Comparative Politics with Endogenous Intra-Party Discipline

September 2010

Micael Castanheira
ECARES, Université Libre de Bruxelles
and
Benoit S Y Crutzen
Erasmus Universiteit Rotterdam

Abstract: Why is intraparty discipline higher in parliamentary than in presidential regimes? Does it matter for policy? We propose a model in which parties choose their ideological position and their internal discipline. We show when institutional constraints induce parties to choose either very high or very low discipline. These results rationalize the patterns observed in parliamentary and presidential regimes. They also highlight a multiplier effect of party discipline: the parties’ best responses amplify changes in institutional constraints. These effects are substantial, as we document for the UK, the US, and France. We also show how intraparty discipline depends on voter preference heterogeneity and how it affects polarization. These results rationalize the dynamics of interparty polarization in the US.

Keywords: parties as brands, political regime, intraparty discipline, polarization

---

1We are grateful to Enriqueta Aragonés, Laurent Bouton, Steve Callander, Tom Cusack, Catherine Dehon, John Ferejohn, Simon Hix, Kai Konrad, Andrea Mattozzi, Antonio Merlo, Alessandro Modica, Abdul Noury, Gérard Roland, Howard Rosenthal, Norman Schofield, Otto Swank, Morgane Tanvé and Barry Weingast for insightful discussions. The comments from the audiences in Barcelona (Autonoma), Brussels (APE-X Conference Part2 and ECARES), Milan (DEFAP and ESEM2008), Namur (FUNDP-PAI), New Orleans (Public Choice Society), Paris (Conference on Changing institutions), Rotterdam (ERASMUS Universiteit), Stanford (APE-X Conference Part1), UPenn and WZB Berlin were also particularly helpful. The usual disclaimer applies.
1 Introduction

In Parliamentary democracies, political parties have gradually increased and then maintained high levels of internal discipline, with legislators sticking closely to the party line. Cox (1987) documents this evolution for Victorian England; Wilson and Wiste (1976) and Huber (1996a) document it for France in the years between the Third and the Fifth Republic. By contrast, intraparty discipline in the US has been loose and reforms typically reinforced candidate freedom. For example, the introduction of direct primaries in basically all US States between 1899 and 1915 did away with the parties' hold on the selection of Congressional candidates. More recent reforms have not reduced these differences: currently, dissent between legislators and party leaders is the exception rather than the rule in the main British parties (Kam 2009, p. 10) whereas discipline remains so low in the main US parties that they are described as “empty vessels” by Katz and Kolodny (1994, p31). The above facts suggest that party discipline is endogenous to the political regime. Yet, we still have a limited understanding of how and why intraparty discipline adapts in this way to the political regime. More generally, why should parties look so different in parliamentary and presidential regimes? Does this matter for policy, for example for the choice of party platforms?

To address these questions, we propose an electoral game in which parties choose both their ideological position and their internal discipline. The first key ingredient of our model is that parties are competing organizations. 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 (1996b) and Diermeier and Feddersen (1998): we simply assume that, even in the absence of party discipline, majority and opposition tend to be more disciplined in a parliamentary system than in a presidential regime.

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 ideological platform (Downs 1957) and their level of

---

2Direct primaries are actually run for all public offices but that of the President. An outstanding account of their introduction is Ware (2002).

3The view that dissent in British parties is the exception rather than the rule has not been dented by the episodes of backbench rebellion that received a lot of attention in the media in the last three decades. For example, Kam (2009, p10) confirms that “[t]he vast majority of the time, parliamentary parties are highly cohesive [...].”

4We concentrate on elections that are run under plurality rule in single-member districts.
internal discipline (Cox and McCubbins 1993, Snyder and Ting 2002). Full intraparty discipline perfectly informs voters about the future policy of a candidate: she cannot deviate from the party platform. With less-than-full discipline, the candidates of a party can put forth policies that together form a cloud around the announced party platform. This leaves voters partly uncertain about future policy decisions.

Building on this function, the extant literature explains why parties exist, but cannot explain why parties look the way they do in different institutional regimes. Our model achieves this by allowing parties to control how much discipline they impose on local candidates. We first show that intraparty discipline introduces a certainty-versus-flexibility trade-off: if party discipline is high, the message sent to voters is very precise but party candidates cannot pander to their local electorate – implying that voters could then prefer voting for a local independent over the party candidate. If discipline is low, legislators may better represent local preferences, but the informational content of the party label is more limited.

This certainty-versus-flexibility trade-off has different implications for the districts close to, and those distant from, the party platform. The former always value high discipline: close districts find that their ideology is well represented by the party platform. Hence, they do not suffer much from limited legislative freedom, while they do benefit from the uncertainty-reduction effect of discipline. By contrast, distant districts always value low discipline: if discipline is “too high”, their legislators are forced to implement policies that they dislike. Thus, high discipline helps winning in close districts, whereas low discipline helps winning in distant districts. A remarkable consequence is that parties never choose intermediate levels of discipline in equilibrium: they always prefer either maximum or minimum discipline.

How does this trade-off interact with institutional constraints to shape the equilibrium? We show that maximal discipline is a dominant strategy in a parliamentary system: parties are always induced to target close districts. This has an impact on equilibrium policy: parties avoid strong levels of polarization. The reason is that only close districts vote for disciplined parties. This constrains parties to locate sufficiently close to the centre, where the bulk of districts are. In a presidential regime instead, parties can appeal to distant districts as well. The impact on policy is that they can polarize more without losing the votes of centrist districts. It is only when district preferences are very homogeneous that they select platforms close to the centre, and switch to maximal discipline.

Our analysis thus shows that if discipline adapts to the political regime, it is through the decisions of parties, and not simply because of the constraint introduced by political institutions.
This identifies a *multiplier effect* of the way parties organize: party leaders may want to switch from very low to very high discipline even when institutional changes are marginal. This is the main contribution of our theory.

In Section 8.1, we document that the multiplier effect was crucial in understanding the evolution of intraparty discipline in: 1) Victorian England, as described by Cox (1987); 2) the US at the end of the nineteenth century, as described by Ware (2002); and 3) France between the Third and Fifth Republics, as described by Wilson and Wiste (1976) and Huber (1996a). We also show how the multiplier effect can be exploited to understand the empirical findings of McCarty et al. (2006): polarization between the two major US parties correlates strongly with income inequality.

The rest of the paper is organized as follows. Section 2 reviews some of the existing literature. Section 3 lays out the model. Sections 4 identifies the effects of party discipline on electoral success, while Sections 5 and 6 solve for the equilibrium of the game in terms of intraparty discipline and platform positions. Section 7 discusses some extensions of the model. Section 8 describes how our findings shed light on a number of stylized observations and, finally, the last section concludes. Most proofs are relegated to the Appendix.

2 Related Literature

Two key ingredients in our model are the parties’ screening technology and the institutional constraints on legislator freedom. We borrow the former from Snyder and Ting (2002), and build on Huber (1996b) and Diermeier and Feddersen (1998) for the latter.

Whereas our goal is to explain why parties look so different across political regimes, the goal of Snyder and Ting (2002, S&T henceforth) 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 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, voters always prefer disciplined parties: if parties could, they would always choose maximal discipline in the S&T setup. This is in no small part 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 level of intraparty discipline. We show that parties may either prefer minimal or maximal discipline 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 candidates 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 candidates 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.

Our paper also contributes to the growing literature on comparative politics. One common element in this field is the marginalization of the role of parties and their internal organization. For example, in Persson, Roland and Tabellini (2000), all players are unitary actors—no clear distinction is made between parties and their candidates. Yet, they assume that the members of the majority coalition are predetermined in the parliamentary regime, and open in the presidential regime, a recognition that intraparty discipline is high in the former but low in the latter. To drive home our point, imagine what would happen in a parliamentary democracy if the executive could not rely on a stable enough majority in the legislature. Similarly, in a US-type presidential system, the checks and balances between the executive and the legislature would lose their effectiveness if the President could impose his will on Congress because of intraparty discipline. A missing link in that literature is thus why party discipline is different across institutional regimes.

To address this issue, we introduce the constraints imposed by these regimes on the feasible level of party discipline: Huber (1996b) shows that the vote of confidence procedure induces high discipline in parliamentary democracies even in the absence of parties. Likewise, Diermeier and Feddersen (1998) rationalize differences in cohesion between individual legislators across

parliamentary and presidential regimes, thus also in the absence of parties. These two analyses demonstrate that a Parliamentary regime produces more cohesive legislative assemblies than a Presidential regime.

We take the shortest possible route between their results and our analysis of party discipline and suppose that a parliamentary regime exogenously produces higher intraparty cohesiveness than a presidential regime. In other words, even in the absence of intraparty discipline, the party label would be more informative in a parliamentary regime than in a presidential one. Here, our contribution is to show that even subtle differences between these institutions can produce opposite party structures in equilibrium, through the \textit{party multiplier effect}. That the way parties organize is correlated with the political regime is well known and is very precisely documented in the studies of Katz and Mair (1992 and 1994) on party organization in 12 Western democracies. Our model also extends these findings to show how the polity’s socioeconomic characteristics influence party discipline.

We should mention a few other important contributions that are close to both Snyder and Ting (2002) and this paper: assuming exogenous party discipline, Eyster and Kittsteiner (2007) show that parties may adopt extreme positions to reduce interparty competition. Conversely, Ashworth and Bueno de Mesquita (2008) focus on how incumbent party members choose intraparty discipline but need exogenous party positions to perform their analysis: like in Snyder and Ting, parties would always maximize their vote share by maximizing discipline and choosing a median platform.

Callander (2005) is another paper that could be seen as rationalizing the evidence of McCarty et al. (2006) that we review in Section 8.2. Building on Palfrey’s (1989) sincere voting setup, Callander (2005) introduces multiple districts and shows that parties may polarize more if interdistrict heterogeneity increases. Yet, polarization would entirely disappear if voters relied on strategic voting, a strategy choice that fits the conventional wisdom about voter behavior in first-past-the-post elections since at least Duverger (1957).\footnote{On the role of strategic voting under first-past-the-post elections, see Morelli (2004) and Myatt (2007).} Further, as we explain in more detail in Section 8.2, his model cannot explain the observed distribution of individual candidate preferences McCarty et al. (2006) use to measure polarization.

To close this literature review, we wish to mention that, if our contribution centers around the choices of party leaders concerning the organization of their legislative troops, other important aspect of leadership deserve to be studied as well. For example, whereas we take as granted that the prefered choices of leaders get implemented by their parties, one should ask what
characteristics leaders should display in order for their troops to follow their lead. This is an important avenue for research that is clearly beyond the scope of this paper but that has received some attention recently, for example by Dewan and Myatt (2007, 2008 and 2009).

3 The Model

The policy space is unidimensional and represented by the real line. Following Snyder and Ting (2002), we model parties as brand names (see Section 2): while the policy of a given candidate is uncertain, party discipline can reduce this uncertainty. Our focus is thus on an electoral game with three types of players: voters, candidates and parties, in which parties are the main character of interest.

The Legislature. The economy is divided into a continuum of districts, and each district elects one legislator under plurality rule. To stack the deck against parties, we assume that each legislator controls a fixed fraction of the decisions that are made during the legislature, independently of her party affiliation. It is easy to check that parties would win more seats if we assumed that their legislators controlled more decisions, but our results would remain qualitatively similar, as long as independent candidates had any strictly positive probability of influencing policy.\(^7\)

Voters. The median voter of district \(i\) has single-peaked and quadratic preferences around \(y_i \in \mathbb{R}\).\(^8\) We assume that the district Condorcet winner is always elected, which is equivalent to assuming that the median voter is always pivotal in his district. In the presence of uncertainty, voters must gauge which candidate provides the highest expected utility, given the information available about the policy preferences of each candidate. That is, the candidate winning in district \(i\) is the one that maximizes:

\[
E[u(y_i, x_c)] = E[-(y_i - x_c)^2],
\]

where \(E\) is the expectation operator on \(x_c \in \mathbb{R}\), the preferred policy of candidate \(c\). Party discipline affects expected utility through its effect on the conditional distribution of \(x_c\); this

\(^7\)Whereas this model is somehow typical of the Downsian tradition in that it sacrifices the legislative bargaining part of the political game in order to focus more on the decisions of party leaders in terms of party positioning and organization, we recognize that the legislative bargaining game may also play an important role in shaping the leaders’ decisions. Yet, dealing with this issue is beyond the scope of the present paper. Nevertheless, we take it up in a companion paper, Castanheira, Crutzen and Modica (2010).

\(^8\)This parametric form of the utility function generates tractable closed-form solutions. Section 8 generalizes the utility function.
relationship is analyzed in Section 4.

**Candidates.** Each candidate’s preferred policy position $x_c$ is private information: alone, a candidate cannot reveal more information about her preferences than that $x_c$ is uniformly distributed on $Y_i \equiv [y_i - 1, y_i + 1]$ (Section 7 generalizes the setup to more general distributions). This distribution is district-specific, which captures the fact that in many democracies candidates must reside in the district in which they run.\(^9\) Some candidates stand for election as independents, whereas others run as party candidates.\(^10\)

**Parties.** Parties maximize their seat share by intermediating between the voters’ demand for and the candidates’ supply of policies. They have two instruments at hand: they announce an *ideological position* $x_P$ and how much *freedom of action* $\phi_P$ they grant to their candidates.\(^11\) These two dimensions of party policy are observed publicly.

The standard Downsian approach assumes that party ideology can be represented by a *point* on the real line. We extend this approach by letting parties define a *range* of admissible policies: they admit candidates with preferences distant up to $\phi_P$ from the party’s ideological position, $x_P$. Thus, any voter knows that a candidate running under the banner of party $P$ must have preferences $x_c$ in the range:

$$x_c \in X_P \equiv [x_P - \phi_P, x_P + \phi_P]. \quad (1)$$

To allow for the party label to play its informational role, we restrict $\phi_P$ to be bounded from above by 1.

This modelling strategy abstracts from the set of other legislative incentives parties develop to ensure legislators vote along party lines. Whereas such an extension would clearly increase the realism of the model, it would not substantially modify its predictions. Indeed, Krehbiel (1999, p832) emphasizes that “primitive preferences account for a large share of legislative

\(^9\) There are two differences with citizen candidate models: first, candidate entry is not strategic as in Besley and Coate (1997) and Osborne and Slivinski (1996). Second, there is an information asymmetry between voters and candidates. In the framework of citizen candidate models, the impact of information asymmetries is studied among others by Casanata and Sand-Zantman (2008) and Großer and Palfrey (2008). These papers abstract from the role of political parties.

\(^10\) We use this modelling approach to follow Snyder and Ting (2002) but independents could alternatively be seen as *potential entrants*. Their role in the model is only to provide an outside option to voters, which allows us to capture the effect of party alienation.

\(^11\) Assuming that parties can directly select $\phi_P$ allows us to save on notation. Snyder an Ting (2002) detail how this level of discipline can result from selection processes that impose costs on candidates. Our results would be identical if we allowed parties to choose this cost function.
behavior”. Thus, even though our modeling choices cut through important realities, they capture a fundamental relationship between party discipline 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 discipline. Finally, as the applications we cover in Section 8 will make clear, the strategic use of candidate selection does play a central role in shaping intraparty discipline.

**Timing.** We consider the following timing:\textsuperscript{12}

- $t = 1$: party leaders $L$ and $R$ select their national platforms, $x_L$ and $x_R$.
- $t = 2$: party leaders select intraparty discipline, $\phi_L$ 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 level of legislative cohesion $\lambda$. The socioeconomic environment is captured by the heterogeneity of voter preferences, $\sigma$.

Legislative cohesion is known to vary substantially across political regimes; it is typically higher when government survival depends on legislative support; see for example Huber (1996b) and Diermeier and Feddersen (1998). While these contributions focus on how legislative institutions impact on the cohesion of legislators in the absence of parties,\textsuperscript{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. We thus assume that, even in the absence of selection or party sticks, legislators are more cohesive in a parliamentary than in a presidential regime:

**Assumption 1** Institutional constraints, identified by the parameter $\lambda(\geq 0)$, determine the feasibility set of party discipline: $\phi_P \in [0, \lambda]$. More precisely, a majority can only operate if

\textsuperscript{12}Reversing the timing between periods 1 and 2 produces the same results.

\textsuperscript{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.
legislator preferences are within distance $\lambda$ of the coalition’s median preference. $\lambda$ is strictly smaller in a Parliamentary than in a Presidential regime.

Of course, beyond these institutional constraints, party leaders can further constrain party members. Since we focus on selection, parties can decide to only accept candidates who have a preference within any distance $\phi_P \leq \lambda$ of the party platform $x_P$. This assumes that imposing discipline is (costless and) independent of the political regime. In reality, imposing tight discipline is more difficult in a presidential regime, for example because the executive cannot dissolve the assembly. Section 7.2 introduces these differences in the model. They actually reinforce our results. The assumption that discipline is costless is thus only meant to clarify the fact that costly discipline is not at the core of our results.

Turning to the economic environment, McCarty et al. (2006, chapter 3) show that economic inequality typically maps into more polarized voter preferences. Castanheira, Crutzen and Sahuguet (2010) 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.\footnote{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 thus use only one parameter to proxy the heterogeneity of both ideological preferences and income inequality across districts. We introduce this parameter as follows:

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

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

in which $\sigma$ proxies preference and income heterogeneity across districts.

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

4 How Does Discipline Impact on Voting? (time 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 \textit{ex ante} identical in the eyes of a voter. The district median’s expected
utility from electing any local independent is:

\[
E u(y_i, x_c|x_c \in \mathcal{Y}_i) = E_{x_c \in \mathcal{Y}_i} \left[-(y_i - x_c)^2\right] = \int_{y_{i-1}}^{y_{i+1}} -(y_i - x_c)^2 \ f(x_c) \ dx_c
= -(y_i - y_i)^2 - 1/3 = -1/3.
\]

Voters have more information about party candidates: first, given that he 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, he 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:\(^{15}\)

\[
\mathcal{P}_i(x_P, \phi_P) \equiv \mathcal{Y}_i \cap \mathcal{X}_P.
\]

It follows that the median voter’s expected utility from electing a candidate of party \(P\) is:

\[
E_i u(y_i, x_c|x_c \in \mathcal{P}_i(x_P, \phi_P)) = -(y_i - \mu_i[x_P, \phi_P])^2 - \sigma_i^2[x_P, \phi_P], \quad (2)
\]
where, by the properties of uniform distributions:

\[
\begin{align*}
\mu_i[x_P, \phi_P] &= \frac{\max[y_i - 1, x_P - \phi_P] + \min[y_i + 1, x_P + \phi_P]}{2}, \\
\sigma_i^2[x_P, \phi_P] &= \frac{(\max[y_i - 1, x_P - \phi_P] - \min[y_i + 1, x_P + \phi_P])^2}{12}. \\
\end{align*}
\]

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 \(x_P\):

- **Define close districts** as the set of 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 the set of districts 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.

This is illustrated in Figure 1.

---

\(^{15}\)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 \(\mathcal{Y}_i\).
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 tighter levels of discipline: this is the *variance-reduction effect* of party discipline. In distant districts instead, tighter discipline also implies that the expected distance between a party candidate and the district median $y_i$ increases. This is the *legislative freedom effect* of party discipline, which reduces the expected utility of electing a party candidate. 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 hump-shaped in discipline. Yet, as the relevant comparison for voters in any district is between the utility from voting for the party candidate and that from voting for the independent, distant districts unambiguously prefer minimal discipline. 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.

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 depends on the party position $x_P$ and on party discipline $\phi_P$:\(^{16}\)

\(^{16}\)Omitted proofs are in the appendix.
**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 $x_P$:

$$E_i u(y_i, x_P) \geq E_i u(y_i, x_I) \iff |y_i - x_P| \leq \kappa(\phi_P),$$

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:** Party $P$’s catchment area depends on its internal structure

Proposition 1 and Figure 2 show how party discipline maps into electoral support. As we said above, intermediate levels of discipline 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 party disciplines minimizes 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 discipline. For relatively low levels of candidate freedom, $\phi_P < 1/2$, the party catchment area contains close districts only. To expand its catchment area, the party thus benefits from further disciplining its candidates, to cash in on the variance-reduction effect of the party label. A local maximum is found when discipline is maximal ($\phi_P = 0$). This is the bottom
part of the figure. By contrast, for relatively high levels of candidate freedom, $\phi_P > 1/2$, the marginal district is distant. In this case, the party has an incentive to further reduce discipline: this increases utility in the marginal district and induces the next district to also prefer the party candidate. The global maximum is found when there is full candidate freedom ($\phi_P = 1$).

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 discipline 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 7 below, in which we discuss how generalizing our assumptions impacts on our results.

5 Equilibrium discipline (time 2)

At time $t = 2$, parties choose their level of intraparty discipline to maximize their seat share, taking as given the ideological positions 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:

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

where $1[\cdot]$ 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 party discipline 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. Then, all districts in party \(P\)’s catchment area prefer the candidate of party \(P\) to that of the other party. Since, by Proposition 1, full discipline maximizes the size of the catchment area, \(\phi_P = 0\) 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. The level of discipline that maximizes the number of seat won in outer districts is still \(\phi_P = 0\). What about centrist districts? Given the institutional constraint \(\phi_P \leq \lambda \leq 1/\sqrt{3}\), 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 discipline. Thus, any district that may potentially elect a candidate of party \(P\) prefers maximal discipline, 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 discipline, 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 intraparty discipline in equilibrium.

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

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

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 intra-party discipline. They reveal a hierarchy of incentives:
1. first, if institutional constraints are tight, parties choose maximal intraparty discipline, irrespective of other considerations. If institutional constraints are loose, then parties face a more complex trade-off. When platforms are highly polarized, parties avoid direct competition. Their primary target is then to maximize the size of their catchment area, which requires granting maximal freedom to their candidates. This is the multiplier effect of party discipline: small institutional changes (\( \lambda \) being slightly smaller or greater than \( 1/\sqrt{3} \)) produce substantially different levels of party discipline (Proposition 2 versus Proposition 3). Section 8 documents the dynamics of party discipline in the UK and in France and shows how a progressive fall in \( \lambda \) did translate into a disproportionate increase in party discipline.

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

6 Equilibrium platforms (time 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 necessarily enforce maximal discipline at time 2. The parties’ vote shares can then be expressed as:

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

(4)

Our next proposition identifies the equilibrium platforms:

**Proposition 4** For \( \lambda \leq 1/\sqrt{3} \), parties always adopt full discipline (\( \phi_P = 0 \)) and the pair of

\[17\] See below for a justification of why platforms are symmetrically distributed around the median.
manifestos is:

$$(-x_L = x_R =) \ x = 0, \ for \ \sigma^2 < 1/(6 \log 2);$$
$$= \sigma \sqrt{2 \log 2} - \sqrt{1/3}, \ for \ \sigma^2 \in [1/(6 \log 2), 2/(3 \log 2)];$$
$$= 1/\sqrt{3}, \ for \ \sigma^2 > 2/(3 \log 2). \quad 18$$

Hence, the median voter theorem holds only for a sufficiently homogeneous polity.

Proposition 4 and Figure 3 show that the two parties choose the median voter’s preferred platform 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 discipline (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 \( f (x_L - 1/\sqrt{3}) \, dx_L \) 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 \( \text{half} \) 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.\(^{19}\)

6.2 Case 2: loose institutions

If institutional constraints are loose, platform choices at stage 1 can affect intraparty discipline at stage 2. Equilibrium platform positions are thus the result of more elaborate strategic considerations. We have:\(^{20}\)

**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 candidate freedom \( (\phi_P = \lambda) \) in equilibrium. In particular, \( \sigma_B (\lambda = 1) = \sqrt{2/ \log 2} \).

\(^{19}\)This also implies that asymmetric equilibria will 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 \( 2\lambda = 2/\sqrt{3} \).

\(^{20}\)Like in the previous case, when \( \sigma \) is large, there exists a neighborhood around the symmetric pair of platform positions where parties can locate. Yet, these equilibria are symmetric insofar as party discipline is concerned and polarization \( (2\lambda) \) is unaffected.
ii) there exists $\sigma_T \equiv 1/\sqrt{\log 2}$ such that $\sigma \leq \sigma_T$ is a sufficient condition for parties to choose centrist platforms $(x_L, x_R) = (0, 0)$ and maximal discipline $(\phi_P = 0)$ 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 discipline or maximal flexibility. 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 stage 2 – remember this is achieved by maximizing candidate freedom. To reach the subgame in which they can take full advantage of candidate freedom, parties must take action at stage 1. Choosing polarized platforms is used for that purpose: it prevents direct competition and sustains maximal flexibility at stage 2.\(^{21}\)

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 stage 1, and maximizing the signalling content of the party label at stage 2 – remember this is achieved by maximizing discipline. Interestingly, this implies that the 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 “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.

\[\textbf{Table 1.} \text{ Summary of the main results.}\]

\(^{21}\)Note that the timing of the game could be reversed without affecting this result. If parties first chose their level of discipline, they would select maximal flexibility at stage 1 as a way to sustain polarization at stage 2.
Preference heterogeneity:

<table>
<thead>
<tr>
<th>Institutional constraints:</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tight:</strong> ( \lambda \leq 1/\sqrt{3} )</td>
<td>( \sigma \leq 1/\sqrt{6\log 2} ) ⇒ Centrist platforms: ( x_P = 0 )</td>
<td>( \sigma^2 &gt; 2/(3\log 2) ) ⇒ Moderate polarization: ( x_R - x_L = 2/\sqrt{3} )</td>
</tr>
<tr>
<td></td>
<td>Maximal discipline: ( \phi_P = 0 )</td>
<td>Maximal discipline: ( \phi_P = 0 )</td>
</tr>
<tr>
<td><strong>Loose:</strong> ( \lambda &gt; 1/\sqrt{3} )</td>
<td>( \sigma \leq 1/\sqrt{6\log 2} ) ⇒</td>
<td>( \sigma &gt; \sigma_B (\lambda) ) ⇒</td>
</tr>
<tr>
<td></td>
<td>Centrist platforms: ( x_P = 0 )</td>
<td>High polarization: ( x_R - x_L = 2\lambda )</td>
</tr>
<tr>
<td></td>
<td>Maximal discipline: ( \phi_P = 0 )</td>
<td>Maximal freedom: ( \phi_P = \lambda )</td>
</tr>
</tbody>
</table>

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 high discipline. 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 switch from their current, low-discipline, organization to one that would mirror the organization 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 discipline when preferences are sufficiently heterogeneous. Polarization is larger when institutions are “looser” precisely because of low discipline. In all cases indeed, the maximal extent to polarization is determined by the tangency of the parties’ catchment areas. Low discipline being an instrument to widen the parties’ catchment areas, it is also the driver of stronger polarization. Surprisingly, 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 Discussion and Extensions

7.1 Preferences of Candidates and Voters

In this section, we show that the two assumptions that (a) candidate preferences are uniformly distributed and (b) voters have quadratic preferences are not necessary for our results to carry through. As will become clear below, they are nonetheless useful to obtain closed form solutions.

Generalizing our setup, suppose that voter preferences are defined by some function \( f \) that
is single-peaked and displays (weak) 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 \( \mathcal{Y}_i = [y_i - 1, y_i + 1] \). 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:

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

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:

\[ U_{iP}(x_P, \phi_P) \equiv \mathbb{E}u(y_i, x_c | \mathcal{P}_i(x_P, \phi_P)) = \int_{x_P-\phi_P}^{x_P+\phi_P} u_i(x) \, g_iP(x) \, dx \text{ in close districts, and} \]

\[ = \int_{x_P-\phi_P}^{y_i+1} u_i(x) \, g_iP(x) \, dx \text{ in distant districts.} \]

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

\[ U_{iP}(x_P, \phi_P) \geq U_I \iff |y_i - x_P| \leq \kappa \]

20
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 strong discipline ($0 \leq \phi_P < 1/2$) and one with maximum candidate freedom ($\phi_P = \lambda$, 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 which $\kappa(\cdot)$ is maximized will be different from 0 and that the value of the expected utilities may not feature tractable closed-formed solutions.

### 7.2 Regime-Specific Restrictions on $\lambda$\(^{22}\)

At the turn of the Twentieth century, American politics underwent what Ranney has described as “the most radical of all the party reforms adopted in the whole course of American history” (Ranney, 1975, p121, quoted by Ware 2002, pp1 and 95) with the introduction of the direct primary for all elected offices (except that of the President of the United States). This reform de facto reduced party control on the candidate selection process.\(^{23}\)

To introduce this loss of control in the model, suppose that parties can only select a value of candidate freedom $\phi_P \in [k, 1]$ with $k > 0$ under the direct primary. The obvious consequence is that party leaders value even less the maximal level of feasible discipline, since the size of the catchment area under maximal discipline $k$ is bound to be smaller than under full discipline:

$\kappa(k) < \kappa(0)$ 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

\(^{22}\)We thank Tom Cusack for drawing our attention on this point.

\(^{23}\)See for example Ware, 2002 for a very clear account of the introduction of direct primaries and Castanheira, Crutzen and Sahuguet (2010) for a theoretical account that rationalizes its introduction.
of a larger value of $\lambda$; it also reduces the party leaders’ capacity to tighten discipline, given the constraints imposed by the direct primary legislation.

8 Applications

This section shows how our theoretical results apply in practice. The first subsection focuses on the evolution of party discipline: we describe evolutions in the UK, the US, and France. It illustrates how institutional changes in the parliamentary regime led to more party discipline (the multiplier effect). In contrast, US parties took steps to decentralize candidate selection. The second subsection focuses on party polarization in the US and provides a theoretical rationale for the dance between polarization and inequality uncovered by McCarty et al (2006).

8.1 Comparative Politics of Intra-Party Discipline

8.1.1 The Evolution on Intraparty Discipline in Victorian England

Contemporary British voters typically vote for a party: 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 identity of their candidate. Cox (1987) 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” (Cox, p51) – and dates the switch around 1868 (p92). What we argue here is that the institutional changes that occurred between 1832 and 1868 also caused the switch in party discipline – 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.

How the nineteenth century electoral Reform Acts led to the birth of Britain’s “efficient secret” is one Cox (1987)’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 contraints (\(\lambda\) falls progressively in our model). Eventually, this induced parties to increase discipline and tighten candidate selection because “the [party] labels themselves became increasingly important as collective symbols [...] since 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” (Cox 1987, p144).

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’” (Cox, p59). The result was a drop in parliamentary cohesion (see Cox, p23, 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’. “ (p60).

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” (p61). The Cabinet took increasing control of the Parliament’s agenda and, by approximately 1850, “came about the efficient secret” (p51). 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” (p85). In this case a motion of confidence should actually reduce discipline. Thirdly, if the threat of dissolution was the sole determinant of discipline, cohesion should have been decreasing as the date of the election neared. Yet, as shown by Cox (p86), this pattern is not in the data.

The multiplier effect identified in Propositions 2 and 3 can instead explain why discipline increased the way it did. We can decompose the increase in party discipline into two steps. First, until about 1870-1880, discipline 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 party discipline (the multiplier effect in the model), largely through candidate selection.

In the first phase, parties used “whipping” to increase discipline: 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 (Cox, Table 3.5, p 24). The effectiveness of whipping on actual

---

24Whips are MPs (or Lords) appointed by each party to make sure that MPs vote the way the party wants.
cohesion 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: “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” (Cox, p143). Indeed, between 1856 and 1868, voters progressively switched from voting for candidates to voting for parties (Cox, Chapter 9). This gave parties the necessary leverage to finally obtain the additional zest of discipline 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” (p78), instead of following the “riskier course” of criticizing the government with the hope of being bought off. This strategy bears its disciplining fruits from 1890-1900 (p79). 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” (p144).

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, i.e. 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” (p136).
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. In contrast with the UK, however, the personal vote is still very important today (see for example Cain, Ferejohn and Fiorina 1984 and Morgenstern and Swindle 2005), and parties do not closely discipline their congressmen and senators. What we show here is that this lack of party discipline is not caused by some type of disciplining impossibility. Historical evolutions clearly show that, when faced with challenges similar to that of UK parties, US parties chose to loosen discipline. They ensured that local candidates would be more independent from Washington, to reinforce the party label and compete independent candidates away.

A cornerstone 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, p65). In other words, polarization ($\sigma$ in the model) increased in the electorate. Massification of politics also produced new informational asymmetries 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

---

25 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 largely linked to the importance of slavery in those States. See Besley, Persson and Sturm (2010) for more on this.
changed radically. [...] A style of politics that worked relatively well in the 1830s was working much less well in the new circumstances” (Ware 2002, p21).

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. 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 local candidates outside the hands of the party:26 the direct primary was approved 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, p121, quoted by Ware 2002, p95).27

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 that of the UK (increasing informational asymmetries and $\sigma$) by totally relinquishing their control over candidate selection to an independent bureaucracy, that is, by awarding local candidates as much freedom as allowed by the rules of the game $- \phi = \lambda$ in the wording of the model. Thus, while Victorian England saw the materialization of the efficient secret and high intraparty discipline, which epitomize parliamentary regimes, the US became the archetypical presidential regime centered around separation of powers and loosely disciplined parties.

8.1.3 Changes between the Third and Fifth Republic in France

Another interesting case is France.28 Under the Third and Fourth Republics (that is, until 1958), French deputies had the reputation of being extremely undisciplined and individualistic, and governments were seldom long-lasting.29 The goal of the reforms introduced by the Fifth

---

26 In Castanheira, Crutzen and Sahuguet (2010), we show when primaries improve the candidate incentives in a moral hazard setup.

27 The actual rules and degree of decentralization of the selection procedure regarding the direct primaries depend on the office for which the election is organized. The only election for which the direct primary does not apply is that for the President of the United States.

28 We thank Howard Rosenthal for drawing our attention to the French case.

29 Sauger (2009, p310) reports that “21 governments were formed within the 12 years of the Fourth republic, compared to only 19 in 50 years for the Fifth Republic. Moreover, most of the Fourth Republic governments were
Republic (in 1958) was to eliminate this instability by tightening legislative constraints, among others by introducing the package vote and the confidence vote (xxzzzz source). These reforms clearly increased the powers of the Prime minister and the President by either not giving legislators any say on proposed laws (the package vote requires legislators to vote in favor or against a proposed law without the right to propose any amendment) or by linking the survival of parliament to the outcome of a vote (through the confidence vote). In the terminology of our model, these reforms severely reduced $\lambda$.

How did parties react to this change? The reaction of the Gaullist UDR is the most interesting case to analyze here, for the following reasons: 1) this was the party whose elite had crafted the Fifth Republic; 2) it was the party with the lowest initial level of internal cohesion, because it had the most diverse set of legislator preferences in 1958; 3) it managed to maintain its control of both the Presidency and the Prime Minister position from 1958 to 1981, when Mitterrand won over Valery Giscard d’Estaing, partly because the latter lost his control of the right wing part of his party.

Even though the Gaullist leaders had crafted the Fifth republic, they quickly realized that they could not free-ride on its institutions, exactly as our model predicts. To the contrary, and in line with the multiplier effect put forth by this paper, after introducing the Fifth Republic in 1958, they built on this reform to tighten intraparty discipline from 1962 onwards:30 “UDR leaders [took] steps to enforce effective discipline: deputies who failed to observe party discipline were subject to immediate exclusion. [...] Even a single refusal of discipline on a key vote could bring expulsion” (Wilson and Wiste 1976, p482). The reaction of the UDR to the introduction of the Fifth republic thus provides a second example in which party discipline was introduced as a direct reaction to a change in the institutional environment. Finally, the Gaullists’ capacity to stay in power for such a long spell of time and the fact that their defeat in 1981 can at least partly be attributed to a significant decline in internal discipline also suggest that the mechanisms and effects the model focuses on are far from being of second order importance.

30 Sauger (2009) argues that 1962, which saw the only dismissal of a government—Pompidou’s—by a no-confidence vote in the Fifth Republic, was the key turning point for the increase in intraparty discipline and cohesion, see Sauger (2009, p. 313)
8.2 The American Polarization ‘Dance’

While Section 8.1.2 focuses on the dynamics of US party discipline in the beginning of the twentieth Century, it is silent about the effect of preference heterogeneity on party positions. Proposition 5 predicts that parties should polarize more when voter preferences become more heterogeneous. With low discipline, this implies that the overlap between the catchment areas of the two parties should decrease (respectively increase) when preference heterogeneity increases (respectively: decreases).

This result provides a theoretical foundation for the findings of McCarty, Poole and Rosenthal (2006) who identify a strong statistical correlation between party polarization and economic inequality in the US. A first visualization of these evolutions is provided by Figure 4 below: it displays the distribution of preferences of the US House representatives in the 93rd and 108th Houses.

Two facts stand out: first, policy preferences within each party form a cloud (roughly around the national party platform), consistently with our theory and contrary to usual Downsian models. Second, the overlap between the two sub-distributions (the catchment areas in the wording of the model) is greater in the 93rd House than in the 108th.\(^{32}\) If our theoretical predictions bear

\(^{31}\)The figures are taken from http://www.ou.edu/special/albertctr/extensions/fall2005/Poole.pdf

\(^{32}\)This is especially evident when one compares across the two Houses the left tail of the Republicans and the right tail of the Democrats.
some water, income inequality—which is a good proxy for preference heterogeneity—should be greater in 2003-04 than in 1973-74. This is indeed what McCarty et al. document. Indeed, they provide a wealth of evidence that shows that there is a systematic tendency for the distribution of legislator preferences to move closer or farther from one another in correlation with income inequality. Figure 5 below indeed depicts vividly what they call the “dance” between inequality and ideological polarization over the course of the XXth century.\footnote{We wish to emphasize that, while previous contributions, such as Callander (2005), may provide a rationale for the positive correlation between preference heterogeneity and the distance between party platforms, they cannot rationalize this evidence. Indeed, like in most Downsian models, parties are assumed to be a single ideological point in Callander (2005). In contrast, the measure of polarization put forth by McCarty et al. (2006) is based on all individual representatives’ liberty to express (and vote for) different positions inside the party. In other words, if there was full discipline, there would be no overlap between the two subdistributions in Figure 4, and polarization would always be maximal in Figure 5.}

![Figure 5: Income Inequality and Political Polarization](http://polarizedamerica.com/)

**Figure 5:** Income Inequality and Political Polarization (Source: http://polarizedamerica.com/)

9 Conclusion

Comparative studies of economic policy across political regimes implicitly rely on parties being highly disciplined in parliamentary regimes and highly flexible in presidential regimes. Yet, these studies systematically disregard parties, and thus cannot explain why parties adopt and maintain these different organizations. We proposed a model that fills this gap. We study an electoral game in which intraparty discipline and ideological platforms are endogenous. 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 discipline is determined both by institutional constraints and by population preference heterogeneity. In turn, discipline influences equilibrium polarization. Our results provide a rationale for the fact that U.S. parties are less centralized and more polarized than, for example, British parties. Our results also provide a novel rationale for why income inequality and the ideological positioning of US parties ‘danced’ together in the Twentieth century.

Where do we go from here? In a companion paper (Castanheira and Crutzen 2010), we exploit the results derived here to relate the internal structure of incumbent parties to the incentives for new parties to enter. This provides a novel rationale for the observation that Duverger’s Law is more likely to hold in US-type presidential regimes than in UK-type parliamentary ones.

The model in these papers falls short of providing predictions regarding public finance and institutional choices. Models of comparative politics along the lines of Persson, Roland and Tabellini (2000) make implicit assumptions about legislative cohesion within the ruling majority. We believe that intraparty discipline is an important institution that must be considered when thinking about the size and composition of public spending and taxation. Along the same lines, institutions are assumed exogenous in these models, as well as in ours. Yet, different party structures may develop different incentives for institutional reforms. For instance, in the U.S., candidate freedom is associated with candidate-centered campaigns. Our model could be used to capture the effect of electoral campaigns by letting the pool of candidate preferences be different across regimes, which may reduce the value of the party label. This in turn suggests that legislators face an incentive to select institutions that sustain, if not reinforce, the candidate-centered nature of campaigns. This clearly influences legislative cohesion and may call for other legislative institutions that reinforce the party label.

The analysis of the mapping from how parties organize to the electoral appeal of candidates could also be carried along different dimensions. For example, Castanheira, Crutzen and Sahuguet (2010) study a moral hazard game in which the competitiveness of the candidate selection procedure impacts on their incentives and is used by the voters to form expectations about
platform quality. They use these results to rationalize the emergence of the American direct primary at the beginning of the nineteenth century and the organizational changes observed in Western European parties since the 1960s.

References


32
10 Appendix: Proofs

Proof of Proposition 1

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

For $|y_i - x_P| \leq 1 - \phi_P$, $E_i u (y_i, x_P) > E_i u (y_i, x_I) \iff |y_i - x_P| \leq \sqrt{\frac{1 - \phi_P^2}{3}}$. (5)

For $|y_i - x_P| \geq 1 - \phi_P$, $E_i u (y_i, x_P) > E_i u (y_i, x_I) \iff |y_i - x_P| \leq \phi_P (> 1/2)$. (6)

(5) can be rewritten as:

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

$$= 1 - \phi_P, \forall \phi_P \geq 1/2.$$

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:

$|y_i - x_P| \leq \sqrt{\frac{1 - \phi_P^2}{3}}$ if $\phi_P \leq 1/2$

$\leq \phi_P$, if $\phi_P \geq 1/2$. 
For \( \phi_P \leq 1/2 \), all the districts within distance \( \sqrt{\frac{1-\phi_P}{3}} \) 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 \( \phi_P \in [0,1/\sqrt{3}] \). We thus need to prove that:

\[
E_i u(y_i, x_P|\phi_P = 0) = -d_{i,P}^2 > -\frac{1+(d_{i,P}-\phi_P)^2+(d_{i,P}-\phi_P)^2}{3} = E_i u(y_i, x_P|\phi_P > 0), \forall \phi_P, d_{i,P} < 1/\sqrt{3}. \quad (7)
\]

Rearranging this inequality yields:

\[
1 + (1 - 2\phi_P) d_{i,P} - 2d_{i,P}^2 - \phi_P + \phi_P^2 > 0. \quad (8)
\]

(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 tightest at the corner value: \( \phi_P = 1/\sqrt{3} \). Hence, \( \phi_P = 0 \) is preferred to any \( \phi_P \in (0,1/\sqrt{3}) \) if it holds in \( \phi_P = 1/\sqrt{3} \):

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

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 \((-x_L = x_R =) x = 1/2\):

\[
E_i u(y_i = 0, x = 1/2 | \phi_P = 0) = E_i u(y_i = 0, x = 1/2 | \phi_P = 1) = -\frac{1}{4}.
\]

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
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:

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

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 \( (x_L, x_R) = (-1/\sqrt{3}, 1/\sqrt{3}) \), we have:

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

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 \( \frac{\partial V_R(\phi_R = 0, x_R; x_L)}{\partial x_R} = 0 \) and similarly for \( L \). 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:

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

Solving this equation yields \( x_R^* = \sigma \sqrt{2 \log 2} - \sqrt{1/3} \). 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:

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

(9)

For any $x_R < x^*$, this derivative is always positive: by the properties of Normal distributions, $f(x_R + 1/\sqrt{3}) > f(x^* + 1/\sqrt{3})$ and $f \left( \frac{x_R + x^*}{2} \right) < f(0)$. 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 $L$:

$$\frac{\partial V_L(\phi_L=\lambda,x_R,x_L=\lambda)}{\partial x_L} = f(\frac{x_R + x_L}{2}) - f(x_L - \lambda) \leq f(0) - f(x_L - \lambda)$$

$$= \frac{1}{\sqrt{2 \pi \sigma}} \exp \left[ -\frac{1}{2} \left( \frac{x_L - \lambda}{\sqrt{2 \pi \sigma}} \right)^2 \right]$$

(10) for $\lambda = 1$.

(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 $V_L < F(\frac{x_L + x_R}{2}) - F(x_L - 1)$. 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 $F(1/\sqrt{3}) - F(-1/\sqrt{3}) \approx 0.226$. 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: $V_L(x_L = -1, \phi_L = 1) \approx 0.381 > 0.226$. 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 $\sigma \rightarrow \infty$, 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 \( \phi_R = \lambda (\geq 1/\sqrt{3}) \), 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:

\[
V_L(x_L = 0, \phi_L = 1) = 2 \left( F(1) - F\left(\frac{1}{2}\right)\right).
\]

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

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

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 4, 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 \( \mathcal{P}_i(x_P, \phi_P) \equiv \mathcal{Y}_i \cap X_P \) 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:

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

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:

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

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:

$$\frac{\partial U_{IP}(x_P, \phi_P)}{\partial \phi_p} \to \begin{cases} [u_i(x_P + \phi_P) - U_{IP}(x_P, \phi_P)] \cdot g_{iP}(x_P + \phi_P) + [u_i(x_P - \phi_P) - U_{IP}(x_P, \phi_P)] \cdot g_{iP}(x_P - \phi_P) & \text{if } \varepsilon \to 0 \\ <0 & g(1) >0 & g(0) \\ >0 & g(1) >0 & g(0), \\ \end{cases}$$

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

QED