Comparative Politics with Endogenous Intra-Party Discipline[1]

October 2009

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. Parties are "brands": they provide information about the preferences of their candidates. In equilibrium, discipline is high in parliamentary regimes, and low in presidential ones. Polarization increases with voter preference heterogeneity, and is higher in presidential regimes. This rationalizes the evolution of intraparty discipline in the U.S., the U.K. and France and rationalizes the positive correlation between polarization and income inequality observed in the U.S.

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

[1]We 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.[2] More recent reforms have not reduced these differences: currently, dissent between legislators and party leaders is the exception rather 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).[3]

More importantly, the above facts suggest that party discipline is *endogenous* to the political regime. Why and how does intraparty discipline adapt to the political regime? Does it matter for policy, for example for the choice of party platforms?

To answer these questions, we propose a novel model of elections in which parties choose both their ideological position and their degree of internal discipline. Parties are thus competing *organizations* that rely on district-specific candidates, each running for a seat.[4] We focus on a 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 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.

Intraparty discipline thus 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

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

[3] The 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 [...]."

[4] We thus concentrate on elections that are run under plurality rule in single-member districts.

local electorate. If discipline is low, voters can expect that their legislator will better represent local preferences, but the informational content of the party label is more limited. Voters also have an "outside option": they can vote for a local independent who is not bound to a party, but whose policy position is highly uncertain. These independents can be seen as potential entrants who limit the parties' ability to catch votes with any type of policy.

We build on the findings of Huber (1996b) and Diermeier and Feddersen (1998) to introduce the institutions of Parliamentarism and Presidentialism in the model. Abstracting from the role of parties, they find that Parliamentarism puts tighter limits on the freedom of individual legislators. We thus model Parliamentarism as a tighter constraint on the feasible levels of party discipline: the minimal level of intraparty discipline needed to operate in a parliamentary system is high, whereas it is lower a presidential regime.

Turning to our results, we find that the certainty-versus-flexibility trade-off has different implications for districts close to the party platform as opposed to more distant ones. The former always value high discipline: the benefits of the uncertainty-reduction effect of party discipline dominates the costs of limited legislative freedom. By contrast, from the standpoint of a distant district, party discipline only ensures that legislators will implement a "distant" policy. If discipline is "too high", these districts actually prefer the highly uncertain but on average closer independent candidate. Thus, high discipline is needed to win a close district, whereas low discipline is needed to win a distant district. 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 also induces them to avoid strong levels of polarization. In a presidential regime instead, parties can more easily escape direct competition. In that case, the equilibrium is such that parties polarize more, and maximize candidate freedom, to target distant districts. Parties need to compete directly only when preferences are very homogeneous across districts, in which case, they select close platforms 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 party structure: party leaders may want to switch from very low to very high discipline even when institutional changes are marginal. In Section 8.1, we

show that this party channel is fully consistent with the historical reports of Cox (1987), Wilson and Wiste (1976) and Huber (1996a) on the evolution of intraparty discipline in England and France. In Section 8.2, we show that our results provide a theoretical rationale for the empirical findings of McCarty, Poole and Rosenthal (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.

Snyder and Ting (2002) show how candidate selection allows party labels to become "brand names" that are valued by voters. In their model, candidates are not district-specific and platforms will typically be median unless the party brand signal is very weak. We extend Snyder and Ting (2002) in *(i)* letting both the party platform and the level of intraparty discipline be endogenously chosen by parties, *(ii)* letting candidates be district-specific and *(iii)* comparing party organizations across political regimes.

Close to both Snyder and Ting (2002) and this paper, Eyster and Kittsteiner (2007) show that parties may adopt extreme positions to reduce interparty competition. Yet, party discipline is exogenous in their setup. Conversely, Ashworth and Bueno de Mesquita (2008) focus on how incumbent party members choose intraparty discipline but they need exogenous party positions to perform their analysis. Thus, they abstract from the link between party structures and polarization, as well as from the role of institutions.

Building on Palfrey's (1989) sincere voting setup, Callander (2005) introduces multiple districts and rationalizes the positive correlation between polarization and inter-district heterogeneity. Yet, polarization would disappear in his setup if voters adopted a more strategic behaviour. Our approach complements his in considering that voters are sufficiently strategic to elect the Condorcet winner in their district. What is more, in Callander (2005) candidates and their

parties are one and the same thing: each party is represented as a point on the real line. By contrast, the measure of polarization constructed by McCarty et al. (2006) is based on the voting behavior of individual legislators. Hence, by construction, we cannot rely on Callander's results to have an appropriate theoretical rationale for the correlation documented by McCarty et al. (2006).

Finally, our paper 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.[5] For example, in Persson, Roland and Tabellini (2000), all players are unitary actors –no clear distinction is made between parties and their candidates– even though, through the assumptions made on the alignment of preferences between the executive and the legislative under a parliamentary and a presidential regime, they recognize that intraparty discipline is high in the former but low in the latter regime. To drive home our point on the importance of differences in intraparty discipline under different political regimes, imagine what would happen in a parliamentary democracy if the executive could not rely on a disciplined (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.

Turning to legislative discipline, Huber (1996b) rationalizes how an institution such as the vote of confidence procedure induces high discipline in parliamentary democracies. Yet, the focus of his analysis is on the bargaining game between the Prime Minister and an *exogenously* given supporting majority in Parliament. He does not analyze how the characteristics of the electorate shape the supporting majority in parliament, intraparty discipline or polarization. Diermeier and Feddersen (1998) rationalize differences in cohesion between *individual* legislators across parliamentary and presidential regimes. Yet, their model abstracts from the role of parties: these are absent from the analysis. In contrast, Katz and Mair (1992 and 1994) do focus on parties. They compare party organization in 12 Western democracies and show that it correlates with the political regime. Our model rationalizes their findings and extends them to show how the polity's socioeconomic characteristics influence the way parties organize.

[5] See for example Persson, Roland and Tabellini (1997 and 2000), Persson and Tabellini (1999, 2000, 2003, 2004a and 2004b), Lizzeri and Persico (2001), Milesi-Feretti, Perotti and Rostagno (2002), Persson, Tabellini and Trebbi (2003), and Acemoglu and Robinson (2006).

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: the policy of a given candidate is uncertain, and party discipline can reduce that 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, each electing one legislator under plurality rule. Once elected, each legislator controls a fraction of the decisions that are taken during the legislature, and implements her preferred policy.[6] This fraction is the same for party members and independents.[7]

Voters. The median voter of district i is always pivotal, which implies that the Condorcet winner in the district wins the election. His preferences are single-peaked and quadratic around $y_i \in \mathbb{R}$.[8] He votes for the candidate that offers the highest expected utility, given his beliefs about the preferred policy of the candidates running in the district. That is, the candidate winning in district i is the one that maximizes:

$$\mathsf{E}\left[u\left(y_i, x_c\right)\right] = \mathsf{E}\left[-\left(y_i - x_c\right)^2\right],$$

where E is the expectation operator on $x_c \in \mathbb{R}$, the preferred policy of candidate c.

Candidates. Each candidate's preferred policy position x_c is private information: alone, a candidate cannot reveal any information about her preferences. It is common knowledge that x_c is a realization from the uniform distribution on $\mathcal{Y}_i \equiv [y_i - 1, y_i + 1]$, which is thus district-specific (Section 7 generalizes the setup to more general distributions). This is meant to capture the fact that in many democracies candidates must reside in the district in which they run.[9]

[6] One should read this assumption differently when thinking about parliamentary systems: the random component is then about who, out of the elected representatives, will be selected as a cabinet member, with powers to set policy on the issues that fall within her portfolio.

[7] Giving more decision powers to party members would not alter the main insights of the model, as long as independents retain some power.

[8] This parametric form of the utility function generates tractable closed-form solutions. Section 8 generalizes the utility function.

[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 information asymmetry between voters and candidates. In the framework of citizen candidate models, the impact of information asymmetries is studied among others by Casamata and Sand-Zantman (2008) and Großer and Palfrey (2008). These papers however abstract from the role of political parties.

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* ϕ_P they grant to their candidates. These two dimensions of party policy are observed publicly. Note that 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 ϕ_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:

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

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

Timing. We consider the following timing:[11]

- $t = 1$: party leaders L and R select their national platforms, x_L and x_R.
- $t = 2$: party leaders select intraparty discipline, ϕ_L and ϕ_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 parameters that define the country's institutional and socioeconomic environment. The country's institutional environment is summarized by its level of *legislative cohesion* λ. The socioeconomic environment is captured by the *heterogeneity of voter preferences*, σ.

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). Given that these two contributions focus on how institu-

[10] Independents could also be seen as potential entrants. Their role in the model is only to provide an outside option for voters, which will allow us to capture the effect of party alienation. We use this modelling approach to follow Snyder and Ting (2002).

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

tions shape legislative cohesion independently of party discipline,[12] we can use their results to identify how a group of legislators would behave in the absence of a party structure – institutional constraints are thus exogenous in our model. Formally:

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 legislator preferences are within distance* λ *of the coalition's median preference, with* λ *being strictly smaller in a Parliamentary than in a Presidential regime. Party discipline can further constrain party members to have a preference within distance* $\phi_P \leq \lambda$ *of the party platform* x_P*.*[13]

Turning to the economic environment, McCarty, Poole and Rosenthal (2006, chapter 3) show that economic inequality typically maps into more polarized voter preferences.[14] Castanheira, Crutzen and Sahuguet (2009) 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. We thus need 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* σ*:*

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

in which σ *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

[12] 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.

[13] For convenience, we normalized the lowest value of ϕ_P to 0 across both political regimes. We discuss below how weakening this restriction would reinforce our findings.

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

with party R. Since voters cannot observe candidate preferences directly, all candidates within one of these sets are *ex ante* identical in the eyes of a voter. The district median's expected utility from electing any local independent is:

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

Voters have more information about party candidates: first, given that 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]

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

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

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

where, by the properties of uniform distributions:

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

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

- 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 $\mathcal{Y}_i$.

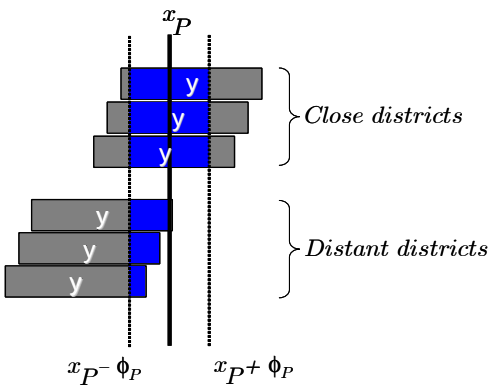

Figure 1: Set of possible platforms for party candidates.

In close districts, the expected position of a party candidate is x_P, independently of ϕ_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 ϕ_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:*

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

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.

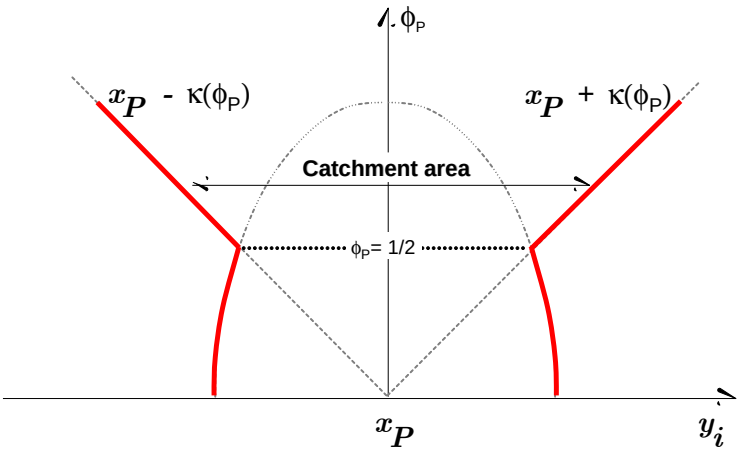

Figure 2: Party P's *catchment area* depends on its internal structure

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 ϕ_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:

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

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

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

5.1 Case 1: tight institutions

We have:

Proposition 2 *When institutional constraints are tight ($\lambda < 1/\sqrt{3}$), maximal 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 $\phi_P \leq \lambda \left(\leq 1/\sqrt{3}\right)$, these districts are at most at distance $1/\sqrt{3}$ from the party platform. The proof of Proposition 2 shows that these districts also prefer maximal 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 $\lambda > 1/\sqrt{3}$, ϕ_P^* depends both on party platforms and on the degree of preference heterogeneity σ:*
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 σ. 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: if institutional constraints are tight, parties choose maximal intraparty discipline, irrespective of other considerations. If institutional constraints are loose, then parties may face a more complex trade-off. When platforms are highly polarized, parties are never in direct competition for any seat. Their primary target is then to maximize the size of their catchment area, which requires granting maximal freedom to their candidates.

Parties face countervailing incentives if platforms are closer to one another. 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. Socioeconomic factors determine the relative benefit of these two strategies. 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:

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

Our next proposition identifies equilibrium platforms in the symmetric equilibrium:

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

$$\begin{aligned} (-x_L = x_R =)\ x &= 0,\ \textit{for } \sigma^2 < 1/\left(6\log 2\right); \\ &= \sigma\sqrt{2\log 2} - \sqrt{1/3}, \textit{for } \sigma^2 \in \left[1/\left(6\log 2\right), 2/(3\log 2)\right]; \\ &= 1/\sqrt{3},\ \textit{for } \sigma^2 > 2/(3\log 2).^{17} \end{aligned}$$

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

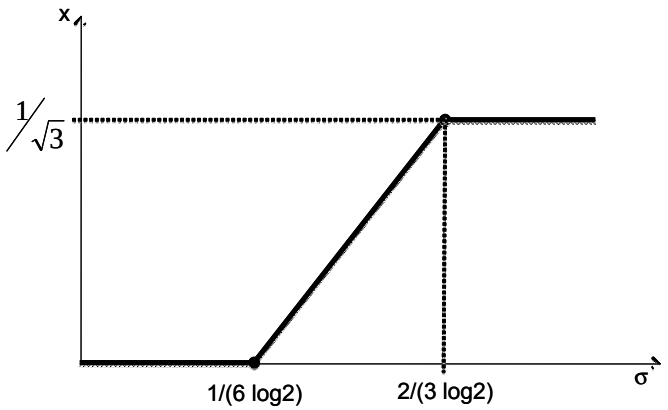

Figure 3: Polarization as function of preference heterogeneity

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 $f\left(x_L - 1/\sqrt{3}\right) dx_L$ seats from the outer left districts, and
2. the gain of $\frac{1}{2} f\left[\left(x_L + x_R\right)/2\right] dx_L$ seats from the centrist districts.

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

Quite clearly, 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 $f\left(y_i\right)$.

Corner solutions involve either full convergence to the median (when σ is sufficiently small) or maximal polarization (when σ 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 σ 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. This implies that asymmetric equilibria will also exist in the neighborhood (the size of which is increasing in σ) of the symmetric equilibrium.

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:[18]

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,* $\sigma_B(\lambda = 1) = \sqrt{2/\log 2}$.
ii) *there exists* $\sigma_T \equiv 1/\sqrt{6\log 2}$ *such that* $\sigma \leq \sigma_T$ *is a sufficient condition for parties to choose centrist platforms* $(x_L, x_R) = (0, 0)$ *and maximal 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.[19]

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

[18] Like in the previous case, when σ 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 (2λ) is unaffected.

[19] 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.

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 (λ small or large) and on whether preference heterogeneity is high or low (σ large or small). Table 1 below summarizes our findings.

Table 1. Summary of the main results.

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

Starting with the first column of the table, we see that institutions have little importance when preference heterogeneity is sufficiently small: independently of the institutional environment, parties want to be strong in centrist districts. This implies the choice of moderate platforms and 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 are seen to affect both polarization and party discipline when preferences are sufficiently heterogeneous. Importantly, 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. Surprizingly, this may also imply that independents (or, for that matter,

additional parties) are less likely to enter the political race when institutions are looser: despite the party label being less informative, parties manage to "cover" a larger part of the ideological spectrum.

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

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

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

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

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

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

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

This implies that:

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

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

Thus, like in the particular case of the uniform distribution and quadratic preferences, the party catchment area is necessarily a compact set centered on x_P. Clearly, the cutoff value κ is still a function of ϕ_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 ϕ_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 wich $\kappa(\cdot)$ is maximized will be different from 0 and that the value of the expected utilities may not feature tractable closed-formed solutions.

7.2 Regime-Specific Restrictions on λ[20]

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"

[20] We thank Tom Cusack for drawing our attention on this point.

(Ranney, 1975, p121, quoted by Ware 2002, pp1 and 95) with the introduction of the direct primary for all elected offices (but that of the President of the United States). This reform de facto reduced parties' control on the candidate selection process.[21]

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 of a larger value of λ; it also reduces the party leaders' capacity to tighten discipline, given the constraints imposed by the direct primary legislation.

8 Applications

The first application relates to the comparison of party structures across political regimes and to the determinants of the evolution of intraparty discipline in England and France. The second application focuses on party polarization in the U.S. 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 Parliamentary versus Presidential Regimes

The contributions by Huber (1996b) and Diermeier and Feddersen (1998) identify a potent effect of the vote of confidence procedure on legislative cohesion, even in the absence of parties. Our analysis identifies a *multiplier effect* that operates through party structures: a marginal tightening of the institutional constraints on legislative freedom can trigger a discrete switch in party structure, whereby legislators are deprived of their freedom of action and forced to stick to the party line. Formally, we identified a critical value for λ, $1/\sqrt{3}$ in the main model, where the trigger operates.

We believe that this party channel complements and reinforces the rationale advanced by Huber and by Diermeier and Feddersen for the observed correlation between institutions and

[21] 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.

legislative cohesion. They show that UK-style Parliamentary regimes are associated with significantly stronger legislative discipline than US-style Presidential regimes. If all the effect could be driven by institutions outside the party, then parties could afford to loosen discipline in parliamentary regimes, to free-ride on such external institutions. In contrast, our model predicts that parties will instead adopt measures that amplify initial differences in institutional constraints. Thus, even minor differences in institutions can trigger widely different cohesion patterns, both inside a given party and between parties of a coalition.

How do parties actually achieve discipline? First, they produce legislation that constrain the candidates' freedom, for example through restrictions on the freedom to levy personal finance. In many parliamentary countries, by law, the bulk of campaign finance must go through the party endowment, and is often public money (see e.g. Katz and Mair 1995 and 2002). By contrast, most campaign finance in the U.S. is levied by the candidates. This makes them much more independent from their party and amplifies their liberty to pander to their constituency. Second, which candidate personalizes the party in each district can either be controlled by the party or by local primaries that fall outside the parties' span of control (as in the US, since the inception of the American Direct Primary in the early 20th century – see Ware 2002). Third, parties can develop means to expel candidates who do not follow the party line sufficiently closely. In the UK, even prominent figures such as Ken Livingstone can be expelled. In the US instead, "there are virtually no sanctions for breaking party ranks" (Katz and Mair, 1994, p40). All these evolutions are endogenous to the party dynamics and fit the model's predictions.

8.1.2 The Evolution on Intraparty Discipline in Victorian England

Of course, many other dimensions differentiate US and UK politics. To isolate the effects of legislative constraints, we must identify a single country in which institutional constraints changed, and check how intraparty discipline evolved. The UK is a first case in point, analyzed by Cox (1987).[22] Consistently with the key finding of our model, he argues that the gradual increase in the level of intraparty discipline in the second half of nineteenth century Victorian England materialized *after* a few important changes to the rules of the political game, two of which are key. First, the passing of the three Reform Acts greatly increased the competitiveness and transparency of electoral races. This raised in turn the incentives of the members of parliament (MPs) to adopt a legislative behavior that was both more visible and more in line with the in-

[22] On top of analyzing how the introduction of institutions such as the vote of confidence impacted on intraparty discipline, Cox (1987) analyzes the extention of the franchise and the importance of changes in the behavior of the electorate.

terests of their constituency (the increase in the number and circulation of newspapers and the invention of the telegraph played a major role in this development too; see Cox 1987, pp.13-15). Second, the materialization of what Walter Bagehot (1865) coined the "efficient secret of the English Constitution " namely "the close union, the nearly complete fusion, of the executive and legislative powers" (Cox 1987, p. 51) with the cabinet as the key connection made it accepted practice to *link the survival of executives or even of Parliament to the passing of crucial government bills*. In short, institutional constraints such as the vote of confidence procedure, that are typical of modern parliamentary systems, materialized around the middle of the nineteenth century (Cox 1987, pp.80-87). These two developments thus put the government and the MPs on a collision course, with the government needing the support of a majority of legislators to operate and individual MPs wanting the legislature to devote more time to issues linked to their individual constituencies. The solution to this dilemma was quickly found by the government: increase intraparty discipline.

Indeed, the data on intraparty discipline reported by Cox (1987,pp.21-31) suggest forcefully that parties increased internal discipline *after*, as a reaction to the developments we described above: the data show a marked increase in discipline from the late 1860s onwards, that is, after the reforms had modified the rules of the political game. The data thus suggest that, instead of free-riding on the vote of confidence procedure, party leaders exploited the threat of dissolution of Parliament after a defeat on any important vote to discipline their troops. They also started refusing to grant dissidents the right to use the party label at the election. On top of this, because party leaders firmly controlled the cabinet, they could increase discipline through the strategic allocation of future ministerial positions (Cox 1987, pp. 77-79).

8.1.3 The Evolution on Intraparty Discipline in France after World War II

Another interesting case is France.[23] Under the Third Republic (i.e. until 1940), French deputies had the reputation of being extremely undisciplined and individualistic. After World War II, the reforms of the Fourth and Fifth Republic (in 1946 and 1958 respectively) went in the direction of a tightening of legislative constraints. In particular, these reforms increased the powers of the Prime minister. Yet, this alone did not create sufficient discipline (see Wilson and Wiste 1976). The second reform game increased powers to the President, and gave political parties and "groups" a new, more prominent, role.

All parties did not react at the same time nor in the same way to these institutional changes.

[23] We thank Howard Rosenthal for drawing our attention to the French case.

The Gaullist UDR is the most interesting case for our purpose: this was the party with initially the lowest internal cohesion because it had the most diverse set of legislator preferences in 1958. With the Fifth republic, thanks to the enhanced powers of party groups, this party could eventually develop ways to improve candidate discipline: "*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). Other parties such as the Communists were instead unified against the Gaullists. This made them cohesive even in the absence of party whips.

The French case thus provides a second example where, party discipline was undoubtedly introduced as a direct reaction to a change in the institutional environment. What is more, we see that a hybrid constitution (the Fourth Republic) produced a highly unstable situation, because the Parliament had enlarged powers but parties could not sufficiently discipline their candidates. According to our analysis, parties are indeed weakest under intermediate levels of discipline. In turn, this provides another rationale for the reform of the Fifth Republic: improved legislative cohesion was needed.

8.2 The American Polarization 'Dance'

A second prediction of our model is that equilibrium party polarization is commensurate to voter preference heterogeneity. We believe that 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 U.S. Their initial finding can be summarized by Figure 4, which is borrowed from their internet site, *http://polarizedamerica.com/*.

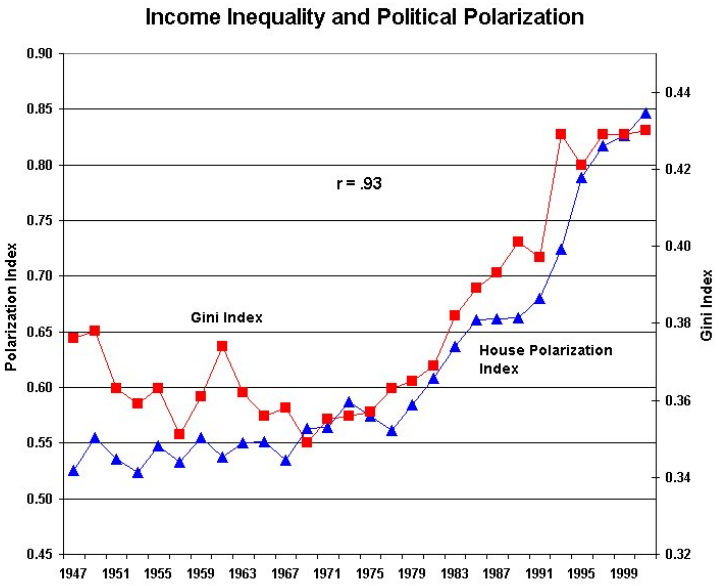

Figure 4: Income Inequality and Political Polarization

This figure illustrates what they call the "dance" between inequality and ideological polarization over the course of that century. We wish to emphasize that previous contributions, such as Callander (2005), do provide a rationale for the positive correlation between income inequality and the distance between party platforms – $2x$ in our model – but cannot rationalize the evidence put forth by McCarty et al. (2006), for the following reason. The measure of polarization that McCarty et al. offer is based on the fraction of *individual legislators* who vote with legislators of their own party as opposed to legislators of the other party. That is, in the terminology of our model, their measure of polarization increases when the candidates' bliss points overlap less across parties. A condition for this measure to operate meaningfully is thus that candidates within a party have sufficiently heterogeneous preferences and that discipline be low. One thus needs a model that captures both the equilibrium distance between the national party platforms and the equilibrium degree of intraparty discipline to rationalize such findings. This is precisely what our model achieves: when institutions are sufficiently loose *and* preference heterogeneity is sufficiently high, parties minimize discipline. The overlap between their catchment areas then reproduces the overlap of legislator preferences, and will decrease in preference heterogeneity.

The following figure on polarization in the US House of representatives provides indirect evidence of this mechanism.[24] The figure suggests quite clearly that the ideological overlap was higher in the 93d Congress than in the 108th one, in line with the sharp increase in inequality

[24] The evidence is indirect for the following two reasons. First, this picture captures probably two effects: changes in polarization and changes in the width of the ideologic spectrum. Secondly, what is mea-

between 1973 and 2003.[25]

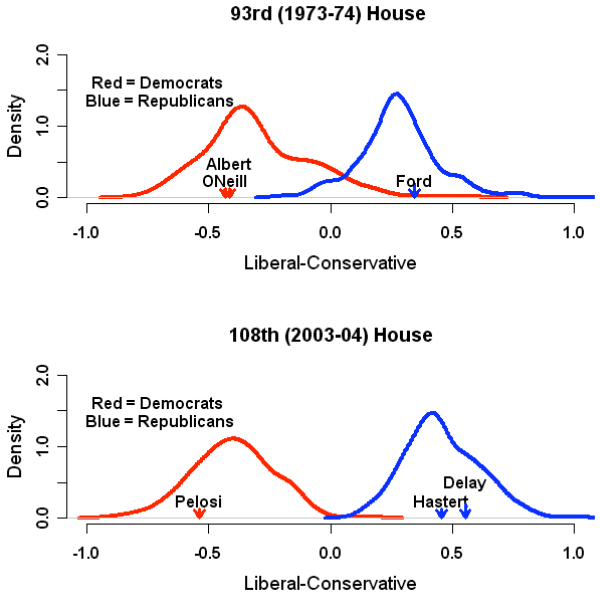

Figure 5: Democrat-Republican polarization dance and overlaps

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

sured actually is not the polity's but, rather, the individual legislators' ideology. The figures are taken from http://www.ou.edu/special/albertctr/extensions/fall2005/Poole.pdf

[25] This is especially evident when one compares across the two Houses the left tail of the Republicans and the right tail of the Democrats.

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 in come 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

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

[2] Ashworth, Scott and Ethan Bueno de Mesquita (2008). "Informative Party Labels with

Institutional and Electoral Variation". *Journal of Theoretical Politics*, 20: 251-273.

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

[4] Callander, Steven (2006). "Electoral Competition in Heterogeneous Districts". *Journal of Political Economy*, 113: 1116-1145

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

[6] Castanheira, Micael and Benoît S.Y. Crutzen. "Intraparty Discipline, Duverger's Law and Public Finance". in preparation

[7] Castanheira, Micael, Benoît S.Y. Crutzen and Nicolas Sahuguet (2010). "Party Organization and Electoral Competition". Forthcoming in the *Journal of Law, Economics and Organization*

[8] Castanheira, Micael, Benoît S.Y. Crutzen and Nicolas Sahuguet (2009). "The impact of party organization on electoral outcomes". Mimeo, Erasmus University Rotterdam

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

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

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

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

[13] Duverger, Maurice (1954). *Political parties: Their organization and activity in the modern state.* London: Methuen.

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

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

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

[17] Huber, John D. (1996a). Rationalizing Parliament: Legislative Institutions and Party Politics in France (Political Economy of Institutions and Decisions). Cambridge University Press

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

10 Appendix: Proofs

Proof of Proposition 1

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

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

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

(5) can be rewritten as:

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

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

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

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

For $\phi_P \leq 1/2$, all the districts within distance $\sqrt{\frac{1-\phi_P^2}{3}}$ of the platform x_P vote for the party. This distance is decreasing in ϕ_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 ϕ_P of x_P vote for the party. QED

Proof of Proposition 2

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

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

Rearranging this inequality yields:

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

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

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

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

Proof of Proposition 3

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

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

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

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

Lemma 2

The following lemma will help us prove Proposition 4.

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

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

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

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

Proof of Proposition 4

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

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

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

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

Conversely, for $\sigma^2 < 2/(3\log 2)$, the first order necessary condition for a pair of platforms $x_L < 0 < x_R$ to be an equilibrium is that $(\frac{\partial V_R(\phi_R=0, x_R; x_L)}{\partial x_R} =) f\left(x_R + \frac{1}{\sqrt{3}}\right) - f\left(\max\left(x_R - \frac{1}{\sqrt{3}}, \frac{x_L+x_R}{2}\right)\right) = 0$. Given that a similar condition must hold for the other party and that the distribution of district medians is symmetric around 0, the two first order conditions imply that we must have $x_L^* = -x_R^*$ in equilibrium, that is, platforms must be symmetric around 0. Exploiting this fact, the first order condition boils down to: :

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

Solving this equation yields $x_R^* \equiv \sigma\sqrt{2\log 2} - \sqrt{1/3}$. Of course, $x_R^* > 0$ requires that $\sigma^2 > 1/\left(6\log 2\right)$. For lower values of σ^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 $\sigma^2 \in \left[1/\left(6\log 2\right), 2/(3\log 2)\right]$, and $x_L = -x^*$, we have:

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

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

QED

Proof of Proposition 5

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

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

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

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

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

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

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

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

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

If $\phi_R = 0$, we have:

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

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

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

Since $V_L(x_L = 0, \phi_L = \lambda)$ is yet smaller for other values of λ, 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 ϕ_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 ϕ_P constant, a similar marginal change in x_P has two effects. It reduces the expected distance between x_c and y_i, which increases expected utility. On the other hand, it increases the variance of x_c, since the length of the subset $\mathcal{P}_i(x_P, \phi_P) \equiv \mathcal{Y}_i \cap \mathcal{X}_P$ increases; this decreases expected utility. The total effect on expected utility is thus ambiguous, and a direct comparison of U_I and U_{iP} is needed. Given that:

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

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

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

QED

10.2 Proof of Proposition 6

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

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

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

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

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

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

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

QED