_R$ a marginal move by $L$ to the right amounts to:

1. the loss of $f\left(x_L - 1/\sqrt{3}\right) dx_L$ seats from the outer left districts, and

2. the gain of $\frac{1}{2} f\left[\left(x_L + x_R\right)/2\right] dx_L$ seats from the centrist districts.

The important difference with the case in which catchment areas do not overlap is that the marginal gain in the center is halved because of direct competition. That is, because of the overlap, each party wins only *half* as many centrist districts as in the absence of an overlap.

The larger is inter-district preference heterogeneity, the lower is the marginal gain of targeting the center, and the higher is the cost. An interior equilibrium is found when the marginal costs and benefits are equalized. Such interior equilibria are therefore characterized by symmetric platform positions, because of the symmetry of the distribution $f(y_i)$.

Corner solutions involve either full convergence to the median (when $\sigma$ is sufficiently small) or maximal polarization (when $\sigma$ is large). To understand the latter case, note that the seat gain from centrist districts is discontinuous at the point where the two catchment areas become tangent: it is reduced by a half. When $\sigma$ is large, this halving makes the net payoff drop from a strictly positive to a strictly negative value. Both parties thus avoid either polarizing or moderating further: they both have an incentive to keep the two catchment areas exactly tangent.[6]

## 6.2 Case 2: loose institutions

If institutional constraints are loose, platform choices at stage 1 can affect the tightness of selection at stage 2. Equilibrium platform positions are thus the result of more elaborate strategic considerations. We have:[7]

**Proposition 5** *For $\lambda > 1/\sqrt{3}$,*
i) *there exists $\sigma_B(\lambda)$ such that $\sigma > \sigma_B(\lambda)$ is a sufficient condition for parties to choose polarized platforms $x_R = -x_L = \lambda$ and maximal ideological heterogeneity $(\phi_P = \lambda)$ in equilibrium. In particular, $\sigma_B(\lambda = 1) = \sqrt{2/\log 2}$.*
ii) *there exists $\sigma_T \equiv 1/\sqrt{6\log 2}$ such that $\sigma \leq \sigma_T$ is a sufficient condition for parties to choose centrist platforms $(x_L, x_R) = (0, 0)$ and maximal homogeneity $(\phi_P = 0)$ in equilibrium.*

Proposition 5 shows that, through preference heterogeneity, the parties' organizational choices become intimately related to their choice of ideological positions. As highlighted in the previous section, 'loose' institutions –that we associate with Presidential regimes– imply that parties may either prefer maximal or minimal homogeneity. Proposition 5 shows that when preference heterogeneity is "large", parties would like to have the possibility of maximizing the size of their catchment area at time 2, by maximizing candidate heterogeneity. To reach the subgame in which they can take full advantage of candidate freedom, parties must take action at time 1. Choosing polarized platforms is used for that purpose: it prevents direct competition and sustains maximal ideological heterogeneity at time 2.[8]

[6] This also implies that asymmetric equilibria also exist in the neighborhood (the size of which is increasing in $\sigma$) of the symmetric equilibrium, but the associated polarization is constant and always equal to $2\lambda = 2/\sqrt{3}$. We thus ignore these equilibria.

[7] As in the previous case, when $\sigma$ is large, there exists a neighborhood around the symmetric pair of platform positions where parties can locate. Yet, as before, these equilibria are symmetric insofar as party ideological homogeneity is concerned and polarization $(2\lambda)$ is unaffected.

[8] Note that the timing of the game could be reversed without affecting this result. If parties first chose

Conversely, when preference heterogeneity is "small", centrist districts are numerous. In that case, parties maximize their seat share by becoming as strong as they can in these districts. This involves choosing a moderate ideology at time 1, and maximizing the signalling content of the party label at time 2, by maximizing homogeneity. Again, Downs' (1957) median voter theorem only holds when preference heterogeneity is sufficiently low: parties then locate at the very center of the preference distribution and impose that all their candidates deliver the same "nationally median message".

## 6.3 Wrap Up

Propositions 4 and 5 identify four cases in total, depending on whether institutions are tight or loose ($\lambda$ small or large) and on whether preference heterogeneity is high or low ($\sigma$ large or small). Table 1 below summarizes our findings.

**Table 1.** Summary of the main results.

| | **Preference heterogeneity:** | |
|---|---|---|
| **Institutional constraints:** | **Low** | **High** |
| **Tight:** $\lambda \leq 1/\sqrt{3}$ | $\sigma \leq 1/\sqrt{6 \log 2} \Rightarrow$<br>Centrist platforms: $x_P = 0$<br>Maximal homogeneity: $\phi_P = 0$ | $\sigma > \sqrt{2/(3 \log 2)} \Rightarrow$<br>Moderate polarization: $x_R - x_L = 2/\sqrt{3}$<br>Maximal homogeneity: $\phi_P = 0$ |
| **Loose:** $\lambda > 1/\sqrt{3}$ | $\sigma \leq 1/\sqrt{6 \log 2} \Rightarrow$<br>Centrist platforms: $x_P = 0$<br>Maximal homogeneity: $\phi_P = 0$ | $\sigma > \sigma_B(\lambda) \Rightarrow$<br>High polarization: $x_R - x_L = 2\lambda$<br>Minimal homogeneity: $\phi_P = \lambda$ |

Starting with the first column of the table, we see that institutions have little importance when preference heterogeneity is sufficiently small: independently of the institutional environment, parties want to be strong in centrist districts. This implies the choice of moderate platforms and maximal homogeneity. When preferences are very homogeneous, both parties locate exactly at the median voter's bliss point. Note that this suggests that even in US-type presidential systems, parties would incfrease the tightness

the tightness of their selectipn process, they would select maximal flexibility at time 1 as a way to sustain polarization at time 2.

of their selection process to mirror that of parties in Westminster-type parliamentary democracies if the polity's heterogeneity of preferences were to shrink sufficiently.

Moving to the second column, institutions affect both polarization and party homogeneity when preferences are sufficiently heterogeneous. Polarization is larger when institutions are "looser" precisely because of high heterogeneity. In all cases indeed, the maximal extent to polarization is determined by the tangency of the parties' catchment areas. High heterogeneity being an instrument to widen the parties' catchment areas, it is also the driver of stronger polarization. Surprizingly, this may also imply that independents (or, for that matter, additional parties) are less likely to enter the political race when institutions are looser: despite the party label being less informative, parties manage to "cover" a larger part of the ideological spectrum.

# 7 Applications

This section shows how our theoretical results can shed light on historical accounts of the evolution of intraparty candidate selection in the U.K. and US. We illustrate how institutional changes in the parliamentary regime led to more party homogeneity (the multiplier effect) in the U.K. Our theory thus offers a theoretical rationale for the emergence of Cox's (1987) 'efficient secret'. We then offer a rationale for why, at the turn of the nineteenth century, the two mainstream US parties took steps to decentralize candidate selection, thereby allowing for more candidate ideological heterogeneity, and why this decentralized candidate selection process is still in full force to this day.

## 7.1 Evolution of Party Cohesion in Victorian England

Contemporary British voters typically vote for a party on the basis of the philosophy or framing put forward by the party manifesto and the personality and views of the party leader – think of the last electoral campaign of the Conservatives led by Boris Johnson against Jeremy Corbin's Labour. The personality of each local candidate bears little weight on the number of votes received because MPs must follow the rule dictated by that their party (Cox 1987, Chapter 9, and Kam 2009). Yet, the situation was opposite in the early 1800s: MPs were quite independent and voters focused primarily on the characteristics of their candidate. Cox argues that voter behavior changed because of the materialization of the "efficient secret" –*the 'nearly complete fusion of the executive and legislative powers' in the Cabinet*" (p. 51)– and dates the switch around 1868 (p. 92). What we argue here is that the institutional changes that occured between 1832 and 1868 also caused the switch in party ideological homogeneity –in line with the party multiplier effect identified by our model: once the efficient secret materialized, parties decided to control who would be given the most important appointments. Only then did legislative behaviour change strongly. In turn, such changes rationalize the shift in British voters's attention towards the philosophy or framing put forward by the party manifesto and the personality and views of the party leader.

How the nineteenth century electoral Reform Acts led to the birth of Britain's "efficient secret" is one of Cox's main focuses. Relating these evolutions to our model, these acts produced an increase in preference heterogeneity in Parliament ($\sigma$ increases in our model). Simultaneously, the birth of the efficient secret produced a gradual tightening of legislative constraints ($\lambda$ falls progressively in our model). Eventually, this induced parties to tighten candidate selection because " *the [party] labels themselves became increasingly important as collective symbols [. . . S]ince their value could be depreciated by indiscipline or incompetence, the parties had a clear incentive to take a larger role in* ***screening candidates*** *who sought to campaign under their banners*" (p. 144; boldface added).

In 1832, the First Reform Act extended the franchise and increased election competitiveness at the local level. As a consequence, each MP became extremely eager to be visible in Parliament: "*as Sir Robert Peel put it, 'there was a great appetite for legislation, and a strong desire among hon. Members to be distinguished as the introducers of new laws'* " (p. 59). The result was a drop in parliamentary cohesiveness (p. 23, Tables 3.1 and 3.2): Parliament became overcrowded with proposals and hot air; meetings ended in the middle of the night, and the parliamentary session extended into late August. *"Thus, an apparent increase in the desire of members to participate was followed by a diminution of their ability to do so meaningfully. [. . . ] The Commons, in other words, faced the 'tragedy of the commons'."* (p. 60).

As a corollary, the importance of the government rose progressively: "*the more or less unwitting beneficiary of this series of procedural crises was the ministry*" (p. 61). The

Cabinet took increasing control of the Parliament's agenda and, by approximately 1850, "*came about the efficient secret*" (p. 51). At this stage, the government or, more precisely, the Prime Minister had to develop tools to overcome the tragedy of the Commons: right after the Second Reform Act of 1867, which further extended the franchise, the survival of the Cabinet was tied to Parliamentary support. The idea was that the threat of new elections should discipline MPs. However, the evolution of legislative cohesion cannot be entirely explained by this motion of confidence procedure: firstly, discipline first rose in the opposition, even though the motion of confidence procedure is available to the executive only. Secondly, because "*when MPs felt that an immediate dissolution would give them or their party a good chance at reelection, they might actually seek it*" (p. 85). In this case a motion of confidence should actually reduce cohesiveness. Thirdly, if the threat of dissolution was the sole determinant of discipline, cohesiveness should have been decreasing as the date of the election neared. Yet, as Cox shows (p. 86), this pattern is simply absent from in the data.

The multiplier effect identified in Propositions 2 and 3 offers an explaination for the above evolution in party discipline, through changes in candidate homogeneity. We can decompose the increase in discipline into two steps. First, until about 1870-1880, the tighteness of candidate selection increased little by little (in the wording of the model, $\phi$ falls in line with $\lambda$). As of 1870-1880 instead – that is soon after the Second Reform Act and the near fusion of powers in the Cabinet – parties introduced new measures that produced a leap in candidate homogeneity (the multiplier effect in the model), largely through candidate selection.

In the first phase, parties used "whipping" to increase discipline:[9] the percentage of whipped votes increased from 49% in 1836 to 67-69% in the period 1850-1869, to 82% in 1871, and to around 90% from 1875 onwards (Table 3.5, p 24). The effectiveness of whipping on actual discipline was however mixed: intra-party cohesion increased markedly only after 1871-1875.

What changed MP behaviour in that later period? As the Cabinet progressively gained power in the 1850s, pledging allegiance to a party became increasingly important:

[9] Whips are MPs (or Lords) appointed by each party to make sure that MPs vote the way the party wants.

"*Candidates unaffiliated with one of the major parties were unpledged on the single most important issue – the control of the cabinet – and increasingly had little chance against candidates who were pledged on this issue [...E]very voter with a clear preference as to control of the executive preferred an appropriately committed candidate to an uncommitted candidate*" (p. 143). Indeed, between 1856 and 1868, voters progressively switched from voting for candidates to voting for parties (Chapter 9). This gave parties the necessary leverage to finally obtain the additional zest of homogeneity that whipping and fusion did not produce. To begin with, parties started selecting candidates for the cabinet strategically: an MP typically had to "*vote with the party whips consistently, speaking in support of his leaders, patiently awaiting his just reward*" (p. 78), instead of following the "riskier course" of criticizing the government with the hope of being bought off. This strategy bears fruits from 1890-1900 (p. 79). Next, the Corrupt and Illegal Practices Act of 1883 deprived candidates from their finance to organize the campaign or buy votes. It constrained them to rely on the party to win a campaign, because it centralized campaign finance in the hands of the party. In parallel, "*it is soon after the third Reform Act* [of 1884]*, according to Berrington (1967-68), that the English parties first began serious and regular efforts to negotiate intra-party differences, rather than carrying them into the division lobbies*", and "*it is precisely in the late 1870s and 1880s that specifically partisan control of nominations [for the general election] began*" (p. 144).

In other words, the increasing dependence of the Cabinet on parliamentary support can be interpreted as a progressive reduction in $\lambda$ in the model. When the critical level was reached ($\lambda = 1/\sqrt{3}$ in the model), parties decided to deeply reorganize their internal procedures to ensure that they could send a homogeneous signal to their voters, that is, reduce $\phi$ as much as possible in the model. Both the timing of the events and the reading of Cox forcefully suggest that party reorganisation is a direct consequence of these institutional changes: the changes in electoral finance or organization occurred *after* the value of the party label rose, as "*postscripts, logical consequences and reinforcements of the fundamental changes in parliamentary procedure and electoral behavior*" (p. 136).

## 7.2 The United States around 1900

The situation in the U.S. around 1800 was similar to that of the UK: only a small fraction of the population was involved in elections, and the personal vote – that is, the importance of the personal characteristics and record of the individual candidates standing for election in any district – played a very important role.[10] In contrast with the UK, however, the personal vote is still very important today (see for example Cain, Ferejohn and Fiorina 1984, Morgenstern and Swindle 2005 and Zittel 2017), and parties decentralize the selection of congressmen and senators. What we show here is that historical evidence clearly shows that, when faced with challenges similar to that of UK parties, US parties *chose* to minimize candidate homogeneity by letting each electoral district choose its favorite candidate. This ensured that local candidates would be more independent from Washington, to reinforce the party label and compete independent candidates away.

A corner stone of the US political regime is full separation of powers between the executive and the legislative branches of government. Indeed, one of the Founding Fathers, Madison, concluded his paper XLVIII by stating: "*The conclusion which I am warranted* [...] *is that a mere demarcation on parchment of the constitutional limits of the several departments is not a sufficient guard against those encroachments which lead to a tyrannical concentration of all the powers of government in the same hands*" (Madison, Hamilton and Jay 1788 (1987), p. 312). In the context of our model, the US institutions were designed from the start to allow the government to function in the absence of a cohesive legislative body ($\lambda$ is high in our model).

The evolutions of the electorate were similar to that of the UK: "*Between the 1820s and 1830s, the United States underwent a major transformation. A system of politics that had been based in the early decades of the republic on social deference and a rather limited popular participation in politics gave way to a political nation. This was a world in which most Americans were partisans and in which partisan politics was one of the central arenas of social life*" (Ware 2002, p. 65). In other words, polarization ($\sigma$ in the model) increased in the electorate. Massification of politics also produced new informational asymmetries

[10] In what follows we ignore the Southern States as the evolution of candidate selection and discipline followed a different path for reasons that probably were largeley linked to the importance of slavery in those States. See Besley, Persson and Sturm (2010) for more on this.

between candidates and voters: "*America consisted of small towns and rural hinterlands* [..., it] *was a face-to-face society in which informal constraints were largely sufficient to regulate the conduct of politics* [...]. *However, in the decades after the emergence of mass party politics in the 1830s, the social base of America changed radically.* [...] *A style of politics that worked relatively well in the 1830s was working much less well in the new circumstances*" (Ware 2002, p. 21).

Still like in the UK, there was an increasingly strong perception among voters that corruption was plaguing the system. New entrants (independents in the wording of the model) could exploit this perception to compete increasingly strongly against the major parties. These parties had to address this problem. The solution found in the UK was the Corrupt and Illegal Practices Act, which deprived candidates from their capacity of campaigning without the support of the party. In the US, as shown by Ware (2002), parties chose to take the *opposite* route: they introduced the "American Direct Primary" which delegated the selection of candidates locally and outside the hands of the party:[11] the direct primary was aproved by all but three States between 1899 and 1915. It introduced the legal obligation for parties to "*choose their candidates through state-administered elections in which any legally qualified person must be allowed to vote*" (Ranney 1975, p. 121, quoted by Ware 2002, p. 95).[12]

This solution, despite being opposite to that of the UK, had a similar effect: it further reinforced the value of the party label in the eyes of the electorate and improved the electoral success of the main parties. Thus, American parties reacted to forces similar to those of the UK (increasing informational asymmetries and $\sigma$) by totally relinquishing their control over candidate selection to local and independent bureaucracies, that is, by awarding local candidates as much freedom as allowed by the rules of the game: $\phi_P = \lambda$ in the wording of the model. Thus, while Victorian England saw the materialization of the *efficient secret* and high intraparty ideological homogeneity, which now epitomize parliamentary regimes, the US became the archetypical presidential regime centered around separation of powers and ideologically heterogeneous parties.

[11] Castanheira, Crutzen and Sahuguet (2010a, b) analyze when primaries improve candidate incentives in a moral hazard setup.

[12] The only election for which the direct primary *does not* apply is that for the US President.

# 8 Discussion and Extensions

## 8.1 Discussion of our modelling assumption

Our modelling strategy abstracts from the set of other legislative incentives parties develop to ensure legislators vote along party lines. Whereas such an extention would clearly increase the realism of the model, it would not substantially modify its predictions. Also, Krehbiel (1999, p 832) emphasizes that "primitive preferences account for a large share of legislative behavior". Thus, even though our modeling choices cut through important realities, they capture a fundamental relationship between party homogeneity and legislative behavior in a parsimonious way. What is more, our focus is on pre-electoral strategies. Separating "primitive preferences" from other means of imposing legislative discipline is thus beyond the scope of this paper and we feel justified in relying on only one variable to proxy intraparty cohesiveness. Finally, as the applications we cover in Section 7 make clear, the strategic use of candidate selection does play a central role in shaping intraparty homogeneity.

Turning to legislative cohesiveness, it is known to vary substantially across political regimes; it is typically higher when government survival depends on legislative support; see for example Huber (1996) and Diermeier and Feddersen (1998). While these contributions focus on how legislative institutions impact on the cohesion of legislators in the absence of parties,[13] we must translate their predictions into how these institutions would influence party cohesion. In a parliamentary regime, too low discipline would mean that the government falls regularly. In a Presidential regime instead, government survival does not depend on legislative support. In the model, parties can decide to only accept candidates whose ideological blisspoint is within distance $\phi_P \leq \lambda$ of the party platform $x_P$. This assumes that tight selection is (costless and) independent of the political regime. In reality, tight selection is more difficult in a presidential regime, for example because the

[13] Diermeier and Feddersen (1998, p611), for instance, look for "an institutional explanation for voting cohesion that relies on the incentives created by the characteristic features of parliamentary constitutions". Our focus is instead on why parties organize the way they do in different environments. Huber (1996b) deals with parliamentary systems only and assumes exogenous size and characteristics of the coalition supporting the executive.

executive cannot dissolve the assembly. We now show that introducing these differences actually reinforces our results. The assumption that tight selection is costless is thus only meant to clarify the fact that costly selection is not at the core of our results.

To introduce a loss of candidate selection control under direct primaries, suppose that under that selection process there is a lower bound for $\phi_P$: $\phi_P$ can only take on values between $k$ and 1, with $k > 0$. The obvious consequence is that party leaders value even less maximal homogeneity, since the size of the catchment area under maximal homogeneity $k$ is bound to be smaller than under full homogeneity: $\kappa(k) < \kappa(0)$ for any $k < 1/\sqrt{3}$ and $\partial\kappa(\cdot)/\partial\phi_P > 0$ for any $\phi_P \geq 1/2$. Yet, if anything, this added restriction would increase the empirical validity of the model in that it provides an additional rationale for why US parties have chosen to organize as "empty vessels". Not only does the presidential regime provide leaders with incentives to favor candidate freedom because of a larger value of $\lambda$; it also reduces the party leaders' capacity to tighten selection, given the constraints imposed by the direct primary legislation.

Turning to the economic environment, McCarty et al. (2006, chapter 3) show that economic inequality typically maps into more polarized voter preferences. Alesina, Stancheva and Teso (2018), Alesina, Miano and Stancheva (2020) and Boxell, Gentzkow and Shapiro (2022) also offer such evidence. Castanheira, Crutzen and Sahuguet (2010b) also illustrate that inequality in the US is associated with increased income dispersion across states, probably because inequality favors the clustering into "rich" and "poor" states.[14] We thus use only one parameter to proxy the heterogeneity of both ideological preferences and income inequality across districts.

## 8.2 Modelling of the preferences of candidates and voters

In this section, we show that two assumptions made in Section 3 are not necessary for our results to carry through even though they are useful to obtain closed form solutions. Before doing this, let us remark that inverting the timing of events in our game would not modify the above equilibria.

[14]The relationship between economic inequality and polarization is reinforced by the clustering of individuals into subgroups that are internally homogeneous. See e.g. Esteban and Ray (1994) for a conceptualization of this argument.

We now relax the assumptions that (a) candidate preferences are uniformly distributed inside a district and (b) voters have quadratic preferences. Let voter preferences be defined by some function $f$ that is single-peaked and displays risk-aversion:

$$u_i(x_c) = u(y_i, x_c) = f(|x_c - y_i|),$$

with $f' < 0$ and $f'' \leq 0$. To maintain comparability with the quadratic case, we normalize $f(0)$ to zero.

Turning to the bliss point of a candidate, $x_c$ is distributed according to some density function $g_i(x_c)$, with mean $y_i$. This district-specific distribution $g_i(\cdot)$ is the *translate* of a distribution $g(\cdot)$, with support $[-1, 1]$:

$$g_i(x_c) = g(x_c - y_i),$$

such that the support in district $i$ is $\mathcal{Y}_i \equiv [y_i - 1, y_i + 1]$. The CDF of candidate preferences is denoted $G_i(x_c)$ with $G_i(y_i - 1) = 0$ and $G_i(y_i + 1) = 1$. Also, for any pair of districts $i$ and $j$ and any $x \in \mathbb{R}$ we have $g_i(x_c - y_i) = g_j(x_c - y_j)$. Finally, $g$ is symmetric: $g(-x) = g(x)$ and quasi-concave: $g'(x) \leq 0 \ \forall x > 0$.

In this generalized setup, voter $i$'s expected utility of electing a local independent is:

$$U_I \equiv \mathsf{E}u(y_i, x_c | x_c \in \mathcal{Y}_i) = \int_{y_i - 1}^{y_i + 1} u_i(x) \ \ g_i(x) \ \ dx.$$

Given a party platform $\{x_P, \phi_P\}$, the bliss point of a party candidate must be in the subset $\mathcal{P}_i(x_P, \phi_P) \equiv \mathcal{Y}_i \cap \mathcal{X}_P$, where $\mathcal{X}_P \equiv [x_P - \phi_P, \ x_P + \phi_P]$. Focusing here on values of $x_P \geq y_i$ (the analysis is symmetric for $x_P < y_i$), through Bayesian updating, voters determine that the bliss point of party candidate is distributed according to the density function $g_{iP}(x_c)$, given by:

$$g_{iP}(x_c) \equiv \frac{g_i(x_c)}{G_i(\min\{y_i + 1, \ x_P + \phi_P\}) - G_i(x_P - \phi_P)}.$$

As before, two subcases must be considered: (i) districts that are "close" to party $P$, such that $x_P + \phi_P \leq y_i + 1$. (ii) districts that are "distant" from party $P$, such that $x_P + \phi_P > y_i + 1$. It follows that the expected utility of electing a candidate of party $P$ is:

$$\begin{aligned} U_{iP}(x_P, \phi_P) \equiv \mathsf{E}u(y_i, x_c | \mathcal{P}_i(x_P, \phi_P)) &= \int_{x_P - \phi_P}^{x_P + \phi_P} u_i(x) \ \ g_{iP}(x) \ \ dx \text{ in close districts, and} \\ &= \int_{x_P - \phi_P}^{y_i + 1} u_i(x) \ \ g_{iP}(x) \ \ dx \text{ in distant districts.} \end{aligned}$$

This implies that:

**Lemma 1** *The set of districts that prefer a candidate of party $P$ to an independent is a compact set centered on $x_P$: there exists some $\kappa > 0$ such that*

$$U_{iP}(x_P, \phi_P) \geq U_I \Leftrightarrow |y_i - x_P| \leq \kappa$$

Thus, like in the particular case of the uniform distribution and quadratic preferences, the party catchment area is necessarily a compact set centered on $x_P$. Clearly, the cutoff value $\kappa$ is still a function of $\phi_P$. Among other things, the following proposition proves under which (mild) conditions on $g$ the size of the party catchment area has a local minimum in $\phi_P = 1/2$ :

**Proposition 6** *(a) For $\phi_P \geq 1/2$ and any distribution $g(\cdot)$ the most distant district in the party catchment area is at distance $\phi_P$ from the party platform $x_P$. That is: $\kappa(\phi_P) = \phi_P$. (b) Moreover, if candidate preferences are sufficiently uncertain,* i.e. *if $g(1)/g(0) > U_I/(u_i(y_i+1) - U_I)$, then $\kappa(\phi_P)$ has a local* minimum *in $\phi_P = 1/2$. In this case, $\kappa(\phi_P)$ has two local maxima: one with high homogeneity ($0 \leq \phi_P < 1/2$) and one with minimal homogeneity ($\phi_P = \lambda$, conditional on $\lambda > 1/2$).*

Thus, the shape of the catchment area in this generalized case is very close to the one we found in Section 4, with a local minimum in $\phi_P = 1/2$, and a global maximum in $\phi_P = 1$. The main difference is that the other value of $\phi_P (< 1/2)$ for wich $\kappa(\cdot)$ is maximized will be different from 0 and that the value of the expected utilities may not feature tractable closed-formed solutions.

# 9 Conclusion

Comparative studies of economic policy across political regimes implicitly rely on parties being homogeneous in parliamentary regimes and heterogeneous in presidential ones. Yet, these studies systematically disregard parties, and thus cannot explain why parties differ across regimes. We proposed a model that fills this gap. We studied an electoral game in which parties can choose their intraparty homogeneity and their ideological platforms to maximize their seat share. Contrary to the usual Downsian assumption, national parties and their local candidates do not coincide. Political parties act as a "brand": they only admit candidates with preferences sufficiently close to the national platform. This selection process provides voters with information about candidate preferences and the amount of information revealed is endogenous: parties can make their message very precise by adopting strict internal discipline, or loose by letting their candidate choose their position more freely. We also endogenized party positions, and therefore polarization.

We showed that equilibrium intraparty homogeneity is determined both by institutional constraints and by population preference heterogeneity. In turn, ideological homogeneity influences equilibrium polarization. Our results provide a rationale for the marked difference in how US and British parties select their candidates and how ideologically homogeneous they are.

# References

[1] Acemoglu, Daron and James Robinson (2006). Economic Origins of Dictatorship and Democracy. Cambridge University Press

[2] Alesina, Alberto, Armando Miano and Stefanie Stantcheva (2020). "The Polarization of Reality" *AEA Papers and Proceedings* 110: 324–328

[3] Alesina, Alberto, Stefanie Stantcheva and Edoardo Teso (2018). "Intergenerational Mobility and Preferences for Redistribution." *American Economic Review* 108: 521-54.

[4] Ashworth, Scott and Ethan Bueno de Mesquita (2008). "Informative Party Labels with Institutional and Electoral Variation". *Journal of Theoretical Politics*, 20: 251-273.

[5] Beath, A., F. Christia, G. Egorov, and R. Enikolopov (2016). "Electoral Rules and Political Selection: Theory and Evidence from a Field Experiment in Afghanistan". *The Review of Economic Studies* 83: 932-968.

[6] Besley, Tim and Steve Coate (1997). "An economic model of representative democracy". *Quarterly Journal of Economics* 112: 85–114.

[7] Besley, Timothy , Torsten Persson and Daniel M. Sturm (2010). "Political Competition, Policy and Growth: Theory and Evidence from the US", *The Review of Economic Studies* 77: 1329-1352

[8] Boxell, Levi, Matthew Gentzkow and Jesse M. Shapiro (2022). "Cross-Country Trends in Affective Polarization". *The Review of Economics and Statistics*

[9] Buisseret, Peter and Carlo Prato (2019). "Competing Principals? Legislative Representation in List PR Electoral Systems". Mimeo, Harris School of Public Policy

[10] Callander, Steven (2005). "Electoral Competition in Heterogeneous Districts". *Journal of Political Economy*, 113: 1116-1145

[11] Casamata, Georges and Wilfried Sand-Zantman (2005). "Citizen Candidacy with Asymmetric Information". *The B. E. Journal of Theoretical Economics (Topics)* 5(1).

[12] Castanheira, Micael, Benoît S.Y. Crutzen and Nicolas Sahuguet (2010a). "Party Organization and Electoral Competition". *Journal of Law, Economics and Organization* Vol 26(1): 212-242

[13] Castanheira, Micael, Benoît S.Y. Crutzen and Nicolas Sahuguet (2010b). "The impact of party organization on electoral outcomes". *Revue Économique* Vol 61(4): 677-696

[14] Cox, Gary W. (1987). *The Efficient Secret. The Cabinet and the Development of Political Parties in Victorian England.* Political Economy of Institutions and Decisions Series, Cambridge University Press

[15] Cox, Gary W. and Matthew D. McCubbins (1993). *Legislative Leviathan.* University of California Press

[16] Crutzen, Benoit S Y and Nicolas Sahuguet (2020). "Comparative Politics with Intraparty Candidate Selection" Mimeo, Erasmus School of Economics

[17] Crutzen, Benoit S. Y., Sabine Flamand and Nicolas Sahuguet (2020). "Prize allocation and incentives in team contests". *Journal of Public Economics* 182

[18] Diermeier, Daniel and Timothy Feddersen (1998). "Cohesion in Legislatures and the Vote of Confidence Procedure". *American Political Science Review* 92: 611-621

[19] Diermeier, Daniel, Hülya Eraslan, and Antonio Merlo (2003). ""A Structural Model of Government Formation", *Econometrica*, 71(1): 27–70

[20] Diermeier, Daniel, Hülya Eraslan, and Antonio Merlo (2007). "Bicameralism and Government Formation". *Quarterly Journal of Political Science*, 2: 227–252

[21] Downs, Anthony (1957). *An Economic Theory of Democracy*, Harper and Row.

[22] Esteban, Joan-Mariá and Debraj Ray (1994). "On the Measurement of Polarization", *Econometrica* 62(4): 819-851

[23] Eyster, Erik and Thomas Kittsteiner (2007). "Party platforms in electoral competition with heterogeneous constituencies". *Theoretical Economics* 2: 41-70

[24] Galasso, Vincenzo and Tommaso Nannicini (2015). "So closed: Political selection in proportional systems". *European Journal of Political Economy* 40: 260-273

[25] Galasso, Vincenzo and Tommaso Nannicini (2017). "Political selection under alternative electoral rules". *Public Choice* 171: 257-281

[26] Großer, Jens and Thomas Palfrey (2008). "A Citizen Candidate Model with Private Information and Unique Equilibrium", Working Paper, California Institute of Technology.

[27] Grofman, B. (2008). "The Impact of Electoral Laws on Political Parties". in The Impact of Electoral Laws on Political Parties, ed. by D. A. Wittman and B. R. Weingast, Oxford University Press.

[28] Grossman, Gene and Elhanan Helpman (2005). "A Protectionist Bias in Majoritarian Politics", *Quarterly Journal of Economics*, Vol. 120, No. 4, Pages 1239-1282

[29] Grossman, Gene and Elhanan Helpman (2008). "Party discipline and pork-barrel politics" in Elhanan Helpman (ed): *Institutions and economic performance*. Cambridge: Harvard University Press

[30] Grossman, Emiliano and Nicolas Sauger (2009). "The End of Ambiguity? Presidents versus Parties or the Four Phases of the Fifth Republic", *West European Politics* 32: 423-437

[31] Huber, John D. (1996). "The Vote of Confidence Procedure in Parliamentary Democracies", *American Political Science Review* 90: 269-282

[32] Invernizzi, Giocvanna M. and Carlo Prato (2020). "Power Sharing, Mobilization, and Party Organization". Mimeo, Columbia University

[33] Kam, Christopher J (2009). *Party Discipline and Parliamentary Politics.* Cambridge University Press

[34] Katz, Richard S. and Robin Kolodny (1994). "Party Organization as an Empty Vessel: Parties in American Politics". In: *How Parties Organize: Change and Adaptation in Party Organizations in Western Democracies,* Katz, Richard S. and Peter Mair (Eds). SAGE Publications

[35] Katz, Richard S. and Peter Mair (1992) (Eds). *Party Organizations: A Data Handbook on Party Organizations in Western Democracies, 1960-1990.* SAGE Publications

[36] Katz, Richard S. and Peter Mair (1994) (Eds). *How Parties Organize: Change and Adaptation in Party Organizations in Western Democracies.* SAGE Publications

[37] Krehbiel, Keith (1999). ""

[38] Lizzeri, Alessandro and Nicola Persico (2001),"The Provision of Public Goods under Alternative Electoral Incentives", *American Economic Review*: 225-239

[39] Madison, James, Alexander Hamilton and John Jay (1987 (1788)). The Federalist Papers. Penguin Books

[40] Matakos, Konstantinos, Riikka Savolaine, Orestis Troumbouni, Janne Tukiainen and Dimitrios Xefteris (2019). "Electoral Institutions and Intraparty Cohesion", working paper 109, VATT institute for economic research

[41] McCarty, Nolan, Keith T. Poole and Howard Rosenthal (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. MIT Press

[42] Milesi-Ferretti, Gian Maria, Roberto Perotti and Massimo Rostagno (2002). "Electoral Systems and Public Spending." *Quarterly Journal of Economics* 117: 609-657

[43] Morelli, Massimo (2004). "Party Formation and Policy Outcomes under Different Electoral Systems". *Review of Economic Studies* 71: 829–853

[44] Morgenstern and Swindle (2005). "Local Voting in Twenty-three Democracies." *Comparative Political Studies*

[45] Myatt, David (2007). "On the Theory of Strategic Voting" *Review of Economic Studies* 74: 255-281

[46] Osborne, Martin and Slivinski, A., (1996). "A model of political competition with citizen-candidates". *Quarterly Journal of Economics* 111: 65– 96.

[47] Palfrey, Tom (1984). "Spatial equilibrium with entry" *Review of Economic Studies* 51: 139-156

[48] Persson, Torsten, Gérard Roland and Guido Tabellini (1997)."Separations of Powers and Political Accountability". *Quarterly Journal of Economics* 112: 1163-1202

[49] Persson, Torsten, Gérard Roland and Guido Tabellini (2000)."Comparative Politics and Public Finance". *Journal of Political Economy* 108: 1121-1161

[50] Persson, Torsten and Guido Tabellini (1999). "The Size and Scope of Government: Comparative Politics with rational Politicians, 1998 Alfred Marshall Lecture". *European Economic Review 43: 699-735*

[51] Persson, Torsten and Guido Tabellini (2000). Political Economics: Explaining Economic Policy. The MIT Press

[52] Persson, Torsten and Guido Tabellini (2003). The Economic Effects of Constitutions. The MIT Press

[53] Persson, Torsten and Guido Tabellini (2004a). "Constitutions and Economic Policy". *Amercian Economic Review* 18: 75-98

[54] Persson, Torsten and Guido Tabellini (2004b). "Constitutional Rules and Fiscal Policy Outcomes". *The Journal of Economic Perspectives* 94: 25-45

[55] Persson, Torsten, Guido Tabellini and Francesco Trebbi (2003). "Electoral rules and Corruption." *The Journal of the European Economic Association* 1: 958-989

[56] Rahat, G., and Cross, W. P. (2018). Political parties and candidate selection. In Oxford Research Encyclopedia of Politics.

[57] Samuels, David J. and Matthew S. Shugart (2010). Presidents, Parties and Prime Ministers. How the Separation of Powers Affects Party Organization and Behavior. Cambridge University Press

[58] Sauger, Nicolas (2009). "Party Discipline and Coalition Management in the French Parliament", *West European Politics* 32: 310-326

[59] Schlozman, D. and Rosenfeld, S. (2019). The hollow parties. Can America govern itself, 120-52.

[60] Shomer, Yael (2014). "What affects candidate selection processes? A cross-national examination". *Party Politics* 20: 533-546

[61] Shomer, Yael (2017). "The Conditional Effect of Electoral Systems and Intra-Party Candidate Selection Processes on Parties' Behavior". *Legislative Studies Quarterly* 42: 63-96

[62] Snyder, James and Michael Ting (2002). "An Informational rationale for Political Parties". *American Journal of Political Science* 46: 90-110

[63] Ware, Alan (2002). The American Direct Primary, Party Institutionalization and Transformation in the North, Cambridge University Press

[64] Wilson, Frank L. and Richard Wiste. (1976). "Party Cohesion in the French National Assembly: 1958-1973". *Legislative Studies Quarterly* 4: 467-490

[65] Zittel, T. (2017). "The personal vote". The SAGE Handbook of Electoral Behaviour. Thousand Oaks, CA: SAGE Publishing, 668-683.

# 10 Appendix: Proofs

**Proof of Proposition 1**

Using (2) and (3), we need to show that:

$$\text{For } |y_i - x_P| \leq 1 - \phi_P,\ \mathsf{E}_i\ u(y_i, x_P) > \mathsf{E}_i\ u(y_i, x_I) \iff |y_i - x_P| \leq \sqrt{\frac{1-\phi_P^2}{3}}, \tag{5}$$

$$\text{For } |y_i - x_P| \geq 1 - \phi_P,\ \mathsf{E}_i\ u(y_i, x_P) > \mathsf{E}_i\ u(y_i, x_I) \iff |y_i - x_P| \leq \phi_P\,(> 1/2)\,. \tag{6}$$

(5) can be rewritten as:

$$\begin{aligned} |y_i - x_P| \leq \min\left[\sqrt{\frac{1-\phi_P^2}{3}}, 1 - \phi_P\right] &= \sqrt{\frac{1-\phi_P^2}{3}},\ \forall \phi_P \leq 1/2 \\ &= 1 - \phi_P,\ \forall \phi_P \geq 1/2. \end{aligned}$$

Similarly, solving for (6) yields the condition : $|y_i - x_P| \in [\phi_P - 1, \phi_P]$, where the lower bound is negative. Combining this with the condition $|y_i - x_P| \geq 1 - \phi_P$ yields: $|y_i - x_P| \in [1 - \phi_P, \phi_P]$, which is an empty set for $\phi_P \leq 1/2$.

These results imply that the party candidate beats the independent in the districts $i$ such that:

$$\begin{aligned} |y_i - x_P| &\leq \sqrt{\frac{1-\phi_P^2}{3}} \text{ if } \phi_P \leq 1/2 \\ &\leq \phi_P, \text{ if } \phi_P \geq 1/2. \end{aligned}$$

For $\phi_P \leq 1/2$, all the districts within distance $\sqrt{\frac{1-\phi_P^2}{3}}$ of the platform $x_P$ vote for the party. This distance is decreasing in $\phi_P$ and has a maximum of $\sqrt{\frac{1}{3}}$ at $\phi_P = 0$. It has a minimum of 1/2 at $\phi_P = 1/2$.

For $1 \geq \phi_P \geq 1/2$, all districts within distance $\phi_P$ of $x_P$ vote for the party. QED

**Proof of Proposition 2**

Let $d_{i,P} \equiv |x_P - y_i|$. From Section 4, we know that all districts with $d_{i,P} < 1 - \phi_P$ prefer $\phi_P = 0$. Here, we show that all districts within distance $d_{i,P} < 1/\sqrt{3}$ prefer $\phi_P = 0$ to any other

$\phi_P \in \left[0, 1/\sqrt{3}\right]$. We thus need to prove that:

$$\mathsf{E}_i\, u\left(y_i, x_P | \phi_P = 0\right) = -d_{i,P}^2 > -\tfrac{1+\left(d_{i,P}-\phi_P\right)+\left(d_{i,P}-\phi_P\right)^2}{3} = \mathsf{E}_i\, u\left(y_i, x_P | \phi_P > 0\right), \forall \phi_P, d_{i,P} < 1/\sqrt{3}. \tag{7}$$

Rearranging this inequality yields:

$$1 + \left(1 - 2\phi_P\right) d_{i,P} - 2d_{i,P}^2 - \phi_P + \phi_P^2 > 0. \tag{8}$$

(8) always holds for the districts such that $d_{i,P} < 1 - \phi_P$. Differentiating with respect to $\phi_P$ also shows that the inequality is tighest at the corner value: $\phi_P = 1/\sqrt{3}$. Hence, $\phi_P = 0$ is preferred to any $\phi_P \in \left(0, 1/\sqrt{3}\right)$ if it holds in $\phi_P = 1/\sqrt{3}$:

$$-2d_{i,P}^2 + \tfrac{\sqrt{3}-2}{\sqrt{3}} d_{i,P} + \frac{4-\sqrt{3}}{3} \geq 0,$$

which is true for any $d_{i,P} \leq 1/\sqrt{3}$. This proves that $\phi_P = 0$ maximizes party $P$'s seat share for any $\lambda \leq 1/\sqrt{3}$. QED

**Proof of Proposition 3**

For $\lambda > 1/\sqrt{3}$, the party must choose whether to adopt the structure that maximizes the size of its catchment area $(\phi_P = \lambda)$ or the one that maximizes voters' utility in close districts $(\phi_P = 0)$. When the two catchment areas cannot overlap, the party must maximize the size of its catchment area which, from Proposition 1, implies that $\phi_P^* = \lambda$.

Now, consider the case in which the catchment areas can overlap. For $\lambda = 1$, the median voter of the median district is indifferent between full flexibility and full discipline if party platforms are $(-x_L = x_R =)x = 1/2$:

$$\mathsf{E}_i\ u\left(y_i = 0, x = \tfrac{1}{2} \mid \phi_P = 0\right) = \mathsf{E}_i\ u\left(y_i = 0, x = \tfrac{1}{2} \mid \phi_P = 1\right) = -\frac{1}{4}.$$

It follows directly that the median district $(y_i = 0)$ prefers maximal candidate freedom $(\phi_P = 1)$ for any $x > 1/2$. That is, $\phi_P = 1$ maximizes seat share. For $x < 1/2$, the median district prefers full discipline $(\phi_P = 0)$, whereas non-centrist districts (districts close to $x_P \pm 1$) prefer $\phi_P = 1$. Hence, switching from $\phi_P = 1$ to $\phi_P = 0$ allows the party to win districts around $y_i = 0$ at the cost of losing the non-centrist ones. Since the ratio $f(0)/f(y)$ is strictly decreasing in $\sigma$ for any $y \neq 0$, the smaller is $\sigma$, the more weight parties put on winning districts around $y_i = 0$; in contrast, the larger is $\sigma$, the more parties put weight on winning in districts close to $x_P \pm 1$. It is easy to check that, for $\sigma \to 0$, full discipline always dominates. For $\sigma \to \infty$, full flexibility dominates. Since $f(0)/f(y)$ is monotonic in $\sigma$, there exists a unique cutoff value $\sigma(x)$ that makes the party indifferent between the two structures. QED

### Lemma 2

The following lemma will help us prove Proposition 4.

**Lemma 2** *For $\lambda < \sqrt{1/3}$, the equilibrium distance between $x_L$ and $x_R$ can never be larger than $2/\sqrt{3}$.*

Whenever $x_L + 1/\sqrt{3} < 0 < x_R - 1/\sqrt{3}$, we have that:

$$\begin{aligned}
\frac{\partial V_L(\phi_L = 0, x_L; x_R)}{\partial x_L} &= f\left(\min\left(x_L + \frac{1}{\sqrt{3}}, \frac{x_L + x_R}{2}\right)\right) - f\left(x_L - \tfrac{1}{\sqrt{3}}\right) > 0 \\
\frac{\partial V_R(\phi_R = 0, x_R; x_L)}{\partial x_R} &= f\left(x_R + \tfrac{1}{\sqrt{3}}\right) - f\left(\max\left(x_R - \tfrac{1}{\sqrt{3}}, \frac{x_L + x_R}{2}\right)\right) < 0.
\end{aligned}$$

Hence, both parties strictly prefer to move their platform in the direction of their opponent, which proves that $|x_L - x_R| > 2/\sqrt{3}$ cannot be an equilibrium. QED

### Proof of Proposition 4

We first show that $-x_L = x_R = 1/\sqrt{3}$ is an equilibrium for $\sigma^2 > 2/(3\log 2)$. Lemma 2 above shows that $x_L < -1/\sqrt{3}$ and $x_R > 1/\sqrt{3}$ can never be profitable deviations from $-x_L = x_R = 1/\sqrt{3}$. It remains to check under which condition $x_L > -1/\sqrt{3}$ and $x_R < 1/\sqrt{3}$ are not profitable either.

Focus on party $R$ (the analysis is symmetric for party $L$): in $(x_L,\ x_R) = \left(-1/\sqrt{3},\ 1/\sqrt{3}\right)$, we have:

$$\begin{aligned}
\frac{\partial V_R(\phi_L = 0, x_R; x_L)}{\partial x_R} &= f\left(x_R + \tfrac{1}{\sqrt{3}}\right) - \frac{1}{2} f(0) \\
&= \frac{\exp\left[-\frac{2}{3\sigma^2}\right] - \frac{1}{2}}{\sqrt{2\pi\sigma^2}}.
\end{aligned}$$

A deviation to a position $x_R < 1/\sqrt{3}$ is only profitable if this derivative is strictly negative. It is immediate to see that this cannot be the case if $\sigma^2 \geq 2/(3\log 2)$.

Conversely, for $\sigma^2 < 2/(3\log 2)$, the first order necessary condition for a pair of platforms $x_L < 0 < x_R$ to be an equilibrium is that $(\frac{\partial V_R(\phi_R=0,x_R;x_L)}{\partial x_R} =) f\left(x_R + \frac{1}{\sqrt{3}}\right) - f\left(\max\left(x_R - \frac{1}{\sqrt{3}}, \frac{x_L+x_R}{2}\right)\right) =$

0. Given that a similar condition must hold for the other party and that the distribution of district medians is symmetric around 0, the two first order conditions imply that we must have $x_L^*= -x_R^*$ in equilibrium, that is, platforms must be symmetric around 0. Exploiting this fact, the first order condition boils down to: :

$$f\left(x_R + \tfrac{1}{\sqrt{3}}\right) - f(0) = 0 \Leftrightarrow \exp\left[-\frac{1}{2}\left(\tfrac{x_R+1/\sqrt{3}}{\sigma}\right)^2\right] = \frac{1}{2}.$$

Solving this equation yields $x_R^* \equiv \sigma\sqrt{2\log 2} - \sqrt{1/3}$. Of course, $x_R^* > 0$ requires that $\sigma^2 > 1/(6\log 2)$. For lower values of $\sigma^2$, we have the corner solution: $x_L^* = 0 = x_R^*$.

This establishes a necessary condition for an equilibrium. It remains to show that adopting any other position would indeed decrease the number of seats won by the party. For $\sigma^2 \in [1/(6\log 2), 2/(3\log 2)]$, and $x_L = -x^*$, we have:

$$\tfrac{\partial V_R(\phi_R=0,x_R;x_L)}{\partial x_R} = f\left(x_R + \tfrac{1}{\sqrt{3}}\right) - \frac{1}{2} f\left(\tfrac{x_R+\sqrt{1/3}-\sigma\sqrt{2\log 2}}{2}\right). \tag{9}$$

For any $x_R < x^*$, this derivative is always positive: by the properties of Normal distributions, $f\left(x_R + 1/\sqrt{3}\right) > f\left(x^* + 1/\sqrt{3}\right)$ and $f\left(\frac{x_R+x_L}{2}\right) < f(0)$. Hence, all $x_R < x^*$ are dominated by $x_R = x^*$. By Lemma 2, $x_R > x^*$ cannot be profitable deviations either.

QED

**Proof of Proposition 5**

We begin by demonstrating that $\phi_L^* = \phi_R^* = 1$ and $(x_L, x_R) = (-1, 1)$ is an equilibrium for $\lambda = 1$ and $\sigma^2 \geq \sigma_B(1) = \frac{2}{\log 2}$. To this end, we show first that these platforms are optimal if parties choose full flexibility at time $t = 2$.

For the same reason as in Lemma 2, parties never deviate towards a platform $x_L < -\lambda$ and/or $x_R > \lambda$. Let us now show that deviating towards a platform $x_L > -\lambda$ or $x_R < \lambda$ is not profitable either. We focus on potential deviations by $L$:

$$\begin{aligned} \tfrac{\partial V_L(\phi_L=\lambda,x_L;x_R=\lambda)}{\partial x_L} &= \frac{f\left(\frac{x_R+x_L}{2}\right)}{2} - f(x_L-\lambda) \leq \frac{f(0)}{2} - f(x_L - \lambda) \\ &= \frac{\frac{1}{2} - \exp\left[-\frac{1}{2}\left(\frac{2}{\sigma^2}\right)^2\right]}{\sqrt{2\pi\sigma^2}} \text{ for } \lambda = 1. \end{aligned} \tag{10}$$

(10) is necessarily non-positive for $\sigma^2 \geq 2/\log 2$. For such values of $\sigma^2$, by the properties of Normal distributions, (10) is strictly negative for any $x_L \in (-1, 0]$. Furthermore, for $x_L > 0$, we have $V_L < F\left(\frac{x_L+x_R}{2}\right) - F(x_L - 1)$. Hence any $x_L > -1$ are dominated by $x_L = -1$ if full flexibility is maintained.

Now, we show that any deviation involving full discipline ($\phi_L = 0$) at stage 2 is also dominated, when $\lambda = 1$ and $\sigma^2 \geq 2/\log 2$. That is, we show that: $\max_{x_L} V_L(x_L, \phi_L = 0) < V_L(x_L = -1, \phi_L = 1)$. To this end, note that $V_L(x_L, \phi_L = 0)$ is necessarily smaller than $F\left(1/\sqrt{3}\right) - F\left(-1/\sqrt{3}\right) \simeq 0.226$. The latter is the maximum fraction of seats won by a party under full discipline in the *absence* of competition by another party. Conversely, for $\sigma^2 = 2/\log 2$, we have: $V_L(x_L = -1, \phi_L = 1) \simeq 0.381 > 0.226$. This is sufficient to establish that $x_L = -1, \phi_L = 1$ dominates any other $(x_L, \phi_L)$ when $\sigma^2 \geq 2/\log 2$ and $\lambda = 1$. This reasoning extends to any other value of $\lambda$ greater than $1/\sqrt{3}$.

For $\sigma \to \infty$, the density of districts tends to a uniform. This implies:

$$\frac{V_L(x_L = -\lambda, \phi_L = \lambda)}{\max_{x_L} V_L(x_L, \phi_L = 0)} > \frac{2\lambda}{2/\sqrt{3}} > 1, \ \forall \lambda > 1/\sqrt{3}.$$

By continuity, this establishes that, for any $\lambda > 1/\sqrt{3}$, there must exist a value $\sigma_B(\lambda)$ such that, $\forall \sigma > \sigma_B(\lambda)$, $-x_L = x_R = \lambda$, $\phi_L = \phi_R = \lambda$ is an equilibrium. This proves point *i*.

To prove point *ii*, note that, by exploiting the steps of the proof of Proposition 4, $x_L = x_R = 0$ are the optimal platforms if $\phi_P = 0$, $P = L, R$. Applying the same steps as in the proof of Proposition 4 for $\phi_L$ and/or $\phi_R = \lambda \left(\geq 1/\sqrt{3}\right)$, it is immediate to see that $x_L = x_R = 0$ is also the equilibrium. This shows that, in equilibrium, the platforms must be $x_L = x_R = 0$. Now, we check that a deviation in party structure cannot be profitable.

If $\phi_R = 0$, we have:

$$V_L(x_L = 0, \phi_L = 1) = 2\left(F(1) - F(1/2)\right).$$

From the tabulated distribution of the Normal, this is strictly smaller than 0.267, $\forall \sigma^2 \leq (6\log 2)^{-1}$. By contrast:

$$V_L(x_L = 0, \phi_L = 0) = F(0) - F\left(-1/\sqrt{3}\right) > 0.38, \forall \sigma^2 \leq (6\log 2)^{-1}.$$

Since $V_L(x_L = 0, \phi_L = \lambda)$ is yet smaller for other values of $\lambda$, comparing these two vote shares demonstrates point *ii*. QED

## 10.1 Proof of Lemma 1

We first show that if $U_{iP}(x_P, \phi_P) \geq U_I$ for some $y_i (\leq x_P)$, then $U_{jP}(x_P, \phi_P)$ must be larger than $U_I$ for any $y_j \in [y_i, x_P]$. By symmetry, this must also be true for districts to the right of $x_P$. Since $u_i(x_c)$ only depends on the distance between $x_c$ and $y_i$, and since all $g_i(x_c)$ are translates of a common distribution $g(x_c)$, it is equivalent to prove that a decrease in $|x_P - y_i|$ cannot decrease $U_{iP}(x_P, \phi_P)$ below $U_I$. We analyze the case of close and distant districts separately.

**(a) close districts**: holding $\phi_P$ constant, a marginal change from $x_P$ to $x'_P$, such that $|x'_P - y_i| < |x_P - y_i|$, shifts probability mass away from $x_P + \phi_P$ towards $x_P - \phi_P$. Noting that $u_i(x_P + \phi_P) < u_i(x_P - \phi_P)$, it is straightforward to check that $U_{iP}$ must strictly increase. This proves that, like in Section **??**, voter preferences in close districts are single peaked in $x_P$.

**(b) distant districts**: holding $\phi_P$ constant, a similar marginal change in $x_P$ has two effects. It reduces the expected distance between $x_c$ and $y_i$, which increases expected utility. On the other hand, it increases the variance of $x_c$, since the length of the subset $\mathcal{P}_i(x_P, \phi_P) \equiv \mathcal{Y}_i \cap \mathcal{X}_P$ increases; this decreases expected utility. The total effect on expected utility is thus ambiguous, and a direct comparison of $U_I$ and $U_{iP}$ is needed. Given that:

$$U_I = \int_{y_i}^{y_i+1} u_i(x_c) \ \frac{g_i(x_c)}{1/2} \ dx_c \ \text{ and } \ U_{iP} = \int_{x_P-\phi_P}^{y_i+1} u_i(x_c) \ \frac{g_i(x_c)}{1 - G_i(x_P - \phi_P)} \ dx_c,$$

it is straightforward to check that $U_{iP} \geq U_I$ iff $x_P - \phi_P \leq y_i$.

Combining (a) and (b), proves that the set of districts that prefer a party candidate to an independent is a compact set. Symmetry in the utility function and in $g_i$ implies that this compact is centered on $x_P$.

QED

## 10.2 Proof of Proposition 6

*(a)* That $\kappa(\phi_P)$ is the identity function for $\phi_P \geq 1/2$ follows directly from part (b) of the proof of Lemma 1, in which we showed that $U_{iP} \geq U_I$ if and only if $x_P - \phi_P \leq y_i$.

*(b)* To show that $\kappa(\phi_P)$ has a local *minimum* in $\phi_P = 1/2$ if $g(1)/g(0) > U_I/(u_i(y_i+1) - U_I)$, we must show that the latter condition implies that $\kappa'(\phi_P) < 0$ for $\phi_P = 1/2 - \varepsilon$ and $\varepsilon \to 0$, given that we already know that $\kappa'(\phi_p) > 0$ for $\phi_p = 1/2 + \varepsilon$.

Consider district $i$ such that $y_i = x_P - \phi_P$. For $\phi_P = 1/2 - \varepsilon$, we have:

$$U_{iP}(x_P, \phi_P) \simeq \int_{x_P-\phi_P}^{x_P+\phi_P} u_i(x) \ g_{iP}(x) \ dx = U_I,$$

where the second equality stems from the fact that $x_P - \phi_P = y_i$ and $x_P + \phi_P = y_i + 1$. Differentiating with respect to $\phi_P$ must take account of two effects: both the bounds of the integral and the density function $g_{iP}(x)$ are a function of $\phi_P$. This yields:

$$\begin{aligned}
\partial\frac{U_{iP}(x_P,\phi_P)}{\partial\phi_P} &= [u_i(x_P+\phi_P)-U_{iP}(x_P,\phi_P)]\ g_{iP}(x_P+\phi_P)+[u_i(x_P-\phi_P)-U_{iP}(x_P,\phi_P)]\ g_{iP}(x_P-\phi_P) \\
&\underset{\varepsilon\to 0}{\longrightarrow} \underbrace{[u_i(y_i+1)-U_I]}_{<0}\ g(1)+\underbrace{[0-U_I]}_{>0}\ g(0),
\end{aligned}$$

which is negative iff $g(1)/g(0) > U_I/(u_i(y_i+1)-U_I)$.

QED