AN ORGANIZATIONAL THEORY OF POLITICAL PARTIES

SPENCER PANTOJA

ABSTRACT. Political parties are ubiquitous features of political systems around the world, topping the lists of the largest and most powerful organizations in the world. Yet, political economy does not use organizational economics to build theories of parties, and organizational economics does not have any models tailored to the unique characteristics of political parties, as opposed to private profit-maximizing firms. In this work, I build an organizational economics style model of political parties, tracing through the emergence of the modern political party through an adapted theory of the firm lens that incorporates political-party-specific features. I show how party leadership can wield effective power even when the members that face their power can always depose them. I then use this model to inspire a reduced-form for a multi-party model of coalition formation. I show how having an "organizational" conception of political party-coalitions changes the logic of ideological conditions which make parties strong and what party systems can be stable.

STANFORD UNIVERSITY, DEPARTMENT OF ECONOMICS

E-mail address: `spantoja@stanford.edu`.

Date: March 23, 2024.

First, many thanks to my advisor Steven Callander for his many helpful comments and frequent meetings on this project. Second, I thank Bob Gibbons for his mentorship, teaching, and the spurring of my interest in organizational economics and theory of the firm. I am also gracious for the feedback from various workshop and conference audience participants. And I would like to thank (in alphabetic order) Yunus Aybas, Dan Baron, Jon Bendor, Peter Buisseret, Martin Castillo Quintana, Sylvain Chassang, Álvaro Delgado-Vega, Ben Golub, Marina Halac, Matt Jackson, Ravi Jagadeesan, Weijia Li, Zhaotian Luo, and Takuo Sugaya.

1. Introduction

Ronald Coase (1937, 1960) and Oliver Williamson (1985) taught us that if there were no transaction costs, then we would not see a world full of organizations. They also emphasized that if organizations were perfect solutions to transaction costs, we would not see markets. Kenneth Arrow (1951) and the "chaos theorems" (McKelvey, 1976; Schofield, 1978) taught us that political outcomes could not generically be determined by preferences and voting rules alone, and almost any decision would be under threat of contestation. In this paper, we ask what happens when both of these problems are relevant?

At this intersection lives the political party, which is our object of interest for this analysis. Our central motivating question is "What are political parties, and why do they exist?" There is a surprising variation in answers to these questions in the political science literature, ranging from "parties do not matter at all" to "the members of a party can be treated as one unified decision maker". Making progress on the theoretical boundary of these questions has a number of important implications. First, it is often contestable how to make sense of correlations in this literature (Krehbiel, 1993)—do party members vote with each other because the party is strong and is forcing them to, or is it just a sign that the party members inherently agree? And if we think consensus fosters party strength, how do we disentangle this? Second, many national party systems are experiencing unprecedented environmental conditions (like the media landscape), and policymakers have been interested in changing political institutions in radical ways (like fostering multiparty competition).

This paper is the first to take the Coasian approach to parties, which comes with a unique set of questions and conditions that distinguishes it from the canonical firm case. **Firstly**, in democracies, parties are not *owned*; instead, the leadership that wields party power is *elected* and can be *deposed* by its own members. **Secondly**, preferences in firm models (with participation constraints) tend to reduce down to "How can I make the most money at the lowest effort?" Political ideology entails a much different preference space, where "Do there even exist goals that we can all agree on?" is a much more demanding question that depends on the identities of the individuals. And **thirdly**, for firms, we consider boundary decisions on a single firm-by-single worker basis, where non-integration means the worker gets absorbed into this large external unmodeled environment called *the market*. If there are 100 senators and one party has 51 of them as members, it matters much more to anticipate what happens to a member after they leave the party (what other party might they join, will they take anyone else with them?). Thus, we need to use the tools of cooperative game theory to understand the *coalitional stability* of this entire system of organizational boundaries.

Building on the **first** and **second** points, our model shows how party leadership can flex its strength even when its authority can be so easily stripped away and the party exhibits a big tent of ideological diversity. Intuitively, legislation is *incomplete in the same way contracts are*—bills are written and passed over time, where it is impossible to write one piece of legislation that specifies what to do in every possible contingency for every possible issue. Thus, it is neither easy nor impossible for a group legislators to form a *long coalition*—that is, cooperate with each other with a mutually-beneficial long-term horizon in mind, in spite of their more myopic tendencies. The range of plausible outcomes neither achieves the Pareto ideal of the Coase theorem (1960) nor constrained by one-shot non-cooperative self-interest. Working within this interior range (a world with transaction costs), the party is a governance structure (formalized in a Williamson (1971) sense) that facilitates long-run cooperation—and so even when many members would like to block a particular decision by leadership, they exhibit forbearance, in fear of shattering the skeleton that their otherwise tenuous coalition is built upon.

Building on the **second** and **third** points, we develop the first model of coalition formation for parties that takes something other than just ideology as a relevant input—organizational capacity. We often take for granted a bijection between the partisan cleavages of society and the ideological ones, but that is often not the case[1]. These counterexamples to the pure ideology models can be understood through organizational capacity—the parties that "should have fractured" were well-oiled machines that have been working together for many decades, while anecdotally those members who "should of deserted" according to ideology were worried about being able to achieve the same level of effectiveness in a new party/coalition. We have a great amount of evidence from the firms literature regarding how seemingly similar organizations can vary widely in how they actually perform, the patterns these performance levels exhibit, and theoretical mechanisms underlying these differences.

Connections to Literature. This paper follows a line of work trying to resolve the paradoxes that the chaos theorems present. The most significant of these is the concept originating from Shepsle (1979), structure-induced equilibrium (SIE). The SIE approach uses institutional rules, such as jurisdictions and committees, to pin down an equilibrium where preferences alone could not induce one. Rather than focusing on the rules of the game, our paper addresses the chaos results essentially by weakening the deviation concept. The concept is still stronger and more cooperative than Nash, but instead of coalitions that can deviate as perfect hive minds, there are inefficiencies to cooperation. In particular, some coalitions may be more efficient at cooperating than others. To this end, we adapt the semi-cooperative solution concept from Ali and Liu (2020).

A work in the SIE literature that is particularly closely related to our paper is the essay of Weingast and Marshall (1988). In it, the authors also propose a theory of the firm approach to legislative bargaining, and in great detail describe the same transaction cost problems that we consider. However, the focus of their paper is on committee structures and procedural rules as a way to address these transaction difficulties. They explicitly choose to exclude parties from the scope of the essay, only in the end suggesting it as a possible avenue of future research. The authors seem to be taking the Greif (2006) perspective on organizations—like institutions, modeled such that players treat them as exogenous "rules of the game"

[1]For three examples, see the "New Deal Coalition" American Democrats (1930-1960) (Noel, 2013), the Peronist Party in Argentina (Levitsky, 2003), and the PRI in Mexico (Langston, 2017).

in the North (1990) sense. Weingast and Marshall (1988) explicitly claim that there is no relevant analogue to vertical integration in the legislative context.

In contrast, our approach mirrors Baker, Gibbons and Murphy (2002) in defining the boundary of the (party) organization through vertical integration and then analyzing possible relational contracts (which are like institutional rules) given an organizational configuration. This approach and our use of an imperfectly cooperative solution concept allows us to analyze how these rules/relational contracts survive even when a majority coalition has the power to unravel them—a self-noted limitation of Weingast and Marshall (1988, p.159) and SIE in general (Aldrich, 1995, Chapter 7). This is an important advantage of our model—in the SIE literature, transaction difficulties are informally used to justify the existence of ex ante exogenous institutional rules, which constitute the starting point of the formal analysis. In our framework, we formalize all the way back to the transaction costs environment, allowing us to endogenize our institutional arrangements (parties and possibly even rules) given known conditions on commitment and information problems.

This paper also connects to the literature on trying to formally model political parties. Snyder and Ting (2002) focus on the outward face of the party—in the electoral arena—with their model of "parties as brands". Snyder and Ting (2002) actually has a close counterpart in the firm theory literature—Kreps (1990) models the firm as a centrally managed "transferable reputation". Our model simply covers a different scope, the inward face of the party.

Two other models of political parties, Jackson and Moselle (2002) and Levy (2004), focus on legislative bargaining and, in particular, parties as a solution to Pareto inefficiency within decisive coalitions. Jackson and Moselle (2002) restricts parties to implementing their Nash bargaining solutions. So, for example, the canonical split-the-dollar game could have stable parties because each party would only be allowed to offer the uniform equal split.

Levy (2004) restricts coalitions to their Pareto frontier by assumption. Levy embeds the game into a citizen-candidate game[2] under plurality rule and sincere voting. This setup does not generically have a stability existence problem, as the author notes (footnote 14). Rather, the focus of the paper is on an interesting necessary condition on the policy space (multidimensionality) for effective parties.

Another particular contrast is that in both Jackson and Moselle (2002) and Levy (2004), their models collapse everything down into one policy decision and vote, with parties creating commitment up to the resolution of said decision. However, empirically, commitment to the party is never permanent, and many decisions must be resolved before other policy questions are even known. In the political science literature, it is common to refer to parties synonymously as "long coalitions" (Aldrich, 1995; Bawn and Rosenbluth, 2006) with their time persistence being a defining property. Thus, it is valuable to capture the dynamic decision of party membership when those opportunities to shed this commitment arise dynamically and policies are determined sequentially.

[2]See Osborne and Slivinski (1996) and Besley and Coate (1997).

The dynamic setting is also important because it helps us justify pervasive transaction costs—if everything is resolved in a single bargaining protocol, it is hard to imagine a strongly Pareto dominated outcome being the final resolution. But if one considers many subdecisions that must be resolved sequentially, there may "locally" not be any Pareto dominance with respect to each subdecision even though "globally" there is when aggregated into one representative decision.

The most prominent theory of party in the political science literature is conditional party government (CPG) theory, developed by Rohde (1991) and Aldrich (1995). It states that a party will be the strongest when these three conditions hold: (i) they strongly oppose the position of the other side/party; (ii) they have homogeneous preferences, such that there are many policies on which they can find unanimous consensus; and (iii) they have the power to influence the agenda, which usually means that they are a majority (Aldrich, 1995, p.204-205). I argue that this is closely related to the logic of Coase (1937, 1960). Conditions (i) and (ii) together essentially state that bad political outcomes are plausible, even though there are policies that Pareto dominate them for the party members. Condition (iii) says that those better outcomes are achievable by the concerted behavior of those party members. Together, they say that transaction costs are a necessary condition for party organizational strength—Coase's thesis.

We present an example of how transaction costs interact with stability, which, from a technical point of view, would be a basic example in the cooperative game theory tradition even by von Neumann and Morgenstern (1944) standards[3]. The only intention is to give new interpretations that correspond to this setting, using labels to suggest what will be an underlying logic for the models to come.

Example—Split-the-Dollar with Transaction Costs. We consider an alternative version of the split-the-dollar game. There are three players $N := \{1, 2, 3\}$, which we refer to as the three legislators, with utility quasi-linear in money. There is a total of one dollar of surplus in this world that feasibly could be allocated to these three legislators. The only action of each legislator i is to simultaneously submit a proposal (P, x_P). First, $P \subseteq N$ is a proposed party that includes i, or $P \cap \{i\} \neq \varnothing$. Second, $x_P \in \mathbb{R}^P$ is an allocation of dollars to the members of the party. A proposal (P, x_P) is implemented if and only if all members of P submit the exact same proposal, otherwise each member that proposed (P, x_P) is in a party by themselves and is allocated nothing $x_i = 0$.

Any party of one does not receive any of the dollar, $|P| = 1 \implies x_P = 0$. Think of this as a normalization, each legislator i has a base utility $\bar{u}_i$ that they can achieve without cooperation $P = \{i\}$, and we have normalized everyone's utility as $u_i' = u_i - \bar{u}_i$. Now, if the legislators join a party together—that is, in this paper, vertically integrate by some sense of the term—then they can reduce their transaction costs and achieve surplus higher than the sum of their $\bar{u}$'s.

[3]Specifically, one could just write this game with a transferable utility characteristic function. A structurally similar example can be found in Example 7.1 from the book of Ray (2007).

In particular, assume that each coalition P such that $|P| \geq 2$ can split a fraction λ_P of the dollar amongst themselves $\sum_{i \in P} x_i \leq \lambda_P$, where $\lambda_P \in [0,1]$. Thus, the difference $1 - \lambda_P$ can be thought of as the *transaction costs of party* P, seeing that for $|P| \geq 2$ no other players get any of the dollar, and thus this difference is akin to money burnt from the system. We define a *core* of this game to be a strategy profile of proposals $\left((P, x_P)_i\right)_{i \in N}$ such that there is no coalition C such that all members of C can benefit from a coordinated cooperative deviation $\left((P', x'_P)_i\right)_{i \in C}$ fixing the strategy profiles of $N \setminus C$. We assume that there exists some $P \subset N : |P| = 2$ with $\lambda_P > 0$.

Theorem. *Denote the three parties of size two as* $P_1 \neq P_2 \neq P_3$ *with* $|P_1| = |P_2| = |P_3| = 2$*, and without loss of generality assume* $\lambda_{P_1} \geq \max\{\lambda_{P_2}, \lambda_{P_3}\}$*. Then, a core to this game exists if and only if* $\lambda_{P_1} \geq \lambda_{P_2} + \lambda_{P_3}$ *or* $\lambda_N \geq (\lambda_{P_1} + \lambda_{P_2} + \lambda_{P_3})/2$*.*

First, to discuss the condition $\lambda_N \geq (\lambda_{P_1} + \lambda_{P_2} + \lambda_{P_3})/2$, assume that the other condition $\lambda_{P_1} \geq \lambda_{P_2} + \lambda_{P_3}$ does not hold. If we want to check if the party $P = N$ is stable, then the transaction costs of the whole party must be much lower than any possible factionalization. This corresponds to the story of our first act, a single larger party attempting to stay together and party leadership trying to impose its will in the face of credible threats of internal sedition. The larger party must be much better at reducing transaction costs than its possible internal factions.

Second, suppose that condition $\lambda_N \geq (\lambda_{P_1} + \lambda_{P_2} + \lambda_{P_3})/2$ does not hold. Thus, our focus is the condition $\lambda_{P_1} \geq \lambda_{P_2} + \lambda_{P_3}$, which underlies the story of our second act. Note that there is no core existence if all two-legislator parties are equally good at economizing their transaction costs $\lambda_{P_1} = \lambda_{P_2} = \lambda_{P_3}$. In that case, we essentially get the traditional split-the-dollar game without core stability. This environment where all coalitions are similarly good/perfect at economizing their transaction costs is a generic example of when the chaos theorems have bite.

However, we do have core existence when some party is far superior at economizing its transaction costs over its competitors, $\lambda_{P_1} \geq \lambda_{P_2} + \lambda_{P_3}$. Think of P_1 as the incumbent party that has been around for a hundred years, with long-lasting relationships and a strong reputation. Reputation and trust are often vital for reducing transaction costs and are far from automatic for entrants, especially when they can hardly operate while the incumbent is in power, as in politics. In the real world, using the U.S. context, P_1 is like the Democratic or the Republican Party, and P_2 or P_3 is like a hypothetical libertarian major party.

If neither of these conditions hold, we have instability. Think about the earliest eras of American politics, when all parties were relatively new entrants without reputations built on long histories of stability. There was a wide variety of parties that separated themselves on different policy dimensions. For example, there were the Federalists (scope of government power), anti-Jacksonians (Democratic-Republicans opposing Andrew Jackson), Anti-Masonics (against conspiratorial secret societies), Know Nothings (nativists), Free Soil (antislavery), Labor party (labor rights), and many more. All who held seats in Congress. One can see in FIGURE 1 how much chaos there was in the US party system early on relative to our recent history. This early era is much more like what one would expect based on the chaos theorem results—constant shifting of alliances and reformulating coalitions.

Persistent Productivity Dispersion. It is well documented in the empirical industrial organization literature that there are large persistent productivity differences between firms even after accounting for a number of on paper factors (Syverson, 2011). For example, Syverson (2004) found that concrete producers in the 90th percentile of productive efficiency had made on average twice as much concrete per the *same amount of inputs* as producers in the 10th percentile. Hsieh and Klenow (2009) find a ratio of around five in China and India for those same 10th and 90th percentiles. If we see such dispersion for profit-seeking organizations, why not for political ones? If something as straightforward and stable as concrete sees this kind of dispersion, one would think that political parties would see even larger gaps.

Gibbons and Henderson (2012) suggest that this could be explained through path dependence of relational contracts–the most (residually) productive firms are the ones with the best relational contracts. Because of equilibria multiplicity, specifically accompanied by problems of clarity and commitment, bad relational contracts are very challenging to move out of. The productivity dispersion is in spite of similar on-paper fundamentals, and thus could be interpreted as heterogeneous transaction costs—the low productive firms are leaving surplus "on the table"[4]. This is certainly the case in Gibbons and Henderson (2012)—their framework considers firms playing the exact same kinds of games with employees but ending up in different, Pareto-rankable equilibria which lead to performance differences.

If a legislator is in a party that is productive, has good relational contracts, and is efficient at economizing transaction costs, it may be risky for them to defect and try to start a new party. The new one they start may not be nearly as efficient, even if on paper it should be just as good or better. Typically, these cooperative game theory approaches to political coalition formation presume a homogeneity of efficiency across coalitions, but this evidence on the distribution of firm productivities suggests that even amongst political agents symmetric in their underlying traits, this is not likely to be the case. Based on these studies, it seems unlikely that each and every possible political coalition would end up splitting "an equal-sized dollar".

In the firms literature, this residual productivity dispersion that is left unexplained by the fundamentals of the environment and actors within the environment is considered a puzzle by many researchers. However, for parties, the presence of this dispersion actually resolves the paradox created by the chaos theorems.

Political Science. There is strong evidence that party power has a significant impact on outcomes. A recent paper by Canen, Kendall and Trebbi (2020) finds that party power contributed to 40% of political polarization in the United States House of Representatives between 1977 and 1986.

The political scientist Hans Noel (2013) wrote a book that highlighted the differences between political parties vs. ideologies, which today many Americans use

[4]Note the productivity differences do not seem to be explained by a shifting of surplus to employees—most papers actually show a *positive* relationship between a workforce's rent share and firm productivity (Card et al., 2018).

to refer to each other interchangeably. For example, he notes

> In April of 1947, the American Institute of Public Opinion Surveys asked a sample of voters a question it had been asking off and on for years: "It has been suggested that we give up the present Republican and Democratic parties and have two new parties—one for the Liberals and one for the Conservatives. Would you favor this idea?" Today, this question might seem absurd...[The question] reminds us that, within some readers' lifetimes, ideology and political parties were two separate ways of organizing political conflict.
>
> —Noel (2013, p.12)

In particular, Noel (2013) highlights the awkward coalition of segregationist southern Democrats and northern New Deal/labor-union supporting Democrats. Noel (2013) provides empirical evidence that the Southern Democrats, along with the Northern Republicans, would have made a more ideologically coherent party along an agenda of conservatism. However, this Democratic Party was wildly successful–from the 73rd Congress (1933-1935) to the 88th Congress (1963-1965)[5], the Democrats won control of the Senate 14 out of 16 times. The only time we have coalitional stability in the chaos theorems is with a unique configuration on unidimensional single-peaked preferences. How could the Democratic Party be so successful and persistent, even though there were much more sensible majority coalitions based on ideological preferences? Why did the Southern Democrats and Republicans not have a profitable coalitional deviation?

Noel (2013) claims that the Southern Democrats refused to work with the Republicans for so long because of the distrust that continued after the Civil War. In game theoretic terms, "distrust" is commonly translated as, for example, two players both playing D in a repeated prisoners dilemma game even when the discount factor is sufficiently high to support (C, C). As we know, (C, C) Pareto dominates (D, D), and so distrust entails significant transaction costs. When it comes to vote trading, the losers of today's policy vote must trust that the other side will compensate them for it with a win later on a different issue. Our paper is the first to be formally able to incorporate concepts like "trust" in a party stability analysis.

Outline of Paper. First, in Sections 2, 3, 4, and 5 we model the single-party problem. We generate our transaction inefficiency through asymmetric information on legislators' preferences and a lack of commitment power on how legislators will vote on future issues. There are three species of actors: legislators, lobbyists, and the whip. The legislators and lobbyists have the canonical interpretation as in the legislative bargaining literature. The Whip is meant to represent "the party organization"—endowed powers and rights from lobbyists and legislators vertically integrated under one central decision maker. As in many Williamson (1985) examples, the Whip is not "technological" in the sense that it does not change what outcomes

[5]Elections that occurred before the passage of the Civil Rights Act, which converted many segregationist Democrats to the Republican Party.

would be implementable by a social planner, but still because of the transaction difficulties, the presence of Whip can make all players better off in equilibrium.

This first setup is similar in style to the models in the organizational economics literature. The main contrast with most organizational economics models is how power is distributed. In particular, instead of principals and agents that are effectively dictators over their own decision domains, we will consider structures like majority rule where there are many possible decisive coalitions that can implement the final outcome. And thus, we apply an appropriate cooperative solution concept for this game.

Second, in Section 6, we take the stylized facts from the previous sections to model an environment with many parties. We analyze the conditions for party power and persistence. We show how limited our predictive power can be when there is uncertainty about the organizational details. This second setup is structured more in the traditions of political economy and cooperative game theory literatures.

2. The One-Period Model

In this section, I present a model of legislative decision-making influenced by a single imperfectly-informed lobbyist. It is a simple and stylized game intended to be embedded into the repeated game framework in Section 3, not to be taken as a standalone model to generalize all kinds of interactions between money and voting.

There is a group of n (odd) legislators, denoted $N \coloneqq \{1, \ldots, n\}$, who vote on policy proposals x located on the real line $\mathbb{R}$. If a policy proposal does not pass, the status quo is implemented, which we normalize to zero $x_{\text{sq}} = 0$. The legislators have single-peaked preferences and are *expressive* voters. This means that their utility only directly depends on how they cast their ballot, not on the final outcome or how any other legislators voted. Voting "No" on any issue x has utility equivalent to voting "Yes" on the status quo $x = x_{\text{sq}} = 0$. We denote legislator i's vote as $v_i \in \{0, 1\}$, with $\mathbf{v} \equiv (v_1, \ldots, v_n)$, where 0 is "No" and 1 is "Yes".

We assume that for any proposal to pass over the status quo, we need unanimity. The interpretation of this assumption is that the n legislators that we are modeling comprise a slim majority in their legislature, and hidden in this game is a $n-1$ opposition whose behavior we have taken as fixed and always voting against x.

The lobbyist L is right biased and outcome oriented, always preferring larger x all else being equal and only getting utility from x if the policy passes. To influence policy making, the lobbyist runs an *issue campaign* $(x, \mathbf{b}) \in \mathbb{R} \times \mathbb{R}^n_{\geq 0}$, where x is the policy that they are promoting, and $\mathbf{b} \equiv (b_1, \ldots, b_i)$ is the amount of money spent on the campaign targeted at each legislator's district. The lobbyist's utility function has the following form:

$$u_L\left(x, \mathbf{b}, \mathbf{v}\right) = \beta x \cdot \mathbb{1}_{\{\mathbf{v}=\mathbf{1}\}} - \sum_{i \in N} b_i, \tag{2.1}$$

for some parameter value $1 < \beta < n$.

The legislators to vote for policies close to their ideal point, and to vote for policies that were promoted in their district, with a legislator i of ideal point $\theta_i \in \mathbb{R}$

having utility function

$$u_i\left(x, \theta_i, b_i, v_i\right) = -|xv_i - \theta_i| + b_i v_i. \tag{2.2}$$

As in Snyder (1991), these ideal points and policy preferences should not necessarily be interpreted as what each legislator believes in their heart of hearts. It is intended to summarize both their internal personal preferences and the many external influences, like the responses of their voting constituencies and effects on their already well-established donor relationships. As for the preferences for promoted policies, this comes from the increased popularity for stating and voting for policies that are now suddenly more popular due to the campaigning of the lobbyist.

At the beginning of the game, Nature draws an *issue topic* from a common prior distribution, which can be summarized as a vector of legislator ideal points $\boldsymbol{\theta} \equiv (\theta_1, \ldots, \theta_n)$, with each ideal point θ_i privately observed by legislator i. The interpretation of this is something happens in the world, whose timing is largely out of the control of the elected officials, that brings a topic to the forefront of everyone's attention and calls for the government to address it. This may be something like a school shooting, one country invading another, or a groundbreaking investigative journalism article about the treatment of immigrants. And for the purposes of this model, all we need to know about the issue is how the legislators feel about it relative to the status quo and each other, or $\boldsymbol{\theta}$.

Public information about the state $\boldsymbol{\theta}$ is represented by a public signal $\mathbf{s} \equiv (s_1, \ldots, s_n)$. We consider a joint distribution $(\boldsymbol{\theta}, \mathbf{s})$ of the following form: there exists $\underline{k}, \bar{k} \in \mathbb{Z}$, $\underline{k} < \bar{k}$ and $\sigma > 0$ such that $s_i \in \{(2k+1)\,\sigma\}_{k=\underline{k}}^{\bar{k}}$ and $\theta_i \mid \mathbf{s} \sim \mathcal{U}\left[s_i - \sigma, s_i + \sigma\right]$. Essentially, we have grid, within each grid interval θ_i is uniformly distributed, and the signal tells you which element of the grid θ_i is in. It is similar to a signal with uniform noise over a uniformly distributed state, but reduces some of the technical casework, like edge effects. We allow $\mathbf{s}$ to have any distribution on this support. This public signal on each legislator's ideal point is meant to be an aggregate of the polling in the legislator's districts, campaign finance records of their major donors, their past public statements on the issue, etc.

The order of events is (1) the issue topic is drawn and privately observed, (2) the public signal is drawn and observed, (3) the lobbyist picks an issue campaign, publicly observed, and finally (4) the legislators vote on the issue.

Now, we are going to do backwards induction. By (2.2), given an issue campaign $(x, \mathbf{b})$, and an issue topic $\boldsymbol{\theta}$, breaking indifference towards voting "Yes", a legislator i should vote "Yes" if and only if

$$b_i \geq |x - \theta_i| - |\theta_i|. \tag{2.3}$$

The lobbyist's objective problem, given (2.1) and (2.3), can be written as

$$\max_{x, \mathbf{b}} \left\{ \beta x \cdot \mathbb{P}\left(b_i \geq |x - \theta_i| - |\theta_i| \text{ for all } i \in N \mid \mathbf{s}\right) - \sum_{i \in N} b_i \right\}. \tag{2.4}$$

Now, we present two lemmas on how the lobbyist optimizes (2.4).

Lemma 2.1. *For all signal realizations* $\mathbf{s}$*, the optimal policy for the lobbyist* $(x, \mathbf{b})$*, if not a corner* $(\mathbf{b} \neq \mathbf{0})$*, will be such that the policy passes with certainty,* $\mathbb{P}(\mathbf{v} = \mathbf{1}) = 1$*.*

Lemma 2.2. *If* $\{s_i\}_i \cap \{\sigma\} = \varnothing$*, there exists an optimal policy for the lobbyist* $(x, \mathbf{b})$ *will be such that the policy passes with certainty,* $\mathbb{P}(\mathbf{v} = \mathbf{1}) = 1$*.*

The intuition of Lemma 2.1 is just based on convexities and complementarities of (2.4)——if there are positive returns to increasing spending on one of the legislators b_i, increasing b_i also increases the probabilities other legislators are pivotal (i would be pivotal if for all $j \neq i$, $v_j = 1$). This increase in pivot probabilities increases expected returns to spending on other legislators, and thus the solution ends up being a corner for any fixed x. Then, Lemma 2.2 puts a floor on the "stakes" of any proposal $x \neq 0$, because the linearity, such that we eliminate the uncertain solution. Our restriction to the certainty solution is for tractability and parsimony, and we conjecture that the broad arguments to follow would hold without it.

Theorem 2.3. *If* $\{s_i\}_i \cap \{\sigma\} = \varnothing$*, then the lobbyist's problem* (2.4) *can be reduced to*

$$\max_x \left\{ \beta x - \sum_{i \in N} \max\{0, x - 2(\max\{s_i - \sigma, 0\})\} \right\}. \tag{2.5}$$

If you let $Z(\mathbf{s}, \beta) \coloneqq \sup\{z : |\{i : s_i < z\}| \leq \beta\}$ *then the optimal solution* $(x^*, \mathbf{b}^*)$ *that corresponds to* (2.5) *is*

$$x^*(\mathbf{s}, \beta, \sigma) = \begin{cases} 2(Z(\mathbf{s}, \beta) - \sigma) & \textit{if } Z(\mathbf{s}, \beta) > 0 \\ 0 & \textit{else.} \end{cases} \tag{2.6}$$

$$b_i^*(\mathbf{s}, \beta, \sigma) = \begin{cases} 2(Z(\mathbf{s}, \beta) - \max\{s_i, \sigma\}) & \textit{if } Z(\mathbf{s}, \beta) > \max\{0, s_i\} \\ 0 & \textit{else.} \end{cases} \tag{2.7}$$

Moreover, this solution is unique if $\beta \notin \mathbb{N}$*.*

Proof. The expression (2.5) follows from Lemma 2.2, by substituting into the spending term $\mathbf{b}$ the unique minimal spending amount such that any policy x passes with certainty. Then, (2.6) follows from the fact that for $x < x^*$, the derivative of (2.5) with respect to x is at least $\beta - \lfloor \beta \rfloor \geq 0$ (which holds strictly if $\beta \notin \mathbb{N}$), and for $x > x^*$, the derivative of (2.5) with respect to x is at most $\beta - (\lfloor \beta \rfloor + 1) < 0$. Intuitively, for any move of x to the right, the lobbyist gets benefit β, but has to compensate each holdout for the move. At x^*, the number of need-to-corrupt holdouts exceeds β. □

The legislators with $0 < \theta_i < x^*$ end up with utility 2σ higher than the baseline in which there is no uncertainty about their type $\sigma \to 0$ with the same θ distribution. And such, we refer to 2σ as "the moderates' information rent". The legislators with

$\theta_i < 0$ are unaffected by σ because absolute value preference means we can treat all $\theta_i < 0$ as-if $\theta_i = 0$ and make them exactly indifferent with probability one. However, legislators with $\theta_i > x^* + 2\sigma$ end up with utility 2σ lower than the baseline without uncertainty. The more ideologically extreme legislators lose from the existence of these information rents, since the rents attenuate the policies that proposed and passed. This latter loss of the ideologues dominates the gain of the moderates if x^* is left of the median ideal point.

3. The Repeated Game Setup

In this section, I extend the one-period model into a repeated game to talk about the role of parties. It is not an exact carbon copy, rather in each period two versions of the one-period game will be present—the party channel and the non-party channel. For parsimony, I assume that only one channel can be active at a time, but this could be relaxed with a little effort on the proof details.

The time environment is infinite-horizon and discrete, with $t = 1, 2, \ldots,$ and all players are expected discounted utility maximizers.

Similarly to Section 2, there is a group of n (odd) legislators, denoted $N :=\{1, \ldots, n\}$, who vote on policy proposals x_t located on the real line $\mathbb{R}$ in each period t. If the policy proposal does not pass, the status-quo is implemented, which we normalize to zero $s = 0$. The legislators again have single-peaked preferences within each period and are expressive voters. We denote legislator i's vote in period t as $v_{it} \in \{0, 1\}$, with $\mathbf{v}_t \equiv (v_{1t}, \ldots, v_{nt})$, where 0 is "No" and 1 is "Yes".

We maintain the unanimity assumption—for a policy x_t to pass, it must be that $\mathbf{v}_t = \mathbf{1}$, all n legislators must vote for it.

There is a single Lobbyist L who is ϕ_t-biased with $\phi_t \in \{-1, 1\}$ and outcome oriented, always preferring larger $\phi_t x_t$ all-else-being-equal and only getting utility from x_t if the policy passes. The Lobbyist has two ways of influencing policy—they can choose to use the non-party channel or the party channel. We assume that ϕ_t is publicly known during the period t.

If the Lobbyist chooses the non-party channel they may run an issue campaign $(\hat{x}_t, \mathbf{b}_t) \in \mathbb{R} \times \mathbb{R}^n_{\geq 0}$, where $\hat{x}_t$ is the policy that they are promoting, and $\mathbf{b}_t \equiv (b_{1t}, \ldots, b_{nt})$ is the amount of money spent on the campaign targeted at each legislator's district. When the Lobbyist chooses the non-party channel, their within-period utility function has the following form

$$u_L(x_t, \mathbf{b}_t, \mathbf{v}_t) = \beta \phi_t x_t \cdot \mathbb{1}_{\{\mathbf{v}_t = \mathbf{1}\}} - \sum_{i \in N} b_{it}, \tag{3.1}$$

for some parameter value $1 < \beta < n$.

In the repeated game, we anoint one of the legislators $W \in \{1, \ldots, n\}$ "the Whip". The Whip only has one additional action in the non-party channel, which is to propose a policy x_t to the floor. We assume that whenever the Whip is indifferent between proposing the Lobbyist's issue campaign x_t and maintaining the status-quo, the Whip proposes x_t, since obviously the Lobbyist could make this indifference strict in their favor with an arbitrarily small change in $(\hat{x}_t, \mathbf{b}_t)$.

Legislators like to vote for policies close to their ideal point and to vote for policies that were promoted in their district. Conditional on the non-party channel being chosen, a legislator i in period t with ideal point θ_{it} has a within-period utility function of

$$u_i\left(x_t, \theta_{it}, b_{it}, v_{it}\right) = -|x_t v_{it} - \theta_{it}| + b_{it} v_{it} \cdot \mathbb{1}_{\hat{x}_t = x_t}. \tag{3.2}$$

The non-constant ideal points θ_{it} are not meant to be interpreted as legislators constantly changing their mind on a single issue. Rather, different issue topics are arising in different periods, and thus $(\theta_{it})_{t\in\mathbb{N}}$ may reflect permanent stable preferences that are defined on many dimensions. The topic in period t may be tax cuts, and the topic in period $t+1$ may be marijuana legalization.

If the Lobbyist chooses the party channel the Lobbyist may donate a lump sum to "the party" of the legislators $D_t \in \mathbb{R}_{\geq 0}$. When the Lobbyist chooses the party channel, their within-period utility function has the following form

$$u_L\left(x_t, D_t, \mathbf{v}_t\right) = \beta \phi_t x_t \cdot \mathbb{1}_{\{\mathbf{v}_t = \mathbf{1}\}} - D_t. \tag{3.3}$$

The Whip takes over the game from here with three kinds of actions. First, the Whip is communicating with all players throughout the game. This is entirely cheap talk, nothing is verifiable, and there is no commitment. Second, the Whip chooses which policy $x_t \in \mathbb{R}$ to bring to the floor. And third, most importantly, after the vote, the Whip chooses an allocation of party dollars $\mathbf{d}_t \equiv (d_{1t}, \ldots, d_{nt}) \in \mathbb{R}^n_{\geq 0}$ within the parties' budget $\sum_i d_{it} \leq D_t$.

Again, legislators like to vote for policies close to their ideal point and to receive party dollars. Conditional on the party channel being chosen, a legislator i in period t with ideal point θ_{it} has a within-period utility function of

$$u_i\left(x_t, \theta_{it}, d_{it}, v_{it}\right) = -|x_t v_{it} - \theta_{it}| + d_{it}. \tag{3.4}$$

Party Channel—Whip Deposition. In the party channel, after the vote has been taken and after an initial allocation of party dollars has been made, the legislators vote to depose the Whip. The Whip is deposed by majority rule. If the Whip is deposed, the legislators determine a new allocation of party dollars $\mathbf{d}'_t$ with $\sum_i d'_{it} \leq D_t$ that supplants the old $\mathbf{d}_t$. This is done through the alternating random proposer game as described in Baron and Ferejohn (1989), with δ (as written in their paper) equal to one. Concretely, each round, a random legislator with a uniform probability is chosen to be the proposer. They propose an allocation $\mathbf{d}'_t$, and the legislators vote whether or not to approve it. If it gets majority approval, the allocation is implemented, and the process ends. If not, we go to another round with a new randomly drawn proposer.

Discounting. The legislators discount utility by factor $\delta \in (0,1)$. The discount rate for the Lobbyist does not matter for any of the technical results; they can be any $\delta_L \in [0,1)$. Interpretatively, it is best to think of the Lobbyist is a sequence of short-run players $\delta_L = 0$, and in some of our justifications for restrictions, we will adopt this convention.

Table 1. Timing of the Stage Game

	Both Channels	
	Non-Party	Party
1.	Nature draws issue topic $\boldsymbol{\theta}_t$ and public signal $\mathbf{s}_t^L$	
2.	$\swarrow$ Lobbyist chooses which channel to use $\searrow$	
3.		Nature draws Whip's signal $\mathbf{s}_t^W$
4.	Lobbyist chooses campaign $(\hat{x}_t, \mathbf{b}_t)$	Lobbyist chooses donation D_t
5.	The Whip chooses proposal x_t, and the legislators vote on it v_{it}	
6.		Whip allocates party dollars $\mathbf{d}_t$
7.		Whip deposition process unfolds

Type Space and Information. At the start of each period t, Nature draws an issue topic $\boldsymbol{\theta}_t \equiv (\theta_{1t}, \ldots, \theta_{nt})$ with each ideal point θ_{it} privately observed by legislator i and the Lobbyist's bias ϕ_t. Public information about the state $\boldsymbol{\theta}_t$ is represented by a public signal $\mathbf{s}_t^L \equiv \left(s_{1t}^L, \ldots, s_{nt}^L\right)$. We assume this draw is independent and symmetric across t, and independent across i—that is $\theta_{it} \mid \theta_{jt'} \cong \theta_{it}$ for all $(i,t) \neq (j,t')$ and $\theta_{it} \cong \theta_{it'}$ for all $t \neq t'$. But we do allow for asymmetry across i, and thus we can make comparisons among the ideologies of different legislators, i.e. if $\theta_{it} \succ_{FOSD} \theta_{jt}$, we can say that "$i$ is more right-wing than j".

We consider a joint distribution $\left(\boldsymbol{\theta}_t, \mathbf{s}_t^L\right)$ of the following form: there exists $\underline{k}, \bar{k} \in \mathbb{Z}$, $\underline{k} < \bar{k}$ and $\sigma^L > 0$ such that $s_{it}^L \in \{(2k+1)\,\sigma\}_{k=\underline{k}}^{\bar{k}}$ and $\theta_{it} \mid \mathbf{s}_t^L \sim \mathcal{U}\left[s_{it}^L - \sigma, s_{it}^L + \sigma^L\right]$. Again, we have grid, within each grid interval θ_{it}^L is uniformly distributed, and the signal tells you which element of the grid θ_{it} is in. We allow $\mathbf{s}_t$ to have any distribution on this support. This has the same interpretation as before of summarizing the public information about the legislator's idea point—campaign finance records, public polling data, and official statements made by the politician. The Lobbyist's bias ϕ_t is independent across time and uncorrelated with any private information, but can be arbitrarily correlated with $\mathbf{s}_t^L$.

Only in the party channel, the Whip receives a private signal $\mathbf{s}_t^W \equiv \left(s_{1t}^W, \ldots, s_{nt}^W\right) \in \mathbb{R}^n$. This is drawn such that $\theta_{it}|\mathbf{s}_t^W = \theta_{it}|\mathbf{s}_t^L, \mathbf{s}_t^W \sim \mathcal{U}\left[s_{it}^W - \sigma^W, s_{it}^W + \sigma^W\right]$ with $\sigma^W = \sigma^L/c$ for some $c \in \{2, 3, \ldots\}$. In other words, the Whip's signal embeds the public information and is of higher precision—in particular, it is the same structure as $\mathbf{s}_t^L$ with a finer grid.

3.1. **Whip's Superior Information and Theory of the Firm.**
Typically in the theory of the firm literature, vertical integration is defined through the transfer of tangible assets. Tangible assets seem much less relevant in parties, but there are *intangible* assets that can fulfill a similar role. Specifically, information and reputation are important assets for parties.

In this paper, we focus on the information side. This has precedents in the theory of the firm literature. Oliver Williamson (1971) wrote "The advantages of internalization reside in the facts that the firm's *ex post* access to the relevant data is superior, it attenuates the incentives to exploit uncertainty opportunistically…". Bengt Holmström (1999) also highlighted the informational transformation across organizational boundaries "measurement costs can play an important role in determining boundaries…[the] control of assets provides control over a number of indirect incentive measures that typically, but not always, can be inaccessible or very costly to use in the market". We model vertical integration as the accumulation of private information of all organizational members under one agent, which we call "the Whip". If a legislator does not transfer their information to the Whip, we say that they are not a part of the party. Much like how a firm is conceptualized as the accumulation of assets under unified ownership, where if a worker maintains their ownership of their relevant assets, they are not an employee, but an outside contractor (Grossman and Hart, 1986; Hart and Moore, 1990; Holmström, 1999).

Now why do whips have superior information? If they are themselves legislators, they have access to classified documents and closed-door meetings. As legislators have allowed their parties to coordinate their electoral campaigns, they may have access to private polling data and other information collected from that operation. And there is historical evidence that they collect information intentionally:

> To be efficient [whips] must know the membership by sight; be on as friendly a footing with them as possible; know where they reside, both in Washington and at home; know their habits, their recreations, their loafing-places; the condition of their health, and that of their families; the number of their telephones; when they are out of the city; when they will return; how they would probably vote on a pending measure; what churches they attend; what theaters they frequent—in short, all about them.
>
> —Former Speaker of the House Champ Clark (1920)

Now it is probably not accidental that this quote is from 1920—today in the United States, importance has shifted more in favor of the parties' "reputational assets" (Kreps, 1990), or "electoral brands" (Snyder and Ting, 2002). But the point of this article is not to proclaim one unifying specific and concrete *raison d'etre* for all parties for all of time—that is impossible. Rather, we construct a single example from the general and flexible framework that theory of the firm provides, designed to connect firms as diverse as Walmart, law firms, and concrete manufacturers. Parties have many roles and faces when you go down to the specific—in the legislative arena, the electoral arena, the lobbying arena, and the mobilization arena.

3.2. **No Party Equilibrium.** For our first result, we characterize an equilibrium in which there is effectively no party. Later, this will be used as a Nash-reversion punishment off-path to support a party equilibrium. Note, from now on, we are assuming that $\mathbb{P}\left(\{s_{it}\}_i \cap \{-\sigma, \sigma\}\right) = 0$ to suppress the borderline case in which it is not clear that the optimal policy passes with probability one.

Proposition 3.1 (No Party Equilibrium)**.** *There always exists a perfect Bayesian equilibrium in which $D_t = 0$ for all t, and for all realizations of $\mathbf{s}_t^L$, the Lobbyist chooses $(\hat{x}_t, \mathbf{b}_t)$ according to the policies from Theorem 2.3 with $\sigma = \sigma^L$, $x_t = \phi_t x^*\left(\phi_t \mathbf{s}_t^L, \beta, \sigma^L\right)$ and $\mathbf{b}_t = \mathbf{b}^*\left(\phi_t \mathbf{s}_t^L, \beta, \sigma^L\right)$.*

Proof. Here, we use the fact that playing a PBE of the stage game every period is an equilibrium of the repeated game. For the off-path in which $D_t > 0$, there are many PBE paths in which the final allocation of party dollars ($\mathbf{d}_t$ or $\mathbf{d}_t'$) is not dependent on the voting outcome $\mathbf{v}_t$. For example, the Whip could refuse to give out the party dollars $\mathbf{d}_t = \mathbf{0}$, get deposed, and they play the equilibrium from Baron and Ferejohn (1989, Proposition 3) in which the first proposer gets $D_t\left(n+1\right)/2n$ and $n-1$ randomly selected other legislators receive D_t/n. Different decisions of the Whip do not lead to different continuation paths, and they do not have preferences for how the party dollars are allocated ceteris paribus, so this is PBE.

Then, since $\mathbf{d}_t$ does not depend on $\mathbf{v}_t$, the legislators vote the same way as they would if $D_t = 0$. The Whip cannot commit to $\mathbf{d}_t$ depending on $\mathbf{v}_t$, and in this equilibrium the legislators do not trust the Whip's cheap talk. Thus, $D_t > 0$ is a pure cost to the Lobbyist with no benefit in voting behavior, which means that $D_t = 0$ is uniquely optimal. With the party channel shut down, Theorem 2.3 gives us the unique equilibrium of the stage game, which implies our result. □

4. Equilibrium Selection

An *equilibrium candidate* is a strategy for each player and their beliefs over each information set. I will break down the solution concept in steps, increasing in significance.

Perfect Bayesian Equilibrium The standard solution concept in the literature.

No Vote-Trading *We assume no strategies are contingent on past voting history—if observed histories h and h' only differ with respect to past policy votes $\mathbf{v}_t \neq \mathbf{v}_t'$ but correspond on every other dimension: $\mathbf{s}_t^P, (x_t, \mathbf{d}_t)$, cheap talk communication, θ_{it}, $\mathbf{s}_t^W$ if $i = W_t$, then the equilibrium must prescribe the same behavior for i at h and h'.*

As discussed in the introduction, intertemporal vote-for-vote trading is hard for many reasons—trust, communication, unforeseen contingencies. With $\delta \to 1$, you can find strategies that condition votes on past behavior such that with the power of the folk theorem you can approximate the first best. We want to consider a setting where it is still hard to construct relational vote trading contracts (for reasons other than patience), and one way to exemplify that is to shut them down completely.

One could relax this *impossible* condition to *difficult*, and the qualitative results should all still hold. And one could microfound some of the difficulties we use to justify this, like imperfectly shared-understandings. We chose to assume the stark case for expositional reasons.

Intratemporal Coalitional Equilibrium This refinement is essentially[6] coincides with that of the *stable convention in the NTU repeated game* of Ali and Liu (2020). The solution concept of this kind was also suggested (but not defined) in Ray (2007). We write an abbreviated definition of our version here, with the details expanded upon in the Appendix (B.1).

Definition 4.1. We say that an assessment (a, μ) is an *Intratemporal Coalitional Equilibrium (ICE)* of a repeated game if at the beginning of any period, either after on-path or off-path history, there cannot exist a profitable one-period deviation[7] for any coalition. That is, for all periods $\tau \in \mathbb{N}$ and histories of length $\tau - 1$, there cannot exist a coalition C and a deviation a' such that a' makes every member of C strictly better off, and a' satisfies (i) for all players $i \notin C$, we must have $a_i = a'_i$ and (ii) for all information sets $\mathcal{I}$ that take place after period τ, we must have $a(\mathcal{I}) = a'(\mathcal{I})$.

So, with this solution concept, we refine PBE to those that are robust to coalitional deviations of this bounded horizon. After the "period" is over, play goes back to what is specified by the original assessment a, although the coalitional deviation clearly takes the game off-path. What we define as a period may consist of many stages, like in our game, and in this solution concept coalitions can commit to their behavior across these many stages all at the start of the period. It does not have to be subgame perfect for coalition members to keep true to their plan, we are assuming there is commitment power, at least within a period.

This restriction should be intuitive. Think of the intuition of Ostrom (1990)—in small-scale communities where players all know each other and interact through many different channels, uncoordinated single-player deviations may not be the appropriate concept. Coordinated coalitions may be feasible through mechanisms outside of knowledge of the modeler—like in Ostrom's theory of polycentric governance. However, these external mechanisms typically require a high degree of common knowledge and specifications for all possible contingencies. These requirements are harder to satisfy under a long time horizon, in particular in politics where power and popular opinion are constantly changing. Thus, it is natural to consider coalitional deviations, limited in how far into the future they prescribe behaviors.

[6]There are only two technical differences. First, they assume finite action spaces while ours are infinite, but all results presented in this paper are proved without reference to those results in Ali and Liu (2020). Second, we have incomplete information in our game, and such we use expected utilities generated from assessments that must satisfy the usual Bayes' consistency criterion.

[7]This is a condition with bite, since the one-shot deviation principle does not extend to these coalitional deviations. Intuitively, where it breaks down is with one-player deviations, we know that one unprofitable deviation + another unprofitable deviation = an unprofitable deviation. But with coalitional deviations, one deviation that is not profitable for all coalition members + another deviation that is not profitable for all members may = a profitable deviation for all members.

Back to our introduction, this kind of solution concept allows us to analyze coalitional stability without assuming that a coalition can automatically achieve its Pareto frontier. If coalitional deviations were able to be infinite-horizon, we would just have the core of the normal-form representation of the repeated game, where we must be on each coalition's Pareto frontier. Our intention is to introduce transaction costs to these coalitional problems, and thus inherently we need a stability concept that still allows for coalitions to be imperfect in their internal transactions. This is an important but subtle point—to study the cooperative properties of an environment with transaction costs, one needs a cooperative concept that reflects those transaction difficulties, rather than being immune to them.

From now on, when we refer to *equilibrium*, we mean equilibrium candidates that are both PBE and ICE. When we say the set of *No Vote-Trading equilibria*, we mean the set of equilibrium that satisfy both PBE, ICE, and the *No Vote-Trading* assumption.

5. The Party Equilibrium

Now we present the main results of this section, the existence of an equilibrium in which the party is active and effective, and a characterization for a constrained optimum under a sufficiently symmetric type distribution.

Theorem 5.1. *If each legislator has the same support of ideal points, there exists a $\underline{\delta} \in (0,1)$ such that there exists a No Vote-Trading equilibrium where the party path is used and every party member does better off than the no party equilibrium.*

The construction of an equilibrium is straightforward. Consider the sets of events in which:

$$\phi_t x^* \left(\phi_t \mathbf{s}_t^L, \beta, \sigma^L\right) < \min\left\{\phi_t s^L_{\text{median},t}, \phi_t s^L_{W,t}\right\} \tag{5.1}$$

This ensures that going from what the lobbyist can achieve without the party, $\phi_t x^* \left(\phi_t \mathbf{s}_t^L, \beta, \sigma^L\right)$, to what the whip can achieve $\phi_t x^* \left(\phi_t \mathbf{s}_t^W, \beta, \sigma^W\right)$ makes (i) the Lobbyist, (ii) the Whip, and (iii) total party surplus better off. In only these events, *might we* use the party path, in all other events, assume the lobbyist uses the non-party path. If the lobbyist deviates to using the party path, assume that legislators vote on the policy as if it were unrelated to the allocation of party dollars, as in the no-party construction. If the party deposes the Whip off path, we go to the no party equilibrium. Intuitively, we can interpret this as the lobbyist losing trust in a dysfunctional party.

If we have symmetric ideal point type distributions, we are done. Future party total surplus is increased by not deposing the Whip and remaining on the current path. By symmetry, that total surplus is split evenly, so it is positive for everyone. Thus, with sufficient patience, it is not profitable for any coalition to depose the Whip and "steal all the party dollars" because it destroys the party in the future which every member of the coalition cares unboundedly much about as $\delta \to 1$.

If it is asymmetric, by finite signal and shared signal support distributions, we can "make it symmetric" by ignoring some "asymmetrically high probability" events of that set with some interior probability and specifying the Non-Party path to be played instead with some probability. For example, if event A has probability .06 and symmetric event B has probability .04, then we can ignore event A with probability 1/3 such that (A, party path) has the same probability as (B, party path). Thus, the future effect of the party path is total surplus positive and surplus symmetric (with this adjustment), and thus again everyone prefers the path in which the party continues to be strong.

Note that everyone is better off in the Party Equilibrium than the No Party Equilibrium—that is, there are transaction costs. In particular, in a world without any transaction difficulties, the legislators who are moderates today would sign court-enforceable contract that trade their vote today for multiple votes in the future on issues they are ideological on (since there are more ideologues than moderates, one-to-many trading is possible). But they cannot write these contracts, so the system depends on these immediate transfers funded by lobbyists to grease the wheels on policy making and vote trading. To make this process even more effective, the legislators and the lobbyists delegate their informational property and donation decision rights to this intermediary, which we call the Whip. However, both sides can take this power back, thereby unraveling it.

So in spite of us having a split-the dollar like coalitional game within each period, we have stability. Intuitively, in terms of our example from the introduction, $\sigma^W < \sigma^L$ means that there is a gap in the future between "how much of the dollar" the party as a whole can offer compared to any revolting faction (which thereby unravels the party). Taking $\delta \to 1$ allows us to make that future gap arbitrarily large, eventually satisfying the "$\lambda_N \geq \left(\lambda_{P_1} + \lambda_{P_2} + \lambda_{P_3}\right)/2$"-esque inequality that is required for the party to be core stable. When $\delta = 0$, donating to the party on path entails a subgame identical to the canonical split-the-dollar game, and thus core stability is impossible to achieve. This gives us our $\underline{\delta}$ condition. We can say something more precise when δ is sufficiently small in the following proposition.

Proposition 5.2. *For each PBE in which there are donations to the party on path, there exists a $\bar{\delta} \in (0, 1]$ such that for all $\delta \in [0, \bar{\delta})$, the equilibrium is not an ICE.*

Without a relevant future, any path that donates to the party entails a subgame that is approximately like split-the-dollar, and the coalitional stability condition coincides with looking for an element of the core which we know is non-existent.

Lemma 5.3. *The no-party equilibrium is an intratemporal coalitional equilibrium for all parameter values.*

Similarly to how repeating a Nash equilibrium of a stage game on all paths is a PBE of the repeated game, we can prove that the no-party equilibrium is a coalitional equilibrium of the normal-form representation of the one period game, thus repeating it on all paths is an ICE of the repeated game.

Theorem 5.4. *If all ideal point distributions are sufficiently symmetric—for some particular $p<1$ we have that for all $s\in S$, i,j, we have $p\mathbb{P}(s_{it}=s)\le\mathbb{P}(s_{jt}=s)\le\frac{1}{p}\mathbb{P}(s_{it}=s)$—then for δ sufficiently high, the following does the best among no-vote trading equilibrium without on-path Whip depositions or Lobbyists with ex-post regret. Given $\mathbf{s}_t^L,\phi_t$*

- *The Lobbyist chooses the party channel as long as the Whip has never been deposed. If the Whip is deposed, then we move to the No-Party equilibrium.*
- *The Whip cheap-talk promises to the Lobbyist the ability to deliver some $x^P\left(\mathbf{s}_t^W,\boldsymbol{\theta}_t,\mathbf{s}_t^L,\phi_t\right):\phi_t x^P(\cdot)\ge\phi_t x^L\left(\mathbf{s}_t^L,\phi_t\right)\equiv\phi_t x^*\left(\phi_t\mathbf{s}_t^L,\beta,\sigma^L\right)$ and*
$$D_t^P\coloneqq\phi_t\beta\left(x^P(\cdot)-x^L(\cdot)\right)+\sum_{i=1}^n b_i^*\left(\phi_t\mathbf{s}_t^L,\beta,\sigma^L\right)$$
is the requested donation, which is fulfilled.
- *The Whip then proposes $x^P(\cdot)$, it passes, the Whip is not deposed, and allocates the party dollars according to some $d^P\left(\mathbf{s}_t^W,\boldsymbol{\theta}_t,\mathbf{s}_t^L,\phi_t\right)$.*
- *If the Whip picks any $x,\mathbf{d}$ such that*
$$(x,\mathbf{d})\notin\left\{\left(x^P,\mathbf{d}^P\right)\left[\hat{\mathbf{s}}_t^W,\hat{\theta}_{Wt},\boldsymbol{\theta}_{-Wt},\mathbf{s}_t^L,\phi_t\right]:\hat{\mathbf{s}}_t^W,\hat{\theta}_{Wt}\text{ are consistent with }\mathbf{s}_t^L\ \&\ \phi_t\hat{\mathbf{s}}_t^W\le\phi_t\mathbf{s}_t^W\right\}$$
this deviation is detectable and the Whip will be deposed from then on.

Proof. The symmetry condition ensures that as long as we increase *total* party surplus from the on-path relative to the off-path punishment phase, then we increase *each* party members expected payoff (since on-path payoffs are lower bounded). Thus, the penal code of the no-party path is an arbitrarily strong punishment as $\delta\to1$, and so we have arbitrary power to punish any off-path deviations. We are extracting all of the additional ex post surplus from the Lobbyist relative to what they could achieve from the no-party path, we could not ask for anymore without violating ex-post regret. Extracting the most possible must weakly be apart of the optimal mechanism, since it adds flexibility about what policies are possible to pass, the only thing relevant for total surplus.

Finally, the reason these particular deviations are undetectable is because of ICE: the party is meant to move policy more "ϕ_t-ideologically". By claiming that the holdouts that are against this are more moderate than they actually are, this will attenuate the move, changing expectations about what policies are "possible" to pass. The attenuated policy will still pass, since the extreme one would pass, and this makes all the holdouts strictly better off—thus undetectable if the holdouts do not snitch on the Whip. Normally, with multiple-sender signalling, we can get senders to snitch on themselves by the belief that assumes many others will snitch either way. But when we check coalitional deviations for ICE, the Whip and the holdouts could pull off the deviation with the coalitional promise of no snitching, making the entire coalition strictly better off. Thus, a necessary condition of ICE is that these deviations are unprofitable even if they are not detected. We do not have to check upward deviations of $\hat{\mathbf{s}}_t^W$, since claims that the holdouts are more ideological than they actually are can only lead toward changes in policy that will not pass the vote in equilibrium, and thus are detectable. □

Proposition 5.5. *Assume $n+\beta$ is an even integer. Whenever $\phi_t\theta_{\frac{n+\phi_t\beta}{2}t} \geq \phi_t x^L(\cdot)$ (assuming we reorder the players according to $\theta_{1t} \leq \cdots \leq \theta_{nt}$) we can truthfully publicly elicit a value θ^* to serve as a policy cap such that it is known that $\theta^* \in \left\{\theta_{\frac{n+\phi_t\beta}{2}t}, \theta_{\left(\frac{n+\phi_t\beta}{2}+\phi_t\right)t}\right\}$ using cheap talk in the equilibrium from Theorem 5.4.*

Proof. Without loss of generality, assume $\phi_t = 1$. Since $n+\beta$ is an even integer, the efficient maximum policy x is anywhere between the $\frac{\beta+n}{2}$th highest θ and the $\frac{\beta+n}{2}+1$th highest θ. Because $s_i^L \notin \left\{-\sigma^L, \sigma^L\right\}$, we know that any policy x such that $\left|B\left(x, \mathbf{s}^L, \sigma^L\right)\right| \geq \frac{\beta+n}{2}$ must be such that we are already past this efficient maximum. Recall the notation $Z\left(\mathbf{s}, \hat{\beta}\right) := \sup\left\{z : |\{i : s_i < z\}| \leq \hat{\beta}\right\}$. Thus, we know that the players $\left\{i \in N : s_i^L \geq Z\left(\mathbf{s}^L, \frac{\beta+n}{2}\right)\right\}$, which has magnitude at least $\frac{\beta+n}{2}+1$, must never be bribed in the equilibrium just from our information on $\mathbf{s}^L$. (They may be based on $x^L\left(\mathbf{s}^L\right)$, but the party should not go any further/change anything relative to this L benchmark).

Thus, we can determine either $\frac{\beta+n}{2}$th highest θ or the $\frac{\beta+n}{2}+1$th highest θ by having the legislators $\left\{i \in N : s_i^L \geq Z\left(\mathbf{s}^L, \frac{\beta+n}{2}\right)\right\} \setminus \{W\}$ sincerely reveal their types θ_i, where the "pivotal i" is the implemented policy "cap". We call this θ^*. Truthtelling is optimal for a median voter logic—left of pivot you can only move the cap rightward and vice versa, if you are the pivot you are getting you "favorite" maximum. Knowing also the Whip's θ_W helps us reveal the *range* of efficient caps, but even without it we can identify at least one point that is in the range (one of the endpoints could be right or left). ☐

We define notation for the set of legislators that need to be bribed, excluding the whip, as $B_{-W}(x, \mathbf{s}, \sigma, \phi) := \{i \in N \setminus \{W\} : 2\phi(s_i - \phi\sigma) < \phi x\}$.

Theorem 5.6 (Partial Characterization of $x^P, \mathbf{d}^P$)**.** *Consider a public signal such that $\mathbf{s}^L \implies \phi\theta_W \notin \left[\phi x^L\left(\mathbf{s}^L, \phi\right), \phi\theta^*\right]$ and assume $n+\beta$ is an even integer. The solution $\left(x^P, \mathbf{d}^P\right)$ to Theorem 5.4 corresponds to the following Bayesian mechanism design problem. Choose $\mathcal{X}$, $f\colon \mathcal{X} \to \mathbb{R}$, and $\tilde{x}\left(\mathbf{s}^W, \theta_W\right)$ to maximize*

$$\mathrm{E}_{\mathbf{s}^W}\left[-\sum_{i\in N}\left|\tilde{x}\left(\mathbf{s}^W, \theta_W\right) - \theta_i\right| + \beta\phi\tilde{x}\left(\mathbf{s}^W, \theta_W\right) \mid \mathbf{s}^L, \theta^*\right] \tag{5.2}$$

subject to

$$0 \leq f(x) \leq \beta\phi\left(x - x^L\left(\mathbf{s}^L, \phi\right)\right) \tag{5.3}$$

$$x^{**}\left(\theta_W, \mathbf{s}^W\right) \in \underset{\hat{x}\in\mathcal{X}\cap\overline{X}(\mathbf{s}^W, f)}{\arg\max} \left\{-|\hat{x} - \theta_W| + f(\hat{x})\right\} \tag{5.4}$$

where

$$\overline{X}\left(\mathbf{s}^W, f\right) \coloneqq \left\{x : \beta\phi\left(x - x^L\left(\cdot\right)\right) - f\left(x\right) \geq \left|\int_{x^L(\cdot)}^{x} \left|B_{-W}\left(\hat{x}, \mathbf{s}^W, \frac{\sigma^L}{c}, \phi\right)\right| \mathrm{d}\hat{x}\right|\right\}. \tag{5.5}$$

The solution satisfies

$$\mathcal{X} \coloneqq \left\{x : \phi x^L\left(\cdot\right) \leq \phi x \leq \phi\theta^*\right\} \tag{5.6}$$

$$x^{**}\left(\theta_W, \mathbf{s}^W\right) = \phi \min\left\{\sup \phi\overline{X}\left(\mathbf{s}^W, f\right), \phi\theta^*\right\} \tag{5.7}$$

where

$$f\left(x\right) = \begin{cases} \left|x - x^L\left(\cdot\right)\right| & \textit{if } \phi\theta_W < \phi x^L\left(\cdot\right) \\ 0 & \textit{if } \phi\theta_W > \phi\theta^* \end{cases}. \tag{5.8}$$

We construct $\mathbf{d}^p$ *according to*

$$d_i\left(\cdot\right) = \begin{cases} b_W^*\left(\phi\mathbf{s}^L, \beta, \sigma^L\right) + f\left(x^P\left(\cdot\right)\right) & \textit{if } i = W \\ \max\left\{\phi x^P\left(\cdot\right) - \max\left\{0, \phi 2\left(s_i^W - \sigma^L/c\right)\right\}\right\} & \textit{else} \end{cases} \tag{5.9}$$

where if there is any leftover party dollars we can distribute it arbitrarily among $B_{-W}\left(x^L\left(\cdot\right), \mathbf{s}^L, \sigma^L, \phi\right)$. *And of course,* $x^P\left(\cdot\right) = x^{**}\left(\cdot\right)$.

Proof. The inequality in expression (5.5) can be interpreted as (relative to the no-party outcome) "the additional money we get from the Lobbyist for promising to deliver this policy x, less the Whip's cut f, must be greater than the additional bribes necessary to pass x". The rest is just a standard translation of the problem.

Because of (5.5), we want f to be as small as possible, subject to satisfy IC in a fixed desired way—large Whip cuts of the party dollars mean less budget to bribe other legislators, making fewer policies feasible. Constraining ourselves to the set $\left\{x : \phi x^L\left(\cdot\right) \leq \phi x \leq \phi\theta^*\right\}$, we have monotonic preferences of increasing ϕx with respect to objective (5.2). When $\phi\theta_W > \phi\theta^*$, we know that we can use the minimal $f\left(x\right) = 0$ to induce the Whip to choose the largest x in the menu $\mathcal{X} = \left\{x : \phi x^L\left(\cdot\right) \leq \phi x \leq \phi\theta^*\right\}$, thus we know that this is an optimal mechanism.

We know that $x^L\left(\cdot\right) \in \mathcal{X}$ because under the worst-case scenario lowest draw of $\mathbf{s}^W = \mathbf{s}^L - \boldsymbol{\sigma}^L + \boldsymbol{\sigma}^W$ (where $\boldsymbol{\sigma} \equiv (\sigma, \ldots, \sigma)$) for all $f \geq 0$ we have $\overline{X}\left(\mathbf{s}^W, f\right) = \left\{x^L\left(\cdot\right)\right\}$. In other words, $x^L\left(\cdot\right)$ must be an option because it might be the only option, and we must have an option implemented that is weakly lobbyist preferred to their outside option of the no-party path (it is without loss to implement $x^L\left(\cdot\right) \in \mathcal{X}$ through the party versus no-party channel).

Thus (5.4) implies that $\theta_W < \phi x^L\left(\cdot\right) \implies f\left(x^{**}\left(\cdot\right)\right) \geq \left|x^{**}\left(\cdot\right) - x^L\left(\cdot\right)\right|$. And similarly to before, the payment scheme $f\left(x\right) = \left|x - x^L\left(\cdot\right)\right|$ can implement any policy choice $x^{**}\left(\cdot\right) \in \left\{x : \phi x^L\left(\cdot\right) \leq \phi x \leq \phi\theta^*\right\}$ in an incentive compatible way, and it cannot be lowered (we only care about how low f is *pointwise* for implemented $x^{**}\left(\cdot\right)$, lowering it for off-path x only affects the IC constraint (5.4) but not feasibility (5.5)). Thus, we know that this must be the optimal—it has the best incentives and we cannot improve feasibility without breaking incentives. □

The reason $\phi\theta_W \in \left[\phi x^L\left(\mathbf{s}^L, \phi\right), \phi\theta^*\right]$ is challenging to characterize is not uncertainty on θ_W. Rather, when θ_W is in the interior (even when it is known), uncertainty regarding $\mathbf{s}^W$ through its effect on $\overline{X}\left(\mathbf{s}^W, f\right)$ becomes important. For example, it may be strictly beneficial to withhold options within the set $\left\{x : \phi x^L\left(\cdot\right) \leq \phi x \leq \phi\theta^*\right\}$ that are "too tempting" for the whip. To know if this is the case, we must know how likely these tempting policies would become relevant, which depends on the distribution of $\overline{X}\left(\mathbf{s}^W, f\right)$.

Note that whenever $\frac{1}{2}\phi x^L\left(\cdot\right) \leq \phi\theta_{Wt} \leq \phi x^L\left(\cdot\right)$, the Whip is strictly profiting from their privileged position as party leader—they are strictly better off than when someone else is the whip and they are just a regular legislator. Because of this, we suspect that we could do strictly better by allowing for Whip deposition, using deposition mixing as a part of our incentives for the Whip, substituting for f. We think this could be reasonable, it was just intractable to characterize—one would have to consider n deposition policies for each possible Whip, where the depose continuation value depend on the anticipated time spent in the other $n-1$ Whip regimes and their within-regime optimal policies.

6. A Model for Multi-Party Realignments

In this section, we present a simple and stylized game, taking some of the central results from the previous model as starting assumptions. For example, we assume that donating through the party reduces lobbying costs. In doing so, we will simplify the model by trimming away those assumptions that justified those facts—like the asymmetric information problem—that no longer importantly interact with the analysis of this section except through the generation of these stylized facts.

This game has $n+1$ players—n legislators $N := \{1, 2, \ldots, n\}$ and one lobbyist L. It is a repeated game in discrete time with an infinite horizon $t = 1, 2, \ldots$. The legislators have common discount factor $\delta \in (0, 1)$ and again it is best to think of the lobbyist as "short-run" $\delta_L = 0$, but it does not matter for the technical results.

At the start of each period, a state $\omega_t \equiv (M_t, L_t, R_t, \phi_t)$ is drawn, publicly observed, where $M_t, L_t, R_t \subseteq N$ such that $M_t \cap L_t = L_t \cap R_t = R_t \cap M_t = \varnothing$ for all t and $\phi_t \in \{-1, 1\}$. We again refer to (M_t, L_t, R_t, ϕ_t) as period t's *issue topic*, with the same interpretation as before. We say that legislator $i \in M_t$ is playing the role of a *moderate*, legislator $i \in L_t$ is a *left ideologue*, and legislator $i \in R_t$ as a *right ideologue*. Finally, if $\phi_t = 1$, the period t lobbyist is *right-biased*, and if $\phi_t = -1$, they are *left-biased*. Although one should not get stuck on the use of this "left-right spectrum"—in all of the following analysis, the state (M_t, L_t, R_t, ϕ_t) is equivalent to $(M_t, R_t, L_t, -\phi_t)$ in every meaningful way, even if you were to only make this swap for one period. Finally, we say that $M_t \cup L_t \cup R_t \subset N$ is the set of legislators *in power* for period t. For simplicity, we can assume constant sizes for these ideological wings, $|L_t| = |R_t| = n_I$ and $|M_t| = n_M$ for some $n_M + 2n_I \leq n$.

We assume that the state is drawn independently and identically across periods, but not necessarily symmetric across the legislators (so unlike before, this i.i.d. assumption is not a particularly strong or important). We denote

$$p_{I;\{J,K\}} := \mathbb{P}\left((I = M_t) \wedge (\{J, K\} = \{L_t, R_t\})\right)$$

as the probability that $I \cup J \cup K \subset N$ are in power with I the moderate wing, and J and K the opposite ideological wings. We assume that ϕ_t is drawn simply as an independent coin flip each period. Note, again, it will not be important to keep track of which ideological wing is on the left vs. right for each issue ex ante, which is why the notation above collapses both cases into one probability.

The lobbyists have a deterministic sequence of budgets $(B_t)_{t\in\mathbb{N}}$. The lobbyist chooses the policy proposal x_t in period t, and bribes the moderate to get their preferred policies. The lobbyist has preferences

$$u_L\left(x_t, \phi_t\right) = \phi_t x_t. \tag{6.1}$$

Thus, they either prefer the policy to be as far right as possible (if $\phi_t = 1$) or as far left as possible (if $\phi_t = -1$). An important difference to the previous model is that they do not have anything against spending money as long as they are under their budget constraint, and thus they always spend the full budget. An interpretation of this lobbyist is that this is a "net-representative lobbyist of this issue topic" after aggregating and canceling out the many competing lobbying interests.

We denote the party structure of period t by $\Pi_t \in \mathcal{P}(N)$, an element within the set of partitions of N. For example, if $n = 7$, $\Pi_t = \{\{1,2,3\},\{4\},\{5,6\},\{7\}\}$ means that in period t, legislators 1, 2, and 3 are in a party together, 5 and 6 are in a party together, and 4 and 7 are independents. Let $\pi_t(i) \in \Pi_t$ denote the party legislator i belongs to, $i \in \pi_t(i)$ for all $i \in N$.

Each moderate has policy preferences $-|x_t|$ plus a utility that is linear in transfers. Thus, if the lobbyist spends their whole budget each period, the total moderate's utility for period t is -$n_M \cdot |x_t| + B_t$, where recall $n_M = |M_t|$. But suppose that, like in the previous sections, each moderate $m \in M_t$ is able to extract positive expected rents from bargaining with the lobbyist. Given a state, party configuration, and policy proposal, a legislator $m \in M_t$ extracts an expected payment of

$$g_m\left(x_t, \Pi_t, \omega_t\right) + |x_t|, \tag{6.2}$$

with g_m weakly-increasing in $|x_t|$. Thus, the moderate's utility within the period nets to $g_m\left(x_t, \Pi_t, \omega_t\right)$. With a lobbyist who spends their entire budget, this implies that

$$|x_t| = \frac{B_t - \sum_{m\in M_t} g_m\left(x_t, \Pi_t, \omega_t\right)}{n_m}. \tag{6.3}$$

Intuitively, as in the previous model, g_m will tend to be such that if the issue direction is favorable to other members of the moderate's party $\pi(m)$, the moderate will extract fewer rents to allow the policy to become more ideological.

An ideologue $l \in L_t$ will always vote for all policies $x_t < 0$ and against all policies $x_t > 0$, and cannot be bribed to change their vote. They have utility

$$u_{it}\left(x_t \mid i \in L_t\right) = -f\left(x_t\right) \tag{6.4}$$

for some f monotonically increasing, $f(x) \geq \left(n_M/n_I\right) x$ for all x, and $f(0) = 0$.

An ideologue $r \in R_t$ will always vote for all policies $x_t > 0$ and against all policies $x_t < 0$, and cannot be bribed to change their vote. They have utility

$$u_{it}(x_t \mid i \in R_t) = f(x_t) \tag{6.5}$$

for that same f as in (6.4).

Where $f(x) \geq (n_M/n_I)\, x$ ensures that there can be gains from intertemporal trade: the marginal benefit to the ideologues dominated the marginal cost to the moderates when policy moves more ideologically extreme. In the single party model, it was due to the fact that the sheer number of legislators playing "ideologue" was greater than the number of "moderates", but it is also equivalent to consider ideologues with more intense marginal preferences.

This game has a similar feature to the previous in that, given a party configuration, the legislators behave quite nearsightedly in how they extract rents and vote on today's policy topic. But the legislators will be more forward-looking with respect to their considerations of party power and persistence. This will be discussed in greater detail, but the essential claim is that "vote for future vote on hypothetical unknown bill" bartering is too challenging even with sufficient patience. The party facilitates with systems of more immediate and tangible compensation, and as a patient and reliable third-party to uphold "agreements".

We define a notion of stability in party configuration that is precisely a particular subcase of *a stable convention in the NTU repeated game* from Ali and Liu (2020).

Definition 6.1. We say that a party configuration $\bar{\Pi} \in \mathcal{P}$ is *stable* if for all $t_0 \in \mathbb{N}$:

i. **No two parties want to merge.** This means that given any arbitrary initial state realization $\omega_0 \equiv (M_{t_0}, L_{t_0}, R_{t_0}, \phi_{t_0})$, there exists no $\pi, \pi' \in \bar{\Pi}$ with $\pi \neq \pi'$, such that the configuration where the two parties merge and everyone else stays put:

$$\Pi' := \{\pi \cup \pi'\} \cup \left\{\bar{\pi}(i) \in \bar{\Pi} : i \in N \setminus (\pi \cup \pi')\right\}$$

does better off for every member of the merging party $i \in \pi \cup \pi'$:

$$\begin{aligned}&(1-\delta)\,\mathrm{E}\left[u_{it_0} \mid \omega_0, \Pi'\right] + \delta \mathrm{E}_{\omega_t}\left[u_{it} \mid \Pi'\right]\\ > &(1-\delta)\,\mathrm{E}\left[u_{it_0} \mid \omega_0, \Pi\right] + \delta \mathrm{E}_{\omega_t}\left[u_{it} \mid \Pi\right].\end{aligned}$$

ii. **No faction wants to become independent from their party.** This means that given any arbitrary initial state realization $\omega_0 \equiv (M_{t_0}, L_{t_0}, R_{t_0}, \phi_{t_0})$, there exists no $\pi \in \bar{\Pi}$ and $\pi' \subset \pi$, such that the configuration where π' breaks off from π and everyone else stays put:

$$\Pi' := \{\pi', (\pi \setminus \pi')\} \cup \left\{\bar{\pi}(i) \in \bar{\Pi} : i \in N \setminus \pi\right\}$$

does better off for every member of the faction $i \in \pi'$:

$$\begin{aligned}&(1-\delta)\,\mathrm{E}\left[u_{it_0} \mid \omega_0, \Pi'\right] + \delta \mathrm{E}_{\omega_t}\left[u_{it} \mid \Pi'\right]\\ > &(1-\delta)\,\mathrm{E}\left[u_{it_0} \mid \omega_0, \Pi\right] + \delta \mathrm{E}_{\omega_t}\left[u_{it} \mid \Pi\right].\end{aligned}$$

iii. **No faction from a party wants to join another party.** This means that given any arbitrary initial state realization $\omega_0 \equiv (M_{t_0}, L_{t_0}, R_{t_0}, \phi_{t_0})$, there exists no $\pi, \pi'' \in \bar{\Pi}$ and $\pi' \subset \pi$, such that the configuration where π' breaks off from π, joins party π'', and everyone else stays put:

$$\Pi' := \{\pi' \cup \pi'', (\pi \setminus \pi')\} \cup \left\{\bar{\pi}(i) \in \bar{\Pi} : i \in N \setminus (\pi \cup \pi'')\right\}$$

does better off for every member of the faction and new party $i \in \pi' \cup \pi''$:

$$\begin{aligned} &(1-\delta)\,\mathrm{E}\left[u_{it_0} \mid \omega_0, \Pi'\right] + \delta \mathrm{E}_{\omega_t}\left[u_{it} \mid \Pi'\right] \\ &> (1-\delta)\,\mathrm{E}\left[u_{it_0} \mid \omega_0, \Pi\right] + \delta \mathrm{E}_{\omega_t}\left[u_{it} \mid \Pi\right]. \end{aligned}$$

This notion of stability coincides with the "Model-δ" notion of stability in Hart and Kurz (1983) with respect to coalition formation and deformation processes. To summarize, in contrast to their alternative "Model-γ" approach, we are assuming that if i leaves coalition $\{ijk\}$, the outcome is $\{\{i\}, \{jk\}\}$ rather than $\{\{i\}, \{j\}, \{k\}\}$. That is, the entire coalition does not disintegrate upon a single defection, and rather the members left behind remain in a coalition together.

In this notion of stability, we are treating coalitional deviations Π' as persistent. This means that one cannot leave the party for the day, just to come back and reform it tomorrow. It also means that two legislators who are from two different parties cannot simultaneously leave their parties and join together in one swift move.

The justification for this assumption comes from our result in the richer model that the lobbyists need to trust a party apparatus for it to function, and instability causes the lobbyist to lose trust in the party. So in this definition of party stability, we are considering a one-shot "grace period" where all parties reformulated will automatically maintain the lobbyist's trust, but this reformulation grace cannot be counted upon in the future. As a robustness check, we could allow these grace periods to occur with some sufficiently low probability or occur deterministically after a sufficiently long fixed waiting period.

6.1. Multi-Party Game—A Baseline Case.

We consider a special case where

$$g_m(x_t, \Pi_t, \omega_t) = \begin{cases} \lambda_I |x_t| & \text{if } x_t \geq 0 \text{ and } |\pi_t(m_t) \cap R_t| \leq |\pi_t(m_t) \cap L_t| \\ \lambda_P |x_t| & \text{if } x_t \geq 0 \text{ and } |\pi_t(m_t) \cap R_t| > |\pi_t(m_t) \cap L_t| \\ \lambda_I |x_t| & \text{if } x_t < 0 \text{ and } |\pi_t(m_t) \cap L_t| \leq |\pi_t(m_t) \cap R_t| \\ \lambda_P |x_t| & \text{if } x_t < 0 \text{ and } |\pi_t(m_t) \cap L_t| > |\pi_t(m_t) \cap R_t| \end{cases} \tag{6.6}$$

with $\lambda_I > \lambda_P \geq 0$. Intuitively, if the policy position is going in the direction of a fellow party member, the moderate extracts fewer rents than they would if they were unaffiliated with that ideologue. In the previous richer model, the party reduced the moderate's rents through the ownership of superior information, but one could imagine many alternative explanations. For simplicity, we assume that for some $\beta > n_M$, $f(x_t) = \beta x_t$.

Proposition 6.2. *The total utility of the legislators is maximized when there are no parties,* $\Pi = \{\{1\}, \{2\}, \ldots, \{N\}\}$.

More generally, effective party activity has a strictly decreasing relationship with total surplus, where party activity can be proxied by the measure on the event[8] $|\pi_t(m_t) \cap R_t| \neq |\pi_t(m_t) \cap L_t|$. Intuitively, when a party is effective at coaxing a moderate closer to the party's larger ideologue wing's ideal, both the moderate and the ideologues at the opposing pole lose. The two ideological wings always cancel out, and so the net utility difference from the party is the loss incurred by the moderate. We remark that party activity is nonmonotonic with respect to the finer-than relation for party configurations—for example, one large party $\Pi = \{\{1, 2, \ldots, N\}\}$ also has zero party activity because the ideological wings of the party will always equalize out $|\pi_t(m_t) \cap R_t| = |\pi_t(m_t) \cap L_t| = n_I$.

Proposition 6.3. *The total utility among any two legislators* i, j *is higher when they form a party* $\{i, j\} \in \Pi$ *than when they remain unaffiliated* $\{i\}, \{j\} \in \Pi$ *holding the party composition of* $N \setminus \{i, j\}$ *constant.*

Proof. This result mirrors that in the richer model when our scope was a single party and half the legislative house. Although the total effect of this party formation $\{\{i\}, \{j\}\} \to \{\{i, j\}\}$ is net negative, it is positive for the legislators that form it i and j . There are two possibilities in period t when i and j form a party. First, the party has all of the power aligned—the lobbyist is bribing a moderate in $\{i, j\}$ towards the position of an ideologue in $\{i, j\}$. In this case, the marginal benefit to the ideologue dominates the marginal loss of the moderate, and the party generates a net benefit to $\{i, j\}$. Or, the party has no impact on outcome. $\square$

Theorem 6.4. *Take* $N = 3$, $n_I = n_M = 1$. *For each pair of legislators* $i, j \in N$ *and any set of parameter values, there exists a* $\underline{\delta} \in (0, 1]$ *such that the party configuration with party* $\{i, j\}$ *satisfies conditions i. and ii. of Definition 6.1 if and only if* $\delta \in (\underline{\delta}, 1)$ *(empty when* $\underline{\delta} = 1$*). Furthermore,* $\underline{\delta}$ *decreases weakly when there is*

- *(i) Greater Party Balance—conditional on when i and j "work together" (one is the moderate, the other is the ideologue), the probability that i is the moderate and j is the ideologue becomes more even with j being the moderate and i the ideologue.*
- *(ii) Acceleration of Money in Politics—the budget becomes* B'_t *where there is some positive monotonically increasing* $(\gamma_t)_{t \in \mathbb{N}}$ *such that* $B'_t = \gamma_t B_t$.

[8] This determines how much "λ_P happens" instead of λ_I.

(iii) Ideological Preferences Become More Intense—β increases.

(iv) More Issues that the Party can Work On—the probability of an $\{i, j\}$ moderate-ideological pairing increases, relative to both being at the ideological poles.

We break it down condition by condition. Greater party balance precisely corresponds to avoiding the asymmetry problem—no one member of the party should always be acting as the moderate without ever being ideological. The acceleration of money in politics, rather than just the magnitude of money in politics, comes from the fact that the amount of money today increases the temptation in the party stability condition, while the amount of money tomorrow increases the consequences of deviation/unraveling the party. So, if money each period increases by the same scalar, the discount factor condition actually does not change. However, having tomorrow's money increase more relative to today's money—that is, acceleration—does make the party more stable. Ideological preferences becoming more intense is intuitive; what the party does is make realized policies more ideological. This is the party's agenda, these policies that are further from the status-quo that they try to push, that are more extreme than what the median-voter theorem might suggest.

Finally, more issues that the party can work on is the simplest intuition that I have yet to discuss—a party with Bernie Sanders and Ted Cruz does not work. There needs to be some common ground between members for the party to be effective. Now we go further than Conditional Party Government Theory (Rohde, 1991; Aldrich, 1995) in that we establish further conditions for parties to be effective relative to the members acting independently based on their homogeneous preferences with conditions (i)-(iii).

Proposition 6.5. *Take $N = 3$, $n_I = n_M = 1$. Under the uniform probability distribution over the finite set of states, there is no party configuration that satisfies all three conditions of Definition 6.1 for any $\delta \in [0, 1)$.*

Proof. If there is a party of two, then in any state where that party wields effective power, the moderate deviates by forming a party of two with the third legislator. By symmetry, the moderate does equally well in the future in this new party, and does better today. The third legislator does better today and in the future.

If there is no party (or equivalently here, a single party of all the legislators), the moderate and the ideologue going against the lobbyist can form a party. Nothing changes today, and by Proposition 6.3, both have a strictly better future, thus again there is a profitable deviation. $\square$

Again, we find that too much symmetry in party capacity leads to chaos.

Theorem 6.6. *For any party and environment, if the preferences among party members become too correlated, the party will become unstable.*

This is a contradiction of Conditional Party Government Theory, which suggests that party strength is monotonically related to preference homogeneity. This result works by a logic related to that of Baker, Gibbons and Murphy (1994)—as the challenges to transactions soften, the fallback option of breaking the party contract becomes too tempting, which crowds out effective party power. As preferences within the party get too homogeneous, the world without the party becomes similar to the one with it, and thus, in those rare moments the party does need to wield power, there is too little incentive for any party members to concede to it. In this result, we are holding fixed the amount of enforcement that is being attempted λ_P, but of course if this is endogenous then instead of instability we just have $\lambda_P \to \lambda_I$ (as if the party wasn't there).

6.2. **Multi-Party Game—A Example with Organizational Capacities.**

Let $PS(N)$ denote the power set generated by the set N. We consider a case where again $f(x) = \beta x$ for $\beta > n_M$ and

$$g_m(x_t, \Pi_t, \omega_t) = \begin{cases} \lambda_I & \text{if } x_t \geq 0 \text{ and } |\pi_t(m_t) \cap R_t| \leq |\pi_t(m_t) \cap L_t| \\ A(m, \pi_t(m), \omega_t) & \text{if } x_t \geq 0 \text{ and } |\pi_t(m_t) \cap R_t| > |\pi_t(m_t) \cap L_t| \\ \lambda_I & \text{if } x_t < 0 \text{ and } |\pi_t(m_t) \cap L_t| \leq |\pi_t(m_t) \cap R_t| \\ A(m, \pi_t(m), \omega_t) & \text{if } x_t < 0 \text{ and } |\pi_t(m_t) \cap L_t| > |\pi_t(m_t) \cap R_t| \end{cases} \tag{6.7}$$

where $A\colon N \times PS(N) \times \Omega \to (-\infty, \lambda_I]$ is referred to as *the collective agenda*, and for a party $\pi \in PS(N)$, the projection $A_\pi \equiv A(\pi, \cdot) : \pi \times \Omega \to (-\infty, \lambda_I]$ is referred to as *party π's agenda*. We denote $K\colon PS(N) \to \mathbb{R}_{\geq 0}$ as the vector of *organizational capacities*. All together, we define a meaning of organizational capacity in this setting.

Definition 6.7. We say that a collective agenda A *respects the vector of organizational capacities* K if for all $\pi \in PS(N)$, the inequality $\lambda_I - K(\pi) \leq A_\pi(m, \omega)$ holds for all $(m, \omega) \in \pi \times \Omega$.

Thus, for lower values of organizational capacity K, a party is less able to get one of their moderates to reduce their inefficient rents "for the good of the party". This could be because of "know-how", like a Whip that is more imprecise at reading the signals of the party members, or it could represent something like trust issues.

Definition 6.8. We say that the pair (Π, A) is *robustly stable with respect to* K, if (i) A respects K, and (ii) for all other A' that respect K and match A on the active parties ($A_\pi = A'_\pi$ for all $\pi \in \Pi$), (Π, A') is stable in terms of Definition 6.1.

In other words, we allow the inactive parties to choose any capacity-respecting agenda when attempting the deviations from Definition 6.1.

Definition 6.9. We say that Π *has a robustly stable agenda with respect to* K, if there exists an agenda A such that (Π, A) is *robustly stable with respect to* K.

Remark. Any party configuration Π must be monotonically more likely to have a robustly stable agenda as (1) the organizational capacities of active parties increase, $K(\pi) \uparrow$ for $\pi \in \Pi$, and (2) the organizational capacities of inactive parties decrease, $K(\pi') \downarrow$ for $\pi' \in PS(N) \setminus \Pi$.

This remark is simply a matter of expanding the set of candidates for robust agendas and shrinking the set of deviations that we must check on each of them.

Theorem 6.10. *Fix any party configuration* Π *and full-support probability distribution over the state space* $p \in \Delta\Omega$. *There exists a* $\underline{\delta}$ *such that for all* $\delta \in (\underline{\delta}, 1)$, *the set of organizational capacities for which* Π *has a robustly stable agenda with respect to and the set of organizational capacities for which it does not are both nonempty.*

That is, with sufficient patience, any party configuration under any preferences might or might not be stable depending on the organizational capacities. Thus, without knowing these capacities, we cannot predict what will be stable.

7. Discussion

Through the Lens of Transaction Cost Economics. In these models, the transaction costs arise from the fact that legislators are unable to make intertemporal vote-for-vote transactions. At every level of this model, if the legislators could write perfect court-enforceable contracts with each other that specify how they must vote on every possible sequence of future issue topic draws, we would not need the party apparatus. Many legislators would like to trade their vote on this issue that they are moderate on today for a vote that they are ideological on tomorrow, but I restrict all the equilibria to those that are invariant to the actual voting outcomes. Thus, they cast issue votes according to myopic incentives.

The justification for this is that politics is too unforeseeable and it is hard to create common knowledge on promises for future hypothetical bills. One legislator may promise to give their vote on the next future immigration bill, but they likely will not accept a deal where they must vote on the bill no matter what it says, i.e. putting children in cages. Where that line is for what is acceptable and reasonable for them to vote on is almost impossible to describe. And then, if that decision ought to condition on factors like public opinion, lobbyist powers, and political power structures, it is both impossible to forecast or describe those conditions far into the future. Thus, long-horizon vote-for-vote trades are too challenging, even with sufficient patience. Weingast and Marshall (1988) elaborate on these transaction difficulties in detail.

The important lesson of this is that *legislation is incomplete for the same reasons contracts are*—a single bill cannot be written that specifies what the government is to do in every possible contingency. Policy making is a continuous, never-ending process. Models that reduce it to a single once-and-for-all vote, in which each and every political outcome in its entirety can be ascribed onto a single proposed bill, implicitly assume completeness. We model this incompleteness through our restrictions of which issues can be resolved and when—-for example, in period $t = 0$, the legislators cannot write and vote on a bill in the form of $\mathbf{x} \in \mathbb{R}^{\prod_{i=1}^{\infty}[z-\Delta, z+\Delta]^n}$ where

$\mathbf{x}((\theta_{t'})_{t'\leq t})$ specifies a committed policy in period t for each possible sequence of draws of $\theta_{t'}$ up to time t. Thus, the policy voted on today is assumed to be incomplete on what is to be done for those future issues. By modeling this incompleteness, we find that the logic of the legislation game drastically changes relative to those previous complete analogues.

Where is the vertical integration in this model? The creature that is "the Whip" which I have created is meant to represent the role of a common ownership structure in political life. In this case, we are integrating the informational access rights of the legislators with the decision rights over donation dollars of the lobbyist.[9] This is analogous to the framing of a firm as the integration of the productive assets of the workers with the decision rights over the dollars of the investors under one common ownership structure (under the purview of a CEO instead of a Whip).

However, one could model the vertical integration of parties in many alternative ways to this one with information. The political party integrates (i) the powers and property of government office holders, (ii) the operations that run political campaigns for elections, and (iii) a donor-politician intermediary that coordinates and optimizes how to spend political contributions. One could even write a Holmström (1999)-esque property rights model of "candidate-centered" electoral campaigns as each candidate owns their own reputation as an asset, and "party-centered" electoral campaigns as the ownership of those reputational assets are centralized. We found the informational story relatively easy and intuitive to model, but our choice of this informational problem is not meant to suggest that we find it to be the uniquely compelling explanation to the nature of political parties.

Firms vs. Parties. So are parties just firms that engage in the production of political outcomes instead of economic outputs? And if not, how do they relate, and how do they differ?

Without going into the semantics of drawing hard definitional lines, there is a fundamental quality that makes the study of parties different from that of firms. In particular, parties arise in institutional settings that preclude permanent ownership of and the freedom to transfer powers and rights. In the democratic context, this is done by law and design. In the nondemocratic context, this is just due to the fragility of power in these settings. For example, in the history of the Communist Party of the Soviet Union, the highest levels of leadership alternated between regulated power sharing agreements (Politburo, the Central Committee, and the Council of Ministers) and individual leaders that were constantly looking over their shoulders for threats to their power (i.e., Stalin and the Great Purge).

This means that in parties, we have to be much more worried about coalitional stability than in firms. For example, in Sections 3, and 5, all we would have to do is replace the rule "the Whip gets deposed by a majority vote" with dictatorship "there is some legislator $i_P \in N$ that can unilaterally decide whether or not the Whip gets deposed and the ensuing reallocation", and now we have precisely a

[9]Integration was taken as a given in Section 5, but clearly all results hold the same if we allow for defections (a legislator "taking back their information"), assuming that defections also lead to the party unraveling like deposing the Whip does.

traditional organizational theory model for a firm. The principal is i_P, who is informally delegating authority[10] to some middle manager called the Whip, the other legislators are workers with hidden types producing political outcomes that the firm is selling to the lobbyists. Just minor tweaks to the labels and preferences and we have a classic firm model.

The drastic difference for analyzing firms vs. parties is when we go from a single principal with unique veto power to a setup where different overlapping coalitions are capable of acting as the principal. As Arrow (1951) showed, the only time a political choice problem does not generally arise is when we have a dictator. The classic theory of the firm tradition covers this special case in which we do have such incontestable veto power. Political parties tend to be organizations in which this condition does not hold, and coalitional stability is of greater concern.

8. Extensions and Open Questions

One-Party Autocratic Systems. The model above considers a democratic system with free entry of political parties and institutions that protect the rights and powers of the players. This is not always the case, even though without these conditions, parties often still form. Some modern examples are in China and Eritrea, and some historical examples are in the Soviet Union and Nazi Germany. One could almost imagine taking the game from Section 5 translating it to an autocratic setting—relabel the Whip as "the autocrat", the legislators as "the elites", the lobbyist as "the masses". Slightly tweak the game from vote buying to some other kind of co-option and violent retaliations. Instead of deposing the leader by some formal majority rule vote, a majority of elites are required to overthrow the leader just by sheer force. Apply the same solution concept (Ali and Liu, 2020).

Where we will find a problem is justifying the interpretation of what the organizational boundaries entail. In the example in Section 5, the interpretation is that being a party member entails transferring their ownership rights of campaign operations or their well-defined/enforced personal privacy rights to the Whip, thereby giving the Whip better information on them. And these are vertically integrated with the legal rights over donation dollars. In general, the four most well-recognized formal theories of the Firm (Gibbons, 2005)—that is Simon (1951), Williamson (1971), Grossman and Hart (1986), Holmström (1999)—all follow the same formula as laid out in Williamson (1985). They take the organization as an *ex ante* power committment to alleviate *ex post* problems with opportunism (Williamson, 1985, p. 48-49). Whether it be transfers of property rights, decision rights, or control, they all follow this recipe even under the most abstract interpretations.

In these weakly-institutionalized settings, how would we justify *ex ante* commitments to transfer power? Without some kind of sufficiently powerful and interested external third party enforcing the agreements between this autocrat and their subordinates, how can this formula of *ex ante* solutions to *ex post* opportunism be coherent? To model parties as organizations in authoritarian settings, one would

[10]As in Baker, Gibbons and Murphy (1999).

have to grapple with a definition of organizational boundaries that does not rely on actions taken today that handcuffs behavior in the future.

One could argue that the two different organizational forms of the firm and the party do not entail related models of boundaries/integration. It could also be that parties in well-institutionalized settings are completely unconnected to parties in weakly-institutionalized setting, at least in respect to boundaries. But it seems intuitive that while there are clear and important differences with these organizational forms, we ought to be able to draw an underlying through line across their theories of existence. And right now it is not clear to me how to do that with the current set of mainstream theories. This raises the question—if indeed there is this missing model for organizational boundaries that can be applied to weakly-institutionalized settings, could it also be something we are not fully capturing in the canonical well-institutionalized firm setting?

Party Building and the Origins of a Party System. We repeatedly note in the multi party model that we are not guaranteed existence of a stable equilibrium—in particular, situations where hypothetical parties are anticipated to all have similar organizational capacities lead to chaos. For old and established party systems, productivity dispersion seems natural—history has led to learning-by-doing, trust being built, and trust being broken. However, this is not the case when a party system is formed for the first time.

One story is the United States. Historical consensus is that our founding fathers were vehemently against the idea of parties and thought their formation would threaten the viability of political order (Hofstadter, 1969; Schattschneider, 1942; Aldrich, 1995). In spite of this opposition, parties formed essentially immediately, and were a necessary step to passing the first ever pieces of legislation, the location of the capital and the structure of state debts. E.E. Schattschneider (1942) went as far to say "Modern democracy is unthinkable save in terms of parties."

In stark contrast to this description of inevitability is the Peru case (Seawright, 2012; Levitsky et al., 2016). During the rise and fall of the antidemocratic populist president Alberto Fujimori, Peru's party system collapsed. Old established parties faded into irrelevance, and since Fujimori, no stable party system has taken root. Chaotically, there are 20+ newly created parties every election, which typically have a lifespan of only 1-2 electoral cycles. There is virtually complete consensus on the desire for stronger parties in Peru, but they still have not taken root.

Under what conditions are parties inevitable in spite of all efforts (like the U.S.), and under what conditions are parties impossible in spite of all efforts (like Peru)? Are parties a necessary condition for stable institutions, or are stable institutions a necessary condition for strong parties, or is it both? An organizational model of party building would help us understand which conditions we must improve for strong parties to naturally develop and take root.

9. Conclusion

In this paper, we model the two faces of political parties at once: the party as an organization and the party as a coalition. The underlying lesson from the chaos theorems (McKelvey, 1976; Schofield, 1978) and games like "split-the-dollar" is that core stability only happens when all of the stars align and cannot generally be expected. Which raises the question—why do we see so much stability empirically?

An underlying assumption in those models is that each decisive coalition can select any political outcome in the policy space, thus generically picking one on their Pareto Frontier. But, as Coase (1937) and Williamson (1985) have stressed, we would not see a world of organizations (like parties) if transaction costs were not pervasive. Additionally, organizations are an imperfect solution to transaction costs—otherwise we would not have markets. Thus, parties are likely far from their Pareto Frontier, and their struggle to get there is foundational to their existence.

We build a model of reputations and intertemporal vote-trading to motivate how a political environment can generically be fraught with nontrivial transaction costs that in turn generate persistent coalitional stability. In a world filled with transaction costs, the stars no longer have to perfectly align for us to get a persistent stable core. Mapping this to real-world examples, we try to explain how it is possible for us to have a system of politics in which stability so often wins out over chaos.

We hope that this demonstrates the fruitful intersection of organization theory and political economy on the topic of political parties. On the political science side, many of the questions asked in regards to political parties already have abundant histories in the organizations literature. The organizations literature could serve as a rich starting point for such an analysis. On the organizations side, political parties and other non-firm organizations present new questions that are not necessarily of first-order relevance to the profit-seeking firms typically studied, like coalitional stability and weakly-institutionalized settings. We think that pursuing these questions not only can help expand our theories of non-firm organizations, but also will force us to reflect on what we may be taking for granted in the firm setting.

References

Aldrich, John H. 1995. *Why Parties? The Origin and Transformation of Political Parties in America. American Politics and Political Economy Series*, Chicago: University of Chicago Press.

Ali, S. Nageeb, and Ce Liu. 2020. "Conventions and Coalitions in Repeated Games." *Working Paper.*

Arrow, Kenneth J. 1951. *Social Choice and Individual Values.* New York: Wiley.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *The Quarterly Journal of Economics*, 109(4): 1125–1156.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1999. "Informal Authority in Organizations." *The Journal of Law, Economics, and Organization*, 15(1): 56–73.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *The Quarterly Journal of Economics*, 117(1): 39–84.

Baron, David P., and John A. Ferejohn. 1989. "Bargaining in Legislatures." *The American Political Science Review*, 83(4): 1181–1206.

Bawn, Kathleen, and Frances Rosenbluth. 2006. “Short versus Long Coalitions: Electoral Accountability and the Size of the Public Sector.” *American Journal of Political Science*, 50(2): 251–265.

Besley, Timothy, and Stephen Coate. 1997. “An Economic Model of Representative Democracy.” *The Quarterly Journal of Economics*, 112(1): 85–114.

Canen, Nathan, Chad Kendall, and Francesco Trebbi. 2020. “Unbundling Polarization.” *Econometrica*, 88(3): 1197–1233.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. “Firms and Labor Market Inequality: Evidence and Some Theory.” *Journal of Labor Economics*, 36(S1): S13–S70.

Clark, Champ. 1920. *My Quarter Century of American Politics.* New York, NY: Harper & Brothers.

Coase, R. H. 1937. “The Nature of the Firm.” *Economica*, 4(16): 386–405.

Coase, R. H. 1960. “The Problem of Social Cost.” *The Journal of Law and Economics*, 3: 1–44.

Gibbons, Robert. 2005. “Four Formal(izable) Theories of the Firm?” *Journal of Economic Behavior & Organization*, 58(2): 200–245.

Gibbons, Robert, and Rebecca Henderson. 2012. “Relational Contracts and Organizational Capabilities.” *Organization Science*, 23(5): 1350–1364.

Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade. Political Economy of Institutions and Decisions*, Cambridge:Cambridge University Press.

Grossman, Sanford, and Oliver Hart. 1986. “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration.” *Journal of Political Economy*, 94(4): 691–719.

Hart, Oliver, and John Moore. 1990. “Property Rights and the Nature of the Firm.” *Journal of Political Economy*, 98(6): 1119–1158.

Hart, Sergiu, and Mordecai Kurz. 1983. “Endogenous Formation of Coalitions.” *Econometrica*, 51(4): 1047–1064.

Hofstadter, Richard. 1969. *The Idea of a Party System: The Rise of Legitimate Opposition in the United States, 1780-1840.* Berkeley, Los Angeles, and London:University of California Press.

Holmström, Bengt. 1999. “The Firm as a Subeconomy.” *Journal of Law, Economics, & Organization*, 15(1): 74–102.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. “Misallocation and Manufacturing TFP in China and India.” *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Jackson, Matthew O., and Boaz Moselle. 2002. “Coalition and Party Formation in a Legislative Voting Game.” *Journal of Economic Theory*, 103(1): 49–87.

Krehbiel, Keith. 1993. “Where’s the Party?” *British Journal of Political Science*, 23(2): 235–266.

Kreps, David. 1990. “Corporate Culture and Economic Theory.” In *Perspectives on Positive Political Economy. Political Economy of Institutions and Decisions*, 90–143. Cambridge:Cambridge University Press.

Langston, Joy Kathryn. 2017. *Democratization and Authoritarian Party Survival: Mexico's PRI.* Oxford:Oxford University Press.

Levitsky, Steven. 2003. *Transforming Labor-Based Parties in Latin America: Argentine Peronism in Comparative Perspective.* Cambridge, UK:Cambridge University Press.

Levitsky, Steven, James Loxton, Brandon Van Dyck, and Jorge I. Domínguez, ed. 2016. *Challenges of Party-Building in Latin America.* Cambridge:Cambridge University Press.

Levy, Gilat. 2004. "A Model of Political Parties." *Journal of Economic Theory*, 115(2): 250–277.

McKelvey, Richard D. 1976. "Intransitivities In Multidimensional Voting Models and Some Implications for Agenda Control." *Journal of Economic Theory*, 12(3): 472–482.

Noel, Hans. 2013. *Political Ideologies and Political Parties in America.* Cambridge:Cambridge University Press.

North, Douglass C. 1990. *Institutions, Institutional Change, and Economic Performance. Political economy of institutions and decisions*, Cambridge:Cambridge University Press.

Osborne, Martin J., and Al Slivinski. 1996. "A Model of Political Competition with Citizen-Candidates." *The Quarterly Journal of Economics*, 111(1): 65–96.

Osborne, Martin J., and Ariel Rubinstein. 1994. *A Course in Game Theory.* Cambridge, MA, USA: MIT Press.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* Cambridge: Cambridge University Press.

Ray, Debraj. 2007. *A Game-Theoretic Perspective on Coalition Formation.* Oxford: Oxford University Press.

Rohde, David W. 1991. *Parties and Leaders in the Postreform House.* Chicago: University of Chicago Press.

Schattschneider, E. E. 1942. *Party Government. American Government in Action Series*, New York:Holt, Rinehart and Winston.

Schofield, Norman. 1978. "Instability of Simple Dynamic Games." *The Review of Economic Studies*, 45(3): 575–594.

Seawright, Jason. 2012. *Party-System Collapse: The Roots of Crisis in Peru and Venezuela.* Stanford, CA:Stanford University Press.

Shepsle, Kenneth A. 1979. "Institutional Arrangements and Equilibrium in Multidimensional Voting Models." *American Journal of Political Science*, 23(1): 27–59.

Simon, Herbert. 1951. "A Formal Theory of the Employment Relationship." *Econometrica*, 19(3): 293–305.

Snyder, James M. 1991. "On Buying Legislatures." *Economics & Politics*, 3(2): 93–109.

Snyder, James M., and Michael M. Ting. 2002. "An Informational Rationale for Political Parties." *American Journal of Political Science*, 46(1): 90–110.

Syverson, Chad. 2004. "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy*, 112(6): 1181–1222.

Syverson, Chad. 2011. “What Determines Productivity?” *Journal of Economic Literature*, 49(2): 326–365.

von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior.* Princeton N.J.:Princeton University Press.

Weingast, Barry R., and William J. Marshall. 1988. “The Industrial Organization of Congress; or, Why Legislatures, Like Firms, Are Not Organized as Markets.” *Journal of Political Economy*, 96(1): 132–163.

Williamson, Oliver E. 1971. “The Vertical Integration of Production: Market Failure Considerations.” *The American Economic Review*, 61(2): 112–123.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting.* New York: The Free Press.

Appendix A. One-Period Model

We consider the one-period model from Section 2 with $s_i \notin \{-\sigma, \sigma\}$.

Preliminaries. We note that when $x \geq 0$, a legislator i is indifferent between the alternative policy x and the status-quo if $\theta_i = \frac{x-b_i}{2}$. For finding the optimal policy of the lobbyist, it is without loss of generality to consider policies $x \geq 0$, since $(x, \mathbf{b}) = (0, \mathbf{0})$ clearly dominates all policies $x < 0$ for all possible voting outcomes. We define $S \colon \mathbb{R} \times \mathbb{R}^n_{\geq 0} \times \mathbb{R}^n \rightrightarrows N$ as the set of legislators whose vote could swing either way given the signals and bribes, for $x \geq 0$:

$$S(x, \mathbf{b}, \mathbf{s}) \coloneqq \left\{ i \in N : \frac{x - b_i}{2} - \sigma \leq s_i \leq \frac{x - b_i}{2} + \sigma \right\}.$$

Similarly, we define the set of legislators $H \colon \mathbb{R} \times \mathbb{R}^n_{\geq 0} \times \mathbb{R}^n \rightrightarrows N$ who will certainly holdout and not vote for the proposal as

$$H(x, \mathbf{b}, \mathbf{s}) \coloneqq \left\{ i \in N : s_i < \frac{x - b_i}{2} - \sigma \right\}.$$

Thus, we can rewrite the lobbyists objective problem (2.4) as

$$\max_{x, \mathbf{b}} \left\{ \beta x \cdot \left[\prod_{i \in S(x, \mathbf{b}, \mathbf{s})} \left(\frac{1}{2} + \left(\frac{2s_i - x + b_i}{4\sigma} \right) \right) \right] \cdot \mathbb{1}_{\{H(x,\mathbf{b},\mathbf{s}) = \varnothing\}} - \sum_{i \in N} b_i \right\}. \tag{A.1}$$

Now, we prove Lemmas 2.1 and 2.2. First, we prove a single intermediate lemma (Lemma A.1), and then prove Lemma 2.1 and Lemma 2.2 immediately after.

Lemma A.1. *The solution to the optimization problem* (A.1)*, if not a corner (*$\mathbf{b} \neq \mathbf{0}$*), takes the form of a probability floor—there exists some* $\underline{p}$ *such that for all legislators such that, when left unbribed, will vote "Yes" on* x *with probability is less* $\underline{p}$*, they will be bribed up to the point their probability of voting "Yes" is exactly* $\underline{p}$*. All other legislators receive no bribe.*

Proof. To see this, first note that obviously any interior solution $\mathbf{b} \neq \mathbf{0}$ must be such that $H(x, \mathbf{b}, \mathbf{s}) = \varnothing$—there is no point in spending positive bribes for a policy that has no chance of winning. Also, there is no point in bribing someone strictly beyond the point they will vote for the policy with certainty, $i \notin S(x, \mathbf{b}, \mathbf{s}) \implies (b_i = 0)$. This just leaves the bribing of legislators in $S(x, \mathbf{b}, \mathbf{s})$. For those legislators $i \in S(x, \mathbf{b}, \mathbf{s})$, the partial derivative of the expected policy benefit in respect to b_i is $\beta x \cdot \mathbb{P}(\mathbf{v}_{-i} = \mathbf{1})$. This value is highest for those legislators that have the lowest probability of voting "Yes" themselves—thus the highest probability everyone besides them is voting "Yes". Since you can always keep the total bribe cost constant by shifting bribes from one legislator to another, it cannot be that both $b_i > 0$ and that there exists some j such that $\mathbb{P}(v_i = 1) > \mathbb{P}(v_j = 1)$. Otherwise, by shifting some of the bribe from i to j, there is a net expected policy-benefit and no change in bribe costs, a contradiction of $\mathbf{b}$ being optimal. ◻

Lemma 2.1. *For all signal realizations* $\mathbf{s}$*, the optimal policy for the lobbyist* $(x, \mathbf{b})$*, if not a corner* $(\mathbf{b} \neq \mathbf{0})$*, will be such that the policy passes with certainty,* $\mathbb{P}(\mathbf{v} = \mathbf{1}) = 1$*.*

Proof. Define $C\colon \mathbb{R} \times [0,1] \times \mathbb{R}^n \rightrightarrows N$ as the set of legislators that do not meet the probability-floor $\underline{p}$ without bribery for a given policy x and signal $\mathbf{s}$:

$$C\left(x, \underline{p}, \mathbf{s}\right) \coloneqq \left\{ i \in N : \frac{1}{2} + \frac{2s_i - x}{4\sigma} \leq \underline{p} \right\}.$$

Define $\widetilde{S}\colon \mathbb{R} \times [0,1] \times \mathbb{R}^n \rightrightarrows N$ as the set of no-bribe swinging legislators that are above the probability-floor:

$$\widetilde{S}\left(x, \underline{p}, \mathbf{s}\right) \coloneqq S\left(x, \mathbf{0}, \mathbf{s}\right) \setminus C\left(x, \underline{p}, \mathbf{s}\right).$$

Then, given Lemma A.1, we can rewrite the objective (A.1) for a given x as:

$$\max_{\underline{p}} \left\{ \beta x \underline{p}^{\left|C\left(x,\underline{p},\mathbf{s}\right)\right|} \left(\prod_{i \in \widetilde{S}\left(x,\underline{p},\mathbf{s}\right)} \left[\frac{1}{2} + \frac{2s_i - x}{4\sigma} \right] \right) - \sum_{j \in C\left(x,\underline{p},\mathbf{s}\right)} \left(x - 2s_j - 2\sigma + 4\sigma \underline{p} \right) \right\}. \tag{A.2}$$

Note this function is continuous and almost-everywhere differentiable, with a finite set of kinks, which are just the jump-discontinuity values of p such $C\left(x, \underline{p}, \mathbf{s}\right)$ starts including new legislators. When it exists, the partial derivative of (A.2) in respect to $\underline{p}$ is

$$\left|C\left(x, \underline{p}, \mathbf{s}\right)\right| \left[\beta x \left(\prod_{i \in \widetilde{S}\left(x,\underline{p},\mathbf{s}\right)} \left[\frac{1}{2} + \frac{2s_i - x}{4\sigma} \right] \right) \left(\underline{p}^{\left|C\left(x,\underline{p},\mathbf{s}\right)\right| - 1} \right) - 4\sigma \right].$$

Note that when this is positive, it is increasing in $\underline{p}$ (including across kinks) which means (A.2) is convex in the range it is increasing. Thus it achieves a maximum at either at $\underline{p} = 0$ or $\underline{p} = 1$. In other words, by Lemma A.1, either there is no bribing or bribing to the point of certainty. $\square$

Lemma 2.2. *If* $\{s_i\}_i \cap \{\sigma\} = \varnothing$*, there exists an optimal policy for the lobbyist* $(x, \mathbf{b})$ *will be such that the policy passes with certainty,* $\mathbb{P}(\mathbf{v} = \mathbf{1}) = 1$*.*

Proof. By Lemma 2.1, it is sufficient to prove this for $\mathbf{b} = 0$. It is without loss of generality to exclude $x < 0$, since the lobbyist being right-biased implies that any such strategy must be dominated by $(0, \mathbf{0})$. Then, if $\min\{s_i : i \in N\} < 0$, $\mathbf{b} = 0$ implies the policy will not pass with certainty, and thus it is equivalent to $(0, \mathbf{0})$ passing with certainty. Thus, we restrict attention to the case in which $\min\{s_i : i \in N\} > 0$, $x > 0$, and $\mathbf{b} = 0$.

First, note that for all $i \in N$, we have that $\mathbb{P}(v_i = 1 \mid x, b_i)$ is weakly decreasing x when $x \geq 0$. This implies that $\mathbb{P}(\mathbf{v}_{-i} = \mathbf{1} \mid x, \mathbf{b})$ also is decreasing in x for all i, as the distributions are independent. For a given $\mathbf{s}$, take $s_l \coloneqq \min\{s_i : i \in N\}$ to be the lowest signal in $\mathbf{s}$. A proposal $x < 2(s_l - \sigma)$ is never optimal, since even at

$\mathbf{b} = \mathbf{0}$, it is possible to acquire $x = 2(s_l - \sigma)$ with probability one, which is strictly better for the lobbyist since they are right-biased. If $\mathbf{b} = \mathbf{0}$, then $x > 2(s_l + \sigma)$ cannot be optimal, since it will fail by the lowest-signal legislator for certain, which again is dominated by policy win $x = 2(s_l - \sigma)$.

Note that the monotonicity of $\mathbb{P}(\mathbf{v}_{-i} = \mathbf{1} \mid x, \mathbf{b})$, given $s_l > \sigma$, implies that

$$\begin{aligned}
&\left(\{0\} = \operatorname*{arg\,max}_{\varepsilon\in[0,2\sigma]} \{2\beta(s_l - \sigma + \varepsilon)\mathbb{P}(v_l = 1 \mid x = 2(s_l - \sigma + \varepsilon), \mathbf{b} = \mathbf{0})\}\right) \\
\implies &\left(\{0\} = \operatorname*{arg\,max}_{\varepsilon\in[0,2\sigma]} \{2\beta(s_l - \sigma + \varepsilon)\mathbb{P}(\mathbf{v} = \mathbf{1} \mid x = 2(s_l - \sigma + \varepsilon), \mathbf{b} = \mathbf{0})\}\right).
\end{aligned} \tag{A.3}$$

The latter expression can only do even worse at the higher $\varepsilon \in (0, 2\sigma]$ since it's the former multiplied by the weakly-decreasing expression $\mathbb{P}(\mathbf{v}_{-i} = \mathbf{1} \mid x, \mathbf{b})$. By the previously-discussed constraints on the optimum x given $\mathbf{b} = \mathbf{0}$, the latter describes the lobbyist's optimization problem at $\mathbf{b} = \mathbf{0}$. Thus, to find a sufficient condition on the former:

$$\begin{aligned}
f(\varepsilon) &\coloneqq 2\beta(s_l - \sigma + \varepsilon)\mathbb{P}(v_l = 1 \mid x = 2(s_l - \sigma + \varepsilon), \mathbf{b} = \mathbf{0}) \\
&= 2\beta(s_l - \sigma + \varepsilon)\left(1 - \frac{\varepsilon}{2\sigma}\right) \\
\implies f'(\varepsilon) &= \frac{\beta}{\sigma}(3\sigma - s_l - 2\varepsilon) \\
\therefore s_l > 3\sigma &\implies \left(\operatorname*{arg\,max}_{\varepsilon\in[0,2\sigma]} \{f(\varepsilon)\} = \{0\}\right).
\end{aligned} \tag{A.4}$$

The condition that $\{s_i\}_i \cap \{\sigma\} = \varnothing$ will imply that the lowest signal will also be such that $s_l > 3\sigma$. Thus, if $\{s_i\}_i \cap \{\sigma\} = \varnothing$, we know from (A.3) and (A.4) that even when $\mathbf{b} = \mathbf{0}$, it will still be optimal for the lobbyist to go for a policy proposal that wins with certainty, $x = 2(s_l - \sigma)$. By Lemma 2.1, this completes the proof. $\square$

Appendix B. Repeated Game

We consider the game setup from Sections 5.

B.1. **Defining ICE.** First, to impose a restriction that depends on periods of time, where a period may consist of many information sets, we must formalize what it means for there to be periods. We adapt the definition and notation of extensive form games with imperfect information from the textbook of Osborne and Rubinstein (1994, Definition 200.1). The only alteration to their definition made comes from the component we call the *period function* and its relation to information sets, and we intentionally keep the other components consistent with how they appear in the text for those readers already familiar with the book.

Definition B.1. An *extensive game with a calendar* has the following components.

- A finite set N (the set of *players*).
- A set H of sequences (finite or infinite) that satisfies the following three properties.
  - The empty sequence $\varnothing$ is a member of H.
  - If $\left(a^k\right)_{k=1,\ldots,K} \in H$ (where K may be infinite) and $L < K$ then $\left(a^k\right)_{k=1,\ldots,L} \in H$.
  - If an infinite sequence $\left(a^k\right)_{k=1}^{\infty}$ satisfies $\left(a^k\right)_{k=1,\ldots,L} \in H$ for every positive integer L then $\left(a^k\right)_{k=1}^{\infty} \in H$.

  (Each member of H is a *history*; each component of a history is an *action* taken by a player.) A history $\left(a^k\right)_{k=1,\ldots,K} \in H$ is *terminal* if it is infinite or if there is no a^{K+1} such that $\left(a^k\right)_{k=1,\ldots,K+1} \in H$. The set of actions available after the nonterminal history h is denoted $A(h) = \{a : (h, a) \in H\}$ and the set of terminal histories is denoted Z.
- A function P that assigns to each nonterminal history (each member of $H \setminus Z$) a member of $N \cup \{c\}$. (P is the *player function*, $P(h)$ being the player who takes an action after the history h. If $P(h) = c$ then chance determines the action taken after history h.)
- A function T that assigns to each nonterminal history an element of the natural numbers $\mathbb{N}$ and satisfies the following three properties.
  - If $(a^k)_{k=1,\ldots,K} \in H \setminus Z$, then for all positive integers $L < K$, it must be that $T\left((a^k)_{k=1,\ldots,L}\right) \leq T\left((a^k)_{k=1,\ldots,K}\right)$.
  - For every $(a^k)_{k=1,\ldots,K} \in H \setminus Z$ and $L < K$, if there exists some $\tau \in \mathbb{N}$ such that $T\left((a^k)_{k=1,\ldots,L}\right) < \tau < T\left((a^k)_{k=1,\ldots,K}\right)$, then there exists some $J \in \mathbb{N}$ with $L < J < K$ such that $T\left((a^k)_{k=1,\ldots,J}\right) = \tau$.

  - The empty set is the unique history with a period of zero, that is $T(\varnothing) = 0$ and $h \neq \varnothing \implies T(h) > 0$.

  (T is the *period function*, $T(h)$ being the *period* in which the action that is taken after the history h occurs in.)

- A function f_c that associates with every history h for which $P(h) = c$ a probability measure $f(\cdot \mid h)$ on $A(h)$, where each such probability measure is independent of every other such measure. ($f_c(a \mid h)$ is the probability that a occurs after history h.)
- For each player $i \in N$ a partition $\mathcal{I}_i$ of $\{h \in H : P(h) = i\}$ with the property that $A(h) = A(h')$ and $T(h) = T(h')$ whenever h and h' are in the same member of the partition. For $I_i \in \mathcal{I}_i$ we denote by $A(I_i)$ the set $A(h)$, by $P(I_i)$ the player $P(h)$, and by $T(I_i)$ the period $T(h)$ for any $h \in I_i$. ($\mathcal{I}_i$ is the *information partition* of player i; a set $I_i \in \mathcal{I}_i$ is an *information set* of player i.)
- For each player $i \in N$ a partition $\mathcal{I}_i^0$ of H^0, where H^0 denotes the set of *period-initiating histories*:

$$H^0 := \left\{ \left(a^k\right)_{k=1}^{K} \in H \setminus Z : T\left(\left(a^k\right)_{k=1}^{K}\right) = T\left(\left(a^k\right)_{k=1}^{K-1}\right) + 1 \right\} \cup \{\varnothing\},$$

  with the property that $T(h) = T(h')$ whenever h and h' are in the same member of the partition. For each $h \in H^0$, we denote $I_i^0(h)$ to be the element of $\mathcal{I}_i^0$ in which h is in, that is $h \in I_i^0(h) \in \mathcal{I}_i^0$. ($\mathcal{I}_i^0$ is the *period-initiating information partition* of player i; a set $I_i^0 \in \mathcal{I}_i^0$ is an *period-initiating information set* of player i.)
- For each player $i \in N$ a preference relation $\succsim_i$ on lotteries over Z (the *preference relation* of player i) that can be represented as the expected value of a payoff function defined on Z.

In summary, we define this concept of a period function, which maps histories to which period they occur in. Those first three properties I refer to as the *time-consistency* properties, which requires that time passes monotonically, we don't skip dates, and we start at period one, where before that is period zero. Secondly, there are the *time-awareness* properties, which have to do with the information sets and period-initiating information sets. Intuitively, we are assuming all players are wearing a watch—that is, they can always tell what time it currently is when they play, and they know exactly when time changes from one period to the next.

Now, the definition of intratemporal coalitional equilibrium, applying the loose definition of weak perfect Bayesian equilibrium (PBE) that requires players to update based on Bayes' rule "whenever possible". Without rewriting the full definitions, recall from Osborne and Rubinstein (1994, Definition 222.1) that an *assessment* (β, μ) consists of a *profile of behavioral strategies* β, and a *belief system* μ that assigns to every information set a probability measure on the set of histories in the information set. And Osborne and Rubinstein (1994, p. 223) define the *outcome* $O(\beta, \mu \mid I)$ *of* (β, μ) *conditional on* I to be the distribution over terminal histories determined by β and μ conditional on information set I being reached. We extend these belief concepts to the period-initiating information sets in the natural way.

Definition B.2. An assessment (β, μ) is a *Intratemporal Coalitional Equilibrium* (ICE) of an extensive game with a calendar $\langle N,\ H,\ P,\ T,\ f_c,\ (\mathcal{I}_i),\ \left(\mathcal{I}_i^0\right),\ (\succsim_i)\rangle$ if there exists no coalition $C \subseteq N$, history $h \in H^0$, and profile of behavioral deviations β' that satisfy the following two properties.

- For each coalition member $i \in C$, they are strictly better off playing β' than β given that $I_i^0(h)$ has been reached, $O\left(\beta', \mu \mid I_i^0(h)\right) \succ_i O\left(\beta, \mu \mid I_i^0(h)\right)$.
- The deviation β' does not change for any players outside of C or for information sets outside of period $T(h)$, that is for all information sets $I \in \cup_i \mathcal{I}_i$

$$(P(I) \notin C) \vee (T(I) \neq T(h)) \implies \left(\beta'_{P(I)}(I) = \beta_{P(I)}(I)\right).$$

Thus, in this refinement of weak perfect Bayesian equilibrium, the equilibrium must be robust to coalitional deviations that last no more than a single period.

Again, this is essentially a specific instance of the concept in Ali and Liu (2020), *stable convention in the NTU repeated game*. There are just a few very technical differences. First, we have infinite action spaces, while they restrict to finite action spaces. Second, we have incomplete information, and thus apply the perfect Bayesian criteria on beliefs and corresponding expected utility assessments.

And third, they already structure the game as a infinitely-repeated game with discounting over the periods and a felicity function defined on each period outcome. There is a simple reason we do not impose this for our game, even though it is a infinitely-repeated game with discounting. We wanted the coalitional planning to take place from the perspective of the legislators already knowing their preferences on the issue, θ_{it}. One of the reasons we argue for this solution concept is that it is much harder to form coalitional plans on hypothetical future bills of unknown content, and so if we are at a point where the coalition can form a plan in regards to a bill, the coalition members ought to know there preferences on said bill.

To achieve this restriction, we will define the period function T such that Nature gets her own period to herself, and thus each period with the legislators starts after Nature has drawn that session's $\boldsymbol{\theta}_t, \mathbf{s}_t^L$. This way of defining the game works well for our purposes, but incorporating the discounting structure into our definition would have us defining one period structure for the deviations (with these "fake" Nature periods) and another period structure for the discounting and felicity (without Nature periods). We thought this might be confusing, so we elected to use the more general form of preferences instead that do not directly build on a time structure.

B.2. **The Party Equilibrium.**

Preliminaries. First, to translate the repeated game from Section 3 into an extensive form game with a calendar structure. The non-calendrical components are all the canonical definitions $\langle N, H, P, f_c, (\mathcal{I}_i), (\succsim_i)\rangle$. As in our discussion of Definition B.2, the period function T will be such that Nature gets her own period to draw the issue topic $\boldsymbol{\theta}_t$ and its public signal $\mathbf{s}_t^L$. Each legislative session period will start after the public signal of the topic is drawn $\mathbf{s}_t^L$, and ends at the last action taken before $\boldsymbol{\theta}_{t+1}$ is drawn. The period initiating information sets $\left(\mathcal{I}_i^0\right)$ will be such that each legislators knows their own type θ_{it} and the realization of the public signal $\mathbf{s}_t^L$, and the lobbyist just knows the public signal $\mathbf{s}_t^L$.

Now we work towards proving Theorem 5.4. First, we prove that the no-party off-path satisfies the conditions for ICE (Lemma 5.3). The intuition is related to an observation from Ali and Liu (2020) regarding their version of the concept. It is an analogous result to how when one considers subgame perfect equilibria in repeated one-stage simultaneous move games, there is always an equilibrium for all $\delta \in [0, 1)$ that just implements a Nash equilibrium of the stage game every period regardless of history. The path which picks a core alternative for every period independent of history is always a *stable convention* in Ali and Liu (2020). For ICE, each period (independent of history) we can pick (if it exists) a strategy profile that is (1) perfect Bayes equilibia of the period-long extensive form game, and (2) has a period-outcome that is in the core of the normal form representation of said extensive form game. The no-party equilibrium satisfies both of these conditions.

Lemma 5.3. *The no-party equilibrium is an intratemporal coalitional equilibrium for all parameter values.*

Proof. By Proposition 3.1, the no-party equilibrium is a perfect Bayes equilibrium, thus is satisfies the first condition of ICE.

First, note that fixing the lobbyist's behavior as is in the no-party equilibrium, there are no externalities across players with actions. The Whip has no additional actions, since the party is not being donated to. Then for each legislator i, their utility depends on how they vote v_{it}, and the choices of the lobbyist b_{it} and x_t, assumed to be fixed. Thus, there optimization problems are separable, and so the fact that they are individually optimizing (Theorem 2.3) means they are achieving the best they could as a coalition. Thus, any coalition we consider must include the lobbyist.

Since the no-party equilibrium is Markovian, it is without loss to consider the one period game since any coalitional deviation has no impact on the continuation path. Since the lobbyist is in the coalition, it must be that $\mathbf{v}_t = \mathbf{1}$, since they receive a strictly positive payoff in the no party equilibrium, and a weakly-negative payoff if $\mathbf{v}_t \neq \mathbf{1}$. For any coalitional deviation that uses the party path with policy x_t', the belief that their policy vote does not effect their party dollar allocation implies that all i such that $-|\theta_{it}| < -|\theta_{it} - x_t'|$ must be in the coalition to vote yes.

Let $x_t^L\left(\mathbf{s}_t^L, \phi_t\right)$ and $\mathbf{b}_t^L\left(\mathbf{s}_t^L, \phi_t\right)$ denote the no party policies from Theorem 2.3 with $\sigma = \sigma^L$ and $\mathbf{s} = \mathbf{s}_t^L$. For any coalitional deviation policy x_t' that is lobbyist preferred $\phi_t x_t' \geq \phi_t x_t^L\left(\mathbf{s}_t^L, \phi_t\right)$, it must be that every legislator such that was receiving positive bribes before $b_{it}^L(\cdot) > 0$ is in the coalition (otherwise they vote no). They will only join the coalition if $-\mathrm{E}\left[x_t'\right] + 2\theta_{it} + \mathrm{E}\left[b_{it}'\right] > -x_t^L + 2\theta_{it} + b_{it}^L$, where $-x_t^L + 2\theta_{it} + b_{it}^L \geq 0$. Minus some mean-zero random noise, these constraints are strictly more restrictive and the payoff function weakly lower than the objective (2.4). Thus, payoff to the lobbyist must be weakly lower, a contradiction. Therefore, there is no profitable coalitional deviation. □

For this next result, I will a minor technical assumption: the behavior of the legislators in the deposition stages never conditions on their private histories, only public histories and possibly a public correlating device drawn right before the deposition process.

This assumption is almost without loss of generality—their private histories are payoff irrelevant at this stage, so conditioning on their private histories is like using a public correlating device. All this does is delay the de facto correlating device's draw until the deposition process, instead of having partial asymmetric private information on the draw's realization during the coalitional planning phase.

Proposition 5.2. *For each PBE in which there are donations to the party on path, there exists a $\bar{\delta} \in (0,1]$ such that for all $\delta \in [0, \bar{\delta})$, the equilibrium is not an ICE.*

Proof. First, we need to put some bounds on time-averaged continuation payoffs for the players in a PBE.

By $\beta < n$, we know that expected total-surplus is bounded. Letting median* denote $\lceil (n + \phi_t \beta)/2 \rceil$ (the effective median counting the lobbyist), the expected maximum surplus is achieved under the following conditions. With probability one for all t, $x_t = \theta_t^{\text{median}^*}$, $\mathbf{v}_t = \mathbf{1}$, and $D_t = \sum_i d_{it}$. However, total D_t and $\mathbf{b}_t$ can take on any values since they are just transfers.

Each legislator can always vote for the status-quo and money cannot be taken away from them even when being minmaxed, and so they have a expected time-averaged continuation value of at least $-\mathrm{E}\left[\theta_{it}\right]$ (finite since the support is finite). The Lobbyist can always play a strategy of $\mathbf{b}_t = \mathbf{0}$ and $D_t = 0$ even when being minmaxed, and so they have a expected time-averaged continuation value of at least 0. Since total-surplus is bounded, these gives us maximum expected time-averaged continuation values for all players as well. Thus, we can put a bound on the difference in payoffs between two continuation paths for a legislator in a PBE—the difference is no greater than the bound on the total surplus plus $\mathrm{E}\left[\theta_{it}\right]$. We denote this as ε.

Consider any period-initiating information set $h \in H^0$ such that there exists some public event $E \in \left(\left\{(2k+1)\frac{\sigma^L}{c}\right\}_{k \in \mathbb{Z}}\right)^n \times \{0,1\}^n$ with the properties that

$$\mathbb{P}\left(\left(\mathbf{s}_{T(h)}^W, \mathbf{v}_{T(h)}\right) = E \mid I_i^0(h)\right) > 0 \text{ for all } i \in N \tag{B.1}$$

and

$$\bar{D} \coloneqq \mathrm{E}\left[D_{T(h)} \mid I_L^0\left(h\right), E\right] > 0. \tag{B.2}$$

The expression (B.1) is reasonably satisfied because the signal space is finite.

Note that the second probability measure is taken from the perspective of the analyst, seeing that within the game, the players only fully observe event E (which includes the vote) after they have perfectly observed $D_{T(h)}$.

This is what we will mean by "on-path party donations", which we have to be a little careful about due to measurability issues and the asymmetric private information. For example, it does not count as "on-path" if $D > 0$ only occurs when a legislator draws a type exactly equal to Euler's constant, where at the period initiating stage all other legislators would take that as a measure zero event. These kinds of perverse strategies seem to be ruled out by subgame optimality anyways, but proving so would greatly complicate the proof for little additional insight gained on a technicality.

The public information is equivalent to the Lobbyist's information $I_L^0\left(\cdot\right)$. Given such an $h \in H^0$, and E, take our coalition to be the $\left(n+1\right)/2$ as the legislators with the lowest values of

$$\mathrm{E}\left[d_{iT(h)} \mid E, I_L^0\left(h\right)\right]. \tag{B.3}$$

By assumption 2, this is equal to $\mathrm{E}\left[d_{iT(h)} \mid E, I_j^0\left(h\right)\right]$ for all j, since the Whip conditions on public information and so do the legislators for determining $\mathbf{d}_{T(h)}$. E and $I_L^0\left(h\right)$ entails all of the public knowledge preceding $D_{T(h)}$ and $\mathbf{d}_{T(h)}$.

Thus, if they make a coalitional plan for whenever E occurs, each coalitional member keeps the amount of party dollars they would have gotten in equilibrium *plus* they divide up whatever party dollars that were supposed to be burnt or go to non-coalition members, they increase each coalition member's average-discounted expected payoffs today by at least

$$\left(1-\delta\right)\frac{n+1}{\left(n-1\right)n}\bar{D}. \tag{B.4}$$

Each coalition member's future loss is at most

$$-\delta\varepsilon. \tag{B.5}$$

Thus, for all δ sufficiently low, this must amount to a profitable deviation for all coalition members. $\square$

Appendix C. Multi-Party Game

Theorem 6.4. *Take $N = 3$, $n_I = n_M = 1$. For each pair of legislators $i, j \in N$ and any set of parameter values, there exists a $\underline{\delta} \in (0, 1]$ such that the party configuration with party $\{i, j\}$ satisfies conditions i. and ii. of Definition 6.1 if and only if $\delta \in (\underline{\delta}, 1)$ (empty when $\underline{\delta} = 1$). Furthermore, $\underline{\delta}$ is weakly-decreases when there is*

- *(i) Greater Party Balance—conditional on when i and j "work together" (one is the moderate, the other is the ideologue), the probability that i is the moderate and j is the ideologue becomes more even with j being the moderate and i the ideologue.*
- *(ii) Acceleration of Money in Politics—the budget becomes B'_t where there is some positive monotonically increasing $(\gamma_t)_{t\in\mathbb{N}}$ such that $B'_t = \gamma_t B_t$.*
- *(iii) Ideological Preferences Become More Intense—β increases.*
- *(iv) More Issues that the Party can Work On—the probability of an $\{i, j\}$ moderate-ideological pairing increases, relative to both being at the ideological poles.*

Proof. It is sufficient to check that neither legislator in the party wants to leave in the periods in which they are the moderate. This is because merging to a party of three is outcome equivalent with more incentive constraints to check, and this event is the most tempting deviation because it is the only one with a strictly positive benefit within the current period (all matching in future periods).

Let p_k be the probability that k is the moderate, which is a sufficient statistic to characterize the distribution when $N = 3$. The inequality that must be satisfied for i to want to remain in a party with j in a period t where i is a moderate, j is the ideologue, and the lobbyist is lobbying in the direction of j is

$$B_t\left(\frac{\lambda_I-\lambda_P}{(1+\lambda_P)(1+\lambda_I)}\right) \le \sum_{\tau=1}^{\infty}\delta^{\tau}\frac{1}{2}B_{t+\tau}\left(\beta p_j - p_i\right)\left(\frac{\lambda_I-\lambda_P}{(1+\lambda_P)(1+\lambda_I)}\right)$$
$$\iff (\beta p_j - p_i) \ge \frac{2B_t}{\sum_{\tau=1}^{\infty}\delta^{\tau}B_{t+\tau}}.$$

Combining this with the symmetric condition for j:

$$\beta\min\{p_j, p_i\} - \max\{p_j, p_i\} \ge \frac{2B_t}{\sum_{\tau=1}^{\infty}\delta^{\tau}B_{t+\tau}}. \tag{C.1}$$

Expression (C.1) characterizes $\underline{\delta}$, with $\beta\min\{p_j, p_i\} - \max\{p_j, p_i\} < 0$ implying that $\underline{\delta} = 1$. All four properties follow immediately from (C.1), where more precisely balancing means raising the minimum or lowering the maximum, and more issues means increasing both probabilities by some scalar. ∎

Theorem 6.10. *Fix any party configuration Π and full-support probability distribution over the state space $p \in \Delta\Omega$. There exists a $\underline{\delta}$ such that for all $\delta \in (\underline{\delta}, 1)$, the set of organizational capacities for which Π has a robustly stable agenda with respect to and the set of organizational capacities for which it does not are both non-empty.*

Proof. First we will prove the following Lemma.

Lemma C.1. *Fix any party configuration Π and capacity K such that all active parties have non-zero capacity, $K(\pi) > 0$ for all $\pi \in \Pi$. If the probability distribution over the state space $p \in \Delta\Omega$ is full-support, then there exists a $\underline{\delta}$ such that for all $\delta \in (\underline{\delta}, 1)$ there exists a capacity-respecting agenda A where each party π is stable with respect to disintegration, i.e. there is no member of $i \in \pi$ that is better off in the world $\Pi' = (\Pi \setminus \{\pi\}) \cup \{\{j\} : j \in \pi\}$.*

Proof. With this capacity construction, the only thing we need to "avoid the Manchin problem" is a full-support probability distribution. The agenda we pick can compensate any probability imbalance by asking the legislators who are most often moderates to give up very little when they are moderates, relative to how much the legislators that most often get to be ideologues give up in the rare moments they happen to be moderates. In particular, since the state-space is finite we can define a $\underline{p} := \min\{p(\omega) \mid \omega \in \Omega\} > 0$. Then, for each active party π, let $A_\pi(m, \omega) = \lambda_I - \left(\underline{p}/p(\omega)\right) K(\pi)$. With this agenda, the members of each party have equal payoffs less their disintegration payoffs. We know that the party has higher total payoff than disintegration from the argument of Proposition 6.3 (result extends immediately to > 2 players). Thus, each party member is better off in the future with the party than without it, and so with sufficient patience these future payoffs deter any one period deviation. $\square$

Then, the result follows from (1) setting $K(\pi) > 0$ for all active parties $\pi \in \Pi$ and $K(\pi') \approx 0$ for all inactive parties $\pi' \in PS(N) \setminus \Pi$, making Π stable or (2) setting $K(\pi) \approx 0$ for all active parties $\pi \in \Pi$ and $K(\pi') > 0$ for all inactive parties $\pi' \in PS(N) \setminus \Pi$, making Π unstable. This is because a party with no capacity is equivalent to disintegration. And Lemma C.1 shows that under this condition, there is some agenda that makes the positive capacity parties (the active/inactive ones, respectively for the two cases) better for all members than the zero capacity parties (inactive/active respectively). Thus, causing stability/instability respectively, proving the result. $\square$

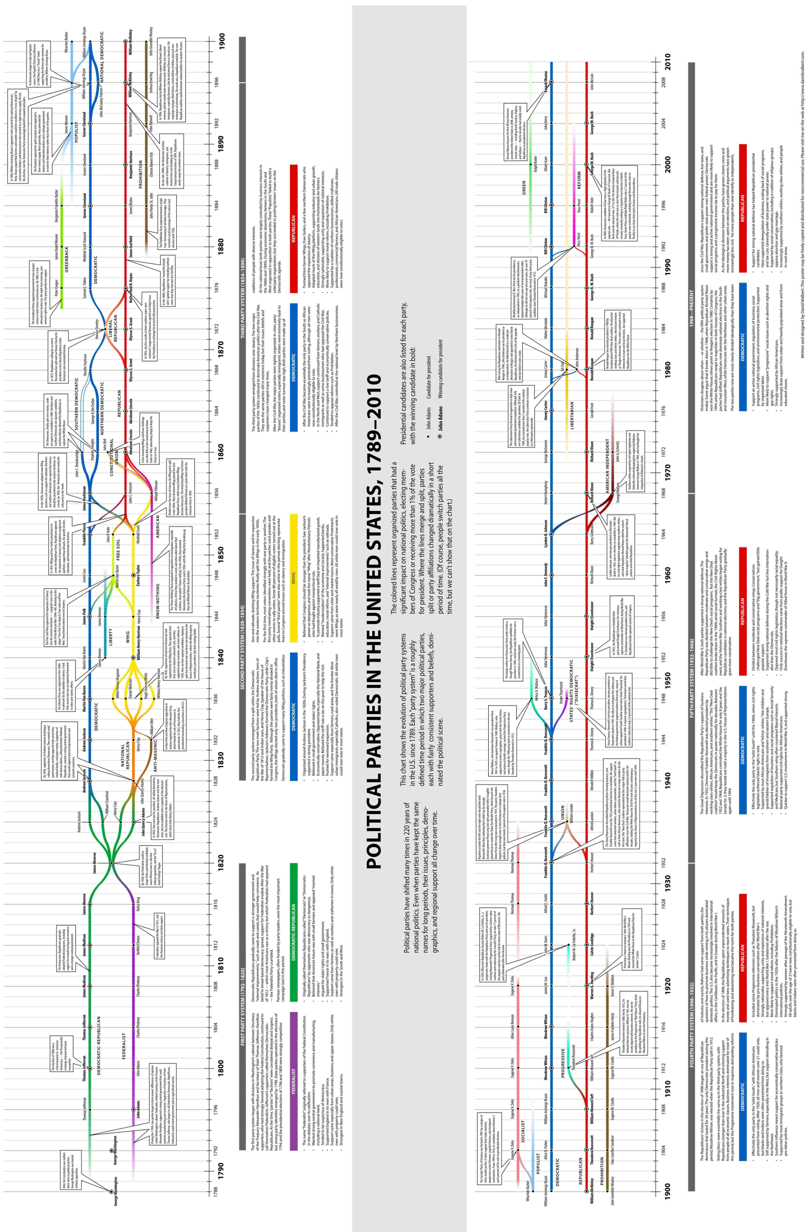

FIGURE 1. History of American Political Parties (Graphic made by David Walbert)