Endogenous Party Discipline with Variable Electoral and Legislative Institutions$^1$

Scott Ashworth$^2$        Ethan Bueno de Mesquita$^3$

First Version: October 21, 2003
This Version: January 14, 2004

$^1$We received helpful comments from seminar audiences at Princeton and Cornell.
$^2$Department of Government, Harvard University. Email: ashworth@fas.harvard.edu
$^3$Department of Political Science, Washington University. Email: ebuenode@artsci.wustl.edu
Abstract

We study a simple model of party formation in which both party discipline and inter-party ideological heterogeneity are jointly determined. Discipline benefits party members because it gives risk-averse voters more confidence in the ideological composition of the party, but this discipline is costly to members who win office. Equilibrium is determined by balancing these forces. We show that this model can account for both comparative facts about parliamentary and presidential systems, and for changes over time in the U.S. Congress.
There is substantial variation in party discipline/cohesion, both across nations and over time. U.S. Congressional party leaders have greater formal authority over legislation than they had prior to 1970 (Rohde [25]), and roll call voting is more cohesive within congressional party caucuses today than before 1970 (Aldrich [1] and Rohde [25]). Comparative research has shown that parties are more cohesive in Westminster systems, such as that used in the U.K., than in Presidential systems like the U.S.’s (Cain, Ferejohn, and Fiorina [9]).

What causes this variation? A clear answer for the cross-country differences is that different institutions (such as Parliamentary vs. Presidential systems) give different incentives for parties to impose discipline and maintain cohesion. However, a satisfactory theoretical explanation should cover the within-constitution variation (such as the over-time trends in the U.S.) as well as this comparative variation. This task requires something beyond a purely institutional account. We propose a formal model in which party discipline is chosen strategically by parties, and show that it can account for both the comparative and historical evidence.

Our formalization also helps to resolve a puzzle about the relationship between party discipline and constituency service activities (e.g., securing geographically targeted spending projects). Traditional political science thinking links weak legislative parties with strong incentives for legislators to engage in behavior that benefits their local constituency. This link is based on the idea that differences in party discipline/cohesion are ultimately caused by the effect of institutions on the tradeoffs voters are willing to make between the party affiliation and individual reputations of candidates. The more institutions encourage a focus on party label, the argument goes, the stronger the parties and the lower the level of particularistic behavior by legislators. We will show that while institutional changes do have this kind of effect, changes in party cohesion can also be produced by changes in the distribution of preferences in the electorate. In a companion paper (Ashworth and Bueno de Mesquita [6]), we show that—unlike institutional changes where strong parties go along with low levels of particularistic behavior—changes in the preferences of voters can lead, simultaneously, to stronger parties and enhanced incentives for particularistic benefits. Thus our theory can explain why party discipline and

\footnote{A Westminster system is a parliamentary form of government in which legislators are elected in plurality rule, single-member districts.}
constituency service look like substitute instruments in cross-sectional comparisons where the causal variation is institutional, and like complementary instruments in U.S. history (both have increased dramatically since the 1960s) where the causal variation is in the electorate’s preferences.

The basic intuitions are simple. Moving from a Presidential to a Parliamentary system changes the marginal voter’s tradeoff between certainty about party ideology and other factors. Increased competitiveness does not change the marginal voter’s tradeoff. Instead, it makes it more likely that the marginal voter will be pivotal for the election. The first change leads an incumbent to shift costs from constituency service to the maintenance of party discipline, while the second leads an incumbent to bear greater costs in both activities. Thus while the comparative evidence (and our comparative model) suggest a tradeoff between cohesion and particularism, changes in district level competitiveness lead to more of both.

The paper is organized as follows. The next section discusses the idea of informative party labels, and locates our work within that literature. Section 2 describes the model, and section 3 describes our concept of a stable party system. Section 4 describes the comparative static results, which are related to some debates in American politics in section 5. Section 6 outlines several extensions to the basic model. Section 7 concludes by informally describing how our companion paper builds on the intuition of this paper to study constituency service.

1 Informative Party Labels

What role do political parties play in elections? A traditional answer is that the party label provides voters with information about the candidates. Knowing that a candidate is a Democrat tells an American voter that the candidate is more likely than average to favor redistribution, abortion rights, etc. This idea has a long history, going back at least to Downs [?]. He pointed out that voters have little incentive to acquire costly information about candidates, so they will rely on low-cost informational shortcuts. If the two parties offer different ideologies on average, this informational use of party labels will be a rational response on the part of voters.

There is substantial evidence that voters do learn about the policy positions of candidates
from their party labels. For example, Snyder and Ting [27] show that party dummies explain a significant part of the variation in voter placements of candidates on a left-right scale, and that these placements vary by party in the intuitive way—Democrats are placed about 2 points (out of 7) to the left of Republicans. Furthermore, conditioning on an estimate of the true ideology (the Poole-Rosenthal scores) adds little to the explanatory power of the regression. This suggests that voters are aware of the differences in candidates ideologies at the gross level of party differences, but that they have little additional information. Alvarez REF shows that voters are risk averse over the policy locations of candidates, so enhancing the informational content of the party labels should give an electoral boost to candidates, all else equal.

Cox and McCubbins [13] have highlighted the role played by the legislative organization of the party in sustaining the electoral value of party labels. To protect the party label, legislators give party leaders the power to impose “discipline”. The leaders use this authority to induce members to vote in ways that support the desired interpretation of the party label, even when they would otherwise vote against these policies (See Aldrich [1] and Calvert and Fox [11] for related models). Parties, on this view, are a type of institutional hand tying.

Snyder and Ting [27] introduced a simple model of informative party labels, formalizing these ideas. In their model, the voters have concave utility over a one-dimensional policy space. They care directly about the ideological preference of their representative, rather than about policy. Since voters do not observe the candidate’s true ideology, they need to find informational cues that will help them vote correctly. If a political party imposes ex-post discipline on its members, then only members whose ideal policies are close to the platform will affiliate with the party. This means the voter can infer something about the candidate’s ideology from the party label. The more disciplined is the party, the more homogeneous are the candidates who affiliate, and thus the more valuable is the label.

Snyder and Ting go on to ask how platform choices are affected by these informative party labels. They find that Downsian parties will converge when information is good, while parties will diverge if information is bad enough. This is because moving away from the other party enhances the informational content of the label.

Snyder and Ting point out that their results do not depend on pressure being applied to
party members during the policy-making process—all that is needed is that the party have an effective “testing” procedure to screen potential members. The same is true of most of our results, as long as more precise testing procedures impose more costs on members of the party than do less precise procedures.

1.1 This Paper’s Contribution

We build on Snyder and Ting to provide a formal model of party discipline, candidate affiliation, and elections that addresses the whole range of empirical findings discussed in the introduction. As in Snyder and Ting, voters use party labels to learn about candidate ideology. The affiliations, and thus the informational content of the labels, depend on the strategically chosen levels of discipline. Thus, party discipline, affiliation, and ideological homogeneity are all determined endogenously within a strategic electoral-legislative setting. Since discipline is endogenous, the model can explain comparative and historical variation in discipline.

In our model, politicians face a trade-off when choosing the level of discipline to impose on their party’s legislative delegation. On the one hand, discipline increases voter confidence in the policy the party will pursue and thereby increases the probability of winning the election. On the other hand, politicians prefer serving in the legislature when they can pursue their own policy agendas, rather than being pressured to toe the party line. The central results will be derived by examining how differences in institutional structures and voter preferences affect this trade-off.

Decisions about party discipline are driven by voter responses to more homogeneous parties. Since these responses are different in Presidential and Westminster regimes, party discipline will differ in these regimes as well. In Presidential systems, the legislature’s policy setting power is constrained by the President. This constraint does not exist in a pure parliamentary system. This difference in legislative power leads voters to care less about uncertainty over their representatives’ policy preferences in a Presidential system than do voters in Parliamentary systems. This is because the President serves as a hedge against legislative extremism. Since the choice of a level of discipline affects affiliation decision and, consequently, the ideological variation in the party, parties will choose different levels of discipline under different institutional
structures. The less the voters value certainty, the lower the level of discipline the party will choose. Thus, parties in Parliamentary systems are more disciplined than those in Presidential systems.

The key variable that we employ to address over-time changes in the U.S. Congress is shifts in the within-district distribution of voter preferences. In the early years of the century, most Congressional districts were relatively uncompetitive. The south was solidly Democratic, while the non-urban parts of the north and west were solidly Republican. This trend reversed in the mid-1960s, due in part to the debate over civil rights. In addition, Cox and Katz [12] show that redistricting in the wake of *Baker v. Carr* (decided in 1962) led to the end of many safe Republican districts in the north. As a result, party competitiveness increased across the board. Within the model, this increase in competitiveness means that the pivotal voter in each district became more likely to be nearly indifferent between the parties. This increased likelihood of indifference implies that a small increase in discipline by one party is more likely to swing the election in favor of that more-disciplined party. This increases the incentives for discipline for both parties, resulting in stronger parties in equilibrium. Thus, historical changes in party strength in the U.S. Congress can be understood with the same model that explains comparative differences in party strength between Westminster and Presidential systems.

2 Basic Model

A polity, divided into $n$ electoral districts, must elect a legislature to set a policy in the one-dimensional policy space, $\mathbb{R}$. A left-wing ($L$) and a right-wing ($R$) party exist, with different, fixed platforms $\pi_L < 0 < \pi_R$. However, the members of a party determine the extent to which the members of that party will adhere to the party’s platform. This will, in turn, determine which politicians are willing to join each party.

There are three dates. At date 0, the two parties simultaneously hold conventions. During the conventions, party affiliations and party discipline are determined. These affiliation decisions, in turn, determine the ideological composition of the parties. At date 1, each legislative district elects a legislator. At date 2, the legislature convenes and sets policy.
A party is characterized by the set of its members, $\mathcal{P}$, and its level of discipline, $\alpha \in [0, \alpha]$. Parties are majoritarian institutions, so if a majority of $\mathcal{P}$ would prefer some other $\alpha'$, they can force a change. We focus on configurations $(\mathcal{P}, \alpha)$ that are immune to such changes. Informally, a cost $\alpha_p$ and set of members $\mathcal{P}$ is stable if no majority in $\mathcal{P}$ prefers a different $\alpha$, taking into account the subsequent changes in affiliations and electoral results. A formal definition is given later.

Our model of discipline follows Snyder and Ting [27]. Discipline, $\alpha$, affects affiliation decisions because a potential legislator whose ideal policy is far from the platform particularly dislikes discipline. A legislator who has ideal point $x$ and is affiliated with party $p$ receives payoff $V(\alpha_p) = B - \alpha_p(x - \pi_p)^2$ if she is in office for one period, while a candidate who does not win office gets 0. $B$ represents the non-policy rewards associated with holding office. A legislator’s utility is decreasing in the divergence between her policy preferences and her party’s platform, and the rate of decrease is greater for members with ideal points further from the platform. One interpretation of this functional form is that, the greater is this distance, the more often she will feel compelled to break ranks with her party. Each time she breaks party discipline she bears a cost, $\alpha$. Thus, when the members of a party choose the level of discipline, they are doing so by threatening to impose costs $\alpha$ on representatives who do not vote with the party. In section 6 we consider the conceptually more satisfying, but analytically more cumbersome, assumption that politicians have preferences over office and the final policy outcome. We show that so long as they are sufficiently office motivated, our results go through.

There is a density of potential politicians for each party, $f^p$, and these densities are symmetric about the platforms $\pi_p$. We keep the model symmetric by assuming that $\pi_L = -\pi_R$ and that $f^L(\pi_L - x) = f^R(\pi_R - x)$ for all $x$. After politicians choose whether or not to affiliate with the parties during these conventions, it is common knowledge that the left party consists of politicians with ideal points $\mathcal{P}_L \subset \mathbb{R}$ and the right party consists of politicians with ideal points $\mathcal{P}_R \subset \mathbb{R}$. Given these affiliation decisions, the voters believe that the ideology of a candidate from party $p$ is a random variable with mean

$$\mu_p = \int_{\mathcal{P}_p} x \, dF^p(x)$$
and variance

$$\sigma_p^2 = \int_{P_p} (x - \mu_p)^2 dF^p(x),$$

The representative voter in district $d$ has preferences represented by $-(x^*_d - x)^2$, where $x$ is a policy and $x^*_d$ is the voter’s ideal point. The candidates and other voters do not know $x^*_d$; their common belief is that $x^* = \gamma_d + \epsilon_d$, where $\epsilon_d$ is a mean zero random variable with an absolutely continuous distribution $F$. The density, $f$, is continuous and log-concave. This is a relatively weak restriction—it is satisfied by most of the usual distributions (e.g., normal, uniform, extreme value, etc.).\(^2\) These ideal points are mutually independent across districts. Finally, we assume that $\sup_x f(x)(\pi_R - \pi_L) < 2$. This assumption requires that there must be a sufficiently large amount of uncertainty regarding the preferences of voters ($f(\cdot)$ must be sufficiently dispersed). The reason this assumption is important is the following. In the model, the direct effect of increasing a party $P$’s discipline is to make a risk-averse voter $i$ more likely to vote for party $P$ because he is more certain of $P$’s ideology. This is the effect in which we are interested. However, there is an indirect effect as well. When $P$ increases discipline, this makes all voters more likely to vote for $P$. Since voters are, on average, centrist, the fact that all other districts are now more likely to vote for party $P$ has the indirect effect of making voter $i$ more likely to vote for $P$’s rival, to balance the expected ideology of the legislature. Assuming that district ideal points have sufficient variance insures that the direct effect is large relative to the indirect effect. This is because, when there is sufficient uncertainty about the preferences of others, voters focus primarily on themselves.

Allowing for $\gamma_d \neq 0$ means that districts might have ideological leanings toward one party. However, we assume that there is no aggregate unbalance between the parties—the $\gamma$ are distributed symmetrically about 0. Denote the voter’s choice of a winning candidate by $w_d \in \{L, R\}$.

After the election, the legislature sets national policy, which is a point on the ideological dimension. The large literature on party pressure and roll-call voting has not reached a consensus on the issue of party effects on policy, conditional on the membership of the legislature. We do not have to take a position on this question here, as our results are robust to the specification

\(^2\)See Bagnoli and Bergstrom [8].
of the legislative stage. In particular, we consider both *majority control*, in which the policy is the average of the ideal points of the majority party, and *floor control*, in which the policy is the average of all of the ideal points.

### 3 Stable Party Systems

The main goal of this section is to characterize *stable party systems*. In later sections, we will use this characterization to derive comparative statics results. These comparative statics will show how party strength and ideological homogeneity differ in Presidential and Westminster systems, as well how changes in ideological divisions in the U.S. electorate caused over-time changes in party strength in Congress.

We solve the game in several steps. First, we derive the equilibrium affiliation decisions for arbitrary levels of party discipline, and use these decisions to derive the voter’s beliefs at the election stage. Then we find the voter’s optimal voting rule in the election stage, given these beliefs. This voting rule is used to construct each potential candidate’s preferences over discipline and affiliations. Then we formally define our stability notion, and analyze the decision over party discipline.

#### 3.1 Affiliations and Beliefs

Consider a potential member of a party who must decide whether or not to affiliate with the party. We focus on the decision of a potential member of party *L*—party *R*’s decision problem is symmetric. At the convention, a potential member of party *L* with ideal point *x* has indirect utility over the level of party discipline (*\(\alpha_L\) ) given by

\[
Pr(w = L|\alpha)(B - \alpha(x - \pi_L)^2).
\]

Write \(V_L(\alpha, x) = B - \alpha(x - \pi_L)^2\). The probability of election depends on \(\alpha\) because party discipline will affect the affiliation decisions, which in turn affect the voters’ beliefs about the candidates’ ideological positions. A potential member affiliates if and only if \(V(\alpha, x) \geq 0\), since the outside option has payoff 0.
Writing \( f^L \) for the density of potential \( L \) member ideal points, we can determine the variance of the ideological positions of party \( L \)'s members as a function of the level of party discipline \((\alpha)\):

\[
\sigma^2_L(\alpha_L) = \int_{V(\alpha, x) \geq 0} (x - \mu_L)^2 f^L(x) \, dx.
\]

It is clear that \( \sigma^2_L(\alpha_L) \) is decreasing in \( \alpha_L \). Because of the symmetry of the distribution of potential members around the party’s platform, the average ideology in party \( L \) is equal to party \( L \)'s platform position, \( \mu_L = \pi_L \).

An example will be instructive. Assume that the distribution of potential ideal points for a party’s candidates is uniform. (This is the main case considered by Snyder and Ting.) If the discipline level is \( \alpha \), then all candidates with ideal points between \( \pi - \sqrt{B/\alpha} \) and \( \pi + \sqrt{B/\alpha} \) will join the party. Thus the conditional density of the ideal points is uniform on that interval, and the variance is

\[
\sigma^2(\alpha) = \int_{\pi - \sqrt{B/\alpha}}^{\pi + \sqrt{B/\alpha}} (x - \pi)^2 \frac{\sqrt{\alpha}}{2\sqrt{B}} \, dx = \frac{B}{3\alpha}.
\]

### 3.2 The Election

The voter chooses which candidate to select in round 2 by comparing the expected utility of each choice. The voter in district \( d \) votes for \( L \) if and only if

\[
-\mathbb{E} \left( (x_{\text{leg}} - x_d^\ast)^2 \mid L \right) \geq -\mathbb{E} \left( (x_{\text{leg}} - x_d^\ast)^2 \mid R \right).
\]

Taking the expectations, this becomes

\[
-(\mu_{\text{leg}|L} - x_d^\ast)^2 - \sigma^2_{\text{leg}|L} \geq -(\mu_{\text{leg}|R} - x_d^\ast)^2 - \sigma^2_{\text{leg}|R}.
\]

Rearrange this to get

\[
x_d^\ast \leq \frac{1}{2}(\mu_{\text{leg}|R} + \mu_{\text{leg}|L}) + \frac{\sigma^2_{\text{leg}|R} - \sigma^2_{\text{leg}|L}}{2(\mu_{\text{leg}|R} - \mu_{\text{leg}|L})}.
\]

To get a feel for what this implies, consider majority control: policy is the mean of the majority party ideal points. In this case, the voter calculates his optimal choice by conditioning
on being the pivotal district: he assumes that if he votes \( L \) the national policy has mean \( \pi_L \) and variance \( 2\sigma_L^2/(n + 1) \), while if he votes \( R \) the national policy has mean \( \pi_R \) and variance \( 2\sigma_R^2/(n + 1) \). Further, since we assume that the two party platforms are themselves symmetric about 0, we can conclude that \( \frac{1}{2}(\mu_L + \mu_R) = 0 \). Thus, the voting rule simplifies to

\[
x^*_{d} \leq \frac{\sigma_R^2 - \sigma_L^2}{(n + 1)(\pi_R - \pi_L)}.
\]

To simplify notation, define the cut-point

\[
c(\alpha_L, \alpha_R) = \frac{\sigma_R^2(\alpha_R) - \sigma_L^2(\alpha_L)}{(n + 1)(\pi_R - \pi_L)}. \tag{1}
\]

The voter prefers \( L \) if \( x^*_{d} < c(\alpha_L, \alpha_R) \). The cut-point, \( c(\alpha_L, \alpha_R) \), is increasing in \( \alpha_L \) and is decreasing in \( \alpha_R \).

It turns out that the conclusions of this analysis also hold in the case of floor control. In both cases, each voter uses a cut-point voting rule: vote \( L \) if and only if \( x^*_{d} \leq c(\alpha_L, \alpha_R) \). Furthermore, these cut-points have the same monotonicity properties as the special case we examined above.

**Proposition 1** Under both majority party control and floor control, every voting subgame has a unique equilibrium. The equilibria are cut-point equilibria, and the cutpoints are increasing in \( \alpha_L \) and decreasing in \( \alpha_R \) for each \( d \).

The proof is in the appendix.

The voter is more likely to vote for \( L \) the more certain he is of the ideology of the left-wing candidate (low \( \sigma_L \)) or the less certain he is of the ideology of the right-wing candidate (high \( \sigma_R \)). This is because both of these scenarios make electing the left-wing candidate relatively less risky, which benefits the risk-averse voter.

### 3.3 Party Conventions

Given these beliefs and the voting strategies, we can solve for the election probabilities as a function of the affiliation decisions. The incentive effects of party discipline are the same for both wings of a given party (that is, potential members to the right and to the left of the
party platform). Because the distribution of potential members is symmetric, when a change is made in the level of discipline symmetric groups from both wings of the party either affiliate or disaffiliate. Consequently, a change in party discipline does not affect the mean ideology of the party \( \mu_p \), but it does affect the variance \( \sigma^2_p \).

A candidate from party \( L \) is elected if and only if:

\[
x^*_d \leq c(\alpha_L, \alpha_R).
\]

We can use the definition of \( x^* \) to rewrite this condition as \( \gamma_d + \epsilon \leq c(\alpha_L, \alpha_R) \), which can again be rewritten \( \epsilon \leq c(\alpha_L, \alpha_R) - \gamma_d \). Since \( \epsilon \sim F \), the probability that the voter votes for the candidate from party \( L \) is \( F(c(\alpha_L, \alpha_R) - \gamma_d) \). Given party \( R \)'s level of discipline, \( \alpha_R \), a politician’s indirect utility from affiliating with party \( L \) can be written:

\[
F(c(\alpha_L, \alpha_R) - \gamma_d) V_L(\alpha, x),
\]

and a politician will affiliate if and only if:

\[
F(c(\alpha_L, \alpha_R) - \gamma_d) V_L(\alpha, x) \geq 0
\]

Memberships for each party can be derived from these affiliation decisions. Let \( P_L \) be the set of politicians who choose to affiliate with party \( L \). This set is determined by the level of party discipline \( \alpha_L \). A pair \((\alpha_L, P_L)\) is stable against \((\alpha_R, P_R)\) if (1) \( P_L \) is exactly the set of people who want to affiliate with party \( L \) given \( \alpha_L \) and \((\alpha_R, P_R)\) and (2) there is no \( \alpha' \) such that majority in \( P_L \) prefer \( \alpha' \) to \( \alpha_L \) given \((\alpha_R, P_R)\) and the new implied affiliation decisions.

Stability represents a natural equilibrium concept for a party that allows free entry and exit and is governed by a majoritarian principle.\(^3\) If there is free exit, then equilibrium requires that no candidate be affiliated if \( V(x, \alpha) < 0 \). Affiliation must be better, from each individual party member’s perspective, than non-affiliation. Similarly, if there is free entry, then no perfect equilibrium can have a non-affiliated candidate with \( V(x, \alpha) > 0 \). No politician will turn down the opportunity to join a party if doing so will make her better off than non-affiliation. Thus,\(^3\) This is similar to the equilibrium concept commonly used to study local public finance with mobility, for example Epple and Romer [16]. Our politicians are sophisticated in Epple and Romer’s terminology.

\(^3\)
free entry and exit, in conjunction with endogenously chosen discipline, suggest stability as the equilibrium concept.

Turning from the internal stability of a single party, we can now look for an equilibrium between the parties. We say that a 4-tuple \((\alpha_L, P_L, \alpha_R, P_R)\) is a stable party system if each party is stable against the other. Further, a symmetric stable party system is a stable party system in which \(\alpha_L = \alpha_R\).

**Proposition 2** Assume that each candidate has a unique favorite level of discipline for each configuration of the opposing party. Then for any value of the parameters there exists at least one stable party system (SPS). Moreover, when there are multiple SPSs, there exists a SPS with the greatest level of party discipline and a SPS with the smallest amount of party discipline.

Notice that the uniqueness requirement is satisfied for the uniform example discussed earlier, since the objective function is strictly logconcave in that case.

The proof of this proposition is in the appendix, since several of the steps are rather involved. However, it will be useful to go over the main points here, since understanding the major components of the proof is the best way to understand how the model works.

An SPS is a pair of discipline levels \((\alpha_L, \alpha_R)\) that reproduce themselves in the following sense: Given the set of candidates who affiliate with party \(p\) given \(\alpha_p\) (called \(P(\alpha_p)\)) and party \(-p\)’s level of discipline \(\alpha_{-p}\), the Condorcet winning discipline level for \(p\) is exactly \(\alpha_p\). The proof uses this “self-generating” idea to build a map whose fixed points are SPSs.

Clearly, a crucial step in this program is showing that Condorcet winners exist. We do this by proving a version of the median voter theorem within each party. This means we can characterize the majority preference over changes in discipline in terms of the median voter’s preferences. To see why this is possible, consider the case where \(\gamma\) is identically 0.\(^4\) Party discipline is more costly for members with ideal points far from the platform. Consequently, potential members have preferences over discipline that are ordered by the distance of their ideal points from the party platform. Politicians whose policy preferences are similar to the party platform prefer more discipline, while those whose policy preferences differ significantly from

---

\(^4\)Variation in the \(\gamma\) is the cause of most of the complications in the proof of the entire proposition.
the party platform prefer less discipline. This gives the parties’ decision problems regarding discipline a natural one-dimensional structure which is needed for the median voter theorem to apply.

The next step is based on a different monotonicity property of candidate preferences. A candidate’s benefit from increasing discipline is increasing in the other party’s level of discipline. Since the median voter in the party is decisive at the convention, this means that the Concorctet winning level of discipline for party $p$ is increasing in $\alpha_{-p}$. This implies that the “best-response” functions are increasing, which is enough to ensure an equilibrium. In addition, this property is important for the comparative statics—any change that leads one party to increase its discipline will lead to feedback effects, increasing discipline for both parties.

To derive the comparative statics results later in the paper, we need a sharper characterization of this median voter’s optimal discipline level. (This characterization can also be used to prove the monotonicity results we appealed to above.) Since the median voter in each party is decisive, her first-order condition must hold with equality at any interior stable party system. This first-order condition is the key to all of our results, so we will go carefully through the intuition. Recall that the median member of party $L$ has utility given by the probability of election (given the stable party system) times the expected benefit of serving in office (given her policy preferences and the level of discipline):

$$F(c(\alpha_L, \alpha_R) - \gamma_d)(B - \alpha_L(x_{med} - \pi_L)^2).$$

Taking logs, we can write the party $L$ median voter’s maximization problem as:

$$\max_{\alpha_L} \{ \log F(c(\alpha_L, \alpha_R) - \gamma_d) + \log (B - \alpha_L(x_{med} - \pi_L)^2) \}.$$

The first-order condition for this maximization problem is:

$$c_L \frac{f(c - \gamma_d)}{F(c - \gamma_d)} = \frac{(x_{med} - \pi_L)^2}{B - \alpha_L(x_{med} - \pi_L)^2}.$$

Similarly, the first-order condition for the party $R$’s median voter is:

$$-c_R \frac{f(c - \gamma_d)}{1 - F(c - \gamma_d)} = \frac{(x_{med} - \pi_R)^2}{B - \alpha_R(x_{med} - \pi_R)^2}.$$
where $c_L$ and $c_R$ are the partial derivative of $c$ with respect to $\alpha_L$ and $\alpha_R$, respectively. Like all first-order conditions, this says that if $\alpha$ is optimal, then the marginal benefit from a small increase in the amount of party discipline must be exactly balanced by the marginal cost of that change. What are the marginal benefits and costs of discipline for the median member of a party? Recall that $c$ is a cut-point $(c(\alpha_L, \alpha_R) = (\sigma_R^2(\alpha_R) - \sigma_L^2(\alpha_L))/(2(\mu_R - \mu_L))$ defining when the pivotal voter in a district will choose the candidate from party $L$. A small increase from $\alpha_L$ to $\alpha_L + d\alpha_L$ moves the cut-point to the right by approximately $d\alpha_L c_L$, and this increases the probability of election by approximately $d\alpha_L c_L f(c - \gamma_d)$. Winning gives the candidate a payoff of approximately $B - \alpha_L (x_{med} - \pi_L)^2$, so the marginal benefit of an increase in discipline is:

$$d\alpha c_L f(c - \gamma_d) (B - \alpha_L (x_{med} - \pi_L)^2).$$

That is, the marginal increase in probability of election multiplied by the benefits of holding office.

The cost of an increase in discipline arises from the fact that, once in office, candidates are more constrained in the policies they can pursue. The extra costs of party discipline associated with a small increase in $\alpha_L$ are approximately $d\alpha_L (x_{med} - \pi_L)^2$. The candidate, of course, only bears these costs if elected, which occurs approximately with probability $F(c - \gamma_d)$. Thus the marginal cost of an increase in discipline is

$$d\alpha (x_{med} - \pi_L)^2 F(c - \gamma_d).$$

The first order condition sets these two expressions equal.

The marginal benefit of an increase in discipline is due to the possibility that a small increase in discipline will swing the election results. Said differently, the benefit of a small increase in discipline comes entirely from its effect on voters who are close to indifference between the parties. Consequently, over-time or cross-institutional variations in the level of discipline must be driven by one of two factors: (1) changes in the responsiveness to discipline of nearly indifferent voters or (2) changes in the number (measure) of such voters. We will show how these two factors can explain variance both between presidential and parliamentary systems and over-time within the U.S. Congress.
Before turning to the comparative statics, we need one more piece of technical apparatus. Note that proposition 2 demonstrates the existence of a stable party system but does not guarantee uniqueness. Indeed, in general there will not be a unique equilibrium. Thus, in order to compare the level of discipline that emerges under various institutional settings we need a way to compare sets of equilibria. We will say that one set of equilibria is greater than a second set if the greatest and least equilibria in the first set are greater than the greatest and least equilibria, respectively, of the second set. Echenique [14] shows that if equilibrium sets are ordered this way, then a broad class of adaptive adjustment processes will converge to greater equilibria whenever a shock increases the equilibrium set.

4 Comparative Statics

4.1 Presidential vs. Westminster Systems

We can now compare party discipline in Presidential and Westminster systems. To focus on the role of the executive, we will consider two systems that are identical except for selection of the executive. There are a number of single-member legislative districts, each of which is contested by the same two parties. In the Westminster system, the winners of these elections make up the parliament and determine policy. In the presidential system, the majority party in the legislature must bargain with the president to set policy. For simplicity, we abstract from heterogeneity among districts for this section, so all of the $\gamma$ are 0. Since this makes everything symmetric, it is natural to focus on equilibria that are symmetric in the sense that (i) all of the voters use the same cut-point in the elections and (ii) the parties choose the same levels of discipline.\(^5\)

For the most part, we will work with a reduced form description of bargaining between the legislature and the president, although we discuss a more detailed bargaining model in section 6.2. Following Alesina and Rosenthal [3], we assume that policy in a Presidential system is a weighted average of the legislature’s proposal and the President’s ideal point, with weight $\beta$ on the legislative proposal: $x = \beta x_{\text{leg}} + (1 - \beta) x_{\text{pres}}$.

\(^5\)Standard arguments can be used to modify propositions 1 and 2 to show that a symmetric equilibrium exists.
Since our focus is on the legislative election, we make the following simplifying assumptions. First, we assume that when casting his legislative vote, a voter assumes that he is not pivotal in the Presidential election. Thus the President’s ideology is a lottery that is not affected by the voter’s choices. Second, we focus on equilibria that treat the parties symmetrically. This means that the lottery over Presidential ideology has mean 0.

Again, we can build some intuition by first considering the case of majority control of the legislature. The Westminster system corresponds exactly to the system we studied in the previous section. Thus, the voter’s cut-point is:

\[ c_{wst}(\alpha_L, \alpha_R) = \frac{\sigma^2_R(\alpha_R) - \sigma^2_L(\alpha_L)}{(n+1)(\mu_R - \mu_L)}. \] (2)

The results of this section will be derived by comparing this cut-point to the cut-point in the presidential system.

In order to determine the optimal voting rule in a presidential system, we mimic the previous analysis, taking into account bargaining between the legislature and the president. Without repeating the algebra, it is clear that the voter in a Presidential system will vote for legislative candidate \( L \) if and only if:

\[ x_d^* \leq \frac{1}{2} \beta(\mu_R + \mu_L) + (1 - \beta)\mathbb{E}x_{\text{Pres}} + \frac{\beta(\sigma^2_R - \sigma^2_L)}{(n+1)(\mu_R - \mu_L)}. \]

Since \( \mu_L = -\mu_R \) and \( \mathbb{E}x_{\text{Pres}} = 0 \) in equilibrium, this simplifies to:

\[ x_d^* \leq \frac{\beta(\sigma^2_R - \sigma^2_L)}{(n+1)(\mu_R - \mu_L)}, \]

so the cut-point is

\[ c_{\text{pres}}(\alpha_L, \alpha_R) = \frac{\beta(\sigma^2_R(\alpha_R) - \sigma^2_L(\alpha_L))}{(n+1)(\mu_R - \mu_L)}. \] (3)

Comparing equations (2) and (3) shows that voters have different incentives under each system. In particular, voters put greater weight on their uncertainty over candidate ideology in the Westminster system. This is because the legislature has less impact on national policy in a Presidential system, since the legislature’s proposal does not determine national policy alone, but rather is averaged with Presidential preferences. The President acts as a hedge against legislative extremism. Voters are more concerned about the possibility of an extreme
legislature in a Westminster system because the legislature is unconstrained by an independent executive branch. Formally, this means that the marginal benefit of party discipline is attenuated in Presidential systems relative to Westminster systems. To see this, recall that the marginal benefit of discipline for party $p$ is proportional to $c$.

Direct calculation shows that $(c_{\text{pres}})_L = \beta (c_{\text{wst}})_L < (c_{\text{wst}})_L$ and $-(c_{\text{pres}})_R = -\beta (c_{\text{wst}})_R < -(c_{\text{wst}})_R$. This intuition suggests the following result.

**Proposition 3** The level of discipline is greater in Westminster systems than in Presidential systems.

**Proof** We begin by recording a fact about the cut-point functions.

**Lemma 1** For all $\alpha_L$ and $\alpha_R$, we have $c_{\text{wst}} L > c_{\text{pres}} L$ and $-c_{\text{wst}} R > -c_{\text{pres}} R$.

**Proof** First, consider the case of floor control. In a symmetric equilibrium, the Westminster cut-point solves the equation

$$
(c_{\text{wst}})_L = \frac{n-1}{2n} \left( F(c_{\text{wst}}) \pi_L + (1-F(c_{\text{wst}})) \pi_R \right) + \frac{\sigma_R^2 - \sigma_L^2}{2n(\mu_R - \mu_L)},
$$

$$
\equiv \phi(c_{\text{wst}}).
$$

The cut-point for the Presidential case can be written

$$
c_{\text{pres}} = \beta \phi(c_{\text{pres}}) + (1 - \beta) x_{\text{Pres}},
$$

where $\phi$ is the function determining the Westminster cut-point. Applying the implicit function theorem, we see that

$$
c_{\text{pres}} L = \frac{\beta \phi L}{1 - \beta \phi_c}.
$$

Since

$$
\phi_c = \frac{n-1}{2n} f(c)(\pi_L - \pi_R) < 0
$$

(all other cutpoints move right, mine moves left), this implies that $c_{\text{pres}} L < c_{\text{wst}} L$. A similar argument works for $\alpha_R$. Finally, in the case of majority control, $\phi_c \equiv 0$, so the result is trivial.

$\square$
Now consider a symmetric stable party system \((\alpha_L, \alpha_R)\). Associated with this pair is a pair of memberships, \(L(\alpha_L)\) and \(R(\alpha_R)\). Rational affiliation implies that 
\[
\mathcal{P}(\alpha_p) = \left\{ x \mid x \in \left[ \pi_p - \sqrt{\frac{B}{\alpha_p}}, \pi_p + \sqrt{\frac{B}{\alpha_p}} \right] \right\}.
\]
Let \(\Delta_p(\alpha_p)\) be the median of \((x - \pi_p)^2\) over \(\mathcal{P}(\alpha_p)\).

In a SPS, each party’s median voter’s first-order condition for \(\alpha\) must be satisfied. Thus a stable party system in a Westminster system is a zero of the system
\[
\begin{align*}
\epsilon_{L}^{wst} \frac{f(c)}{F(c)} - \frac{\Delta_L(\alpha_L)}{B - \alpha_L \Delta_L(\alpha_L)} &= 0, \\
-\epsilon_{R}^{wst} \frac{f(c)}{1 - F(c)} - \frac{\Delta_R(\alpha_R)}{B - \alpha_R \Delta_R(\alpha_R)} &= 0,
\end{align*}
\]
while a stable party system in a presidential system is a zero of the system
\[
\begin{align*}
\epsilon_{L}^{pres} \frac{f(c)}{F(c)} - \frac{\Delta_L(\alpha_L)}{B - \alpha_L \Delta_L(\alpha_L)} &= 0, \\
-\epsilon_{R}^{pres} \frac{f(c)}{1 - F(c)} - \frac{\Delta_R(\alpha_R)}{B - \alpha_R \Delta_R(\alpha_R)} &= 0.
\end{align*}
\]
The lemma implies that the LHS of the first system is point-wise greater (in the usual vector order) than the LHS of the second. Furthermore, these functions are continuous in \((\alpha_L, \alpha_R)\) since \(c, f,\) and \(\Delta\) are all continuous, and party \(p\)’s first-order condition is increasing in \(\alpha_p\), since the objective functions are logsupermodular in \(\alpha_p\) and \(\alpha_{-p}\). Thus theorem 4 of Milgrom and Roberts [24] implies that the set of solutions to the first system is larger than that of the second. \(\square\)

Party discipline is weaker in the Presidential system because the pivotal voter is less responsive to changes in the variance of ideal points within a party. This means that an increase in discipline has a relatively small marginal impact on the probability of winning in a Presidential system, since the voters do not care as much about the national policy preferences of their representatives as do the voters in a Westminster system. Thus, the parties benefit less from imposing discipline in Presidential systems than they are in Westminster systems.

Importantly, the responses of the parties are greater than they would be in a non-strategic setting. Since levels of discipline are strategic complements between the parties, any initial
reaction by a party to an increase in the responsiveness of voters to discipline is subject to a multiplier effect—the parties increase their level of discipline in response to the greater responsiveness. This leads to a cycle of increasing discipline: greater discipline by party \( p \)'s incentive to raise its own level of discipline, leading to a reinforcing sequence of increases. Thus competition between the parties exaggerates the impact of a change in voter responsiveness.

4.2 More Competitive Districts

Institutional changes are not the only ones that lead to more disciplined parties—changes in the distribution of voter preferences can do so as well. To capture increased district-level competitiveness, we consider changes in the distribution of the partisan leanings of the districts. Specifically, we consider a change in which all of the \( \gamma \) become closer to 0, and we ask what happens to the equilibrium levels of party discipline. In order to account for changes in party strength over-time in Congress, we must show that increases in the competitiveness of elections leads to an increase in the each party’s median member’s most preferred level of discipline. We will further assume that the median voter in each party is from a district controlled by that party.

An across the board increase in competitiveness means that the districts became less biased towards one party or the other. Southern voters were more willing to consider electing a Republican while non-urban Western and Northern voters were willing to consider Democratic candidates. In the model this constitutes a decrease in the dispersion of the ideal points of the district-level median voters. That is, the ideal point of median voter in each district moved toward the center. In Republican-leaning districts the median voter became more willing to consider a Democrat, and in Democratic-leaning districts the median voter became more willing to consider a Republican.

What is the effect of a decrease in dispersion of district median voter ideal points on our model of party organization? The benefit of party discipline is to improve voter confidence in the national policy agenda that a candidate will pursue once in power. This benefit is only realized by a given candidate when the level of discipline actually persuades the pivotal voter in
a district to vote for that candidate when he otherwise would have voted for the other party’s
candidate. However, the party members bear the costs of discipline regardless of whether the
level of discipline swings the election. Consequently, the marginal benefit of party discipline
is proportional to the likelihood that an extra amount of discipline will swing the election,
while the marginal costs are fixed with respect to this likelihood. Discipline can only swing an
election if the decisive voter is close to being indifferent between the two parties. The cause
of an increase in competitiveness of district elections is precisely a greater likelihood that this
near indifference obtains. Thus, more competitive elections increase the marginal benefit of
discipline, creating incentives for greater levels of party discipline. We formalize this intuition
in the following proposition.

Proposition 4 Consider the case of majority party control of the legislature. If the competi-
tiveness increases between the parties, in the sense that each district’s expected median moves
closer to the midpoint between the parties, then the level of party discipline increases.

Proof The party medians choose the levels of discipline. At an interior equilibrium, a
median voters’ optimal levels of discipline satisfy the first-order condition:

\[
\begin{align*}
\frac{c_L}{f(c(\alpha_L, \alpha_R) - \gamma)} - \frac{(x_{\text{med}} - \pi_L)^2}{B - \alpha_L(x_{\text{med}} - \pi_L)^2} &= 0 \\
-\frac{c_R}{1 - F(c(\alpha_L, \alpha_R) - \gamma)} - \frac{(x_{\text{med}} - \pi_R)^2}{B - \alpha_R(x_{\text{med}} - \pi_R)^2} &= 0,
\end{align*}
\]

where \(c_L\) is the partial derivative of \(c\) with respect to \(\alpha_L\). A SPS is a zero of this system.

Consider first the median member of party \(L\)’s first-order condition. If the district expected
median increases to be closer to the midpoint between the parties (from \(\gamma\) to \(\gamma'\) with \(\gamma < \gamma' < 0\)),
then the argument of the hazard rate decreases. Since logconcavity is the same as a decreasing
hazard rate, this means the left hand side of the first order condition increases.

We proceed similarly for party \(R\)’s median voter. The argument of the hazard rate is
increasing in \(\gamma\). Thus as the expected district median decreases toward the parties’ midpoint,
the left hand side of the first order condition increases. Thus, the left hand side of the system
with less competitive parties is point-wise greater than the left hand side of the system with
more competitive parties in the usual vector order. Theorem 4 of Milgrom and Roberts [24]
therefore implies that the set of solutions to the first system is larger than that of the second.

Thus, according to proposition (4), increases in the competitiveness of district-level elections provide a national-level incentive for greater party discipline. The intuition for this result is different than the one for the comparative case. Here, the marginal voter (the type who is exactly indifferent between the candidates) does not change his tradeoff between the variance and idiosyncratic payoff. Instead, he is more likely to change the outcome of the election if he changes his vote. At the same time, the expected cost of discipline is falling, because the candidate is winning less often in a competitive district. Thus the marginal benefit of discipline increases and the marginal benefit falls, leading to more discipline in equilibrium.

Does this result extent to the case of floor control? The complication is that the cutpoints in the election stage will generally depend on the entire distribution of the $\gamma$. The comparative static result will hold as long as the change in the biases does not move the cutpoints in the wrong direction by “too much”.

We can say definitively what happens in the special case where $F$ is the uniform distribution on $[-\tau, \bar{\tau}]$. In this case, the equilibrium cutpoints are additively separable in the $\gamma$ and the $\alpha$, so the second effect is absent.

Here are some details of the argument. District $d$’s cut-point is

$$c_d = \frac{1}{2n} \sum_{k \neq d} \left( \pi_R + (\pi_L - \pi_R) \frac{C_k - \gamma_k}{2\bar{\tau}} \right) + \frac{\sigma^2_R - \sigma^2_R}{2n(\pi_R - \pi_L)}.$$  

Rewrite this as

$$c_d + \left( \frac{\pi_R - \pi_L}{4n\bar{\tau}} \right) \sum_{k \neq d} C_k = -\frac{n-1}{n} \pi_R + \frac{1}{4n\bar{\tau}} \sum_{k \neq d} \gamma_k - \frac{\sigma^2_R - \sigma^2_R}{2n(\pi_R - \pi_L)}.$$

Stacking these equations gives us

$$Xc = g(\alpha_L, \alpha_R) + h(\gamma).$$

At a solution to this system, the cutpoints satisfy

$$c = X^{-1}g(\alpha_L, \alpha_R) + X^{-1}h(\gamma).$$

Thus the derivatives of $c$ with respect to $\alpha_L$ and $\alpha_R$ do not depend on the $\gamma$.  

21
4.3 Comparison of the Results

We have shown that parliamentary government and more competitive districts both increase the equilibrium level of party discipline. It is important to notice that these effects work through different channels. A parliamentary system gives the marginal voter different incentives than does a presidential system. More competitive districts do not affect the marginal voter’s tradeoff; rather they increase the likelihood that the marginal voter is pivotal in the election. This difference means that these two changes have different implications for other aspects of legislative politics. As discussed more fully in the conclusion, this lets discipline and particularistic service look like substitutes comparatively at the same time they look like complements historically.

5 The Return of the Congressional Party

The middle of the 20th century was the height of the “textbook Congress”—a term sometimes used to describe the Congressional institutions familiar from works such as Mayhew’s Congress: The Electoral Connection [23]. It was characterized by the unwillingness of its members to let their parties exert much discipline in voting. Committees, and their chairs, were the true nexus of power (Shepsle [26]).

In retrospect, Mayhew’s classic description came at the end of the institutional equilibrium he was describing. The aftermath of Watergate saw a variety of reforms, such as making committee chairs accountable to parties, that strengthened parties in Congress. These reforms have been followed by increases in the cohesiveness of parties in legislative votes. There is a substantial body of empirical work that is consistent with these ideas. To take a recent example, Aldrich, Berger, and Rohde [2] examine both Houses of Congress between 1877 and 1994. They show that there is a strong correlation between measures of preference homogeneity, such as (minus) the standard deviation of Poole-Rosenthal ideal point estimates or the $R^2$ of a regression of those estimates on party dummies, and qualitative measures of the power of party leaders. In particular, they show that both measured preference homogeneity and party strength have increased dramatically since the mid-1960s.
We demonstrate how changes in the electorate in the 1960s account for over-time changes in the Congress within the context of our model. For much of the 20th century, many Congressional districts were relatively uncompetitive. The south was solidly Democratic, while the non-urban parts of the north and west were solidly Republican. This trend reversed in the 1960s, due in part to the debate over civil rights. In addition, Cox and Katz [12] show that redistricting in the wake of Baker v. Carr (decided in 1962) led to the end of many safe Republican districts in the north. As a result, party competitiveness increased across the board.

Today the parties are much more competitive in most states. Erikson, Wright, and McIver [17] find that public opinion is roughly balanced between the two parties in most states, and Ansolabehere and Snyder [4] find that the portion of vote shares explained by the states partisan leanings (“the normal vote”) has declined dramatically since mid-century. They write:

The normal vote accounts for 53 percent of the variation in the vote in the 1940s. It’s importance drops substantially in the 1950s, to 40 percent of total variance in the vote. And it collapses in the 1960s, explaining only 20 percent of the variance in the vote in the 1960s and 1970s. The decline of the normal vote as an explanatory factor continues in the 1980s, falling to 10 percent in the 1980s and 1990s.

Our model says that this change should lead to stronger parties, as can be observed in the historical record. Consequently, the model is consistent with the empirical finding that U.S. Congressional parties became stronger beginning in the late 1960s. According to the model, this institutional change was a rational response to shifting electoral conditions brought about by the civil rights movement and redistricting.

Notice that our argument does not depend on the south being “liberal”. An alternative interpretation of γ comes from a valence advantage for the L party. Assume that the voter’s payoff to voting \( L \) is \( E - (x_L - x^*)^2 + \delta \). Then the cut-point rule will be: vote \( L \) if \( x^* \leq c(\alpha_L, \alpha_R) + \delta/(2(\mu_R - \mu_L)) \). Thus a district with a median voter who gets positive utility from the \( L \) party independent of policy is just like a district with a more left-leaning median voter.
5.1 Conditional Party Government and its Critics

In the lively debate over party effects in Congress, changes in the ideological heterogeneity within parties plays a key theoretical role. The leading explanation for the changes in party strength in the U.S. Congress is Aldrich’s and Rohde’s idea of conditional party government (CPG). They argue that parties delegate authority to leaders only when there is sufficient ideological homogeneity among the members. Members of homogeneous parties have no reason to fear that powerful leaders will force them to support legislation they oppose, since their interests and the party’s interests are aligned. Thus, Aldrich [1] and Rohde [25] argue that the observed increase in party cohesiveness in roll call votes in the U.S. was caused by an increase in the ideological homogeneity in Congressional parties. Furthermore, they posit the following causal chain: increased ideological homogeneity led party members to grant the party leadership more disciplinary power which, in turn, led to greater cohesiveness in roll call votes.

Cox and McCubbins [13] link the CPG theory with informative party labels, suggesting that parties and their leaders serve as “cartels” that prevent free-rider problems that could degrade the party label. Building on Aldrich’s [1] and Rohde’s [25] argument, Cox and McCubbins contend that ideological homogeneity causes a convergence of interests that make cartel-like parties particularly attractive. Hence, consistent with the earlier CPG theorists, Cox and McCubbins predict that parties will be strong when they are ideologically homogeneous.

Although the empirical work documenting the relationship between ideological heterogeneity and party voting is impressive, the CPG interpretation has inspired several theoretical critiques. Krehbiel [22] and Calvert and Dietz [10] argue that, if the CPG account is correct, then parties are not actually fulfilling any important role in the legislature. In particular, the CPG model predicts that parties will become powerful precisely when they are least needed. Parties become strong when there is ideological homogeneity. But strong parties are not needed to control legislation or to protect the party label if the membership is ideologically homogeneous. The members will not be tempted to deviate from a party platform with which they agree. Thus, according to Krehbiel [22] and Calvert and Fox [11], it is precisely at those times when they are least needed that CPG predicts that parties will be strongest.
Furthermore, Krehbiel points out that one cannot conclude that ideological homogeneity causes strong parties simply by observing that legislative votes are more cohesive in parties that are ideologically homogeneous. When a party is ideologically homogeneous the membership agree with one another on which policies are desirable. Consequently, cohesive votes should be expected among ideologically homogeneous legislative delegations with or without strong parties.

Our model makes three contributions to the debate over party discipline and ideological heterogeneity. First, membership is endogenous, so the model itself explains why parties become more or less ideologically homogeneous. In particular, we establish the role of electoral concerns in the party formation process. Both Aldrich-Rohde and Krehbiel ultimately trace changes in legislative cohesion back to changes in the composition of parties, in spite of their differences over the precise channel of this effect. Second, there are clear comparative statics results—more legislative control of policy and more competitive elections lead to both more homogeneous parties and more discipline imposed by leaders. These features allow for a better understanding of the joint determination of the variables studied in the empirical debate over party strength in Congress. Finally, our model shows that it is possible to understand both historical variance in the power of U.S. parties and comparative variance in the power of parties across different institutional structures within a unified theoretical framework.

6 Extensions

6.1 Candidates with Preferences over Policy Outcomes

A potential weakness of the model we have considered so far is that candidates do not care about the final policy choice. Now we show that our results continue to hold as long as candidates put sufficient weight on the “ego rents” of office.

For this section, we assume that each legislator takes position $\alpha \pi + (1 - \alpha) x$ and the policy is the average of these positions. Thus discipline actually compels members to support the platform $\alpha$ percent of the time.

Consider a politician who is choosing whether to affiliate with party $L$. Her payoffs condi-
tional on her decision and the electoral outcome in her district are described by the following table:

<table>
<thead>
<tr>
<th></th>
<th>$L$ wins</th>
<th>$R$ wins</th>
</tr>
</thead>
<tbody>
<tr>
<td>affiliate</td>
<td>$B + u(S)$</td>
<td>$u(R)$</td>
</tr>
<tr>
<td>not</td>
<td>$u(L)$</td>
<td>$u(R)$</td>
</tr>
</tbody>
</table>

where $u(S)$ is her expected utility for the policy lottery if she runs and wins, $u(L)$ is her expected utility over the policy lottery if another member of her party wins the seat in her district, and $u(R)$ is her expected utility if the $R$ party wins in her district. Once she is affiliated, only the top row is relevant for her preferences over $\alpha$. When choosing an affiliation, only the first column is relevant.

She will affiliate if and only if $U = B + u(S) - u(L) \geq 0$. $u(S)$ is decreasing in $\alpha$, since she will be able to give her favorite policy less support as party discipline increases. Further, $u(L)$ is increasing in $\alpha$, since increased discipline decreases the risk associated with allowing another politician (with a different ideal point) to serve as the district’s representative. Thus $U$ is decreasing in $\alpha$, as before.

Now consider her preferences over $\alpha_L$ if she affiliates. The payoff in the top left cell is

$$
\mathbb{E} \left( B - (p - x)^2 \right) = B - \left( \mathbb{E}p - x \right)^2 - \sigma_p^2
$$

$$
= B - \left( \frac{n - 1 + \alpha}{n} \pi + \frac{1 - \alpha}{n} x - x \right)^2 - \sigma_p^2
$$

$$
= B - \left( \frac{n - 1 + \alpha}{n} \right)^2 (\pi - x)^2 - \sigma_p^2.
$$

This is almost the same as before. The only difference is the term $\sigma_p^2$. The analysis from before goes through if the changes in this variance with $\alpha$ do not upset the logsupermodularity of the function.

The numerator of the cross-partial of the log of the payoff is $-B + \sigma^2 - \alpha \sigma_p^2$. For the previous results, this should be negative. This will be true so long as $B$ is sufficiently large. Thus, as long as politicians care enough about holding office, our results are robust to allowing legislators care about final policy outcomes, rather than the policy for which they individually vote.
6.2 Veto Bargaining

Much of the literature on congressional-presidential relations uses a bargaining protocol that more closely conforms to the actual strategic relationship between the President and Congress: veto bargaining. Under veto bargaining, there is a status quo policy ($x_S$) and the Congress proposes an alternative ($x_{leg}$). The President then chooses her favorite alternative from the set \{\(x_S\), \(x_{leg}\)\}. The equilibrium policy is

\[
x = \arg \max_{\hat{x} \in \{x_S, x_{leg}\}} |\hat{x} - x_{pres}|
\]

Variation in $x_{leg}$ now leads to variation in the equilibrium policy only when the status quo falls in the range where the legislature’s ideal policy would be chosen. This can be used to demonstrate that our qualitative results comparing Presidential and parliamentary systems still hold.

Imagine that there is pre-electoral uncertainty over the status quo, perhaps because of uncertainty over which issues will be important in a given legislative session. When the status quo is realized, after the election, changes in the variability of legislature ideal points will have a smaller impact on the variability of policy under veto bargaining than under a system where policy is fully determined by the legislature (as in a parliamentary system). This is because, under veto bargaining, there is only variance in policy when the status quo is such that legislature’s proposal will not be vetoed. That is, veto bargaining (like our earlier model of presidential systems) gives the legislature less control over policy. Consequently, voters are less concerned about uncertainty over their legislators’ policy preferences under veto bargaining than in a parliamentary system. As a result increases in discipline have less marginal impact on the probability of election under veto bargaining than in parliamentary systems and so less discipline is imposed. Thus, our qualitative results are insensitive to the exact nature of the bargaining and, in particular, are consistent with veto bargaining.
6.3 Costly Affiliation

Now assume that a candidate must bear some cost $\delta > 0$ if she affiliates with a party. In this case, she will affiliate with $L$ if and only if

$$\text{Pr}(w = L | \alpha)(B - \alpha(x - \pi_L)^2) \geq \delta.$$ 

The LHS of this inequality is decreasing in $\gamma_d$ and in $(x - \pi_L)^2$. This means that in more left-leaning districts, candidates who have ideal points far from the platform are more willing to affiliate. This means that the party label has more informational content in districts that lean toward the other party.

This result also implies that an absolute majority of the members of a party will be from districts that lean toward that party, even though the potential members are spread uniformly across districts.

6.4 Party Polarization

The model we have presented misses an important aspect of reality, and of the conditional party government thesis—the role of changes in the polarization of the parties. Largely for technical reasons, we have treated party platforms as fixed. If platforms and discipline are both choice variables for parties, then the parties have two-dimensional strategy spaces. In our current model of party decision making, we would run into the problem of generic nonexistence of a Condorcet winner. Of course, this problem is not unique to our work.

Our guess is that our results will be robust so long as the parties diverge in equilibrium. There are several models that suggest divergence might be an equilibrium outcome in the environment we study. First, Snyder and Ting show that, with fixed levels of discipline, a sufficiently strong incentive for transmitting information can lead to divergence. Second, many models of platform competition with policy-motivated parties have divergence in equilibrium.

6.5 Institutional Change

Finally, our results compare the stable party systems in different political environments. A shock to the political structure will lead to an adjustment period. Integrating our equilibrium
story with a strategic account of the process of institutional change awaits future research.

7 Conclusions

In this paper, we presented a model in which party memberships and party discipline are jointly determined in equilibrium. The model is able to explain differences in party strength, both across different political systems and over time in the U.S.

While we believe that the results we have presented regarding party strength and ideological make-up are interesting in their own right, this paper is part of a larger project that attempts to explain the covariation a variety of legislative and electoral outcomes. While the project is too large to be presented in a single paper, it is useful to see how the logic developed above can be applied to other empirical findings in the study of comparative legislatures.

Among the most interesting empirical findings in comparative legislative research have to do with the relationship between party strength, the level of constituency service in which legislators engage, and the size of the personal incumbency advantage. Comparisons between presidential systems like the U.S.’s and parliamentary systems like the U.K.’s have shown presidential systems have less cohesive parties (Cain, Ferejohn, and Fiorina [9]), more constituency service (Cain, Ferejohn, Fiorina [9]), and a stronger incumbency advantage (Katz and King [21], Gelman and King [20]). Indeed, as a result of such comparisons, it is often thought that cohesive parties preclude high levels of constituency service and strong personal incumbency advantages. However, compared to 1950, today’s U.S. Congress has more cohesive parties (as discussed in this paper) and it has more constituency service (Fiorina [19], Fenno [18]) and stronger incumbency advantages (Gelman and King [20]). This presents a puzzle for traditional thinking.

In order to address these empirical regularities, in other work we extend our model to include the provision of constituency service (Ashworth [5], Ashworth and Bueno de Mesquita [6] and [7]). Politicians are endowed with different levels of skill and part of a voters’ decision calculus depends on his assessment of his incumbent politician’s ability relative to a challenger. Legislative outcomes provide voters with information about an incumbent’s ability, and they only reelect incumbents whom they believe to be high ability. This has two effects. First it
gives legislators an incentive to provide constituency service. Second, over time voters become more confident in their incumbents, giving rise to the incumbency advantage.

The causal mechanism underlying our account of differences in party strength also explains patterns of constituency service and the incumbency advantage. First, consider the comparison between presidential and parliamentary systems. The voter has to consider two factors when deciding whether to reelect an incumbent: the incumbent’s ideology (as in the current model) and the voter’s assessment of the incumbent’s ability to provide constituency service. As in the model in the current paper, voters put more weight on ideological concerns than on constituency service in a parliamentary system than in a presidential system because legislators have more control over policy in a parliamentary system. Consequently, an increase in constituency service has a greater effect on the probability of reelection in a presidential system. Thus, consistent with empirical findings, this argument predicts that there is more constituency service in presidential systems. Similarly, the incumbency advantage exists in this model because of voter concern over ability to provide constituency service. As such, again consistent with empirical findings, the incumbency advantage is expected to be larger in presidential systems than parliamentary systems.

We argue that the same model that accounts for the comparison between presidential and parliamentary systems can also account for the seeming contradictory over-time empirical trend in the U.S. Congress where strong parties have coincided with a high level of constituency service and a large incumbency advantage. The analysis in this paper showed that increases in competitiveness of district races account can explain heightened party strength. As competitiveness increased, so did the probability that greater discipline would swing an election. Similarly, when elections become more competitive a marginal increases in constituency service is more likely to swing an election because the decisive voter is more likely to be swayed by an increase in his beliefs about the incumbent’s ability. Hence, competitive elections imply an increases in constituency service and, for similar reasons, the incumbency advantage.
A Proof of Proposition 1

We have already show this for majority control; now we establish the result for floor control. All best responses are cutpoints since the voters’ payoffs are supermodular. As before, the voter in district \(d\) uses the cut-point

\[
  c_d = \frac{1}{2}(\mu_{\text{leg}|R} + \mu_{\text{leg}|L}) + \frac{\sigma^2_{\text{leg}|R} - \sigma^2_{\text{leg}|L}}{2(\mu_{\text{leg}|R} - \mu_{\text{leg}|L})}.
\]

Since the district-level elections are independent, we have

\[
  \sigma^2_{\text{leg}|R} - \sigma^2_{\text{leg}|L} = \frac{1}{n^2}(\sigma^2_R - \sigma^2_L)
\]

and

\[
  \mu_{\text{leg}|R} - \mu_{\text{leg}|L} = \frac{1}{n}(\mu_R - \mu_L).
\]

These quantities do not depend on the cutpoints used in other districts. These cutpoints do matter for the term \((1/2)(\mu_{\text{leg}|R} + \mu_{\text{leg}|L})\). This can be written as

\[
  (1/2)(\mu_{\text{leg}|R} + \mu_{\text{leg}|L}) = \sum_{k \neq d} F(c_k - \gamma_k)\pi_L + (1 - F(c_k - \gamma_k))\pi_R.
\]

Let

\[
  C = \left[\frac{n - 1}{2n} \pi_L + \frac{-\sigma^2_L(\bar{\pi})}{2n(\pi_R - \pi_L)}, \frac{n - 1}{2n} \pi_R + \frac{\sigma^2_R(\bar{\pi})}{2n(\pi_R - \pi_L)}\right].
\]

This set is convex and compact, and it contains every cut-point that a rational voter might use.

Let \(\phi : C^n \to C^n\) be the map given by

\[
  \phi_d(c, \alpha_L, \alpha_R) = \sum_{k \neq d} [F(c_k - \gamma_k)\pi_L + (1 - F(c_k - \gamma_k))\pi_R] + \frac{\sigma^2_R(\alpha_R) - \sigma^2_L(\alpha_L)}{2n(\pi_R - \pi_L)}.
\]

This defines a map \(c \mapsto \phi(c)\) whose fixed points are equilibria of the voting stage under floor control. \(\phi\) is continuous, so a fixed point exists by Brouwer’s theorem.

Let \(\alpha\) be the vector of party disciplines \((\alpha_L, \alpha_R)\). The equilibrium cutpoints, \(c^*(\alpha)\) satisfy the equation \(\phi(c^*(\alpha), \alpha) = c^*(\alpha)\). By the implicit function theorem, we have

\[
  (I - D_c\phi)D_\alpha c^* = D_\alpha \phi.
\]
The Jacobian $D_c\phi$ has a zero diagonal and negative off diagonal terms, and the matrix $I - D_c\phi$ has a positive dominant diagonal, since our assumptions imply that
\[
\frac{\partial \phi_d}{\partial c_k} < \frac{1}{n}
\]
for $d \neq k$. This has two implications. First, equation 4 has a unique solution given by
\[
D_\alpha c^* = (I - D_c\phi)^{-1}D_\alpha\phi.
\]
Second, $(I - D_c\phi)$ positive dominant diagonal implies that $(I - D_c\phi)e \gg 0$, where $e$ is the vector of all 1s. Thus $D_c\phi$ is a productive Leontief matrix, and the proof of Proposition 5.AA.1 of Mas-Colell, Whinston, and Green [?] implies that $(I - D_c\phi)^{-1}$ has all nonnegative elements. Since each $\phi_d$ is increasing in $\alpha_L$ and decreasing in $\alpha_R$, the proof is complete.

B Proof of Proposition 2

We will use the following result from Milgrom and Roberts [24].

**Theorem** Let each $\phi_i(x_i, x_{-i}, t) : [0, 1]^N \times T \to [0, 1]$ be continuous but for upward jumps in $x_i$ and nondecreasing in $x_{-i}$ and $t$, where $T$ is any partially ordered set. Then there exist greatest and least fixed points of $\phi$ for each $t \in T$, and these fixed points are nondecreasing in $t$.

The strategy of the proof is to construct a function whose fixed points are stable party systems, and then show that this function satisfies the conditions of the theorem. This will imply the result.

This function will be based on the optimal levels of discipline of the members, so we record some facts about those first. Fix an arbitrary $\alpha_R$. Let $\alpha^*(\Delta, \gamma)$ be the level of discipline most preferred by a candidate whose ideal policy’s squared distance from the party platform is $\Delta$ and who is running in a district where the median voter’s expected ideal point is $\gamma$. That is
\[
\alpha^*(\Delta, \gamma) = \arg\max_{\alpha_L} \left\{ F(c(\alpha_L, \alpha_R) - \gamma)(B - \alpha_L\Delta) \right\},
\]
where $\Delta = (x - x_L)^2$. Berge’s theorem of the maximum implies that $\alpha^*$ is an upper-hemicontinuous correspondence. The objective function of the candidate is differentiable and strictly logsupermodular in $(\alpha, -\Delta, \gamma)$. (To see this, just compute the cross-partial derivatives.) Thus the
strict monotonicity theorem of Edlin and Shannon [15] implies that if \((\Delta', \gamma') > (\Delta, \gamma)\) then 
\(\alpha^*(\Delta', \gamma') > \alpha^*(\Delta, \gamma)\).

The next step is to show that, for any discipline level \(\alpha'_{-p}\) and any potential coalition \(\mathcal{C}'\), there is a Condorcet winning level of discipline for party \(p\). Notice that this is the only step of the proof for which singleton-valuedness of \(\alpha^*\) is needed.

Let \(\alpha_{\text{med}}\) be the median of the \(\{\alpha^*(\Delta, \gamma)\}\). Now consider some \(\alpha < \alpha_{\text{med}}\). Let \(h = \{(\Delta, \gamma) \mid \alpha^*(\Delta, \gamma) = \alpha_{\text{med}}\}\). Since \(\alpha^*\) is singleton-valued and upper hemicontinuous, it is continuous. This implies that every \((\Delta', \gamma')\) is ordered (in the usual vector order) with respect to some \((\Delta, \gamma) \in h\). It is greater if and only if \(\alpha^*(\Delta', \gamma') < \alpha_{\text{med}}\) and is less if and only if \(\alpha^*(\Delta', \gamma') > \alpha_{\text{med}}\). (This is logsupermodularity.) If it is less, then we have

\[
  u(\alpha_{\text{med}}, \Delta', \gamma') - u(\alpha, \Delta', \gamma') > u(\alpha_{\text{med}}, \Delta, \gamma) - u(\alpha, \Delta, \gamma) \geq 0,
\]

so all candidates with \(\alpha^* > \alpha_{\text{med}}\) prefer \(\alpha_{\text{med}}\) to \(\alpha\), and \(\alpha_{\text{med}}\) is a Condorcet winner.

Now we construct our function. Define \(\mathcal{BR} : [0, \overline{\alpha}]^2 \to [0, \overline{\alpha}]^2\) by

\[
  \mathcal{BR}(\alpha'_L, \alpha'_R) = (\mathcal{BR}_L(\alpha'_L, \alpha'_R), \mathcal{BR}_R(\alpha'_L, \alpha'_R)),
\]

where

\[
  \mathcal{BR}_p(\alpha'_L, \alpha'_R) = \text{med}\{\alpha^*(\Delta, \gamma) \mid (\Delta, \gamma) \in \mathcal{P}(\alpha'_L, \alpha'_R)\}.
\]

By the previous result, this correspondence picks the Condorcet winning levels of \(\alpha\) for each party, given the affiliations implied by the \(\alpha'\). At a fixed point of this correspondence, the optimal choices recreate the status quo, which is stability. All that’s left is to argue that the correspondence has a fixed point.

Now we show that \(\mathcal{BR}\) is continuous but for upward jumps in \(\alpha_p\). It’s clear from Topkis’s theorem that \(\mathcal{BR}_p\) is increasing in \(\alpha'_{-p}\), so we can restrict attention to sublattices with fixed \(\alpha'_{-p}\).

To derive a contradiction, assume that there is an interval \(A = (\underline{\alpha}, \hat{\alpha})\) such that \(\inf \mathcal{BR}_p(A) > \mathcal{BR}_p(\hat{\alpha})\). This means there is a sequence \(\{\alpha^n\}\) such that \(\alpha^n \to \hat{\alpha}\) and \(\alpha^n \in A\) for all \(n\) and an \(\epsilon > 0\) such that \(\mathcal{BR}_p(\alpha^n) - \mathcal{BR}_p(\hat{\alpha}) > \epsilon\) for all \(n\). From the definition of \(\mathcal{BR}\), this means that
\( \alpha_{\text{med}}(\alpha^n) - \alpha_{\text{med}}(\hat{\alpha}) > \epsilon \) for all \( n \). Since the ideal \( \alpha \)'s lie in the compact set \([0, \overline{\alpha}]\), we can choose the sequence to be convergent. Finally, observe that \( \mathcal{P}(\hat{\alpha}) \subset \mathcal{P}(\alpha^n) \) for all \( n \).

Define \( \mu \) to be the restriction of the measure induced by \( f^p \) to \( \mathcal{P}(\hat{\alpha}) \), and let \( \mu^n \) be the restriction of this measure to \( \mathcal{P}(\alpha^n) \setminus \mathcal{P}(\hat{\alpha}) \). Let \( \phi = \mu \circ (\alpha^*)^{-1} \) be the push-forward measure on ideal levels of discipline, and define \( \phi^n \) similarly. The median ideal discipline level at \( \alpha^n \) satisfies

\[
\phi(\alpha_{\text{med}}^n) + \phi(\alpha_{\text{med}}^n) = \frac{1}{2} (\phi(\overline{\alpha}) + \phi^n(\overline{\alpha})).
\]

Let \( \eta^n = \frac{1}{2} \phi^n(\overline{\alpha}) - \phi^n(\alpha_{\text{med}}^n) \). Then

\[
\lim_{n \to \infty} \phi(\alpha_{\text{med}}^n) = \lim_{n \to \infty} \frac{1}{2} \phi(\overline{\alpha}) + \eta^n = \frac{1}{2} \phi(\overline{\alpha}),
\]

where the second equality follows from \( \phi^n(\alpha) \leq \phi^n(\overline{\alpha}) \to 0 \). We claim that \( \phi \) is continuous on \([0, \overline{\alpha}]\), which contradicts \( \lim_{n \to \alpha_{\text{med}}^n} \alpha^n_{\text{med}} > \alpha_{\text{med}} \).

Now we prove the claim. To see that \( \phi \) is continuous on \([0, \overline{\alpha}]\), note that each interior \( \alpha^* \) satisfies the Edlin-Shannon conditions for strict monotonicity. We claim that \( (\alpha^*)^{-1}(\alpha) \) has Lebesgue measure 0 for any interior \( \alpha \). Given this claim, the result follows since there can be no interior atoms. To prove the claim, observe that if such a set had positive measure, then it would contain an open set, and that open set would contain two points that are strictly ordered, contradicting the strict comparative statics result.

Finally, the result is trivial if \( \alpha_{\text{med}} = \overline{\alpha} \).
References


