

UNBUNDLING POLARIZATION

NATHAN CANEN, CHAD KENDALL, AND FRANCESCO TREBBI

ABSTRACT. This paper investigates the determinants of political polarization, a phenomenon of increasing relevance in Western democracies. How much of polarization is driven by divergence in the ideologies of politicians? How much is instead the result of changes in the capacity of parties to control their members? We use detailed internal information on party discipline in the context of the U.S. Congress – whip count data for 1977-1986 – to identify and structurally estimate an economic model of legislative activity in which agenda selection, party discipline, and member votes are endogenous. The model delivers estimates of the ideological preferences of politicians, the extent of party control, and allows us to assess the effects of polarization through agenda setting (i.e. which alternatives to a status quo are strategically pursued). We find that parties account for approximately 40 percent of the political polarization in legislative voting over this time period, a critical inflection point in U.S. polarization. We also show that, absent party control, historically significant economic policies would have not passed or lost substantial support. Counterfactual exercises establish that party control is highly relevant for the probability of success of a given bill and that polarization in ideological preferences is more consequential for policy selection, resulting in different bills being pursued.

Date: August 28, 2019.

Canen: University of Houston, Department of Economics (ncanen@uh.edu).

Kendall: University of Southern California, Marshall School of Business (chadkend@marshall.usc.edu).

Trebbi: University of British Columbia, Vancouver School of Economics, Canadian Institute for Advanced Research and National Bureau of Economic Research (francesco.trebbi@ubc.ca).

We thank Matilde Bombardini, Josh Clinton, Gary Cox, Ernesto Dal Bó, Andrew Hall, Jeffery Jenkins, Keith Krehbiel, as well as seminar participants at various institutions for comments. We are grateful for funding from CIFAR and for hospitality at the Graduate School of Business at Stanford University during part of the writing of this paper.

1. INTRODUCTION

We focus on a set of open questions in the political economy literature on political polarization, a phenomenon that has taken a sharply increasing tack since the mid-1970s in the United States.¹ Other OECD countries have experienced similar trajectories recently, and deeply antagonistic political environments are commonplace across Western Europe today. To many observers, polarization has been linked to heightened policy uncertainty over government spending, regulation and taxes, with consequences for the pricing of financial assets and sovereign debt market volatility (Baker et al., 2014, 2016; Pastor and Veronesi, 2012; Kelly et al., 2016). Critically, this segmentation of legislatures across party lines may be the result of more than just exogenous shifts in the ideologies of elected representatives. The goal of this paper is to present a credibly identified method for unbundling polarization in votes into its constituent determinants: polarization in ideologies and party control. We also quantitatively analyze the differential effects of these underlying mechanisms on expected equilibrium policy outcomes in the U.S. Congress.

A first question is how much of political polarization in votes is the result of more ideologically polarized legislators and how much is due to party leaderships forcing rank-and-file members to toe the party line.² The question of whether or not the current political polarization in Congress can be solely attributed to changes in the ideological composition of the legislative chambers, for example due to the progressive replacement of moderate representatives with extreme ones³, remains unsettled (Theriault, 2008; Moskowitz et al., 2017).⁴ Political parties, through changes in institutional rules and in their system of internal leadership, may have contributed to polarization in votes across party lines by allowing parties to more effectively steer members in support of strategically set agendas.⁵

¹For discussions of political polarization in the electorate and U.S. Congress, see for instance Gentzkow (2016); McCarty et al. (2006).

²See Ban et al. (2016) for a discussion of whether political polarization is the result of better internal enforcement by party leaders.

³Following a large literature in ideal point estimation, we consider a politician's ideology to be fixed. In this context, polarization in ideologies can only be driven by replacement.

⁴To answer this question, one must first deal with the primitive problem of assessing the ideal points of politicians, a long-standing issue in the political economy and political science scholarship focused on the behavior of national legislatures (Levitt, 1996; Poole and Rosenthal, 2001; McCarty et al., 2006; Mian et al., 2010). Showing where politicians' preferences are located, absent any equilibrium disciplining by parties on floor votes, requires recovering the unbiased distribution of within-party individual ideologies, a problem subject to severe identification issues (Krehbiel, 2000; Snyder and Groseclose, 2000). Levitt (1996) specifically offers an early decomposition of Congressional voting focused on isolating individual ideology from other determinants of the voting decision.

⁵Seminal work from Cox and McCubbins (1993), Cox and McCubbins (2005) and Aldrich (1995) emphasizes the importance of parties for the functioning of Congress. It focuses on how parties use the available institutions to coordinate and set policies to their benefit, as well as how party leaders work towards their goals with their party members. Cox and McCubbins emphasize institutional mechanisms by which majority parties get their policies on the floor, blocking the minority's policies. They discuss incentives to do so, including the "brand" value of a party, increasing re-election chances for politicians, increasing the coordination of policies that politicians may be unsure of, setting policy positions, as well as helping to enforce and coordinate policies and votes. Evidence, such as in Forgette (2004), has shown that these

A second question is how polarization in the legislature affects the policies that are pursued. Polarization may affect not only the details of the bills proposed, but also which status quo policies are contested in the first place (and which are instead left unpursued). Policy alternatives, including tax cuts, healthcare reforms, trade policy or tariffs bills, are endogenous and presented strategically based upon the likelihood that a given proposal will pass. The different drivers of polarization may affect the policy alternatives chosen ex-ante by the agenda setter, who, based on how the equilibrium probability of bill passage varies, may respond differently to changes in the technology of party control relative to shifts in the ideological composition of fellow legislators.

The first contribution of this paper is to provide an economic model of legislative activity for a two-party system. The model is designed to capture strategic considerations on multiple nested dimensions. The first dimension is which issues (and for a given issue, which specific policy alternatives) are selected by proposing parties. Policies that are not sufficiently valuable vis-à-vis a specific status quo, or too difficult to pass given the extant chamber composition, may not be pursued at all. The second dimension is whether or not, once a certain alternative to a status quo is proposed, the leadership decides to invest in acquiring extra information about the prospects of that specific policy alternative (i.e. “*to whip count*” a bill). Policies that appear unpromising once more information is acquired may not be pursued further (i.e. not brought to the floor for an official vote). The 2017 repeal attempt of the Affordable Care Act is a salient example. A third dimension for consideration is, if a bill is eventually brought to the floor for a vote, which legislators can be disciplined (i.e. “*whipped*”) in order to maximize the likelihood of passage. As our economic model formalizes, member voting decisions (the observable output of the model) are ultimately endogenous to all of these previous phases of the process. Quantitative approaches based on sincere voting or abstracting from party control, as in the vast majority of the political economy literature, overlook these important dimensions.

The second contribution of the paper is empirically unbundling the multiple elements of this process. We identify and estimate our model structurally. We are able to resolve the identification problems previous researchers have faced thanks to the use of new data that supplements standard floor voting (“*roll call*”) information, thus decoupling true individual ideological positions (before any party control is exerted) from party discipline targeted towards members on the fence

mechanisms of policy positioning and agenda-setting are present, as measured by the attendance rates and transcripts from party caucuses, and affect legislative roll call voting. Aldrich (1995) and his Conditional Party Government theory proposes that parties play an important role in pushing policies of interest to the rank and file. Lawrence et al. (2006), instead, focus on implications of agenda selection models. Economists such as Caillaud and Tirole (2002) have also taken a similar stance to party organization, emphasizing internal control issues, but with a focus on electoral success.

of support for a bill.⁶ We make use of a complete corpus of whip count votes compiled from historical sources by Evans (2018) for the U.S. House of Representatives. Whip counts are private records of the voting intentions of party members, used by party leaders to assess the likelihood of success of specific bills under consideration before they are voted.⁷ Our sample period includes the 95th to 99th Congress (years 1977 to 1986). These Congresses occur at the inflection point of contemporary U.S. polarization dynamics (McCarty et al., 2006), allowing us to observe how ideological differences across parties and party discipline evolve over this critical time period. Section 2 presents background information on the whip system and institutional context useful for the framing of our model.

Members' responses at the whip count stage are useful for recovering the true ideological positions of politicians *before* party control is exerted. Our argument is three-fold. First, the information revelation value of whip counts resides in the repeated interaction between members and the leadership, limiting the ability of rank-and-file politicians to systematically lie or deceive their own party leaders. These interactions are frequent and the stakes are typically high. Second, by a revealed preference argument, the fact that costly whip counts are systematically employed by the party leadership to ascertain the floor prospects of crucial bills bears witness to their usefulness and informational value. It is unclear why leaders would spend valuable time on these counts otherwise. In the model, this information revelation is achieved in equilibrium as legislators are atomistic and cannot individually influence a party's aggregate information or decision. Third, as we model explicitly, certain designated party members (called *whips*), who are responsible for ensuring some subset of members toe the party line, maintain constant relationships with their delegation and know their districts. These relationships make private preferences at least partially observable, reducing the ability of members to misreport their ideological positions (Meinke, 2008).⁸

⁶The main difficulty lies in being able to compare outcomes with parties to those without. In a series of works, Keith Krehbiel (Krehbiel, 1993, 1999, 2000) has argued that the previous literature failed to address the confounding issues of whether parties are effective, or whether they are only a grouping of like-minded politicians. This identification problem comes from using outcomes such as roll call votes, party cohesion, or party unity scores that are a combination of politician preferences and of party effects. Moreover, politicians from the same party are likely to share similar ideologies, so could be voting in the same way regardless of party discipline. The paradox, as stated by Krehbiel (1999), is that parties appear strongest when members are most homogeneous ideologically (and hence, when parties are needed the least). That, in turn, leads to an empirically difficult problem: how does one separate individual ideology measurements from party effects? In particular, how does one estimate party effects when ideology measures confound both parties and individual ideologies?

⁷The data structure of whip counts has been explored occasionally in the past, as in the works of Ripley (1964) and Dodd (1979) for example, but with different objectives. In both papers, the data was collected when the authors worked within the Whip Offices (as American Political Science Association Congressional Fellows). Two works in particular have looked at whip counts in the context of parties and party discipline. Burden and Frisby (2004) look at 16 whip counts and their roll calls and find that most of the switching of votes has gone in the direction of party leaders. Evans and Grandy (2009) also use whip counts, and provide an extensive survey of whipping in the House of Representatives and the Senate, drawing attention to some historical examples.

⁸Scholarship discussed in the next section, particularly Evans (2018), support this view. Multiple assistant and regional whips are part of the party leadership hierarchy and are typically appointed or elected within a delegation. As further

In addition to providing information about politicians' true ideological positions, the whip count data offers identifying variation for assessing party discipline and agenda setting. Concerning party discipline, switching behavior in Yes/No between the whip count stage and the roll call stage provides the variation necessary to pin down the extent of whipping – how much control the party is able to exert. Concerning agenda setting, we exploit the fact that not all bills that are voted on the floor are whip counted, and that certain bills that are whip counted are subsequently dropped without a subsequent floor vote.⁹ By explicitly modeling this selection process, we theoretically identify thresholds determining which bills are voted on and/or whip counted. Together with flexible assumptions on the distribution of latent status quo policies, these thresholds allow us to recover information on policies that are never proposed and never voted.

This paper establishes several findings. Our results show that standard approaches to the estimation of ideal points based on random utility models that employ roll call votes alone, such as the popular DW-Nominate approach (Poole and Rosenthal, 2001), miss important density in the middle of the support of the ideological distribution. These methods, which conflate party control with the estimation of individual ideologies (Snyder and Groseclose, 2000), show a polarization level of ideal points much larger than the actual one based upon our unbiased estimates. Across the 95th-99th Congresses, we find that the distance between party medians is on average about 60% of that based upon standard DW-Nominate estimates. According to our estimates, the share of traditional DW-Nominate ideological polarization which actually stems from party discipline varies from 34 percent in the 95th Congress to 44 percent in the 99th Congress. Importantly, these results do not rely on arbitrary assumptions about which bills may be whipped or not by the party (we operate under the assumption that parties can discipline votes on any bill) or the omission of any floor votes from the analysis, including lopsided or unanimous votes.

In terms of agenda-setting, we show that for every 100 issues that the majority party (Democrats in our sample) could potentially deliberate within a congressional cycle, on average, 7 are never voted because they are not sufficiently valuable for the leadership; 86 are brought directly to the floor where they are whipped and voted; and 7 are whip counted. Of the 7 bills that are whip counted, 2 are subsequently dropped, while 5 are brought to the floor, where they are whipped and voted.

With our structural estimates in hand, we show that party discipline matters substantially and has proven crucial for the passage of important bills. Eliminating party discipline in the form

testimony of the value of whips' activities, the Majority and Minority Whips, who organize these counts, are ranked second or third in importance within the party hierarchy. See Section 2.

⁹For a recent important example, consider early 2017 efforts to repeal the Affordable Care Act by the Republican leadership in the House. These attempts were repeatedly whip counted, but not voted.

of whipping is precisely rejected relative to a model with party discipline using standard model selection tests. The extent of party discipline is statistically different from zero, quantitatively sizable, and growing between 1977 and 1986.

Given the specific time period over which the whip count data is available, we are also able to assess, through counterfactuals, the role of parties in steering particularly salient economic bills in the early 1980s, including the two Reagan Tax Reforms of 1981 and 1984, several Social Security Amendments and Debt Limit Increase Acts, the National Energy Act of 1977, and the implementation of the Panama Canal Treaty in 1979. Some of these bills would not have passed or would have substantially lost support absent party discipline. In counterfactual exercises that focus on agenda setting, we also establish that party control is highly relevant for the equilibrium probability of success of a given policy alternative against the status quo. Polarization in the ideological preferences of legislators is instead more consequential for setting the policy alternative for each status quo, resulting in substantially different bills being pursued.

This paper contributes to three strands of literature. First, it is concerned with the polarization of political elites. The empirical literature on political polarization has a rich history (Poole and Rosenthal, 1984), and has experienced a recent resurgence in interest due to glaring increases in partisanship in voting (McCarty, 2017, but also media reports¹⁰). Rising political polarization has been detected not only in legislator ideology assessments based on roll calls, but in candidate survey responses (Moskowitz et al., 2017), congressional speech scores (Gentzkow et al., 2017), and campaign contributions measures (Bonica, 2014). Considerations on polarization from an economic perspective, related to the seemingly increasing policy gridlock after the 2008 financial crisis, are offered in Mian et al. (2014). We contribute to this discussion from an empirical perspective by quantitatively unbundling some of the deep determinants of polarization. In this respect our work complements other recent attempts, such as Moskowitz et al. (2017), and the decomposition exercise of Levitt (1996) earlier on, but it differs in terms of theory, identification strategy, and in the use of a structural approach.

A second, closely related, literature considers the problem of separating politician's ideological preferences from party discipline. At the heart of the problem is the observation by Krehbiel (1999, 1993) that party unity in floor voting may not necessarily be conclusive evidence of discipline. This observation is, at its core, an identification critique. Politicians from the same party are likely to share a similar ideology, and hence may vote similarly even absent party control. Exemplifying

¹⁰See, for instance, Philip Bump, December 21, 2016, "Farewell to the most polarized Congress in more than 100 years!" *Washington Post*.

one of the most popular existing procedures used to estimate legislator ideology¹¹, [McCarty et al. \(2006\)](#) offers a broad discussion of this research area and links it to parallel relevant phenomena, such as the co-determined evolution of U.S. income inequality ([Piketty and Saez, 2003](#)).

Decomposition efforts in problems of political agency are rooted in an older literature that seeks ways to separate a politician's true policy preferences from that of the party, by focusing on situations in which one or the other factor would not be present. [Snyder and Groseclose \(2000\)](#) propose one such method of separating party effects from politician ideology, which has been widely used and adapted (e.g. [McCarty et al., 2001](#); [Minozzi and Volden, 2013](#)). Their argument is that parties concentrate their efforts on results that they can influence, such as close legislative votes. Seemingly, expected lopsided votes would not attract nor need party intervention. Absent party effects on lopsided votes, [Snyder and Groseclose \(2000\)](#) argue in favor of estimating individual ideologies from a first stage on lopsided roll calls alone. After recovering estimates of individual preferences, in a second stage they study close votes to recover party effects, given the previously estimated legislator true preferences. There are two main methodological obstacles to this approach. First, which vote is lopsided and which is contested is endogenous to the choice of policy alternative by the agenda setter (see the discussion in [Bateman et al., 2017](#)). This selection mechanism is explicit in our framework. Secondly, [McCarty et al. \(2001\)](#) note that this method provides poor identifying variation due to minimal differences in vote choices within a party for lopsided votes. In contrast, our paper does not rely on an arbitrary selection of votes where parties are assumed to be inactive.¹² [Ansolabehere et al. \(2001a\)](#) use a survey directly targeted at candidate ideology (NPAT, also used in [Moskowitz et al., 2017](#)) to estimate ideal points, hence moving away from roll calls. Also, [Ansolabehere et al. \(2001b\)](#) find a decline in the responsiveness of Congress members to constituents in 1970s and 1980s, consistent with our findings.¹³ Previous work has also discussed

¹¹Among the standard approaches to estimation are [Poole and Rosenthal \(1997\)](#); [Clinton et al. \(2004\)](#); [Heckman and Snyder \(1997\)](#).

¹²In Appendix D, we explore a specification inspired by [Snyder and Groseclose \(2000\)](#) in which parties do not discipline lopsided votes in order to allow for different strategies in party discipline. However, in contrast to their work, in our robustness check all bills are informative about ideologies, guaranteeing stronger identification. While party discipline is estimated to be about 7 percentage points smaller when parties do not whip lopsided bills, our main quantitative messages continue to hold in this case.

¹³Other closely related papers such as [Clinton et al. \(2004\)](#), who use Bayesian methods to estimate ideal points, also employ lopsided bills to recover party discipline. Another approach looks at politicians who change party to see how their voting behavior changes. As [Nokken \(2000\)](#) finds, congressmembers who switch party do change voting patterns, suggesting that ideology is not their sole decision factor. Our model microfound this change in behavior. An interesting historical approach is presented by [Jenkins \(2000\)](#). By studying congressmembers who initially served in the U.S. House and then served in the Confederate House during the American Civil War, he finds striking differences in the estimated ideologies for the same politician from voting behavior in the different Houses. Since the legislators were the same, and in very similar institutional settings, he concludes (with further evidence) that differences were due to agenda setting and party discipline rather than mere ideology. Finally, [Lee \(2009\)](#) discusses how party competition may be a key driver of polarization in Congress.

how polarization and agenda setting may interact (Clinton et al., 2014; Bateman et al., 2017), a point that our model clarifies.

Evans (2018) is a closely related, but less formal foray into the study of internal party organization as a tool for identifying party discipline. The central contribution of his volume remains its compendious treatment of the whip system, historical and contemporary.

A final literature to which we contribute deals with the consequences of polarization for the behavior of legislatures. Mian et al. (2014) offers a discussion of the effects of political polarization on government gridlock and lack of reform. They also discuss how gridlock may be particularly damaging in the contexts of the aftermath of deep economic crises, where political stalemate may trigger secondary adverse events (e.g. sovereign debt crises following banking crises). The relationship between slowdowns in legislative productivity and polarization is also a topic frequently discussed in political science (e.g. Binder 2003 and references therein). None of these works, however, offers a theory for the analysis of the role of polarization in the context of strategic party control efforts and endogenous agenda setting decisions.

2. INSTITUTIONAL BACKGROUND: PARTY LEADERSHIP AND WHIPS

This Section provides a brief primer on the leadership structure and whip organization of the Democratic and Republican Parties in the United States. It does not attempt an exhaustive review, but rather a sufficiently accurate synthesis of the main institutional features necessary to guide the reader in the following sections, which rely on such features. Evans (2018) offers a comprehensive discussion of the whip system.¹⁴

The internal organization of modern party apparatuses requires both the transmission of information within the hierarchy and the allocation of both rewards and punishments across rank-and-file members. Historically, British and American legislative bodies developed the whip system¹⁵ to serve such purposes. Although different in terms of their form of government, Westminster systems and the U.S. Congress are characterized by parties where the role of “Whip” is a recognized tier of their formal leadership structure.¹⁶ The United States Congress glossary defines whips as “*Assistants to the floor leaders who are also elected by their party conferences. The majority and minority whips*

¹⁴The internal organization of parties in the modern U.S. Congress is the subject of a large literature, see for instance Cox and McCubbins (1993, 2005); Poole and Rosenthal (2011).

¹⁵The term originates from the “whipper-in” who keeps the pack of hounds tight during the hunting of foxes on behalf of the huntsman.

¹⁶For instance in the contemporary Congress, Majority or Minority Whip are third and second respectively in the official party ranking. In the United Kingdom the chief Whip not only officiates in the legislative chamber, but is customarily appointed to a cabinet position and participates to the executive. The official role is Parliamentary Secretary to the Treasury, a junior ministerial position in the British Government, with only nominal association to the Treasury.

*(and their assistants) are responsible for mobilizing votes within their parties on major issues. In the absence of a party floor leader, the whip often serves as acting floor leader.”*¹⁷

The Majority or Minority Whip is aided by a set of assistant, deputy or regional whips, either appointed or elected. Table F.1 in Appendix reports their number by Congress and party in our sample. The number of Democratic whips trends upward from 35 in the 95th Congress to 64 in the 99th, and, for Republicans, from 16 in the 95th Congress to 25 in the 99th. Assistant whips play the role of the eyes and ears of the leadership across congressional delegations, and steer members in the direction of the party leadership. The latter activity may involve the provision of incentives, which may take the form of valuable committee appointments, floor time, or leadership political action committee campaign funds.¹⁸ Assistant whips may also play a role in communicating more forcefully the importance of certain issues to selected members, affecting their stance on a vote. We refer to the activity of selectively providing incentives to toe the party line as “whipping” (Meinke, 2008).

The chief Whip, in conjunction with the party leadership, also conducts straw polls to elicit the extent of support among the rank-and-file on certain bills.¹⁹ Such head counts are costly and strategically employed in about 6 percent of all bills in our sample. Typically, support for the party position on a legislative issues is elicited, requiring an indication of yes, no, undecided or other. In practice any position that does not provide firm support of the leadership’s stance can be interpreted as not supportive.²⁰ This straw poll activity is what we refer to as “whip counting”.

The issue of truthfulness of the information elicited at the whip count stage is worthy of attention. Evans (2011) notes: “One common question about whip counts is whether the responses of members can be trusted. Are there any incentives for them to overstate their opposition to the party program, potentially securing favors in exchange for their support? Four points are worth mentioning in response. First, the whip process is a “repeated game” and members develop reputations. There

¹⁷https://www.senate.gov/reference/glossary_term/whips.htm

¹⁸Parties also sanction their whipped members. The House Democratic leadership did not allow Representative Phil Gramm (D-TX) to retain his seat on the Budget Committee after he was unresponsive to whip pressure against President Reagan’s economic program in 1983. See Baker (1985).

¹⁹In the words of current Senate Majority Leader Mitch McConnell, a former Majority Whip, such counts play a crucial role: “Producing an accurate vote count is the most important function of the Whip. Accordingly, the question posed to members on the whip card must be phrased with precision, so as not to distort the accuracy of the tabulation. Moreover, the question has to be presented in as fair and accurate a way as possible; otherwise leadership could wind up repelling wavering senators by seeming to be heavy-handed. After the whip cards are distributed, the whip collects the data, and based on that information, the party leader determines how to proceed on a matter. For me, as party leader, it is crucial that the whip count be accurate. If it is not, our leadership team might be embarrassed, and precious floor time could be wasted on a failed measure.” McConnell and Brownell II (2019) (p.190). For a discussion on the Democratic side, specifically about Tip O’Neill’s role as a leader and his efforts in improving whip count accuracy see Meinke (2016), p.90.

²⁰ “Member decisions not to respond with one of the standard [yes/no] categories were far from random, in other words, and there are obvious signs of strategic behavior. Often, members who were disinclined to support the party simply refused to respond to leadership entreaties about their views, or otherwise were unwilling to take a clear position.” Evans (2018), p.112

are incentives for them to be truthful. Second, congressional leaders generally know a lot about the constituencies of rank-and-file members and can be very difficult to fool. Third, in a sense it does not matter. If a member claims that she will oppose a bill or amendment unless she receives some concession, then that essentially becomes her position and the polled question and the concession are for all practical purposes inseparable. Fourth, and most important, participants in the whip process believe that whip poll responses are accurate, which is precisely why they base strategic decisions on the results.” (p.13).

This perspective is not merely academic²¹, but finds support in statements offered by practitioners.²² It is common to find evidence of the importance and reliance on the internal whip system by party leaders. For example, current House Speaker Nancy Pelosi, also a former Minority Whip, is known to have often asserted reliance on accurate internal counts and focused effort on them.²³ For the remainder of this article, it is important to assert that we will not be assuming that whip counts are perfectly truthful about the position of each party member on each vote. Based on the discussion above, we will assume, however, that whip count responses are truthful on average. This is tantamount to ruling out systematic deception and gaming of the leadership by rank-and-file members.

3. MODEL

With the institutional context of Section 2 in mind, we present a model with two main features: (i) party discipline, and (ii) agenda-setting. Two parties compete for votes on a series of issues that make up a congressional term. Each party employs a subset of their legislators (the whips) to discipline their members (including other whips). For a given status quo policy, a randomly-selected proposing party chooses the alternative policy (if any) to be voted upon, accounting for both parties’ abilities to discipline their members, and on the value and likelihood of passage of the alternative policy. Because floor votes are costly, not all status quo policies will be pursued. If an alternative is pursued, the proposing party can employ a formal whip count, which allows it to obtain additional information about a bill’s probability of success before a floor vote, and to drop bills that are unlikely to pass conditional on the count.²⁴ Whether the proposing party chooses to

²¹It dates back at least to Ripley (1964).

²²“[About lying to leadership] a Republican aide surmised: ‘some of that happens’ but ‘it doesn’t happen as much as people think it would.’” Green and Harris (2019), p.19.

“[...] members of Congress have strong reasons to refrain from lying about their vote choice. [...] there is strong norm in Congress about violating one’s commitments and deceiving fellow representatives (Barber (1965); Weingast (1979))”, Green and Harris (2019), p.19.

²³See Kathryn L. Pearson “Nancy Pelosi victorious – why the California Democrat was reelected speaker of the House”, The Conversation, January 3, 2019.

²⁴The party not setting the agenda may also conduct a whip count, but this occurs less frequently in our data so we do not model its reason for doing so.

conduct a formal whip count depends upon its option value relative to the fixed cost of undertaking this process.

3.1. Preliminaries. Party members vote on a series of policies at times $t = 1, 2, \dots, T$ with the majority vote determining the winning policy. Each party, $p \in \{D, R\}$, has a mass of N_p members whose underlying ideologies, θ , are continuously distributed with cumulative distribution functions (CDFs), $F_p(\theta)$, in a single-dimensional space. We assume that the corresponding probability distribution functions (PDFs), $f_p(\theta)$, have unbounded support. The median member(s) of a party are identified by θ_p^m and represent the preference of the party overall. We assume without loss that $\theta_D^m < \theta_R^m$.

In each period, party D is randomly recognized with probability γ , allowing it to set the policy alternative, x_t , to be put to a vote. With the remaining probability, $1 - \gamma$, party R is recognized. The recognized party draws a status quo policy, q_t , from a continuous CDF, $W(q)$, with corresponding PDF, $w(q)$, which is also assumed to have unbounded support.²⁵

3.2. Preferences. There are three sets of actors for each party: non-whip members, whip members, and the party itself.

Whips are a ‘technology’ that a party uses to discipline its members. We take the mass and ideologies of whips as given and assume an exogenous matching of whips to members for which they are responsible, such that each member is controlled by exactly one whip. Whips acquire information from members and are rewarded for obtaining votes that the party desires.

All party members (whips and non-whips) derive expressive utility from the policy, $k_t \in \{q_t, x_t\}$, that they vote for. This utility is given by $u(k_t, \omega_t^i)$, where $\omega_t^i = \theta^i + \delta_{1,t}^i + \delta_{2,t}^i + \eta_{1,t} + \eta_{2,t}$ determines their position on a particular bill. We assume a symmetric, strictly concave utility function: $u(k_t, \omega_t^i) = u(|k_t - \omega_t^i|)$ with $u(\omega_t^i, \omega_t^i) = u_k(\omega_t^i, \omega_t^i) = 0$, $u_{kk}(k_t, \omega_t^i) < 0$.

θ^i is a member’s fundamental ideology, a constant trait of i .²⁶ A member’s position on a particular bill is determined by this ideology, two idiosyncratic shocks, $\delta_{1,t}^i$ and $\delta_{2,t}^i$, and two aggregate shocks, $\eta_{1,t}$ and $\eta_{2,t}$. Multiple shocks are required to model the information acquisition problem of the proposing party, as will become clear below. The aggregate shocks are common across all members of both parties and are independent draws from a Normal distribution with mean zero and standard deviation, σ_η . The idiosyncratic shocks $\delta_{1,t}^i$ and $\delta_{2,t}^i$ are identically and independently

²⁵In our application, D is the majority party. We do not model how the frequency of recognition is determined by the leadership of both parties.

²⁶In this regard, we follow the discussion and evidence from Lee et al. (2004) and Moskowitz et al. (2017). As a result of constant ideology, polarization due to changing ideologies within a chamber can only arise from the replacement of moderate politicians over time. The issue of replacement has been discussed in great detail in the recent debate on polarization: see Moskowitz et al. (2017).

distributed across i and t according to the continuous, unbounded, and mean zero CDF, $G(\delta)$ with corresponding PDF, $g(\delta)$.

Whip members, in addition to their utility from voting, receive a payment of r_p (which may differ across parties) for each member i for whom the whip is responsible and that votes with the party. r_p may represent, for example, improved future career opportunities within the party hierarchy.²⁷ We model whip influence over the members for whom she is responsible as an ability to persuade a member to change his position on a particular bill. To influence a member's position by an amount, y_t^i (i.e. to move his ideal point to $\omega_t^i + y_t^i$), a whip bears an increasing cost, $c(y_t^i)$ ($c' > 0$), which can be thought of, most simply, as an effort cost.²⁸ We assume $c(0) < r_p$ so that a whip optimally exerts a non-zero amount of influence. The contribution to a whip's utility from whipping is therefore given by $\sum_i (r_p I(i \text{ votes with party}) - c(y_t^i))$, where $I(\cdot)$ is the indicator function and the summation is over all members for whom she is responsible. Whips are allowed to whip any bill, independently of the party promoting the bill, this way also capturing active obstruction by the minority of majority-proposed bills. Whips are not allowed, however, to cross-discipline or entice members of the other party. Such behavior is not completely infrequent, but dominated by whipping within one's own party, so we omit it.

Each party derives utility from that of its median member, $u(k_t, \theta_p^m)$ where $k_t \in \{q_t, x_t\}$ is the winning policy. For simplicity, we assume that the party's position, represented by their median member, is not subject to idiosyncratic or aggregate shocks.²⁹ Because the party does not directly bear the cost of whipping its members, whipping is costless to the party (and thus both parties' whips are engaged on every vote).

3.3. Information and Timing. At each time t (see Figure 1):

- (1) The proposing party is randomly recognized and a status quo policy, q_t , is drawn.
- (2) Whip count stage:
 - (a) The proposing party chooses the policy x_t as an alternative to the status quo q_t and decides whether or not to conduct a whip count at a cost, $C_w > 0$.³⁰

²⁷Rewarding the whip only if she switches a member's vote does not change the results.

²⁸Having the shocks and influence operate on the ideological bliss point rather than as changes in utility (i.e. $u(k_t, \theta^i) + \delta_{1,t}^i + \delta_{2,t}^i + \eta_{1,t} + \eta_{2,t} + y_t^i$) simplifies the model in two ways. First, it ensures that the maximum influence exerted by a whip (see Section 4.2) is a constant, independent of the locations of the policies and the distance between them. Second, it ensures the expected number of votes monotonically decreases in the extremeness of the alternative policy, x_t (see the proof of Proposition 1), which need not be the case for utility shocks.

²⁹This assumption rules out the possibility that an aggregate shock causes the proposing party to prefer the status quo over the alternative they themselves proposed.

³⁰We assume a closed agenda setting rule: x_t cannot be modified after observing the outcome of the whip count. Empirically, minor changes are captured by the aggregate shocks, η . Changes that target individual legislators, such as certain earmarks, can be captured in our set-up by the transfers, y_t^i . When changes to x_t become truly substantial, the issue typically translates into a new vote, which we then examine as a distinct t . Substantial changes to the alternative

- (b) The first aggregate and idiosyncratic shocks, $\eta_{1,t}$ and $\delta_{1,t}^i$, are realized and observed noisily: each member observes his idiosyncratic shock, $\delta_{1,t}^i$, and the policy he prefers, $u(x_t, \theta^i + \delta_{1,t}^i + \eta_{1,t}) \leq u(q_t, \theta^i + \delta_{1,t}^i + \eta_{1,t})$, but not the realization of $\eta_{1,t}$.³¹
 - (c) If a whip count is undertaken, each member makes a report, $m_t^i \in \{Yes, No\}$, to his whip, answering the question of whether or not they intend to support the alternative policy, x_t . The outcome of the whip count is common knowledge.
 - (d) The proposing party (conditional on the whip count, if taken) decides whether or not to proceed with the bill, taking it to a roll call vote at a cost, $C_b > 0$.
- (3) Roll call stage:
- (a) The second aggregate and idiosyncratic utility shocks, $\eta_{2,t}$ and $\delta_{2,t}^i$, are realized and observed as in the case of the first shocks: each member observes his idiosyncratic shock, $\delta_{2,t}^i$, and the policy he prefers $u(x_t, \omega_t^i) \geq u(q_t, \omega_t^i)$, but not the realization of $\eta_{2,t}$.
 - (b) Similar to a whip count, whips communicate with their members to learn the sum of the aggregate shocks, $\eta_{1,t} + \eta_{2,t}$. They then communicate this sum back to each member.³²
 - (c) Whips learn the sum of the idiosyncratic shocks, $\delta_{1,t}^i + \delta_{2,t}^i$ of the members for whom they are responsible and choose the amount of influence to exert, y_t^i , over each member.
 - (d) The roll call vote occurs.

The information structure (who knows what and when) is a formalization of the role that whips play in obtaining and aggregating information by keeping close relationships with the rank-and-file members for which they are responsible. Information about individual member positions is important for determining (i) which members will be most easily persuaded to toe the party line, and (ii) the aggregate position on a bill, which is important for determining the likelihood that a particular bill is going to pass the roll call.

policy are treated as different votes by the House of Representatives itself, which registers different roll call votes not just for final passage votes, but also for amendments, motions, etc.

³¹The particular information structure is not critical, but it is important that no single member observes the aggregate shock perfectly because, were it the case, a whip count would not be necessary. In the absence of aggregate shocks, due to the fact that we have a continuum of members, the outcome of a vote would be known ex ante with probability one.

³²We assume whips communicate truthfully without modeling it explicitly. In reality, party leadership also communicates with members so that if the whip count were not reported truthfully, the whip would likely suffer severe consequences (i.e. lose their position).

4. ANALYSIS

We solve the model via backward induction. In Sections 4.1 and 4.2, we determine the decisions of members and whips. These decisions are the same for each party, so we drop the party label for convenience. In Sections 4.3 through 4.5, we turn to the decisions unique to the proposing party: which alternative policy to pursue, if any, and whether or not to conduct a whip count and a floor vote.

4.1. Roll Call Votes. Prior to the roll call vote, whips communicate with the members for whom they are responsible in order to learn the value of $\eta_{1,t} + \eta_{2,t}$, which is necessary for deciding how much influence to exert (see Section 4.2). To do so, each whip asks each member whether or not they intend to vote for the alternative policy, x_t . Integrating across politicians, this process reveals the aggregate shocks as in the case of a whip count (see Section 4.3). Whips then communicate the values of the aggregate shocks to all members, so that they have full information at the time of their vote.

A member votes for x_t if and only if $u(x_t, \omega_t^i + y_t^i) \geq u(q_t, \omega_t^i + y_t^i)$, where $\omega_t^i + y_t^i$ is the member's ideological bliss point after whip influence.³³ It is convenient to define the marginal voter as the ideological position of the voter who is indifferent between the two policies. Given symmetric utility functions, this voter is located at $MV_t = \frac{x_t + q_t}{2}$, absent party discipline and aggregate shocks. At roll call time, after both aggregate shocks, we define the *realized* marginal voter, $\tilde{MV}_{2,t} \equiv MV_t - \eta_{1,t} - \eta_{2,t}$ (similarly, we define the realized marginal voter at whip count time, $\tilde{MV}_{1,t} \equiv MV_t - \eta_{1,t}$).

4.2. Whipping Decisions. Just prior to roll call, each whip has full information about the ideological position of his members. He therefore knows whether or not a given (conditional) transfer induces a vote for a party's preferred policy or not, and so either exerts the minimal influence necessary to make the member indifferent between policies, or exerts no influence at all. The maximum influence he is willing to exert, y_p^{max} , is such that the cost of exerting this influence is equal to its benefit, $r_p = c(y_p^{max})$. y_p^{max} is strictly greater than zero because we assume that the cost of exerting no influence is less than the reward of successfully whipping a member ($c(0) < r_p$).

Given y_p^{max} , Lemma 1 establishes that only members who would not otherwise vote for the party's preferred policy, and are within a fixed distance of the marginal voter are whipped (see Figure 2 for an illustration).

³³Ties have measure zero due to the continuous nature of the shocks and therefore the vote tie-breaking rule is immaterial.

Lemma 1: *Assume a party strictly prefers policy k_t over policy k'_t . Then, only members, i , whose realized ideologies are on the opposite side of MV_t from k_t and such that $|\omega_t^i - MV_t| \leq y_p^{max}$ are whipped.*

4.3. The Whip Count. If a whip count is conducted, whips receive reports, $m_t^i \in \{Yes, No\}$, from each member for whom they are responsible and subsequently make these reports public. If each member reports truthfully, he reports $m_t^i = Yes$ if $u(x_t, \theta^i + \delta_{1,t}^i + \eta_{1,t}) \geq u(q_t, \theta^i + \delta_{1,t}^i + \eta_{1,t})$ and $m_t^i = No$ otherwise. Given the continuum of reports, $\{m_t^i\}$, and knowing θ^i , q_t , and x_t , a whip is able to infer $\hat{\eta}_{1,t}$, the realized value of $\eta_{1,t}$, with probability one by application of the law of large numbers.

All members reporting truthfully forms part of an equilibrium strategy of the overall game because no single member can influence beliefs about $\hat{\eta}_{1,t}$, and hence cannot influence the eventual policy outcome by misreporting.³⁴ We therefore assume in what follows that members play a truth-telling strategy.³⁵

We formalize these claims in Lemma 2.

Lemma 2: *Truth-telling at the whip count stage forms part of an equilibrium strategy. Under truth-telling, the realization of the first aggregate shock, $\hat{\eta}_{1,t}$, is known with probability one.*

4.4. Optimal Policy Choices. After observing q_t , the proposing party can choose to do one of three things. One, it can decide not to pursue any alternative policy. Two, it can choose an alternative policy to pursue, x_t , without conducting a whip count. In this case, the party pays the cost, C_b , of pursuing the bill to the roll call stage. Three, the party can choose an alternative policy to pursue and conduct a whip count at a cost, C_w . In this case, after observing the results of the whip count, the party can decide whether or not to continue with the bill at a cost of C_b . Choosing to undertake the whip count is analogous to purchasing an option: the option to save the cost of pursuing the bill should the initial aggregate shock $\eta_{1,t}$ turn out unfavorably.

For status quo policies to the left of the proposing party's ideal point, θ_p^m , the alternative policy pursued (if any) must lie to the right of the status quo: any policy to the left of q_t is less preferred than q_t and q_t can be obtained at no cost. Similarly, for status quo policies to the right of θ_p^m , the proposed alternative policy must lie to the left of the status quo. In choosing how far from the status quo to set the alternative policy, the proposing party faces an intuitive trade-off: policies

³⁴In addition, misreporting does not change the amount of influence a member's whip exerts because the whip learns the member's true position before exerting influence.

³⁵As usual, there also exists an equilibrium of the whip count subgame in which each member babbles, so that nothing is learned about $\hat{\eta}_{1,t}$. This equilibrium is not empirically plausible because in this case no costly whip count would ever be conducted.

closer to its ideal point are more valuable, should they be successfully voted in, but are less likely to obtain the necessary votes to pass.

To formalize this intuition, define the number of votes that x_t obtains as $Y(\tilde{M}V_{2,t})$, which is stochastic only because of the random aggregate shocks – the idiosyncratic shocks average out because of a continuum of members. Using these definitions, the proof of Lemma 3 shows that more preferred policies obtain less votes on average.

Lemma 3: *The number of votes that the alternative policy, x_t , obtains, $Y(\tilde{M}V_{2,t})$, strictly decreases with the closeness between x_t and the proposing party's ideal point.*

The result of Lemma 3 guarantees that the alternative policy proposed must lie between the party's ideal point and the status quo policy. An alternative policy on the opposite side of the ideal point from the status quo is dominated by $x_t = \theta_p^m$, which is both more preferred and obtains more votes in expectation.

For the remainder of the analysis we present the case in which party D is the proposer – the case of party R is symmetric. Given the whipping technologies available to each party (defined by the maximum influence their whips are willing to exert, y_R^{max} and y_D^{max}), we can define the position of the marginal voter when the alternative policy is such that it obtains exactly half of the votes. Denote this position, $\hat{M}V_{i,j}$, where the subscripts $i, j \in \{L, R\}$ indicate the directions of the policy that parties D and R whip for, respectively.³⁶ Each $\hat{M}V_{i,j}$ is then given by $Y(\hat{M}V_{i,j}) = \frac{N_R + N_D}{2}$.

In the absence of a whip count, if party D pursues an alternative policy, the alternative policy x_t must maximize

$$EU_D^{no\ count}(q_t, x_t) = Pr(x_t\ wins)u(x_t, \theta_D^m) + Pr(x_t\ loses)u(q_t, \theta_D^m) - C_b$$

where the cost of proceeding with the bill, C_b , is paid with certainty.

For status quo policies to the left of θ_D^m , since $x_t \in (q_t, \theta_D^m]$, both parties prefer and whip for x_t , the rightmost policy. Because $Y(\tilde{M}V_{2,t})$ is monotonically decreasing in x_t , and therefore in $\tilde{M}V_{2,t}$, x_t wins if and only if $\tilde{M}V_{2,t} < \hat{M}V_{R,R}$ so that $Pr(x_t\ wins) = Pr(\tilde{M}V_{2,t} < \hat{M}V_{R,R})$. The sum of the aggregate shocks, $\eta_{1,t} + \eta_{2,t}$, is normally distributed with a variance of $\sigma^2 = 2\sigma_\eta^2$, so that we can write $Pr(x_t\ wins|x_t > q_t) = 1 - \Phi\left(\frac{\tilde{M}V_{2,t} - \hat{M}V_{R,R}}{\sigma}\right)$, where Φ denotes the CDF of the standard Normal distribution.

For status quo policies to the right of θ_D^m , we have $x_t \in [\theta_D^m, q_t)$. Party D therefore whips for the leftmost policy, x_t , but party R may whip for either policy depending on where q_t and x_t lie

³⁶Each $\hat{M}V_{i,j}$ is a function of many parameters of the model, so we suppress their dependencies for convenience. Note, however, that each is independent of q_t and x_t .

with respect to θ_R^m . As a simplification, we assume party R always whips for q_t in this case.³⁷ Under this assumption, x_t wins if and only if $\tilde{M}V_{2,t} > M\hat{V}_{L,R}$, so that $Pr(x_t \text{ wins} | x_t < q_t) = \Phi\left(\frac{\tilde{M}V_{2,t} - M\hat{V}_{L,R}}{\sigma}\right)$. Figure 3 illustrates this case, showing how moving the alternative policy closer to party D 's ideal point lowers the probability that it passes.

Conducting a whip count provides the option value of dropping the bill and avoiding the cost, C_b , if the first aggregate shock makes it unlikely the bill will pass. After conducting the whip count, party D continues to pursue the bill if and only if

$$Pr(x_t \text{ wins} | \eta_{1,t} = \hat{\eta}_{1,t}) (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) + u(q_t, \theta_D^m) - C_b \geq u(q_t, \theta_D^m)$$

where $\hat{\eta}_{1,t}$ is the realized value of $\eta_{1,t}$ and $u(q_t, \theta_D^m)$ is the party's utility from the outside option of dropping the bill. $Pr(x_t \text{ wins} | \eta_{1,t} = \hat{\eta}_{1,t})$ is easily shown to be strictly monotonic in $\hat{\eta}_{1,t}$, so that we can define cutoff values of $\eta_{1,t}$, $\underline{\eta}_{1,t}$ and $\bar{\eta}_{1,t}$, such that party D continues to pursue the bill if and only if $\eta_{1,t} > \underline{\eta}_{1,t}$ (for status quo policies to the left of θ_D^m) or $\eta_{1,t} < \bar{\eta}_{1,t}$ (for status quo policies to the right).

Given these continuation policies, prior to the whip count, party D chooses x_t to maximize

$$\begin{aligned} EU_D^{count}(q_t, x_t) = & Pr(\eta_{1,t} > \underline{\eta}_{1,t}) \left[Pr(x_t \text{ wins} | \eta_{1,t} > \underline{\eta}_{1,t}) (u(x_t, \theta_D^m) - C_b) \right. \\ & \left. + (1 - Pr(x_t \text{ wins} | \eta_{1,t} > \underline{\eta}_{1,t})) (u(q_t, \theta_D^m) - C_b) \right] + Pr(\eta_{1,t} < \underline{\eta}_{1,t}) u(q_t) \end{aligned}$$

for status quo policies to the left of θ_D^m and

$$\begin{aligned} EU_D^{count}(q_t, x_t) = & Pr(\eta_{1,t} < \bar{\eta}_{1,t}) \left[Pr(x_t \text{ wins} | \eta_{1,t} < \bar{\eta}_{1,t}) (u(x_t, \theta_D^m) - C_b) \right. \\ & \left. + (1 - Pr(x_t \text{ wins} | \eta_{1,t} < \bar{\eta}_{1,t})) (u(q_t, \theta_D^m) - C_b) \right] + Pr(\eta_{1,t} > \bar{\eta}_{1,t}) u(q_t) \end{aligned}$$

for status quo policies to the right of θ_D^m .

We define x_t^{count} and $x_t^{no \text{ count}}$ to be the optimal alternative policies pursued (if any alternative is pursued) when a whip count is conducted and when it is not, respectively. Proposition 1 shows that, provided that the cost of pursuing a bill, C_b , is not too large, these optimal policies are unique and bounded away from the party's ideal point. Furthermore, alternative policies pursued with whip counts are closer to the party's ideal policy. Intuitively, the fact that a whip count allows the

³⁷Similarly, if party R proposes an alternative to a status quo policy, $q_t < \theta_R^m$, we assume party D always whips for the status quo. We can solve the model without these assumptions, and the results are qualitatively similar.

party to drop bills that are unlikely to pass after observing the first aggregate shock allows it to pursue policies that are more difficult to pass.

Proposition 1: *There exists a strictly positive cutoff cost of pursuing a bill, $\hat{C}_b > 0$, such that for all $C_b < \hat{C}_b$, the optimal alternative policies, x_t^{count} and $x_t^{no\ count}$, are unique and contained in (q_t, θ_D^m) for $q_t < \theta_D^m$, contained in (θ_D^m, q_t) for $q_t > \theta_D^m$, and equal to θ_D^m for $q_t = \theta_D^m$.*

The requirement in Proposition 1 that C_b be sufficiently small is for analytical purposes only. Numerically, we have been unable to find a counterexample for which the conclusion of the proposition does not hold.

4.5. The Whip Count and Bill Pursuit Decisions. To complete the analysis, we determine for which status quo policies alternative policies are pursued and, when they are pursued, whether or not a whip count is conducted. Define the value functions, $V_D^{count}(q_t) = EU_D^{count}(q_t, x_t^{count}) - u(q_t, \theta_D^m)$ and $V_D^{no\ count}(q_t) = EU_D^{no\ count}(q_t, x_t^{no\ count}) - u(q_t, \theta_D^m)$, as the gains from pursuing an alternative policy with and without conducting a whip count, respectively (note that these definitions account for the cost of pursuing a bill, C_b , but ignore the cost of the whip count, C_w). Lemma 4 characterizes the value functions as a function of the status quo policy.

Lemma 4: *Fix $C_b < \hat{C}_b$ such that the optimal alternative policies, x_t^{count} and $x_t^{no\ count}$, are unique. Then, for all $q_t \neq \theta_D^m$, the value of pursuing an alternative policy with a whip count, $V_D^{count}(q_t)$, strictly exceeds that without, $V_D^{no\ count}(q_t)$. Furthermore, both value functions strictly decrease with $|q_t - \theta_D^m|$, but the difference between them, $V_D^{count}(q_t) - V_D^{no\ count}(q_t)$ strictly increases.*

Intuitively, both value functions decrease as the status quo approaches the proposing party's ideal point because there is less to gain from an alternative policy. More interestingly, the difference between the value functions increases as the status quo approaches the party's ideal point because the whip count is an option that allows the proposing party to initially pursue a bill, but drop it if the initial aggregate shock turns out to be unfavorable (thus avoiding the cost, C_b). This option value is always positive because the party could always ignore the result of the whip count. It increases as the status quo nears the party's ideal point because passing an alternative policy becomes more difficult (fixing x_t , as q_t approaches θ_D^m , the marginal voter approaches θ_D^m , resulting in a lower probability of passing). Therefore, exercising the option becomes more likely, and hence more valuable.

Using the nature of the value functions, Proposition 2 shows which bills are pursued with and without a whip count, accounting for the fact that whipping is costly.

Proposition 2: Fix $C_b < \hat{C}_b$ such that the optimal alternative policies, x_t^{count} and $x_t^{no\ count}$, are unique and fix the cost of a whip count, $C_w > 0$. Then, we can define a set of cutoff status quo policies, $\underline{q}_l, \bar{q}_l, \underline{q}_r$, and \bar{q}_r , with $\underline{q}_l \leq \bar{q}_l < \theta_D^m < \underline{q}_r \leq \bar{q}_r$ such that:

- (1) for $q_t \in [-\infty, \underline{q}_l] \cup [\bar{q}_r, \infty]$, the optimal alternative policy, $x_t^{no\ count}$, is pursued without conducting a whip count.
- (2) for $q_t \in (\underline{q}_l, \bar{q}_l] \cup [\underline{q}_r, \bar{q}_r)$, the optimal alternative policy, x_t^{count} , is pursued and a whip count is conducted.
- (3) for $q_t \in (\bar{q}_l, \underline{q}_r)$, no alternative policy is pursued.

We illustrate Proposition 2 via an example in Figure 4. For status quo policies nearest to party D 's ideal policy, alternative policies are never pursued because the value of such an alternative over the existing status quo is small. For status quo policies farther away, alternative policies may be pursued with or without a whip count, but when both are possible (as in the empirically relevant case illustrated), it is always policies farthest from the party's ideal policy that are pursued without a whip count, because they have a higher probability of passing ex ante (lower option value).

5. DATA

We use data from two main sources. The whip count data was compiled from historical sources by Evans (Evans, 2012, 2018), and the roll call voting data come from VoteView.org (Poole and Rosenthal, 1997, 2001).

The whip count data collected by Evans is a comprehensive set of whip counts retrieved from a variety of historical sources, mostly from archives that hold former whip and party leaders' papers. Evans (2012) describes the data collection procedure in depth. We use data from 1977-1986, as whip count data for other Congresses are not as comprehensive and complete as those for the 95th-99th Congresses, mainly due to idiosyncratic differences in the diligence of record-keeping by the Majority and Minority Whips. Importantly, however, the period under analysis is interesting because, according to most narratives, it sits at the inflection point of modern political polarization in U.S. politics (e.g. McCarty et al., 2006).

For the Republican Party, we have data from 1977-1980, originating from the Robert H. Michel Collection, in the Dirksen Congressional Center, Pekin, Illinois, Leadership Files, 1963-1996. This part of the data "appears to be nearly comprehensive about whip activities on that side of the partisan aisle, 1975-1980" (Evans, 2012). Data for the Democratic Party covers 1977 to 1986, and originates from the Congressional Papers of Thomas S. Foley, Manuscripts, Archives and Special Collections

Department, Holland Library, Washington State University, Boxes 197-203. Although John Brademas was the Majority whip from 1977 to 1980, his papers are collected within the Thomas Foley Collection (his successor).³⁸

We rely on the matching of Evans (2012) to associate each whip count with a bill voted on the floor (if the latter was sufficiently close to the one that had a whip count). In total, we have 340 bills with whip counts covering the period of 1977 to 1986, of which 238 can be directly associated with a subsequent floor vote in the House. 70 of the whip counts are Republican and the remaining 270 are Democratic. For each whip count, we have data on the Yes or No responses of each member to the party's particular question. Several bills include further whip counts (i.e. a second, third whip count), in which case we use the first whip count, as it is most representative of a member's position pre-whipping.

In these whip counts, votes from party members are predominantly recorded as: "Yes, Leaning Yes, Yes if Needed, Undecided, Leaning No, No, Expected to be Absent for Vote" (94% of the sample). We categorize these answers into the coarser groups of "Yes" or "No", which we can then compare to the leadership's position. The coded "Yes" votes (44.2% of the sample) and "No" votes (9.8%) are immediate to be classified. Among the other groups, first we deem that "Leaning Yes" and "Yes if Needed" (together, 7.2% of the sample) are "Yes" votes. Similarly, we treat "Leaning No" and "No" (together, 12.5% of the sample) as "No" votes. Finally, we take the position that "Undecided" (16.7% of the sample), "No Response" (13.0% of the sample), and "Expected to be Absent" (0.8% of the sample) are "No" votes, for two reasons. First, questions in the whip counts are generally phrased in support of the party (i.e. "Will you vote with the leadership for/against..."), so these responses suggest the member does not yet support the party's position. Second, as discussed in footnote 20, Congressional scholars have taken the position that such answers are strategically ambiguous and reflect a negative stance.³⁹

Next, we construct variables that indicate whether or not a member voted with the party leadership, as well as make Yes or No votes comparable between whip counts and roll calls (whip count questions may be framed opposite to that of the roll call).⁴⁰ To do so, we use party leadership votes

³⁸According to Evans (2012), "the Brademas records are extensive and very well organized, and I am confident that they are nearly comprehensive. For that matter, I also have a similar sense of the archival file from Foley's time in the position".

³⁹There are other categories in the data, although they make up only a small sample of the data. One group of categories include "No Comment" and "Other Response" (0.06% of the sample). For these observations, we maintain our previous definition that they are a "No" vote. Another group of responses, which we pool together as "Missing", are dropped from the analysis. These include "Missing sheets", "Ill or out of Town at the time of the poll" and "Unclear or Ambiguous". They constitute only 3.30% of the data, and could not be coded from archival records. We also drop records for the Speaker (coded separately in the raw data; 2.08% of the sample) when he doesn't take a position, as he rarely votes.

⁴⁰For example, often for the minority party, but not always, a whip count is framed in the negative, "Will you vote against...?" .

to assign the party's preferred direction on a particular whip count/roll call. In order of priority, we use the (majority/minority) party leader's vote, the (majority/minority) party Whip's vote, and, for the small set of votes for which neither are available, the direction in which the majority of the party voted.

For each roll call vote, we also need a proxy for the party that proposed the bill, in order to both determine in which region the status quo must lie and the directions each party whips. We again rely on the direction in which party leadership votes. For the majority of bills, this revealed preference, together with guidance from the theory, pins down the whipping directions. In particular, if the two party leaderships vote differently, we know from the theory that the status quo must have originated between the party's preferred positions. In this case, each party whips in the direction its party leadership prefers. If the leadership of both parties votes Yes, then the status quo could either be left of both medians with the Democrats proposing, or right of both medians with the Republicans proposing. In the former case, we expect a greater fraction of Republicans to support the bill, and vice versa in the latter case. Therefore, when the party leaderships both vote Yes, we assign the proposing party to the party that has the *least* support for the bill. Finally, a small minority of bills are supported by neither party, which cannot be reconciled with our theory. In order to avoid any selection issues, we include them by treating them as a tremble by one of the party leaderships, assigning the proposing party to be that with greater support of the bill and assuming the parties whip in opposite directions.

To demonstrate the differences between whip counts and roll calls in the raw data, Figure 5 plots the distribution of individual vote choices aligned with the party leadership at each phase (for bills proposed by the majority party that have both whip count and roll call votes). The number of members voting with the leadership dramatically increases at roll call time - a shift from approximately 160 votes with leadership at whip count time to 218 at roll call time. Notice that 218 is the simple majority threshold for the chamber - what is needed to pass a bill at roll call. Around 58 members are persuaded to toe the party line on average, moving in the direction supported by the party leaders, in accordance with our theory.

Table F.2 in Appendix provides aggregate statistics on the number of bills for which we have: (i) whip counts only (subsequently dropped), (ii) whip counts and roll calls, and (iii) roll calls only. Key bills in our time-frame address a variety of questions about economic policy, foreign aid, and domestic policy, among others. Examples include the Reagan Tax Reforms of 1981 and of 1984, the National Energy Act of 1977, the Healthcare for the Unemployed Act of 1983, the Contra affair in Nicaragua of 1984, the implementation of the Panama Canal Treaty in 1979, and multiple votes for increasing the debt limit. We revisit such key bills in our counterfactuals.

6. IDENTIFICATION AND ESTIMATION

6.1. Identification. We provide a formal proof of identification in Appendix B. Here, we state the necessary assumptions and provide intuition about the identifying variation.

The first assumption provides a normalization of the location of ideal points:

Assumption 1 (Ideal Point Locations): *We normalize the ideal point of one member (without loss of generality, member ‘O’), $\theta^0 = 0$.*

As with a discrete choice model, we must choose the distribution, G , for the idiosyncratic shocks, δ_t . The ‘scale’ of the ideal points is pinned down by a normalization of the variance of this distribution. We assume G is standard Normal so that the convolution of the two shocks, $\delta_1 + \delta_2$, which we denote G_{1+2} , is a Normal distribution with a variance of two.⁴¹

Assumption 2 (Ideal Point Scale): *G is standard Normal, with CDF denoted by $\Phi(\cdot)$.*

The following two assumptions (Assumptions 3 and 4) are needed solely for the analysis of agenda setting and are not required for our theory or for the estimation of ideal points and party discipline.

In order to be able to determine the mass of status quo policies that are never pursued (which we do not observe), we must make a parametric assumption about the distribution of status quo policies, $W(q)$. We assume a Normal distribution, $\mathcal{N}(\mu_q, \sigma_q^2)$ for the status quo policies themselves, but note that the resulting distribution of marginal voters (as determined by the proposing party) is generally different from Normal. For the purpose of allowing the status quo distribution to change over time, we allow $W(q)$ to vary by Congress.

Assumption 3 (Status Quo Distributions): *The distribution of status quo policies is $W(q) \sim \mathcal{N}(\mu_q, \sigma_q^2)$. μ_q and σ_q^2 may vary by Congress.*

Lastly, in order to determine the optimal alternative policy and hence marginal voter, we assume each party has a quadratic loss utility function around its ideal point.

Assumption 4 (Utility): *The utility a party derives from a policy, k_t , is given by a quadratic loss function around the ideal point of its median member, $u(k_t, \theta_p^m) = -(k_t - \theta_p^m)^2$.*

Under Assumption 2, the probability that a member of party D votes Yes at the whip count is given by

⁴¹A Normal distribution, while not essential, is convenient because it has a simple closed form for the convolution G_{1+2} .

$$\begin{aligned}
P(Yes_t^i = 1 \mid X_{i,t}) &= P(\delta_{1,t}^i + \theta^i \leq MV_t - \eta_{1,t} \mid X_{i,t}) \\
&= P(\delta_{1,t}^i \leq \tilde{M}V_{1,t} - \theta^i \mid X_{i,t}) \\
(6.1) \qquad &= \Phi(\tilde{M}V_{1,t} - \theta^i),
\end{aligned}$$

where $X_{i,t}$ denotes a matrix of dummy variables (for each individual i and each vote t , at both the whip count and roll call stages). The covariates, $X_{i,t}$, are common across whip counts and roll calls, because they are at the politician and bill level only.

The probability of voting Yes at roll call time it is given by

$$\begin{aligned}
P(Yes_t^i = 1 \mid X_{i,t}) &= P(\delta_{1,t}^i + \delta_{2,t}^i \leq MV_t - \eta_{1,t} - \eta_{2,t} - \theta^i \pm y_D^{max} \mid X_{i,t}) \\
&= P(\delta_{1,t}^i + \delta_{2,t}^i \leq \tilde{M}V_{2,t} - \theta^i \pm y_D^{max} \mid X_{i,t}) \\
(6.2) \qquad &= \Phi\left(\frac{\tilde{M}V_{2,t} - \theta^i \pm y_D^{max}}{\sqrt{2}}\right).
\end{aligned}$$

In (6.2), the sign with which y_D^{max} enters depends upon the direction that party D whips (see Section 6.2). We seek to identify the parameter vector:

$$\Theta = \{\{\{\theta_p^i\}, y_p^{max}, \underline{q}_{l,p}, \bar{q}_{l,p}, \underline{q}_{r,p}, \bar{q}_{r,p}\}_{p \in \{D,R\}}, \gamma, \mu_q, \sigma_q, \{\tilde{M}V_{1,t}\}, \{\tilde{M}V_{2,t}\}, \sigma_\eta\}.$$

As is standard in ideal point estimation, the member ideal points, $\{\theta_p^i\}$, are identified relative to each other by the frequencies at which the members vote Yes and No over a series of whip count votes. Namely, they are proportional to their probabilities of voting Yes over the same set of bills. Their absolute positions are then pinned down by the normalization assumptions (Assumptions 1 and 2). Given the ideal points, the realized marginal voter at each whip count, $\{\tilde{M}V_{1,t}\}$, is then identified as the ‘cutpoint’ that divides the Yes and No votes.

At roll call time, each party has a different cutpoint (because of different party discipline parameters) given by $\{\tilde{M}V_{2,t}\} \pm y_p^{max}$. The two cutpoints are identified by the locations that best divide Yes and No votes within a party. We determine the sign of the party discipline parameter using a proxy for the whipping direction, as described in Section 5. With whip count data, we can separately identify each party discipline parameter by the average change in votes between the whip count and roll call.⁴² In Congresses for which we have whip count data for only one of

⁴²To identify the individual party discipline parameters from the change between whip count and roll call requires that the aggregate shock between these stages to be mean zero or at least of known mean. Knowledge of the mean of $\eta_{1,t}, \eta_{2,t}$ allows us to separate whether systematic changes in votes from whip counts to roll calls originate from the aggregate

the parties, identification is guided by the fact that some members are present in Congresses for which we have whip counts for both parties. Next, because the estimated cutpoint at roll call time within a party is given by $\{\tilde{M}V_{2,t}\} \pm y_p^{max}$, we can recover the realized marginal voters, $\{\tilde{M}V_{2,t}\}$. The variance in the second aggregate shock, η_2 , is given by the variance of the differences between realized marginal voters at whip count and at roll call.

Identification of the parameters governing agenda-setting, $\{\gamma, \mu_q, \sigma_q, \{\underline{q}_{l,p}, \bar{q}_{l,p}, \underline{q}_{r,p}, \bar{q}_{r,p}\}_{p \in \{D,R\}}\}$, requires the distributional assumption, Assumption 3. Under this assumption, the status quo distribution that the parties draw from is Normal, which, from the theory, means that the bills with only roll calls are drawn from a truncated Normal.⁴³ The resulting distribution of marginal voters is pinned down by the relationship between status quo policies and optimal alternative policies (Lemma A1 in the Appendix shows that the relationship between status quo and marginal voter is one-to-one), assuming each party has a quadratic loss utility function around its ideal point (Assumption 4). Convolving the distribution of marginal voters with those of the first and second aggregate shocks (whose variances have already been identified) provides a distribution over the realized marginal voters, $\{\tilde{M}V_{2,t}\}$, which we then match to the data.

Intuitively, the mean, variance, and cutoffs of the truncated Normal distribution all provide independent effects on the distribution of realized marginal voters for bills with roll calls only. Once the status quo distribution is identified, the cutoffs, $\bar{q}_{l,p}$ and $\underline{q}_{r,p}$, that determine the range of status quo policies for which whip counts are conducted, are pinned down by the number of whip counted bills. Finally, the probability that D proposes a bill, γ , is determined by our proxy for the party proposing the bill augmented by information on the extent of selection of status quos, as discussed in the following subsection. We verify this intuition with extensive Monte Carlo simulations, reported in Appendix E.

The probabilities and likelihoods shown in the next section are also conditional on $X_{i,t}$, but to keep notation light we omit this dependency from the main equations.

6.2. Two Step Estimation. We observe votes for both parties, $p \in \{D, R\}$, at both the whip count stage (denoted $Y_{est,p}^{i,wc}$) and at the roll call stage (denoted $Y_{est,p}^{i,rc}$), for each politician $i \in \{1, \dots, N\}$ and period $t \in \{1, \dots, T\}$. We estimate the model in two steps.

shocks or from party discipline. In addition, we have a second source of identification for the party discipline parameters. This comes from the two parties agreeing on some proposals (whipping in the same direction), but disagreeing on others (whipping in opposite directions). The difference between their cutpoints for any bill may be either the difference or the sum of the individual discipline parameters.

⁴³For computational reasons, we estimate the status quo cutoffs directly rather than the cost parameters, C_b and C_w , that determine them. The cutoffs are complex, implicit functions of the cost parameters making it infeasible to calculate them within the optimization loop. By allowing the cutoffs to be different on either side of each party's median, we are implicitly allowing the costs to be potentially different in each case. This assumption therefore allows the cost of pursuing a bill to depend upon whether or not parties agree or disagree over the alternatives.

In the first step, we take the distribution of status quo policies as given, which is possible because we estimate the realized marginal voters as fixed effects. We estimate the set of parameters, $\Theta_1 = \{\{\theta_p^i\}, y_p^{max}\}_{p \in \{D, R\}}, \{\tilde{M}V_{1,t}\}, \{\tilde{M}V_{2,t}\}, \sigma_\eta\}$, by Maximum Likelihood, allowing the party discipline parameters, y_p^{max} , to vary by Congress.

Replacing the conditional probability of observing a Yes vote at roll call given a Yes vote at whip count by its unconditional probability, we can define the pseudo-likelihood for the first step:

$$(6.3) \quad \mathcal{L}(\Theta_1; Yes_{t,p}^{i,wc}, Yes_{t,p}^{i,rc}) = \prod_{p \in \{D, R\}} \prod_{t=1}^T \prod_{n=1}^{N_p} P(Yes_{t,p}^{i,wc} = 1)^{Yes_{t,p}^{i,wc}} P(Yes_{t,p}^{i,wc} = 0)^{1-Yes_{t,p}^{i,wc}} \times P(Yes_{t,p}^{i,rc} = 1)^{Yes_{t,p}^{i,rc}} P(Yes_{t,p}^{i,rc} = 0)^{1-Yes_{t,p}^{i,rc}}$$

Using the pseudo-likelihood as opposed to the more cumbersome original likelihood has no effect on consistency of the estimation (Gourieroux et al., 1984; Wooldridge, 2010), because our model is identified despite the nuisance of the dependence between the roll call and the whip count stages.

For the Democratic Party, we can use equations (6.1) and (6.2), together with our parametrization to re-express the likelihood of a series of votes by member of party D in (6.3) as:

$$(6.4) \quad \mathcal{L}_D(\Theta_1; Yes_{t,p}^{i,wc}, Yes_{t,p}^{i,rc}) = \prod_{t=1}^T \prod_{n=1}^{N_D} \Phi(\tilde{M}V_{1,t} - \theta^i)^{Yes_{t,p}^{i,wc}} \left(1 - \Phi(\tilde{M}V_{1,t} - \theta^i)\right)^{1-Yes_{t,p}^{i,wc}} \times \Phi\left(\frac{\tilde{M}V_{2,t} - \theta^i \pm y_D^{max}}{\sqrt{2}}\right)^{Yes_{t,p}^{i,rc}} \left(1 - \Phi\left(\frac{\tilde{M}V_{2,t} - \theta^i \pm y_D^{max}}{\sqrt{2}}\right)\right)^{1-Yes_{t,p}^{i,rc}}$$

using $P(Yes_{t,p}^{i,stage} = 1) = 1 - P(Yes_{t,p}^{i,stage} = 0)$, for $stage \in \{wc, rc\}$. An analogous expression for the likelihood of votes by member of party R holds (see Appendix B).

We estimate (6.3), subject to $\theta^0 = 0$ (Assumption 1), then obtain an estimate of σ_η^2 from the variance of the difference between the realized marginal voters at whip count and roll call (for those bills which have both).

In the second step, we estimate the remaining parameters,

$\Theta_2 = \{\gamma, \mu_q, \sigma_q, \{\underline{q}_{l,p}, \bar{q}_{l,p}, \underline{q}_{r,p}, \bar{q}_{r,p}\}_{p \in \{D, R\}}\}$, using both the realized marginal voters, $\{\tilde{M}V_{2,t}\}$, for

bills with only roll calls and the number of whip counts (whether pursued to roll call or not).⁴⁴ In each period, we observe either a whip count ($WC_t = 1$) or the realized marginal voter for a roll call without whip count ($RC_t = 1$) so that the likelihood can be written

$$\mathcal{L}^{second\ step}(\Theta_1; \tilde{WC}_t, \tilde{MV}_{2,t}) = \prod_{t=1}^T P(WC_t)^{WC_t} P(\tilde{MV}_{2,t})^{RC_t}$$

The probability of observing a whip count is simply the probability that a status quo is drawn from the appropriate interval of the q support. Because for some status quo policies (those between $\bar{q}_{l,p}$ and $\underline{q}_{r,p}$) we observe neither a whip count nor a roll call, we must condition on the probability that we observe either. For example, for a whip count for a status quo to the right of a party's median, we have, using Proposition 2:

$$P(WC_t) = \frac{\Phi\left(\frac{\bar{q}_{r,p} - \mu_q}{\sigma_q}\right) - \Phi\left(\frac{\underline{q}_{r,p} - \mu_q}{\sigma_q}\right)}{P(WC_t \cup RC_t)}$$

where

$$P(WC_t \cup RC_t) = \gamma \left(\Phi\left(\frac{\bar{q}_{l,D} - \mu_q}{\sigma_q}\right) + 1 - \Phi\left(\frac{\underline{q}_{r,D} - \mu_q}{\sigma_q}\right) \right) + (1-\gamma) \left(\Phi\left(\frac{\bar{q}_{l,R} - \mu_q}{\sigma_q}\right) + 1 - \Phi\left(\frac{\underline{q}_{r,R} - \mu_q}{\sigma_q}\right) \right)$$

A realized marginal voter can come from a range of status quo policies. For example, the probability of observing a particular realized marginal voter for a status quo drawn from the right of the Democrats median (conditional on observing either a whip count or roll call) is:

$$P(\tilde{MV}_{2,t}) = \int_{\bar{q}_{r,D}}^{\infty} \phi\left(\frac{\tilde{MV}_{2,t} - MV(q_t)}{\sigma}\right) \frac{\phi\left(\frac{q_t - \mu_q}{\sigma_q}\right)}{P(WC_t \cup RC_t)} dq_t$$

The term, $\frac{\phi\left(\frac{q_t - \mu_q}{\sigma_q}\right)}{P(WC_t \cup RC_t)}$, is the conditional probability of drawing a particular q_t . A given q_t determines the marginal voter, $MV_t = MV(q_t)$, through the first-order condition.⁴⁵ The term, $\phi\left(\frac{\tilde{MV}_{2,t} - MV(q_t)}{\sigma}\right)$ is then the probability of observing a particular realized marginal voter, $\tilde{MV}_{2,t}$,

⁴⁴Although the first step also recovers the realized marginal voters at the time of the whip count, $\{\tilde{MV}_{1,t}\}$, they are a function of the unobserved cost parameter, C_b , and so are not easily incorporated into the likelihood function. They are not necessary, however, as the number of whip counts themselves are sufficient to recover the associated cutoffs.

⁴⁵Importantly, the first-order condition in case of no whip count does not depend on the unobserved cost parameters. For each Congress, we calculate the optimal policy alternatives for each party using estimates of the party medians, the standard deviation of the sum of the aggregate shocks, and the $\hat{MV}_{i,j}$ parameters calculated from the estimates obtained in the first step.

for the given MV_t . Integrating over all possible q_t 's that could generate the observed realized marginal voter gives the probability.⁴⁶

In estimating the second step likelihood, we allow the cutoff status quo policies, $\{q_{l,p}, \bar{q}_{l,p}, q_{r,p}, \bar{q}_{r,p}\}_{p \in \{D,R\}}$ and the distribution (μ_q and σ_q) to vary by Congress, but hold the probability that the Democrats propose the bill, γ , constant.⁴⁷ As such, we are implicitly allowing the costs, C_b and C_w , to vary by Congress.

Appendix E presents the results of extensive Monte Carlo simulations of both the first and second steps, demonstrating good finite sample performance in each step. It also discusses the validity of our asymptotic inference in this context.

7. RESULTS

7.1. First Step Estimates: Ideologies and Party Discipline. Table 1 presents our first step, Maximum Likelihood estimates. In this step, we recover the estimated ideologies, θ^i , for 711 members of Congress from 315 whip counts and 5424 roll call votes. We report the party medians for each congressional cycle. We also recover the party discipline parameters, y_D^{max} and y_R^{max} , for each Congress, and the standard deviation of the aggregate shocks, σ_η . All parameters are precisely estimated.

In our first main result, Table 1 shows that both party discipline parameters, y_D^{max} and y_R^{max} , are positive and statistically different from zero in each Congress, rejecting the null of a model without party discipline (i.e. with no whipping). This party discipline results in additional polarization in votes, above and beyond that due to ideological polarization itself. Under standard methods that use roll calls only and assume sincere voting by politicians, this additional polarization in votes incorrectly loads on the ideologies, producing *perceived* ideological polarization that is too large. In fact, party discipline results in the party medians being exactly $y_D^{max} + y_R^{max}$ too far apart when party discipline is ignored.⁴⁸ To illustrate this fact, Figure 6 plots kernel densities of the estimated legislator ideologies, θ^i , by party and over time from our full model (solid lines). For comparison

⁴⁶To estimate the second step likelihood, we need to identify for each whip count and realized marginal voter, the associated range of status quo policies. For roll call votes, we do so based on our proxy for which party proposed the bill as described in Section 5. For whip counts with subsequent roll calls, we identify the associated range of status quo policies for the whip counts based upon the corresponding range of status quo policies associated with the roll call. For whip counts without roll calls, we have no way to determine the leadership stance of the party that did not conduct a whip count. The natural assumption is that a party is more likely to conduct a whip count when it expects opposition from the other party, so we assume that the party conducting the whip count is the proposer and that the status quo is right of the party median for Democratic proposals and left of the party median for Republican proposals.

⁴⁷We estimated a specification that allowed γ to vary by Congress, but rejected this specification through a likelihood ratio test. The values of γ in each Congress were very similar.

⁴⁸One may hypothesize that party discipline results in a ‘hollowing out’ of the middle of the distribution. However, party discipline simply shifts the cutpoint between Yes and No (see equation 6.2), which, under the assumption of unbounded idiosyncratic shocks, affects the estimates of all ideologies in the same way.

purposes, it also plots the corresponding ideological distributions (dashed lines) which result from estimates of a misspecified model in which we impose no party discipline, $y_D^{max} = 0$ and $y_R^{max} = 0$.

Differences in our methodology from standard methods (i.e. DW-Nominate random utility, optimal classification scores, Heckman-Snyder linear probability model scores, or Markov Chain Monte Carlo approaches) are not driving our results.⁴⁹ As evidence, Figure 7 compares the estimated ideologies from our full model (right panel) and misspecified model with no party discipline (left panel) to the standard DW-Nominate estimates. The misspecified model and DW-Nominate estimates are very close, demonstrating that the two methods produce comparable results. Our full model, however, reveals a gap in density over the ideological middle ground, driven by DW-Nominate's loading of party discipline on legislator ideology. This misspecification results in a sizable bias in DW-Nominate estimates, amounting to around 0.20 DW-Nominate units.

Figure 8 combines all three models (our main estimates, the misspecified model without party discipline, and DW-Nominate) to show graphically the impact of ignoring party discipline. Tracing across Congresses, party polarization, defined in terms of the distance between party medians $\theta_R^m - \theta_D^m$, widens over time, as can also be seen in Table 1. Thus, even controlling for party discipline, we confirm the established view that ideologies appear to be diverging across party lines.

However, Figure 9 illustrates that party discipline is also becoming more important over time for both parties: the trend in y_p^{max} for each party is clearly positive, tracing an increase in the reach of party leaders over rank-and-file members. The null hypothesis of a constant y_p^{max} across Congresses is rejected via a likelihood ratio test after obtaining estimates from the constrained model (see Table F.3 in Appendix F for details).⁵⁰

The perceived ideological polarization in a misspecified model, visualized in Figure 8 for instance, increases not only because of actual increases in ideological polarization, but also due to stronger party discipline. Table 2 shows that party discipline accounts for 34 to 44 percent of perceived ideological polarization, and is increasing in importance over time. This measure is well defined as both y^{max} and ideologies are defined and measured in the same ideological space.

This rise in party discipline in the mid 1970s coincides with large reforms conducted in the House of Representatives, in particular among the majority Democratic party. During this period, power was heavily concentrated in the party leadership's hands. Among the changes, leaders became

⁴⁹For a discussion of optimal classification and maximum score estimators and their properties, see Appendix G. Combining the discussion in this section with that in Appendix G should make clear that using a nonparametric estimator does not, by itself, solve identification issues related to party discipline.

⁵⁰As specified in Section 6.1, given the rest of the data available, our estimates of y_p^{max} do not depend on whether whip count data is available for the Republican party in the 97th-99th Congresses or not. y_R^{max} , in particular, is identified. In Figure 9, while a jump in polarization happens in Congress 98 when whip count data for Republicans is absent, we do not observe a break in the trend of y_R^{max} between the 96th and 97th Congresses.

responsible for committee assignments (including the Rules Committee), the Speaker gained larger control of the agenda progress, new tactics emerged (such as packaging legislation into ‘megabills’), and the Democratic Steering and Policy Committee was formed. The latter met regularly to gather information and determine tactics and policies, with the leadership controlling half of the votes. One strong motivation for these reforms appears to be policy: to prevent more liberal policies from being held back by Committee chairmen. See Rohde (1991) for a thorough description of some of these reforms and their motivation.⁵¹

In Table 3, we report in-sample model fit: individual vote choices correctly predicted by the model. The overall fit for roll call votes (with and without whip counts) is 85.5 percent. For whip count votes, the fit is lower, at 63 percent, due to the fact that whip count votes are much fewer in number and Maximum Likelihood weighs whip count votes and roll call votes equally. Overall, the fit of the model is very good, especially considering that we do not drop any roll call (we include both lopsided and close votes). This approach differs from extant approaches that condition on (occasionally hard to justify) selected subsamples of votes. For comparison, over our sample, the DW-Nominate prediction rate is 85.9 percent, but for reasons that were not immediate to us the procedure drops 892 roll calls, that we instead include.

Lastly, our first step produces an estimate of the size of the aggregate shock between whip count and roll call, $\eta_{2,t}$. In the theory, we assume that $\eta_{2,t}$ follows a mean-zero Normal distribution. In practice, we recover the distribution of $\eta_{2,t}$ semi-parametrically. Figures 10a and 10b show, via a histogram and a QQ-plot, that a Normal distribution with mean 0 and variance $\hat{\sigma}_\eta^2 = 0.859$ fits the recovered distribution of these aggregate shocks very well, providing empirical support for our assumption.⁵²

7.2. Second Step Estimates: Agenda Setting. Table 4 presents the results of Maximum Likelihood Estimation of the second step. We find that the means of status quo policy distributions, $W(q)$, lie between the party medians, with a standard deviation similar to the estimated distance between the party medians.⁵³

⁵¹One can also observe polarization in votes in the Senate, starting in the mid to late 1970’s. Although the Senate did not face institutional changes as extensive as those in the House of Representatives, their leaders also adopted “technological innovations” such as megabills, omnibus legislation, and time-limitation agreements, allowing more control over their party members and the agenda. See Deering and Smith (1997) for a discussion.

⁵²The estimated shocks have slightly larger tails than a Normal, which is expected given that the shocks are convoluted with estimation error.

⁵³We do not model explicitly intertemporal linkages across Congresses in terms of policy alternatives today that become tomorrow’s status quo policies, or any dynamic considerations in this respect on the part of party leaders. These extensions appear completely intractable. However, our parametric time-varying distribution of status quo policies allows the model to capture these dynamic considerations across Congresses, to a reasonable extent.

Our theoretical framework makes clear predictions about which status quo policies, q_t , are: (i) never brought to the floor; (ii) whip counted and then brought to the floor with a corresponding alternative, x_t , and (iii) brought directly to the floor with a corresponding alternative. In particular, as illustrated in Figure 4, the model predicts that status quo policies closest to a party's median are not pursued at all, the next closest are pursued with a whip count, and those furthest away proceed directly to roll call. We partially test this implication of the model in Table 5, by comparing the average absolute distance of the realized marginal voters among policies that were whip counted (whether they proceeded to roll call or not) to those brought directly to roll call. Because status quo policies closer to the party median result in realized marginal voters closer to the party median (on average), we expect realized marginal voters to be closer for policies with whip counts than for those that proceed directly to roll call. The results of Table 5 strongly confirm this prediction of our theory with either the Democrats or the Republicans as the proposing party.

We illustrate the unobservable ‘missing mass’, those status quo policies that are never pursued, in Figures 11 and 12. Status quo policies brought directly to the floor are indicated by dashed lines and those shaded in gray are preceded by whip counts. The gaps in the distributions around the party medians represent estimates of the missing mass. As reported in Table 6, the fraction of missing mass hovers around 15 percent across Congresses for the minority party and ranges from 5 to 11 percent for the majority party.⁵⁴ Bills that are first whip counted may also never see a floor vote, a form of agenda setting made explicit in our model. In the data, across all Congresses, on average 2 out of 7 whip counted bills are abandoned before reaching the floor (Table F.2 in Appendix). Overall, our results suggest substantial censoring of the status quo policies pursued, indicating selection is an important role of parties in legislative activity.

Lastly, agenda setting works not only through selection, but also through the choice of policy alternatives to pursue in the first place. In Appendix Figures F.1 and F.2, we report the implied distributions of marginal voters based upon the estimated status quo distribution and the optimal policy alternatives, x_t^* , from theory.⁵⁵ Each graph illustrates both parties' efforts to move policy closer to their ideal points across the entire distribution of status quo policies. The reduction in the variance of the marginal voter distribution relative to that of the status quo policies is substantial,

⁵⁴Note that our estimates of the missing mass do not directly relate to counts of the number of proposed bills that never make it to the whip count stage (for example, dropped in committee). These counts would include bill proposals that neither party ever intended to pursue.

⁵⁵This is preferable than plotting the distribution of alternative policies, x_t^* , because the latter is a non-monotone function of q_t which is difficult to depict graphically.

indicating sizable changes in policy. In addition, the variance in the marginal voter distribution narrows over time, consistent with the finding that parties are increasingly able to discipline members, and can thus pursue policy alternatives closer to their ideal points.

7.3. Robustness. Appendix D presents three additional estimates of the first stage of our model. First, we re-estimate our model on the subsample of final passage votes alone (as opposed to amendments or motions), showing that our results are largely unchanged in a subset of salient votes. Second, we implement a procedure akin to [Snyder and Groseclose \(2000\)](#) in which we assume lopsided votes are not whipped (relaxing the assumption that all roll calls are subject to whipping). Although the average party discipline decreases by about 7 percentage points in this specification (as the focus on divisive bills magnifies defection from the party line), this modification of our approach does not substantially alter our main quantitative conclusions. Third, we explore the possibility that our results are overly dependent on imposing a single dimensional policy space. In the time period we study, other approaches, such as DW-Nominate, find evidence supporting a second latent component of ideology, along the civil rights dimension ([Poole, 2007](#)). We re-estimate our model dropping bills that split Northern and Southern Democrats using data from David Rohde’s PIPC-University of Oklahoma repository. We do not find significant changes on the ideological distributions. This suggests that while a second dimension may not be unimportant, precluding it does not radically distort our findings.

In Appendix D, we also address an aspect not accounted for in the model - the salience of a bill might provide a reason for it to be whip counted. There, we note that only a small fraction of whip counted votes are included among the Congressional Quarterly’s ‘Key Votes’ series, a standard reference for highly visible votes. Only 32.5 percent of these Key Votes are whip counted, suggesting that a vote being salient is neither necessary nor sufficient for it to be whip counted. This fact, combined with our result that whip counted votes tend to have marginal voters closer to the party median (Table 5), suggests that the informational value of a whip count may provide at least as good a rationale of why whip counts are conducted as the visibility of a bill to electors.

Finally, our estimate of γ in Table 4 suggests that the majority party does not appear to have a majority of proposal opportunities, counter to what one might expect. The main reason for this result is that our sample of roll call votes is highly heterogeneous and includes both important final passage votes and a vast number of smaller amendments and motions typically originating from the minority party. Among final passage bills in which the two parties oppose each other, the Democrats propose more than 94 percent. Thus, among these highly significant votes, the majority does in fact propose most of them. But, as pointed out by [Jenkins et al. \(2014\)](#), among the far greater

number of amendment and motion votes, the minority party is actively involved, negotiating with the majority party, using as bargaining chip threats of obstruction. Relatedly, [Jenkins et al. \(2014\)](#) uses bill sponsorship data to identify the proposing party. In Appendix D, we limit our sample to those votes in which our proposal proxy (described in Section 5) matches that of [Jenkins et al. \(2014\)](#) based on the identity of the bill sponsor. In this subsample (where there is little doubt about the identity of the proposing party) we find $\hat{\gamma} = 0.56$, which is larger than our main estimate. For further discussion, we direct the reader to Appendix C and Appendix D.

8. COUNTERFACTUALS

8.1. Salient Bills. In our first counterfactual exercise, we analyze the role of party discipline for the approval of historically salient legislation, focusing on a series of economically consequential bills from our sample. To do so, we maintain the policy alternatives to be voted on as they were proposed in Congress (including realized aggregate shocks), but assume that parties cannot discipline members' votes: legislators vote solely according to their ideologies. Specifically, we calculate the predicted votes for a bill setting $y_D^{max} = y_R^{max} = 0$.

Among the bills we consider are the lifting of the arms embargo to Turkey, the Panama Canal Treaty, several increases to the Debt Limit, the Social Security Amendments of 1983, and the Reagan Tax Reforms of 1981 and 1984. The first and second columns of Table 7 show that our baseline model fits these votes well. The third column presents the results of the counterfactual exercise, showing that party discipline is quantitatively important for the outcomes of these bills as, in some cases, their passage would have been reverted. In particular, a lack of party discipline would have reversed the approval of increases to the Debt Limit and significantly decreased support for the Social Security Amendments of 1983 and the 1984 Reagan Tax bill.⁵⁶

Although many bills lose support, Table 7 shows that others actually gain votes, a subtle consequence of differences in the location of the marginal voter and the directions each party whips their members. Consider H.R. 5399 banning aid to the Contras. For this bill, the Democrats whipped in favor and the Republicans against. The estimated marginal voter at roll call time is 0.288, right of both party medians.⁵⁷ Shutting down the ability of Democrats to whip for support of this bill changes a limited number of votes, as very few Democrats lie to the right of the marginal voter. On the other hand, shutting down the ability of the Republicans to whip against the bill increases its support substantially, because many Republican ideal points lie near the marginal voter. Thus,

⁵⁶Selecting such bills as the 1984 Tax Reform is motivated by their use in economics to study consumer decisions, labor supply and labor and income elasticities [Auerbach and Slemrod \(1997\)](#); [Hausman and Poterba \(1987\)](#); [Soules \(2002\)](#). We provide further historical details about other bills in Appendix H.

⁵⁷This number rationalizes the large number of both Democrats and Republicans voting Yes, even if the Republican leadership voted against it.

absent party discipline by either party, the number of Yes votes actually increases. An analogous argument, with opposite signs, leads to a decrease in support for the National Energy Act and for the 1984 Tax Reform. As a final example, H.R. 9290 which increased the temporary debt limit in the 95th Congress, loses about 35 Yes votes absent whipping, changing the outcome of the vote. The estimated marginal voter is -1.20 , a point sufficiently to the left that only a small minority of politicians would have voted Yes without both parties whipping for its support.

The results in this section point to the quantitative importance of party discipline in determining policy outcomes. Our exercise here is, however, only a partial equilibrium one: absent the ability to discipline members, the equilibrium policy alternatives would have also changed. We consider the full equilibrium effects of a lack of ability to discipline in the following section.

8.2. Agenda Setting.

8.2.1. No Party Discipline. We consider now a counterfactual exercise with no whipping ($y_D^{max} = y_R^{max} = 0$), but unlike in the previous section, we allow the proposing party to re-optimize. This entails choosing which status quo policies to pursue, whether to perform a whip count or not, and selecting the optimal alternative policy, x_t . Because we cannot identify the status quo associated with a particular bill (due to aggregate shocks), in this section we focus on averages across simulated bills. In particular, we calculate the average probability that a bill will pass and the average distance between the status quo and the proposed alternative, focusing on status quo policies that lie between the party medians (as estimated with our main model).

Table 8 reports these two measures for the estimates from our model, as well as under the counterfactual of no whipping. From these results, we see that party discipline impacts the probability of approval of a bill by over 10 percent, while the magnitude of its effects on the policy alternative is very close to 0. For bills proposed by the Democrats, we observe a decrease in the passage rate of approximately 5 percentage points on average, relative to a baseline probability of 43 percent. For Republicans, however, when neither party whips there is an increase in bill approval of approximately 4 percentage points on a baseline of 22 percent. The Republicans benefit from a lack of whipping by both parties, but the Democrats suffer, because the Democrats exert more discipline (see first step estimates in Table 1) and are the majority party. For both reasons, when discipline is shut down for *both* parties, the Democrats lose more votes than Republicans, making proposals by Republicans more likely to pass and proposals by Democrats less so.

The lack of ability to discipline also impacts the size of the mass of bills that are never pursued (see Table 6). For the Democrats, we observe small increases in the missing mass, consistent with it being more difficult for them to pass legislation, lowering the value of pursuing a policy alternative.

For the Republicans, the opposite occurs. The value of pursuing a bill increases because bills are passed more easily, enlarging the set of status quo policies that Republicans pursue and reducing their missing mass.

8.2.2. Increased Ideological Polarization. Our final counterfactual considers the effects of an increase in ideological polarization. In particular, holding everything else constant, we shift the Democratic party median to the left and the Republican party median to the right, increasing the distance between medians by $\frac{y_D^{max} + y_R^{max}}{2}$. We consider the same measures as in the previous section: probability of bill approval, distance between alternative and status quo policies, and the extent of the missing mass. Table 8 presents the results for the first two measures and Table 6 reports the missing mass results.

We find that an increase in ideological polarization has very different effects from changes in party discipline. The probability that a bill passes is relatively unchanged, but alternative policies are now set further left by Democrats and further right by Republicans.⁵⁸ Hence, the polarization in ideologies translates directly to polarization in the bills pursued. The magnitudes of these changes are quantitatively significant, ranging from 6 to 15 percent of the distance between the party medians relative to where they would have been, an order of magnitude larger than the changes resulting from a lack of party discipline. Interestingly, the missing mass changes go in the opposite direction to those under the counterfactual of no party discipline. The missing mass decreases for the Democrats and increases for the Republicans, suggesting that the value of pursuing a policy alternative increases for the majority party, but decreases for the minority party as ideological polarization increases.

Taken together, our counterfactual results suggest that an increase in polarization, either through an increase in party discipline or through ideological polarization, increases the value of pursuing an alternative policy for the majority party (lowers the missing mass for the Democrats), but decreases the value for the minority party (increases the missing mass for the Republicans). The results therefore suggest that increases in polarization via either channel benefit the majority party at the expense of the minority party. However, the channel matters - ideological polarization produces more polarized policies, while party discipline affects mainly the probability of bill approval. The benefit of explicitly modeling party discipline, optimal policy selection, and bill pursuit decisions simultaneously is that it demonstrates the subtle interactions between these factors. Omitting any single factor could lead to different and potentially biased conclusions.

⁵⁸These average effects mask heterogeneity with respect to the status quo as shown in Appendix Figures F.3 - F.4. The counterfactual effects depend upon whether or not the status quo is between the party medians. For example, no party discipline is beneficial for the minority for bills within the medians (as the majority no longer whips against it), but it is harmful outside of it (where party discipline by the majority goes in the direction of the minority).

9. CONCLUSION

Polarization of political elites is an empirical phenomenon that has recently reached historical highs. It has consequential implications, ranging from heightened policy uncertainty (with consequences for investment and trade) to gridlock and inability of political elites to respond to shocks and crises. Extant literature has suggested competing views of the drivers of polarization and what can be done to counter this phenomenon. Some researchers point squarely at the ideological polarization of legislators, arguing that it is a result of more polarized electorates electing extremists. Under this view, polarization is the result of deep drivers, linked to secular trends in the electorate, for which policy response seems arduous, if at all warranted. Other researchers caution about the role of individual ideology and instead emphasize changes in the rules of controlling the legislative agenda, tightening of the leadership's grip over policy, and the capacity of parties to more precisely reward and punish their rank-and-file. Differently from ideology, these drivers appear more the product of political tactics and amenable to reversal.

Separating these different drivers, both of which we show are at play, requires an effort to solve extant political economy problems speaking to the internal organization of parties – particularly related to the internal aggregation of information from the rank-and-file and party discipline. Our empirical analysis provides an identification strategy useful for the quantitative assessment of the role of preferences and parties over the initial phase of modern congressional polarization and our theoretical setting rationalizes these issues within an internally coherent structure. A series of counterfactual exercises indicate a relevant role for party discipline, almost as important as legislator ideology, in explaining polarization dynamics. Future research should pursue the possibility of extending our estimation methodology to time periods well beyond the 99th Congress, where identifying information as precise and comprehensive as that we employ here is not available.

REFERENCES

- Aldrich, J. H. (1995). *Why parties?: The origin and transformation of political parties in America*. University of Chicago Press.
- Ansolabehere, S., Snyder, J. M., and Stewart III, C. (2001a). The effects of party and preferences on congressional roll-call voting. *Legislative Studies Quarterly*, pages 533–572.
- Ansolabehere, S., Snyder Jr, J. M., and Stewart III, C. (2001b). Candidate positioning in us house elections. *American Journal of Political Science*, pages 136–159.
- Auerbach, A. J. and Slemrod, J. (1997). The economic effects of the tax reform act of 1986. *Journal of Economic Literature*, 35(2):589–632.

- Baker, R. K. (1985). Party and institutional sanctions in the us house: The case of congressman gramm. *Legislative Studies Quarterly*, pages 315–337.
- Baker, S. R., Bloom, N., Canes-Wrone, B., Davis, S. J., and Rodden, J. (2014). Why has us policy uncertainty risen since 1960? *American Economic Review*, 104(5):56–60.
- Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.
- Ban, P., Moskowitz, D. J., and James M. Snyder, J. (2016). The changing relative power of party leaders in congress. mimeo.
- Barber, J. D. (1965). *The lawmakers: Recruitment and adaptation to legislative life*, volume 11. New Haven: Yale University Press.
- Bateman, D. A., Clinton, J., and Lapinski, J. S. (2017). A house divided? political conflict and polarization in the u.s. congress, 1877- 2011. *American Journal of Political Science*, 61(3):698–714.
- Binder, S. (2003). *Stalemate: Causes and consequences of legislative gridlock*. Brookings DC.
- Bonica, A. (2014). Mapping the ideological market place. *American Journal of Political Science*, 58(2):367–386.
- Burden, B. C. and Frisby, T. M. (2004). Preferences, partisanship, and whip activity in the us house of representatives. *Legislative Studies Quarterly*, 29(4):569–590.
- Caillaud, B. and Tirole, J. (2002). Parties as political intermediaries. *The Quarterly Journal of Economics*, 117(4):1453–1489.
- Clinton, J., Jackman, S., and Rivers, D. (2004). The statistical analysis of roll call data. *American Political Science Review*, 98(2):355–370.
- Clinton, J., Katznelson, I., and Lapinski, J. (2014). Where measures meet history: Party polarization during the new deal and fair deal. *Governing in a Polarized Age: Elections, Parties, and Representation in America*.
- Cox, G. W. and McCubbins, M. D. (1993). *Legislative Leviathan: Party Government in the House*, volume 23. Univ of California Press.
- Cox, G. W. and McCubbins, M. D. (2005). *Setting the agenda: Responsible party government in the US House of Representatives*. Cambridge University Press.
- Deering, C. J. and Smith, S. S. (1997). *Committees in congress*. Sage.
- Dodd, L. C. (1979). Expanded roles of the house democratic whip system - 93rd and 94th congresses. In *Congressional Studies - A Journal of the Congress*, volume 7 (1), pages 27–56. US Capitol Historical Society, 200 Maryland AVE NE, Washington, DC 20515.

- Evans, C. L. (2011). Growing the vote: Majority party whipping in the us house, 1955-2002. In *10th Annual Congress and History Conference, Brown University, June*, pages 9–10.
- Evans, C. L. (2012). Congressional whip count database. In *College of William and Mary, mimeo (Online)*.
- Evans, C. L. (2018). *The Whips: Building Party Coalitions in Congress*. University of Michigan Press.
- Evans, C. L. and Grandy, C. E. (2009). The whip system of congress. In *Congress Reconsidered*, volume 9. CQ Press Washington, DC.
- Fernández-Val, I. and Weidner, M. (2016). Individual and time effects in nonlinear panel models with large n , t . *Journal of Econometrics*, 192(1):291–312.
- Forgette, R. (2004). Party caucuses and coordination: Assessing caucus activity and party effects. *Legislative Studies Quarterly*, 29(3):407–430.
- Gentzkow, M. (2016). Polarization in 2016. *Toulouse Network of Information Technology white paper*.
- Gentzkow, M., Shapiro, J. M., and Taddy, M. (2017). Measuring polarization in high-dimensional data: Method and application to congressional speech. *NBER Working Paper No. 22423*.
- Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica*, 52(3):681–700.
- Green, M. N. and Harris, D. B. (2019). *Choosing the Leader: Leadership Elections in the US House of Representatives*. Yale University Press.
- Hausman, J. A. and Poterba, J. M. (1987). Household behavior and the tax reform act of 1986. *Journal of Economic Perspectives*, 1(1):101–119.
- Heckman, J. J. and Snyder, J. M. (1997). Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *The RAND Journal of Economics*, 28.
- Jenkins, J. A. (2000). Examining the robustness of ideological voting: evidence from the confederate house of representatives. *American Journal of Political Science*, pages 811–822.
- Jenkins, J. A., Monroe, N. W., and Provins, T. (2014). Minority-party influence in the u.s. congress: Floor proposals and minority leverage in the house and senate. *The Rights of the Political Minority in America Conference, Rothermere American Institute, University of Oxford.*, USC, mimeo.
- Kelly, B., Pástor, L., and Veronesi, P. (2016). The price of political uncertainty: Theory and evidence from the option market. *The Journal of Finance*, 71(5):2417–2480.
- Krehbiel, K. (1993). Where’s the party? *British Journal of Political Science*, 23(2):235–266.
- Krehbiel, K. (1999). Paradoxes of parties in congress. *Legislative Studies Quarterly*, pages 31–64.

- Krehbiel, K. (2000). Party discipline and measures of partisanship. *American Journal of Political Science*, pages 212–227.
- Lawrence, E. D., Maltzman, F., and Smith, S. S. (2006). Who wins? party effects in legislative voting. *Legislative Studies Quarterly*, 31(1):33–69.
- Lee, D., Moretti, E., and Butler, M. (2004). Do voters affect or elect policies? evidence from the u.s. house. *Quarterly Journal of Economics*, 119(3):807–860.
- Lee, F. E. (2009). *Beyond ideology: Politics, principles, and partisanship in the US Senate*. University of Chicago Press.
- Levitt, S. D. (1996). How do senators vote? disentangling the role of voter preferences, party affiliation, and senator ideology. *The American Economic Review*, 86(3):425–441.
- Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics*, 3(3):205–228.
- Manski, C. F. (1988). Identification of binary response models. *Journal of the American statistical Association*, 83(403):729–738.
- McCarty, N. (2016-2017). Polarization, congressional dysfunction, and constitutional change symposium. *Indiana Law Review*, 50:223.
- McCarty, N., Poole, K. T., and Rosenthal, H. (2001). The hunt for party discipline in congress. *American Political Science Review*, 95(3):673–687.
- McCarty, N., Poole, K. T., and Rosenthal, H. (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge: MIT Press.
- McConnell, M. and Brownell II, R. E. (2019). *The US Senate and the Commonwealth: Kentucky Lawmakers and the Evolution of Legislative Leadership*. University Press of Kentucky.
- Meinke, S. (2016). *Leadership organizations in the House of Representatives: Party participation and partisan politics*. University of Michigan Press.
- Meinke, S. R. (2008). Who whips? party government and the house extended whip networks. *American Politics Research*, 36(5):639–668.
- Mian, A., Sufi, A., and Trebbi, F. (2010). The political economy of the us mortgage default crisis. *American Economic Review*, 100(5).
- Mian, A., Sufi, A., and Trebbi, F. (2014). Resolving debt overhang: Political constraints in the aftermath of financial crises. *American Economic Journal: Macroeconomics*, 6(2):1–28.
- Minozzi, W. and Volden, C. (2013). Who heeds the call of the party in congress? *The Journal of Politics*, 75(3):787–802.
- Moskowitz, D. J., Rogowski, J., and Jr., J. M. S. (2017). Parsing party polarization. mimeo.

- Neyman, J. and Scott, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrica: Journal of the Econometric Society*, pages 1–32.
- Nokken, T. P. (2000). Dynamics of congressional loyalty: Party defection and roll-call behavior, 1947-97. *Legislative Studies Quarterly*, pages 417–444.
- Pastor, L. and Veronesi, P. (2012). Uncertainty about government policy and stock prices. *The Journal of Finance*, 67(4):1219–1264.
- Piketty, T. and Saez, E. (2003). Income inequality in the united states 1913 - 1998. *Quarterly Journal of Economics*, 118(1):1–39.
- Poole, K. T. (2000). Nonparametric unfolding of binary choice data. *Political Analysis*, 8(3):211–237.
- Poole, K. T. (2007). Changing minds? not in congress! *Public Choice*, 131(3/4):435–451.
- Poole, K. T. and Rosenthal, H. (1984). The polarization of american politics. *Journal of Politics*, 46(4):1061–1079.
- Poole, K. T. and Rosenthal, H. (1997). *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.
- Poole, K. T. and Rosenthal, H. (2001). D-nominate after 10 years: A comparative update to congress: A political-economic history of roll-call voting. *Legislative Studies Quarterly*, pages 5–29.
- Poole, K. T. and Rosenthal, H. L. (2011). *Ideology and congress*, volume 1. Transaction Publishers.
- Richardson, J. and Nordhaus, R. (1995). The national energy act of 1978. *Natural Resources & Environment*, 10(1):62–88.
- Ripley, R. B. (1964). The party whip organizations in the united states house of representatives. *American Political Science Review*, 58(3):561–576.
- Rohde, D. W. (1991). *Parties and Leaders in the Postreform House*. The University of Chicago Press.
- Rosenthal, H. and Voeten, E. (2004). Analyzing roll calls with perfect spatial voting: France 1946–1958. *American Journal of Political Science*, 48(3):620–632.
- Shull, S. A. and Vanderleeuw, J. M. (1987). What do key votes measure? *Legislative Studies Quarterly*, pages 573–582.
- Skidmore, D. (1993). Foreign policy interest groups and presidential power: Jimmy carter and the battle over ratification of the panama canal treaties. *Presidential Studies Quarterly*, pages 477–497.
- Smith III, T. J. and Hogan, M. J. (1987). Public opinion and the panama canal treaties of 1977. *Public Opinion Quarterly*, 51(1):5–30.
- Snyder, J. M. and Groseclose, T. (2000). Estimating party influence in congressional roll-call voting. *American Journal of Political Science*, pages 193–211.

- Souleles, N. S. (2002). Consumer response to the reagan tax cuts. *Journal of Public Economics*, 85(1):99–120.
- Spirling, A. and McLean, I. (2006). The rights and wrongs of roll calls. *Government and Opposition*, 41(4):581–588.
- Theriault, S. M. (2008). *Party Polarization in Congress*. New York: Cambridge University Press.
- Weingast, B. R. (1979). A rational choice perspective on congressional norms. *American Journal of Political Science*, pages 245–262.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

10. TABLES AND FIGURES

FIGURE 1. Timeline

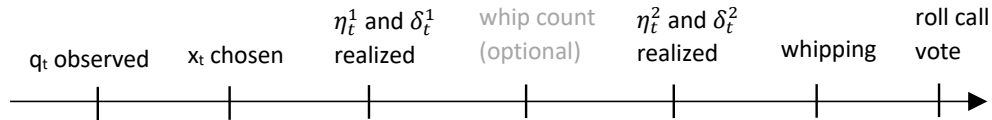
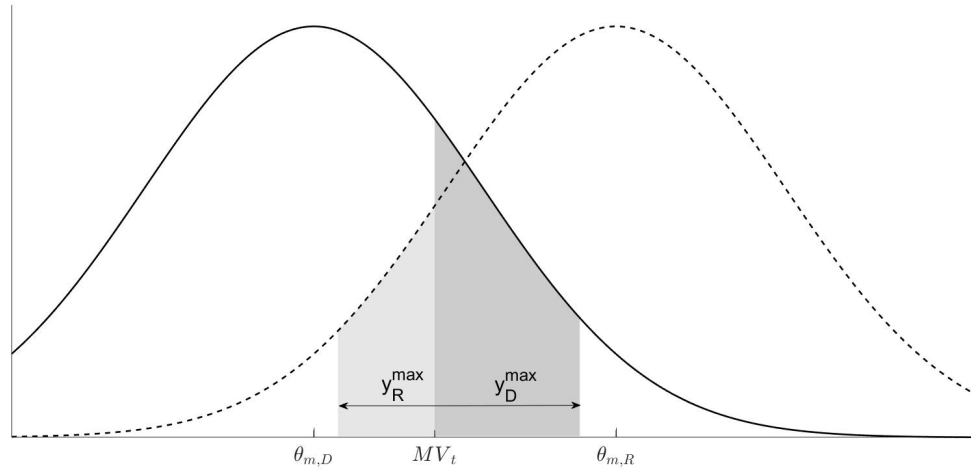
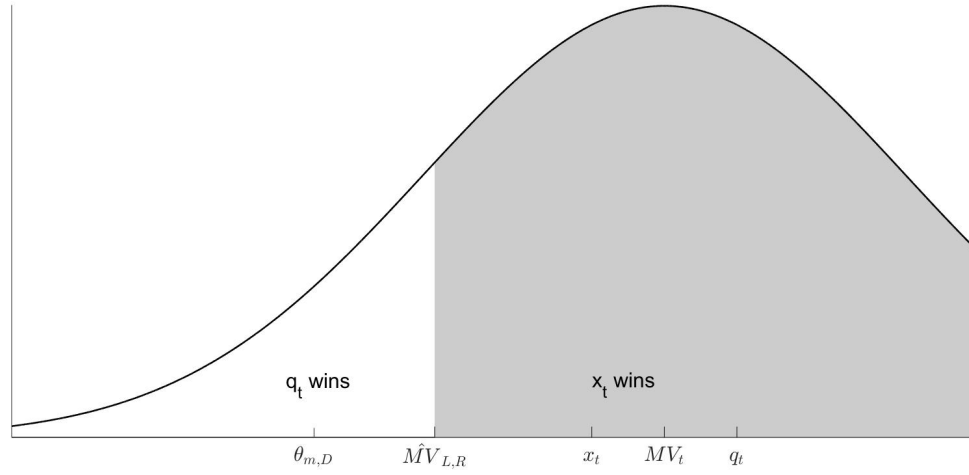


FIGURE 2. Whipping



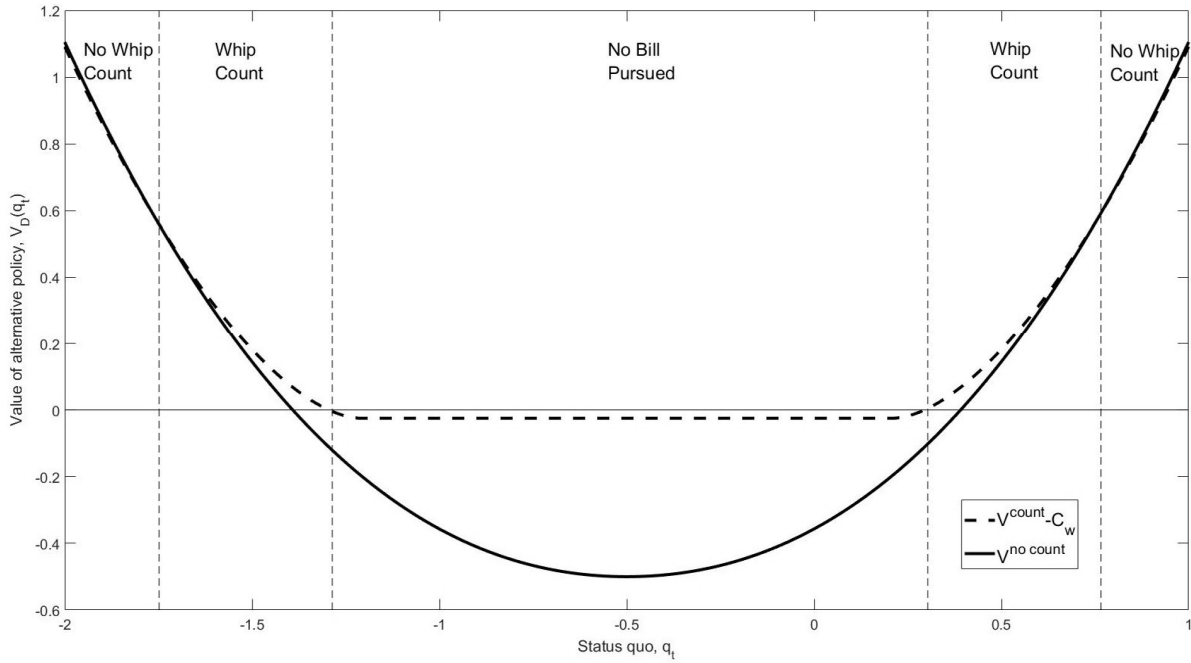
Notes: All Democrats whose realized ideal points, ω_t^i , are within a distance of y_D^{max} , and to the right of the marginal voter, MV_t , are whipped. Similarly, all Republicans within a distance of y_R^{max} , and to the left of the realized marginal voter, $\tilde{M}V_{2,t}$, are whipped.

FIGURE 3. Optimal Policy Alternative



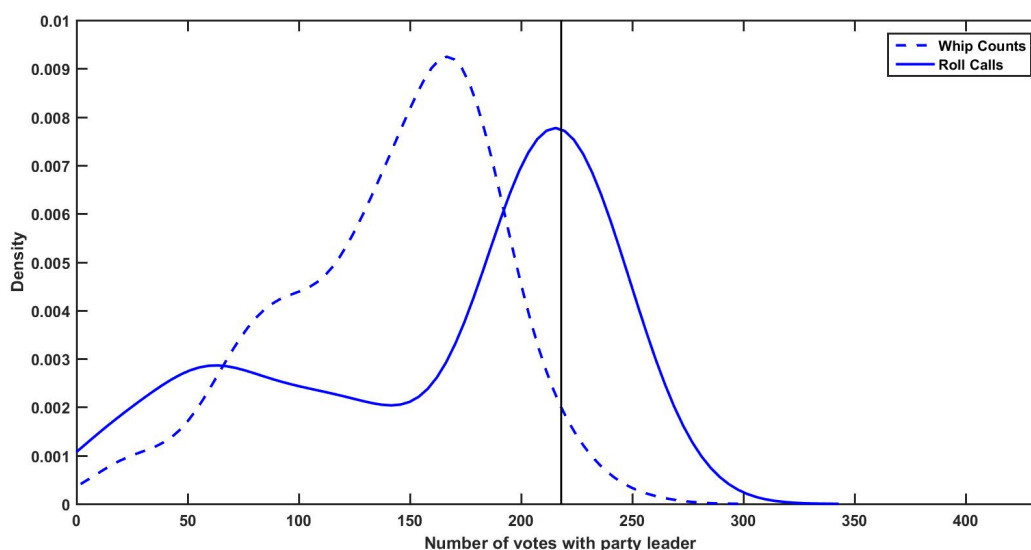
Notes: Optimal policy selection by the Democratic party for a status quo, q_t , right of their ideal point, $\theta_{m,D}$, for a bill that goes directly to roll call. The shaded area is the probability that the policy alternative, x_t , wins. x_t wins if the sum of the aggregate shocks is such that the realized marginal voter lies to the right of $\hat{M}V_{L,R}$, the position of the marginal voter for which votes are equally split between q_t and x_t . A policy alternative chosen closer to the Democratic ideal point is preferred, but is less likely to pass because as it shifts left, the marginal voter, MV_t , also shifts left, reducing the size of the shaded area.

FIGURE 4. Example of Value Functions



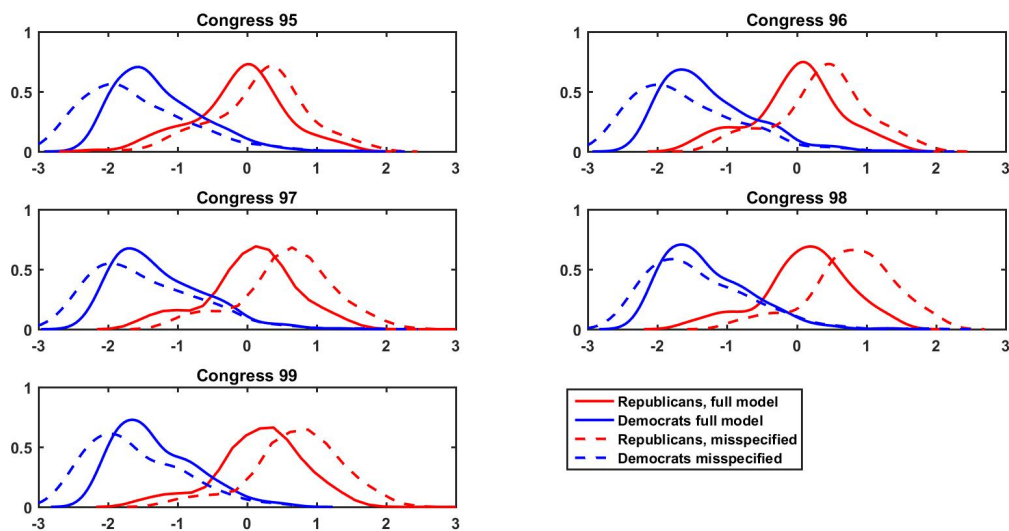
Notes: Value functions of pursuing an alternative policy with and without a whip count. Party D is the proposing party. The value functions are simulated using $\theta_D^m = -0.5$, $\theta_R^m = 0.5$, $\hat{M}V_{R,R} = \hat{M}V_{L,R} = -0.5$, $\sigma_\eta = 1$, $C_b = 0.5$, $C_w = 0.025$, and quadratic utility.

FIGURE 5. Majority Party Votes with Leadership



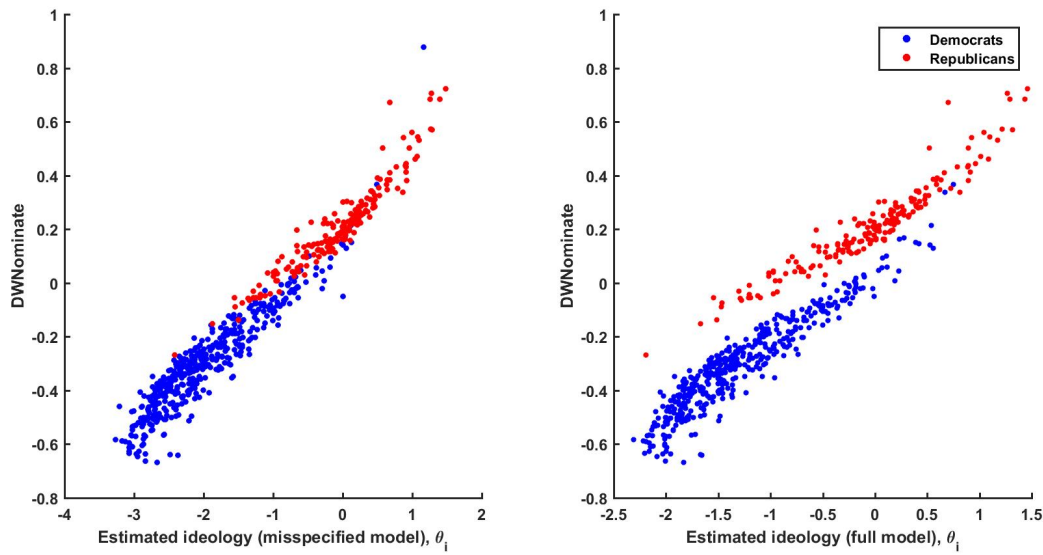
Notes: Kernel densities of the number of Democratic votes with their party leadership at the whip count and roll call stages. Includes only bills with both whip counts and roll calls. The vertical line at 218 indicates the majority needed to pass a bill in the House of Representatives.

FIGURE 6. Estimates of Ideological Points



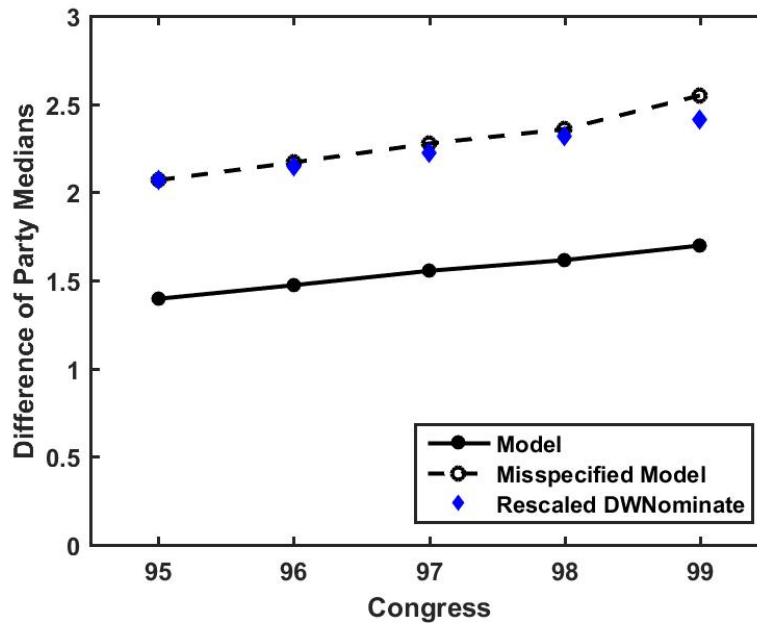
Notes: Each graph (one per Congress) provides the kernel density of the estimated ideological points for each party (solid lines). For comparison (dashed lines), the graphs show the kernel density estimates under a misspecified model that assumes no party discipline.

FIGURE 7. Estimated Ideologies Compared to DW-Nominate Estimates



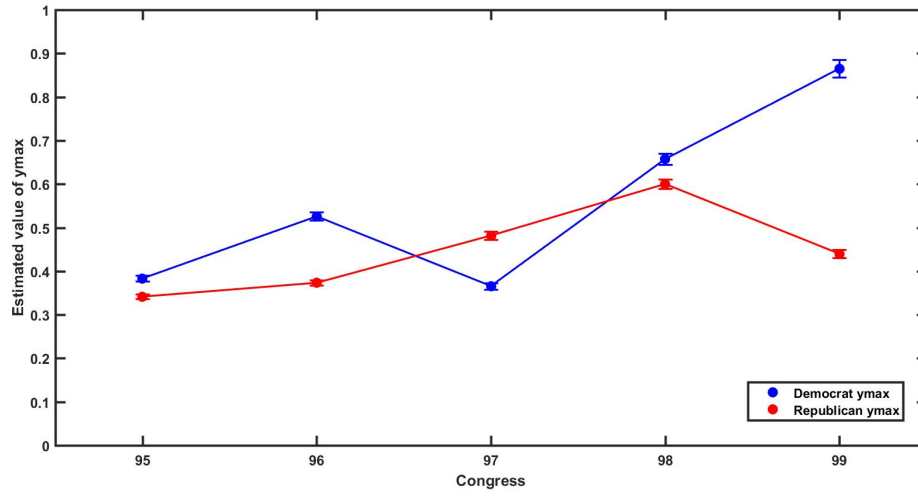
Notes: Correlations between our estimates of ideologies to those of DW-Nominate. In the left panel, the estimates are for a misspecified model with no party discipline (correlation = 0.976). In the right panel, the estimates are for the full model (correlation = 0.957).

FIGURE 8. Trends in Polarization Across Model Estimates



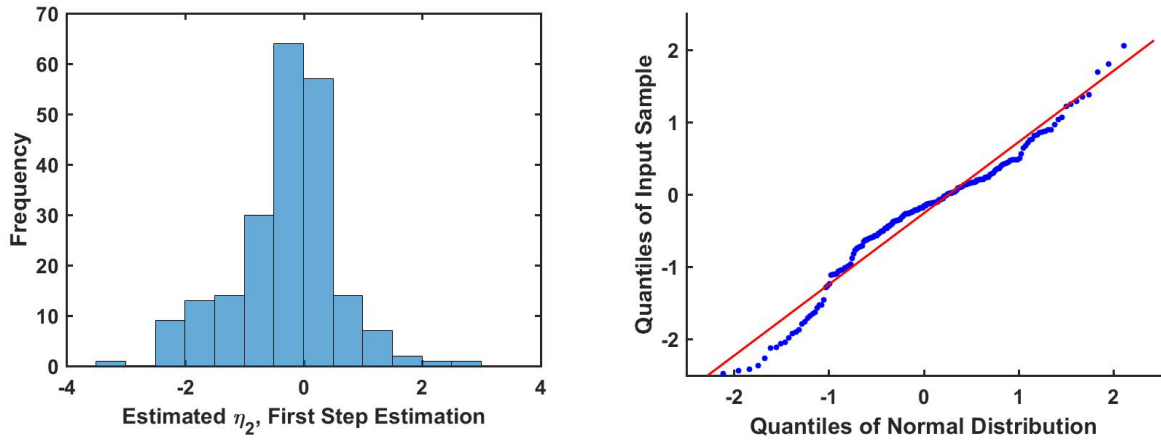
Notes: We present the time trends of ideological polarization across three models: our main model (black line), a misspecified model that assume no party discipline (dashed black), and DW-Nominate (rescaled to have the same mean and support as the misspecified model).

FIGURE 9. Estimates of Party Discipline



Notes: Time series of the estimates of the party discipline (whipping) parameters for each party. Each parameter is in units of the single-dimension ideology.

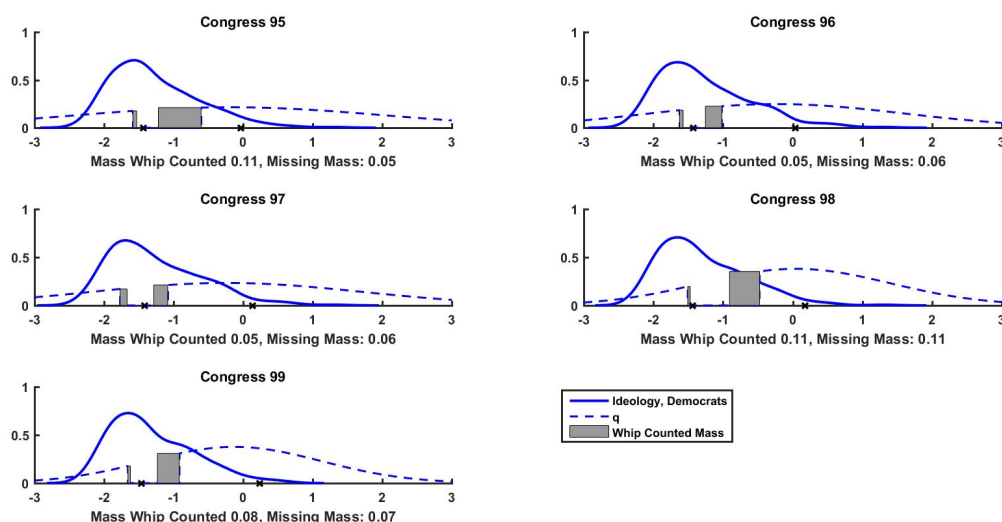
FIGURE 10. Estimated Aggregate Shocks



(A) Histogram of Estimated $\hat{\eta}_{2,t}$ (B) QQ-Plot (Empirical vs. Theoretical Distributions)

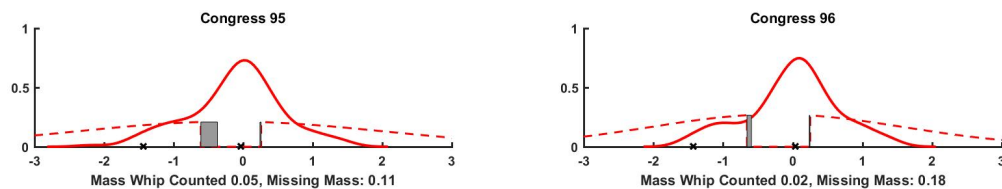
Notes: We present two subgraphs for the estimated aggregate shocks between whip count and roll call. The first is a histogram. The second is a QQ-plot which compares the empirical distribution (dotted line) to the one we theorize (a Normal distribution with mean 0 and standard deviation estimated at $\hat{\sigma}_{\eta} = 0.859$; solid line).

FIGURE 11. Pursued Status Quo Policies: Democrats



Notes: Estimated status quo distributions by Congress (dashed lines). Status quo policies that are pursued by the Democrats with whip counts are shown in gray. The remaining gap in the distribution is the ‘missing mass’ of status quo policies that are not pursued by the Democrats at all. For reference, the ideologies of Democrats are shown as solid lines.

FIGURE 12. Pursued Status Quo Policies: Republicans



Notes: Estimated status quo distributions by Congress (dashed lines). Status quo policies that are pursued by the Republicans with whip counts are shown in grey. The remaining gap in the distribution is the ‘missing mass’ of status quo policies that are not pursued by the Republicans at all. For reference, the ideologies of Republicans are shown as solid lines.

TABLE 1. First Step Estimates

Parameter	Congress				
	95	96	97	98	99
Party Discipline y^{max} , Democrats	0.383 (0.003)	0.526 (0.005)	0.366 (0.004)	0.658 (0.007)	0.865 (0.010)
Party Discipline y^{max} , Republicans	0.342 (0.003)	0.373 (0.003)	0.482 (0.005)	0.600 (0.006)	0.440 (0.005)
Standard Deviation of Aggregate Shock σ_η			0.859 (0.953)		
Party Median - Democrats, θ_D^m	-1.431 (0.069)	-1.431 (0.052)	-1.420 (0.055)	-1.435 (0.053)	-1.462 (0.102)
Party Median - Republicans, θ_R^m	-0.036 (0.060)	0.042 (0.142)	0.134 (0.144)	0.181 (0.049)	0.236 (0.061)
N: 711					
T: 315 Whip Counted bills, 5424 Roll Called bills					

Notes: Estimates of the first step parameters. Asymptotic standard errors are in parentheses. Non time-varying parameters are centered in the table and apply to all five Congresses.

TABLE 2. Decomposition of Polarization

	95	96	Congress 97	98	99
Implications of Table 1 for Polarization					
A: Polarization due to ideology ($\theta_R^m - \theta_D^m$)	1.395	1.473	1.554	1.615	1.698
B: Polarization due to whipping ($y_R^{max} + y_D^{max}$)	0.725	0.899	0.848	1.258	1.305
C: Share of Perceived Ideological Polarization due to whipping (B/(A+B))	0.342	0.379	0.353	0.438	0.435
D: Share of Change in Perceived Ideological Polarization Explained by Change in Party Discipline		0.198	-0.058	0.464	0.054

Notes: Decomposition of perceived polarization (polarization in ideologies from a misspecified model that ignores party discipline) into that due to ideological polarization and that due to party discipline, by Congress. The last row reports the changes in perceived polarization from Congress to Congress that are explained by changes in party discipline.

TABLE 3. Model Fit

Model	Variable	% Correctly Predicted Votes ("Yes/No")
Full Model	Roll Call Votes	0.855
	Whip Count Votes	0.628

Notes: Fraction of correctly predicted votes at the whip count and roll call stages.

TABLE 4. Second Step Estimates

	95	96	Congress 97	98	99
Probability Democrat is Proposer, γ			0.427 (0.020)		
Status Quo Distribution (Mean), μ_q	-0.188 (0.243)	-0.227 (0.298)	-0.237 (0.250)	0.045 (0.218)	-0.125 (0.189)
Status Quo Distribution (Standard Deviation), σ_q	2.222 (0.456)	1.816 (0.532)	1.937 (0.322)	1.354 (0.217)	1.252 (0.151)

Notes: Estimates of the second step parameters. Asymptotic standard errors, accounting for estimation error from the first step, in parentheses. Standard errors are computed by drawing 100 samples from the asymptotic distribution of first step estimates, recomputing the second step estimates, and using the Law of Total Variance.

TABLE 5. Distance from Marginal Voter to Party Median

	Whip count	Roll call	p-value
Democrats	0.479	1.234	(0.000)
Republicans	0.910	1.163	(0.010)

Notes: Average absolute distance from marginal voter to party median across all whip counts (left column) and bills that go directly to roll call (middle column). The rightmost column provides unpaired t-tests of the means.

TABLE 6. Missing Mass

	95	96	Congress		
			97	98	99
Democrats					
<i>Main Model</i>	0.048	0.057	0.064	0.110	0.071
<i>Counterfactual: No Whipping</i>	0.049	0.058	0.064	0.115	0.076
<i>Counterfactual: Polarized Ideologies</i>	0.046	0.052	0.061	0.088	0.048
Republicans					
<i>Main Model</i>	0.109	0.180	-	-	-
<i>Counterfactual: No Whipping</i>	0.106	0.170	-	-	-
<i>Counterfactual: Polarized Ideologies</i>	0.113	0.195	-	-	-

Notes: Mass of status quo policies ('missing mass') that are not pursued by the party at all. For the counterfactuals, C_b and C_w are determined from the second step estimates and held fixed, allowing new thresholds to be calculated.

TABLE 7. Counterfactual: Voting Outcomes on Salient Bills

Bill	Yes Votes (Data)	Yes Votes (Model Predicted)	Yes Votes (Counterfactual, No Whipping)
Security, International Relations and Other Policies			
Aid to Turkey/Lifting of Arms Embargo (H.R. 12514, Congress 95)	212	193	147
Foreign Intelligence Surveillance Act of 1978 (H.R. 7308, Congress 95)	261	283	280
National Energy Act, 1978 (H.R. 8444, Congress 95)	247	271	258
Panama Canal Treaty, 1979 (H.R. 111, Congress 96)	224	243	180
Contra Aid, 1984 (H.R. 5399, Congress 98)	294	279	343
Economic Policies			
Increase of Temporary Debt Limit, (H.R.9290, Congress 95)	221	242	185
Increase of Temporary Debt Limit, (H.R.13385, Congress 95)	210	235	201
Increase of Temporary Debt Limit, (H.R.2534, Congress 96)	220	239	208
Depository Institutions Deregulation and Monetary Control Act of 1980, (H.R. 4986, Congress 96)	369	404	391
Increase of Public Debt Limit, Make it part of Budget Process (H.R. 5369, Congress 96)	225	244	217
Economic Recovery Tax Act of 1981 (H.R. 4242, Congress 97)	284	329	276
Garn-St. Germain Depository Institutions Act of 1982 (H.R.6267, Congress 97)	263	279	327
Social Security Amendments of 1983 (H.R.1900, Congress 98)	282	299	230
Tax Reform Act of 1984 (H.R. 4170, Congress 98)	319	370	292

Notes: Counterfactual vote outcomes on certain key bills absent party discipline (whipping). The policies are assumed fixed.

TABLE 8. Counterfactual: Agenda Setting

	95	96	Congress 97	98	99
<i>Panel A: Average Change in the Probability of Bill Approval</i>					
Democrats					
Baseline Probability (Main Model)	0.357	0.467	0.421	0.431	0.544
Main Model - No Whipping	0.032	0.060	0.009	0.054	0.011
Main Model - Polarized Ideology	-0.005	-0.011	0.010	-0.013	-0.024
Republicans					
Baseline Probability (Main Model)	0.240	0.220	-	-	-
Main Model - No Whipping	-0.034	-0.042	-	-	-
Main Model - Polarized Ideology	0.028	0.032	-	-	-
<i>Panel B: Average Change in Pursued Policies, x_t</i>					
Democrats					
Main Model - No Whipping	-0.011	-0.018	-0.003	-0.024	-0.042
Main Model - Polarized Ideology	0.085	0.161	0.107	0.163	0.285
Republicans					
Main Model - No Whipping	-0.011	-0.016	-	-	-
Main Model - Polarized Ideology	-0.057	-0.048	-	-	-

Notes: Estimated and counterfactual probabilities of bill approval and average distance between the proposed policy alternative and the status quo, for status quo policies that lie between the party medians.

APPENDIX A. (FOR PUBLICATION) PROOFS

Proof of Lemma 1:

Consider first $k_t > k'_t$. Given the increasing cost of exerting influence, a whip exerts the minimum amount of influence necessary to ensure a vote for k_t , provided this amount is less than or equal to y_p^{max} . The minimum amount of influence is such that the member is indifferent, $u(k_t, \omega_t^i + y_t^i) = u(k'_t, \omega_t^i + y_t^i)$ or $|\omega_t^i + y_t^i - k_t| = |\omega_t^i + y_t^i - k'_t|$. This equality is satisfied if and only if $\omega_t^i + y_t^i = MV_t = \frac{k_t + k'_t}{2}$. If $\omega_t^i \geq MV_t$, the required influence is weakly negative (absent influence, the member votes for k_t) and so no influence is exerted. If $\omega_t^i < MV_t$, a positive amount of influence, $y_t^i = MV_t - \omega_t^i > 0$ is required which increases linearly in $MV_t - \omega_t^i$. Therefore, a member is whipped if and only if their ideology is such that $MV_t - y_p^{max} \leq \omega_t^i < MV_t$. For $k_t < k'_t$, the argument is reversed: only members for which $MV_t < \omega_t^i \leq MV_t + y_p^{max}$ are whipped. \square

Proof of Lemma 2:

Consider the mass, $f(\theta)$, of members at some θ , each of whom has an independent signal of $\hat{\eta}_{1,t}$ due to their independent ideological shocks. The average number of Yes reports from the N members at θ is given by $\lim_{N \rightarrow \infty} \frac{f(\theta)}{N} \sum_{i=1}^N I(u(x_t, \theta + \delta_{1,t}^i + \hat{\eta}_{1,t}) \geq u(q_t, \theta + \delta_{1,t}^i + \hat{\eta}_{1,t}))$ where $I(\cdot)$ represents the indicator function. By the law of large numbers, as $N \rightarrow \infty$, this average converges to:

$$\begin{aligned} f(\theta)E[I(u(x_t, \theta + \delta_t^1 + \hat{\eta}_{1,t}) \geq u(q_t, \theta + \delta_t^1 + \hat{\eta}_{1,t}))] &= f(\theta)Pr(u(x_t, \theta + \delta_t^1 + \hat{\eta}_{1,t}) \geq u(q_t, \theta + \delta_t^1 + \hat{\eta}_{1,t})) \\ &= f(\theta)Pr(\theta + \delta_t^1 + \hat{\eta}_{1,t} \geq MV_t) \\ &= f(\theta)(1 - G(MV_t - \theta - \hat{\eta}_{1,t})). \end{aligned}$$

Therefore, after observing the number of Yes reports for a given θ , $\hat{\eta}_{1,t}$ is known with probability one. \square

Proof of Lemma 3:

Consider $x_t > q_t$. Let $G_{1+2}(\cdot)$ denote the cdf of $\delta_{1,t}^i + \delta_{2,t}^i$ (with corresponding pdf, $g_{1+2}(\cdot)$). For a given $\tilde{M}V_{2,t}$, the number of votes for x_t from a given party's members is known with probability one due to independent idiosyncratic shocks and a continuum of members. To see this fact, consider the continuum of party p 's members located at each θ , each with independent shocks, $\delta_{1,t}^i$ and $\delta_{2,t}^i$. With N voters at θ , the average number of votes from these members is given by $\lim_{N \rightarrow \infty} \frac{f(\theta)}{N} \sum_{i=1}^N I(\theta + \delta_{1,t}^i + \delta_{2,t}^i \geq \tilde{M}V_{2,t} \pm y_p^{max})$, where the sign with which y_p^{max} enters depends upon the direction that party p whips. By the law of large numbers, as $N \rightarrow \infty$, this average converges to:

$$\begin{aligned} f(\theta)E[I(\theta + \delta_t^1 + \delta_t^2 \geq \tilde{M}V_{2,t} \pm y_p^{max})] &= f(\theta)Pr(\theta + \delta_t^1 + \delta_t^2 \geq \tilde{M}V_{2,t} \pm y_p^{max}) \\ &= f(\theta)(1 - G_{1+2}(\tilde{M}V_{2,t} \pm y_p^{max} - \theta)). \end{aligned}$$

Using this fact, the number of votes for x_t from party D 's members is given by $Y_D(\tilde{M}V_{2,t}) = N_D \left[\int_{-\infty}^{\infty} \left(1 - G_{1+2}(\tilde{M}V_{2,t} - \theta \pm y_D^{max}) \right) f_D(\theta) d\theta \right]$. The corresponding expression for party R is $Y_R(\tilde{M}V_{2,t}) = N_R \left[\int_{-\infty}^{\infty} \left(1 - G_{1+2}(\tilde{M}V_{2,t} - \theta \pm y_R^{max}) \right) f_R(\theta) d\theta \right]$. The total number of votes for x_t is then given by $Y(\tilde{M}V_{2,t}) \equiv Y_D(\tilde{M}V_{2,t}) + Y_R(\tilde{M}V_{2,t})$.

$Y(\tilde{M}V_{2,t})$ is strictly decreasing in x_t . To see this, consider the votes from party D 's members, $Y_D(\tilde{M}V_{2,t})$:

$$\begin{aligned} \frac{\partial Y_D(\tilde{M}V_{2,t})}{\partial x_t} &= \frac{1}{2} \frac{\partial}{\partial \tilde{M}V_{2,t}} N_D \left[\int_{-\infty}^{\infty} \left(1 - G_{1+2}(\tilde{M}V_{2,t} - \theta \pm y_D^{max}) \right) f_D(\theta) d\theta \right] \\ (A.1) \quad &= -\frac{N_D}{2} \int_{-\infty}^{\infty} g_{1+2}(\tilde{M}V_{2,t} - \theta \pm y_D^{max}) f_D(\theta) d\theta \end{aligned}$$

(A.1) is strictly less than zero given that that ideological shocks are unbounded, independent of the (finite) amount or direction of whipping. The same is true of the derivative of $Y_R(\tilde{M}V_{2,t})$, ensuring $Y(\tilde{M}V_{2,t})$ strictly decreases in x_t for $x_t > q_t$. For $x_t < q_t$, we have $Y_D(\tilde{M}V_{2,t}) = N_D \left[\int_{-\infty}^{\infty} G_{1+2}(\tilde{M}V_{2,t} - \theta \pm y_D^{max}) f_D(\theta) d\theta \right]$ and $Y_R(\tilde{M}V_{2,t}) = N_R \left[\int_{-\infty}^{\infty} G_{1+2}(\tilde{M}V_{2,t} - \theta \pm y_R^{max}) f_R(\theta) d\theta \right]$ so that $Y(\tilde{M}V_{2,t})$ increases in x_t . Since for $q_t < \theta_p^m$ we must have $x_t > q_t$ and for $q_t > \theta_p^m$ we must have $x_t < q_t$, we see that the number of votes for x_t strictly decreases the closer it gets to the proposing party's ideal point. \square

Proof of Proposition 1:

For $q_t = \theta_D^m$, clearly $x_t^{count} = x_t^{no\ count} = \theta_D^m$ are the unique optimal alternative policies because party D can do no better than its ideal point.

In the case of no whip count, and $q_t < \theta_D^m$ so that $x_t > q_t$, we can rewrite party D 's expected utility as

$$EU_D^{no\ count}(q_t, x_t) = \left(1 - \Phi \left(\frac{MV_t - \hat{M}V_{R,R}}{\sigma} \right) \right) (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) + u(q_t, \theta_D^m) - C_b$$

The derivative with respect to x_t is given by

$$\left(1 - \Phi \left(\frac{MV_t - \hat{M}V_{R,R}}{\sigma} \right) \right) u_x(x_t, \theta_D^m) - \frac{1}{2\sigma} \phi \left(\frac{MV_t - \hat{M}V_{R,R}}{\sigma} \right) (u(x_t, \theta_D^m) - u(q_t, \theta_D^m))$$

where $\phi(\cdot)$ denotes the pdf of the standard Normal distribution. At $x_t = q_t$, the derivative is strictly positive given $q_t < \theta_D^m$ and the fact that $\hat{M}V_{R,R}$ is finite. At $x_t = \theta_D^m$, it is strictly negative given $u(q_t, \theta_D^m) < 0$. Together these facts ensure an interior solution, which we now show is unique. Any interior solution must satisfy the first-order condition,

$$(A.2) \quad \begin{aligned} & \left(1 - \Phi \left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma} \right) \right) u_x(x_t^{no\ count}, \theta_D^m) \\ & - \frac{1}{2\sigma} \phi \left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma} \right) (u(x_t^{no\ count}, \theta_D^m) - u(q_t, \theta_D^m)) = 0 \end{aligned}$$

Defining $z_t^{no\ count} \equiv \frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma}$, we can re-write the first-order condition as:

$$(A.3) \quad \frac{1 - \Phi(z_t^{no\ count})}{\phi(z_t^{no\ count})} = \frac{1}{2\sigma} \frac{u(x_t^{no\ count}, \theta_D^m) - u(q_t, \theta_D^m)}{u_x(x_t^{no\ count}, \theta_D^m)}$$

The left-hand side of (A.3) is the inverse hazard rate of a standard Normal distribution and so is strictly decreasing in $z_t^{no\ count}$ (and therefore $x_t^{no\ count}$ since $x_t^{no\ count}$ strictly increases in $z_t^{no\ count}$). The sign of the derivative of the right-hand side with respect to $x_t^{no\ count}$ is given by $u_x(x_t^{no\ count}, \theta_D^m)^2 - u_{xx}(x_t^{no\ count}, \theta_D^m) (u(x_t^{no\ count}, \theta_D^m) - u(q_t, \theta_D^m))$ which is strictly positive because $u_{xx}(x_t^{no\ count}, \theta_D^m) < 0$ and $u(x_t^{no\ count}, \theta_D^m) > u(q_t, \theta_D^m)$. Thus, the right-hand side is strictly increasing in $x_t^{no\ count}$. Together, these facts guarantee a unique solution, $x_t^{no\ count} \in (q_t, \theta_D^m)$.⁵⁹

In the case of a whip count and $q_t < \theta_D^m$, we can rewrite the party's expected utility:

$$\begin{aligned} & EU_D^{count}(q_t, x_t) \\ &= Pr(\eta_{1,t} \geq \underline{\eta}_{1,t}) \left(Pr(x_t \text{ wins} | \eta_{1,t} \geq \underline{\eta}_{1,t}) (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) + u(q_t, \theta_D^m) - C_b \right) \\ & \quad + Pr(\eta_{1,t} < \underline{\eta}_{1,t}) u(q_t, \theta_D^m) \\ &= Pr(\eta_{1,t} \geq \underline{\eta}_{1,t}, x_t \text{ wins}) (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) - Pr(\eta_{1,t} \geq \underline{\eta}_{1,t}) C_b + u(q_t, \theta_D^m) \\ &= \int_{\underline{\eta}_{1,t}}^{\infty} \left(1 - \Phi \left(\frac{MV_t - \hat{M}V_{R,R} - \eta}{\sigma_\eta} \right) \right) \frac{1}{\sigma_\eta} \phi \left(\frac{\eta}{\sigma_\eta} \right) d\eta (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) \\ & \quad - \left(1 - \Phi \left(\frac{\underline{\eta}_{1,t}}{\sigma_\eta} \right) \right) C_b + u(q_t, \theta_D^m) \end{aligned}$$

Taking the derivative with respect to x_t yields:⁶⁰

⁵⁹The second-order condition at $x_t^{no\ count}$ is also easily checked, but must be satisfied given that marginal expected utility is increasing at $x_t = q_t$, decreasing at $x_t = \theta_D^m$ and the solution is unique.

⁶⁰The necessary conditions for applying the Leibniz Integral Rule with an infinite bound are satisfied. Specifically, the integrand and its partial derivative with respect to x_t are both continuous functions of x_t and η , and it is possible to find integrable functions of η that bound the integrand and its partial derivative with respect to x_t .

$$\begin{aligned}
\frac{dEU_D^{count}(q_t, x_t)}{dx_t} &= -\frac{d\underline{\eta}_{1,t}}{dx_t} \frac{1}{\sigma_\eta} \phi\left(\frac{\underline{\eta}_{1,t}}{\sigma_\eta}\right) \left(1 - \Phi\left(\frac{MV_t - \hat{M}V_{R,R} - \underline{\eta}_{1,t}}{\sigma_\eta}\right)\right) (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) \\
&\quad - \frac{1}{2\sigma_\eta^2} \int_{\underline{\eta}_{1,t}}^\infty \phi\left(\frac{MV_t - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) \\
&\quad + \frac{1}{\sigma_\eta} u_x(x_t, \theta_D^m) \int_{\underline{\eta}_{1,t}}^\infty \left(1 - \Phi\left(\frac{MV_t - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
&\quad + \frac{1}{\sigma_\eta} \frac{d\underline{\eta}_{1,t}}{dx_t} \phi\left(\frac{\underline{\eta}_{1,t}}{\sigma_\eta}\right) C_b \\
&= \frac{1}{\sigma_\eta} u_x(x_t, \theta_D^m) \int_{\underline{\eta}_{1,t}}^\infty \left(1 - \Phi\left(\frac{MV_t - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
&\quad - \frac{1}{2\sigma_\eta^2} \int_{\underline{\eta}_{1,t}}^\infty \phi\left(\frac{MV_t - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta (u(x_t, \theta_D^m) - u(q_t, \theta_D^m))
\end{aligned}
\tag{A.4}$$

where the second equality uses the fact that $\underline{\eta}_{1,t}$ satisfies

$$\left(1 - \Phi\left(\frac{MV_t - \hat{M}V_{R,R} - \underline{\eta}_{1,t}}{\sigma_\eta}\right)\right) (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) = C_b
\tag{A.5}$$

Consider the limit as $C_b \rightarrow 0$. From (A.5), we can see that, provided x_t is bounded away from q_t so that $u(x_t, \theta_D^m) - u(q_t, \theta_D^m) > 0$ (which we subsequently confirm), we must have $\underline{\eta}_{1,t} \rightarrow -\infty$ as $C_b \rightarrow 0$. But, as $\underline{\eta}_{1,t} \rightarrow -\infty$, the party always continues to pursue the bill after the first aggregate shock. In this case, the optimal alternative policy is identical to the case of no whip count. Formally,

$$\begin{aligned}
\lim_{\underline{\eta}_{1,t} \rightarrow -\infty} \frac{dEU_D^{count}(q_t, x_t)}{dx_t} &= \frac{1}{\sigma_\eta} u_x(x_t, \theta_D^m) \int_{-\infty}^\infty \left(1 - \Phi\left(\frac{MV_t - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
&\quad - \frac{1}{2\sigma_\eta^2} \int_{-\infty}^\infty \phi\left(\frac{MV_t - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) \\
&= u_x(x_t, \theta_D^m) \left(1 - \Phi\left(\frac{MV_t - \hat{M}V_{R,R}}{\sigma}\right)\right) \\
&\quad - \frac{1}{2\sigma} \phi\left(\frac{MV_t - \hat{M}V_{R,R}}{\sigma}\right) (u(x_t, \theta_D^m) - u(q_t, \theta_D^m))
\end{aligned}
\tag{A.6}$$

where the equality follows from the fact that the convolution of two standard Normal distributions is a Normal distribution with the sum of the variances, and using $\sigma^2 = 2\sigma_\eta^2$. Comparing (A.6) with (A.2), we can see immediately that, in the limit, the first-order condition for the whip and no whip cases are identical, and it therefore follows that x_t^{count} is unique and interior as in the no whip case. This fact ensures that $u(x_t, \theta_D^m) - u(q_t, \theta_D^m) > 0$ in the limit, confirming that we must have $\underline{\eta}_{1,t} \rightarrow -\infty$ as $C_b \rightarrow 0$.

We now show that x_t^{count} is unique and interior for strictly positive C_b . From (A.4), we see that $\frac{dEU_D^{count}(q_t, x_t)}{dx_t}$ is strictly positive at $x_t = q_t$ and strictly negative at $x_t = \theta_D^m$, ensuring an interior optimum, x_t^{count} which must satisfy the first-order condition⁶¹

$$(A.7) \quad \frac{\int_{\underline{\eta}_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta}{\frac{1}{2\sigma_\eta} \int_{\underline{\eta}_{1,t}}^{\infty} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta} = \frac{(u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m))}{u_x(x_t^{count}, \theta_D^m)}$$

As in the case of no whip count, the right-hand side of (A.7) strictly increases in x_t^{count} . It remains to show that, in the limit as $C_b \rightarrow 0$, the left-hand side of (A.7) strictly decreases in x_t^{count} , which, by continuity of the left-hand side in C_b , ensures there exists a strictly positive value of C_b , $\hat{C}_b > 0$, such that for all $C_b < \hat{C}_b$, the left-hand side continues to strictly decrease. It then follows that x_t^{count} is unique for all $C_b < \hat{C}_b$. The sign of the derivative of the left-hand side of (A.7) with respect to x_t^{count} , is determined by⁶²

$$(A.8) \quad \begin{aligned} & -\frac{d\underline{\eta}_{1,t}}{dx_t^{count}} \phi\left(\frac{\underline{\eta}_{1,t}}{\sigma_\eta}\right) \left(1 - \Phi\left(\frac{MV_t - \hat{M}V_{R,R} - \underline{\eta}_{1,t}}{\sigma_\eta}\right)\right) \frac{1}{2\sigma_\eta} \int_{\underline{\eta}_{1,t}}^{\infty} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\ & + \frac{d\underline{\eta}_{1,t}}{dx_t^{count}} \frac{1}{2\sigma_\eta} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \underline{\eta}_{1,t}}{\sigma_\eta}\right) \phi\left(\frac{\underline{\eta}_{1,t}}{\sigma_\eta}\right) \int_{\underline{\eta}_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\ & - \left(\frac{1}{2\sigma_\eta} \int_{\underline{\eta}_{1,t}}^{\infty} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta\right)^2 \\ & - \frac{1}{4\sigma_\eta} \int_{\underline{\eta}_{1,t}}^{\infty} \phi'\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \int_{\underline{\eta}_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \end{aligned}$$

By the implicit function theorem, $\frac{d\underline{\eta}_{1,t}}{dx_t}$ must satisfy (from (A.5))

$$\begin{aligned} & -\phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \underline{\eta}_{1,t}}{\sigma_\eta}\right) \frac{1}{\sigma_\eta} \left(\frac{1}{2} - \frac{d\underline{\eta}_{1,t}}{dx_t^{count}}\right) (u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)) \\ & + \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \underline{\eta}_{1,t}}{\sigma_\eta}\right)\right) u_x(x_t^{count}, \theta_D^m) = 0 \end{aligned}$$

or

⁶¹These statements require $\underline{\eta}_{1,t} < \infty$, which, by continuity, is true for C_b sufficiently small given that $\underline{\eta}_{1,t} \rightarrow -\infty$ as $C_b \rightarrow 0$.

⁶²Again, the necessary conditions for applying the Leibniz Integral Rule with an infinite bound are satisfied.

$$(A.9) \quad \frac{d\eta_{1,t}}{dx_t^{count}} = \frac{1}{2} - \frac{\sigma_\eta \left(1 - \Phi \left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta_{1,t}}{\sigma_\eta} \right) \right) u_x(x_t^{count}, \theta_D^m)}{\phi \left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta_{1,t}}{\sigma_\eta} \right) (u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m))}$$

In the limit as $C_b \rightarrow 0$, $\eta_{1,t} \rightarrow -\infty$, in which case the second term of (A.9) approaches zero because x_t^{count} is bounded away from q_t and θ_D^m , and the inverse hazard rate of a standard Normal random variable approaches zero as its argument approaches infinity.⁶³ The limit of (A.8) as $C_b \rightarrow 0$ is then determined by the limit of its second two terms because the first two terms approach zero. Defining $z_t^{count} \equiv \frac{MV_t^{count} - \hat{M}V_{R,R}}{\sigma}$, this limit is given by

$$\begin{aligned} & \lim_{\eta_{1,t} \rightarrow -\infty} - \left(\frac{1}{2\sigma_\eta} \int_{\eta_{1,t}}^{\infty} \phi \left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta} \right) \phi \left(\frac{\eta}{\sigma_\eta} \right) d\eta \right)^2 \\ & - \frac{1}{4\sigma_\eta} \int_{\eta_{1,t}}^{\infty} \phi' \left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta} \right) \phi \left(\frac{\eta}{\sigma_\eta} \right) d\eta \int_{\eta_{1,t}}^{\infty} \left(1 - \Phi \left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta} \right) \right) \phi \left(\frac{\eta}{\sigma_\eta} \right) d\eta \\ & = - \left(\frac{1}{2\sigma_\eta} \int_{-\infty}^{\infty} \phi \left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta} \right) \phi \left(\frac{\eta}{\sigma_\eta} \right) d\eta \right)^2 \\ & - \frac{1}{4\sigma_\eta} \int_{-\infty}^{\infty} \phi' \left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta} \right) \phi \left(\frac{\eta}{\sigma_\eta} \right) d\eta \int_{-\infty}^{\infty} \left(1 - \Phi \left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta} \right) \right) \phi \left(\frac{\eta}{\sigma_\eta} \right) d\eta \\ & = - \left(\frac{1}{2\sigma} \phi \left(\frac{MV_t^{count} - \hat{M}V_{R,R}}{\sigma} \right) \right)^2 - \frac{1}{4\sigma^2} \phi' \left(\frac{MV_t^{count} - \hat{M}V_{R,R}}{\sigma} \right) \left(1 - \Phi \left(\frac{MV_t^{count} - \hat{M}V_{R,R}}{\sigma} \right) \right) \\ & = - \left(\frac{1}{2\sigma} \phi(z_t^{count}) \right)^2 - \frac{1}{4\sigma^2} \phi'(z_t^{count}) (1 - \Phi(z_t^{count})) \\ & = - \left(\frac{1}{2\sigma} \phi(z_t^{count}) \right)^2 + \frac{1}{4\sigma^2} z_t^{count} \phi(z_t^{count}) (1 - \Phi(z_t^{count})) \\ & < - \left(\frac{1}{2\sigma} \phi(z_t^{count}) \right)^2 + \frac{1}{4\sigma^2} \phi(z_t^{count})^2 \\ & = 0 \end{aligned}$$

where the second equality uses properties of the convolution of Normal distributions, and the inequality follows from the fact that, for a standard Normal random variable, $x(1 - \Phi(x)) < \phi(x)$.

For $q_t > \theta_D^m$ so that $x_t < q_t$, we assume party R whips against the bill (supports q_t). In case of no whip count, we can write party D 's expected utility as

$$EU_D^{no\ count}(q_t, x_t) = \Phi \left(\frac{MV_t - \hat{M}V_{L,R}}{\sigma} \right) (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) + u(q_t, \theta_D^m) - C_b$$

With a whip count, it is

⁶³ $\lim_{x \rightarrow \infty} \frac{1 - \Phi(x)}{\phi(x)} = \lim_{x \rightarrow \infty} \frac{-\phi(x)}{\phi'(x)} = \lim_{x \rightarrow \infty} \frac{-\phi(x)}{-x\phi(x)} = 0$ where the first equality uses L'Hôpital's rule.

$$\begin{aligned}
& EU_D^{count}(q_t, x_t) \\
&= \int_{-\infty}^{\bar{\eta}_{1,t}} \Phi\left(\frac{MV_t - \hat{M}V_{L,R} - \eta}{\sigma_\eta}\right) \frac{1}{\sigma_\eta} \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) \\
&\quad - \Phi\left(\frac{\bar{\eta}_{1,t}}{\sigma_\eta}\right) C_b + u(q_t, \theta_D^m)
\end{aligned}$$

Using these expressions, the optimal policy candidates, x_t^{count} and $x_t^{no\ count}$, can be shown to be unique (provided C_b is not too large) as in the previous case. \square

To prove Lemma 4, we first define and prove Lemma A1.

Lemma A1: Fix $C_b < \hat{C}_b$ such that the optimal alternative policies, x_t^{count} and $x_t^{no\ count}$, are unique. Then, the alternative policies that satisfy the first-order conditions with and without a whip count ((A.7) and (A.3) are such that:

- (1) For $q_t \neq \theta_D^m$, the optimal alternative policy with a whip count, x_t^{count} , lies strictly closer to party D's ideal point, θ_D^m , than that without, $x_t^{no\ count}$.
- (2) $MV_t^{count}(q_t)$ and $MV_t^{no\ count}(q_t)$ strictly increase for $q_t < \theta_D^m$ and strictly increase for $q_t > \theta_D^m$.

Proof of Lemma A1:

Part 1. Consider the case of $q_t < \theta_D^m$. We can write the first-order condition in the case of no whip count as an integration over the second aggregate shock (as in the case of the whip count):

$$\begin{aligned}
& \int_{-\infty}^{\infty} \left[1 - \Phi\left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right. \\
& \quad \left. - \frac{1}{2\sigma_\eta} \phi\left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \left(\frac{u(x_t^{no\ count}, \theta_D^m) - u(q_t, \theta_D^m)}{u'(x_t^{no\ count}, \theta_D^m)} \right) \right] \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta = 0
\end{aligned}$$

Consider the left-hand side of this expression, evaluated instead at x_t^{count} :

$$\begin{aligned}
& \int_{-\infty}^{\infty} \left[1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right. \\
& \quad \left. - \frac{1}{2\sigma_\eta} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u'(x_t^{count}, \theta_D^m)} \right) \right] \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
& = \int_{\underline{\eta}_{1,t}}^{\infty} \left[1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right. \\
& \quad \left. - \frac{1}{2\sigma_\eta} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u'(x_t^{count}, \theta_D^m)} \right) \right] \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
& \quad + \int_{-\infty}^{\underline{\eta}_{1,t}} \left[1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right. \\
& \quad \left. - \frac{1}{2\sigma_\eta} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u'(x_t^{count}, \theta_D^m)} \right) \right] \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
& = \int_{-\infty}^{\underline{\eta}_{1,t}} \left[1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right. \\
& \quad \left. - \frac{1}{2\sigma_\eta} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u'(x_t^{count}, \theta_D^m)} \right) \right] \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta
\end{aligned} \tag{A.10}$$

where the last equality follows from the fact that x_t^{count} satisfies the first-order condition for the case of a whip count. Consider the sign of the integrand in (A.10):

$$\begin{aligned}
& \left[1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) - \frac{1}{2\sigma_\eta} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u'(x_t^{count}, \theta_D^m)} \right) \right] \phi\left(\frac{\eta}{\sigma_\eta}\right) \geq 0 \\
& \iff \frac{1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)}{\frac{1}{2\sigma_\eta} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)} - \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u_x(x_t^{no\ count}, \theta_D^m)} \right) \geq 0
\end{aligned}$$

The left-hand side of this inequality is a strictly increasing function of η , so that there is at most one value of η at which the integrand is zero. As $\eta \rightarrow \infty$, the integrand approaches 1. Thus, to satisfy the first-order condition for the case of a whip count at x_t^{count} , the integrand evaluated at $\underline{\eta}_{1,t}$ must be strictly negative so that the single zero-crossing is contained in $[\underline{\eta}_{1,t}, \infty)$ (otherwise the integrand is positive over the whole range and cannot integrate to zero). Thus, the integrand in (A.10) must be strictly negative over $[-\infty, \underline{\eta}_{1,t}]$ so that the integral is strictly negative: the marginal expected utility for the case of no whip count must be negative when evaluated at the optimal alternative policy for the case of a whip count. But, then we must have $x_t^{no\ count} < x_t^{count}$ to ensure that the first-order condition for the case of no whip count is satisfied (given that $x_t^{no\ count}$ is the unique optimum, for every $x_t < x_t^{no\ count}$, the marginal expected utility is positive). The case of $q_t > \theta_D^m$ can be shown similarly.

Part 2. Consider the case of $q_t < \theta_D^m$ when a whip count is conducted. MV_t^{count} is determined implicitly by the first-order condition, (A.7). Taking its derivative with respect to q_t , we have

$$\frac{\partial}{\partial q_t} \left[\frac{\int_{\underline{\eta}_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta}{\frac{1}{2\sigma_\eta} \int_{\underline{\eta}_{1,t}}^{\infty} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta} - \frac{(u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m))}{u_x(x_t^{count}, \theta_D^m)} \right] = 0$$

$$\begin{aligned}
& \Leftrightarrow \frac{\partial}{\partial MV_t^{count}} \left(\frac{\int_{\eta_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta}{\frac{1}{2\sigma_\eta} \int_{\eta_{1,t}}^{\infty} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta} \right) \frac{\partial MV_t^{count}}{\partial q_t} \\
& \quad - \frac{\partial}{\partial x_t^{count}} \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u_x(x_t^{count}, \theta_D^m)} \right) \frac{\partial x_t^{count}}{\partial q_t} = 0 \\
& \Leftrightarrow \frac{\partial}{\partial MV_t^{count}} \left(\frac{\int_{\eta_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta}{\frac{1}{2\sigma_\eta} \int_{\eta_{1,t}}^{\infty} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta} \right) \frac{\partial MV_t^{count}}{\partial q_t} \\
& \quad - \frac{\partial}{\partial x_t^{count}} \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u_x(x_t^{count}, \theta_D^m)} \right) \left(2 \frac{\partial MV_t^{count}}{\partial q_t} - 1 \right) = 0 \\
& \Leftrightarrow \frac{\partial MV_t^{count}}{\partial q_t} \left[\frac{\partial}{\partial MV_t^{count}} \left(\frac{\int_{\eta_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta}{\frac{1}{2\sigma_\eta} \int_{\eta_{1,t}}^{\infty} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta} \right) \right. \\
& \quad \left. - 2 \frac{\partial}{\partial x_t^{count}} \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u_x(x_t^{count}, \theta_D^m)} \right) \right] \\
& \quad - \frac{\partial}{\partial x_t^{count}} \left(\frac{u(x_t^{count}, \theta_D^m) - u(q_t, \theta_D^m)}{u_x(x_t^{count}, \theta_D^m)} \right) = 0
\end{aligned}$$

As shown in the proof of Proposition 1, the term in brackets on the left-hand side is strictly negative for $C_b < \hat{C}_b$, and the last term on the left-hand side is also strictly positive so that we must have $\frac{\partial MV_t^{count}}{\partial q_t} > 0$. Similarly, $\frac{\partial MV_t^{no\ count}}{\partial q_t} > 0$. For $q_t > \theta_D^m$, we can similarly establish $\frac{\partial MV_t^{count}}{\partial q_t} < 0$ and $\frac{\partial MV_t^{no\ count}}{\partial q_t} < 0$. \square

Proof of Lemma 4:

$V_D^{count}(q_t) > V_D^{no\ count}(q_t)$ because, for C_b sufficiently small, $\eta_{1,t} < \infty$ and $\bar{\eta}_{1,t} > -\infty$ (see footnote 61) so that an alternative policy is pursued for a non-zero measure of the support of $\eta_{1,t}$. Therefore, for the same alternative policy, party D 's expected utility with a whip count must strictly exceed that without because over this support of $\eta_{1,t}$, the cost, C_b , is avoided and the probability of the alternative passing is the same. If party D pursues a different alternative policy with a whip count (which it generally does), then it must because it does even better.

Consider the case of $q_t < \theta_D^m$. We claim both value functions decrease with q_t , but the difference $V_D^{count}(q_t) - V_D^{no\ count}(q_t)$ increases. By the envelope theorem, the derivative of the value function for the case of no whip count with respect to q_t is given by

$$\begin{aligned}
\frac{\partial V_D^{no\ count}(q_t)}{\partial q_t} &= - \left(1 - \Phi\left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma}\right) \right) u_q(q_t, \theta_D^m) \\
&\quad - \frac{1}{2\sigma} \phi\left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma}\right) (u(x_t^{no\ count}, \theta_D^m) - u(q_t, \theta_D^m)) \\
&= - \left(1 - \Phi\left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma}\right) \right) u_q(q_t, \theta_D^m) \\
&\quad - \left(1 - \Phi\left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma}\right) \right) u_x(x_t^{no\ count}, \theta_D^m) \\
&= - \left(1 - \Phi\left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma}\right) \right) (u_q(q_t, \theta_D^m) + u_x(x_t^{no\ count}, \theta_D^m))
\end{aligned}$$

where the first equality follows from applying the first-order condition. With unbounded aggregate shocks and $q_t, x_t^{no\ count} < \theta_D^m$, the marginal utilities are strictly positive so that the overall derivative is negative.

In a similar manner, for the case of a whip count, we have

$$\begin{aligned}
\frac{\partial V_D^{count}(q_t)}{\partial q_t} &= - \frac{1}{2\sigma_\eta^2} \int_{\underline{\eta}_{1,t}}^{\infty} \phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta (u(x_t, \theta_D^m) - u(q_t, \theta_D^m)) \\
&\quad - \frac{1}{\sigma_\eta} u_q(q_t, \theta_D^m) \int_{\underline{\eta}_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
&= - \frac{1}{\sigma_\eta} (u_q(q_t, \theta_D^m) + u_x(x_t^{count}, \theta_D^m)) \int_{\underline{\eta}_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta
\end{aligned}$$

which is also strictly negative, given $\underline{\eta}_{1,t} < \infty$.

Finally, consider the marginal difference of the value functions:

$$\begin{aligned}
&\frac{\partial (V_D^{count}(q_t) - V_D^{no\ count}(q_t))}{\partial q_t} \\
&= - \frac{1}{\sigma_\eta} (u_q(q_t, \theta_D^m) + u_x(x_t^{count}, \theta_D^m)) \int_{\underline{\eta}_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right) \right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
&\quad + (u_q(q_t, \theta_D^m) + u_x(x_t^{no\ count}, \theta_D^m)) \left(1 - \Phi\left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma}\right) \right)
\end{aligned}$$

From the first part of Lemma A1, $x_t^{no\ count} < x_t^{count}$, which ensures $u_x(x_t^{no\ count}, \theta_D^m) > u_x(x_t^{count}, \theta_D^m)$. Furthermore,

$$\begin{aligned}
& 1 - \Phi\left(\frac{MV_t^{no\ count} - \hat{M}V_{R,R}}{\sigma}\right) \\
> & 1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R}}{\sigma}\right) \\
= & \frac{1}{\sigma_\eta} \int_{-\infty}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
> & \frac{1}{\sigma_\eta} \int_{\underline{\eta}_{1,t}}^{\infty} \left(1 - \Phi\left(\frac{MV_t^{count} - \hat{M}V_{R,R} - \eta}{\sigma_\eta}\right)\right) \phi\left(\frac{\eta}{\sigma_\eta}\right) d\eta \\
> & 0
\end{aligned}$$

given $\underline{\eta}_{1,t} < \infty$. Therefore, the difference in expected utility strictly increases with q_t .

For $q_t > \theta_D^m$, we can establish that both value functions increase in q_t , but their difference decreases, in an identical manner. \square

Proof of Proposition 2:

Assume $C_b < \hat{C}_b$ so that, from Proposition 1, x_t^{count} is unique. Consider $q_t < \theta_D^m$. We first show that as $q_t \rightarrow \theta_D^m$, $V_D^{no\ count}(q_t) \rightarrow -C_b$ and $V_D^{count}(q_t) \rightarrow 0$. The first follows from simple inspection of $EU_D^{no\ count}(q_t, x_t)$, noting that $x_t^{no\ count}$ must approach θ_D^m as $q_t \rightarrow \theta_D^m$ because it is contained in the interval, (q_t, θ_D^m) , by Proposition 1. Similarly, inspecting $EU_D^{count}(q_t, x_t)$, we see that $V_D^{count}(q_t) \rightarrow -\left(1 - \Phi\left(\frac{\underline{\eta}_{1,t}}{\sigma_\eta}\right)\right) C_b$. But, as $q_t \rightarrow \theta_D^m$, we can see from (A.5) that $\underline{\eta}_{1,t}$ must approach infinity such that $\Phi\left(\frac{\underline{\eta}_{1,t}}{\sigma_\eta}\right) \rightarrow 1$.

Given these facts, strictly positive costs, and the result of Lemma 4 that both value functions strictly decrease with $|q_t - \theta_D^m|$, there exists a status quo cutoff, $\bar{q}_l < \theta_D^m$, such that for all $q_t \in (\bar{q}_l, \theta_D^m)$, no alternative policy is pursued. Specifically, \bar{q}_l is given by the larger of the two policies, q_1 and q_2 which satisfy $V_D^{no\ count}(q_1) = 0$ and $V_D^{count}(q_2) = C_w$, respectively.

For $q_t < \bar{q}_l$, there are two possibilities. If $q_1 > q_2$, then set $\underline{q}_l = \bar{q}_l = q_1$ so that $V_D^{count}(q_1) < C_w$ and $V_D^{no\ count}(q_1) = 0$. In this case, for any $q_t < q_1$, an alternative policy is pursued without a whip count: by Lemma 4, over this range, $V_D^{no\ count}(q_1) > 0$ so that an alternative policy without a whip count is preferred over not pursuing an alternative policy and, as q_t decreases from q_1 , $V_D^{count}(q_t) - V_D^{no\ count}(q_t)$ decreases so that not conducting a whip count remains more valuable than conducting one.

If $q_1 < q_2$, then set $\bar{q}_l = q_2$ and define $\underline{q}_l < \bar{q}_l$ to be the policy for which $V_D^{count}(\underline{q}_l) - C_w = V_D^{no\ count}(\underline{q}_l)$. Such a point must exist because, by Lemma 4, as q_t decreases from \bar{q}_l , $V_D^{count}(q_t) - V_D^{no\ count}(q_t)$ decreases and so must eventually approach zero. Thus, for q_t sufficiently small, $V_D^{count}(q_t) - C_w < V_D^{no\ count}(q_t)$. With these cutoffs, for $q_t \in (-\infty, \underline{q}_l]$, an alternative policy is pursued without a whip count because $V_D^{no\ count}(q_t) > V_D^{count}(q_t) - C_w > 0$ for all $q_t < \underline{q}_l$. For $q_t \in (\underline{q}_l, \bar{q}_l]$, an alternative policy is pursued with a whip count because $V_D^{count}(q_t) - C_w > 0$ and,

by Lemma 4, $V_D^{count}(q_t) - V_D^{no\ count}(q_t)$ increases with q_t over this range so that $V_D^{count}(q_t) - C_w > V_D^{no\ count}(q_t)$.

Symmetric arguments establish cutoffs, \underline{q}_r and \bar{q}_r , for the bill pursuit decisions over the range $q_t > \theta_D^m$. \square

ONLINE APPENDIX

APPENDIX B. (NOT FOR PUBLICATION) IDENTIFICATION AND ESTIMATION SUPPLEMENTARY MATERIAL

B.1. Formal Treatment of Identification. We provide a more formal treatment of the proof of identification of the parameters governing voting decisions (member ideal points, party discipline, and the variances of the aggregate shocks) for the first step parameters. Identification is based on the joint distribution of voting decisions at roll call and whip count ($Yes_{t,p}^{i,wc}, Yes_{t,p}^{i,rc}$), and dummy variables (for individual i and vote t , at both the whip count and roll call stages). The latter are our covariates, denoted here as a matrix $X_{i,t}$.⁶⁴

Knowledge of the joint distribution of covariates and voting decisions at the whip count stage implies knowledge of $\mathbb{E}[Yes_{t,p}^{i,wc} | X_{i,t}] = P(Yes_{t,p}^{i,wc} = 1 | X_{i,t})$. Similarly, we know $\mathbb{E}[Yes_{t,p}^{i,rc} | X_{i,t}] = P(Yes_{t,p}^{i,rc} = 1 | X_{i,t})$. Let us focus on the case for party D , as the case for R is analogous, and identification holds within each party.

For any $t = 1, \dots, T$, given an independent draw q_t , $\tilde{M}V_{1,t}$ is given and can be treated as a parameter to be identified and estimated - it is a bill fixed effect (parameter multiplying the appropriate dummy variable). The vector of individual ideal points $\theta = \{\theta^i\}$ for $i = 1, \dots, N$ is a set of individual fixed effects to be identified. Similarly, for any $t = 1, \dots, T$, $\tilde{M}V_{2,t}$ is also given and can be treated as a parameter. The vector of discipline parameters $\{y_p^{max}\}_{p \in \{D,R\}}$ is a set of constant parameters to be identified.

From equation (6.1), we have that, at the time of the whip count, for every i and t :

$$(B.1) \quad \Phi^{-1}(\mathbb{E}(Yes_{t,p}^{i,wc} = 1 | X_{i,t})) = \tilde{M}V_{1,t} - \theta^i.$$

The difference of equation (B.1) across politicians i and 0 in period t is:

$$(B.2) \quad \Phi^{-1}(\mathbb{E}(Yes_{t,p}^{0,wc} = 1 | X_{i,t})) - \Phi^{-1}(\mathbb{E}(Yes_{t,p}^{i,wc} = 1 | X_{i,t})) = \theta^i,$$

where we have used that $\theta^0 = 0$ (Assumption 1). Because θ^i is known, we have that $\tilde{M}V_{1,t}$ is known for an arbitrary t as it is the bill fixed effect (seen in equation (B.1)). As such, ideologies are identified by the average “Yes” votes at the whip count stage relative to a normalizer (which

⁶⁴ $X_{i,t}$ is common across whip counts and roll calls because the dummy variables are at the politician and bill level only.

pins down location). The realized marginal voter is identified from the average “Yes” votes given ideologies.

At roll call, equation (6.2) can be rewritten as

$$(B.3) \quad \Phi^{-1}(\mathbb{E}(Yes_{t,p}^{i,rc} = 1 \mid X_{i,t})) = \frac{\tilde{M}V_{2,t} - \theta^i \pm y_D^{max}}{\sqrt{2}},$$

for every i, t .

$\theta^i \pm y_D^{max}$ is then identified, as it is the individual fixed effect at the roll call (and we have that $\eta_{2,t}$ is mean zero which pins down location).⁶⁵ Thus, the party discipline parameters are identified up to their sign, which is pinned down by the direction of whipping (known from the theory). In summary, y_p^{max} is identified from the average switching behavior of a politician i across roll calls and whip counts. It follows that $\tilde{M}V_{2,t}$ is then identified for all t .

Finally, notice that by definition of the realized marginal voters we obtain:

$$(B.5) \quad \tilde{M}V_{1,t} - \tilde{M}V_{2,t} = \eta_{2,t}$$

the distribution of which is semiparametrically identified, allowing us to recover its variance, σ_η^2 .

We can also formally demonstrate the criticality of the whip count data. In its absence, y_D^{max} is not identified (the essence of Krehbiel’s critique (Krehbiel (1993))). From (6.2), if we do not know θ^i but instead had to estimate it from roll call data only, we could redefine $\tilde{\theta}_i = \theta^i \pm y_D^{max}$ so that:

$$(B.6) \quad \begin{aligned} P(Yes_{t,p}^{i,rc} = 1 \mid X_{i,t}) &= \Phi\left(\frac{\tilde{M}V_{2,t} - \theta^i \pm y_D^{max}}{\sqrt{2}}\right) \\ &= \Phi(\tilde{M}V_{2,t} - \tilde{\theta}_i). \end{aligned}$$

Hence, with roll call data alone, we cannot separate a shift in everyone’s (true) ideology from the party discipline effect due to whipping.

B.2. Governing Equations for Party R .

In our description of the theory and estimation, we focused on party D . Here we provide the key equations for party R , beginning with the probabilities of observing a member of party R voting Yes (corresponding to (6.1) and (6.2) for party D). The difference stems from the fact that, when

⁶⁵More formally, using equations (B.1), (B.3) and (B.5), we have that:

$$(B.4) \quad \mathbb{E}_t \left(\Phi^{-1}(\mathbb{E}(Yes_{t,p}^{i,wc} = 1 \mid X_{i,t})) - \sqrt{2}\Phi^{-1}(\mathbb{E}(Yes_{t,p}^{i,rc} = 1 \mid X_{i,t})) \right) = \pm y_D^{max},$$

where we take the expectation over all bills t on both sides and use $\eta_{2,t}$ is mean zero.

the two parties prefer different policies, members of D to the left of the marginal voter vote Yes while members of R to the left vote No. At the whip count stage:

$$\begin{aligned}
 P(Yes_{t,p}^{i,wc} = 1 \mid X_{i,t}) &= P(\delta_{1,t}^i + \theta^i \geq MV_t - \eta_{1,t} \mid X_{i,t}) \\
 (B.7) \qquad \qquad \qquad &= 1 - \Phi(\tilde{M}V_{1,t} - \theta^i).
 \end{aligned}$$

At the roll call stage,

$$\begin{aligned}
 P(Yes_{t,p}^{i,rc} = 1 \mid X_{i,t}) &= P(\delta_{1,t}^i + \delta_{2,t}^i + \theta^i \geq MV_t - \eta_{1,t} - \eta_{2,t} \pm y_R^{max} \mid X_{i,t}) \\
 (B.8) \qquad \qquad \qquad &= 1 - \Phi\left(\frac{\tilde{M}V_{2,t} - \theta^i \pm y_R^{max}}{\sqrt{2}}\right),
 \end{aligned}$$

The likelihood of a sequence of votes by members of party R is therefore derived from (6.3) by substituting these expressions for the probabilities.

The other key equation is that which governs the optimal policy alternative chosen by party R in case of no whip count (corresponding to (A.3) for party D). For a status quo policy to the left of party R 's median, party R chooses an alternative further to the right so that the first-order condition is identical to (A.3) except that $\hat{M}V_{R,R}$ is replaced by $\hat{M}V_{L,R}$ because the parties whip in opposite directions. For a status quo policy to the right of party R 's median (so that the alternative is left of the status quo and both parties whip left), it is given by

$$(B.9) \qquad \frac{-\Phi\left(\frac{MV_t^{no \ count} - \hat{M}V_{L,L}}{\sigma}\right)}{\phi\left(\frac{MV_t^{no \ count} - \hat{M}V_{L,L}}{\sigma}\right)} = \frac{1}{2\sigma} \frac{(u(x_t^{no \ count}, \theta_R^m) - u(q_t, \theta_R^m))}{u_x(x_t^{no \ count}, \theta_R^m)}.$$

APPENDIX C. (NOT FOR PUBLICATION) ON THE SHARE OF BILLS PROPOSED BY THE MAJORITY

In our main estimates for the second step (Table 4), we find that the share of bills proposed by the majority party is $\gamma \approx 0.43$, or 43%. This finding might appear surprising given a literature on agenda setting that suggests majority parties should maintain full control over the proposal of bills in Congress (either through positive agenda control, as in the Conditional Party Government of Aldrich (1995)), or through negative agenda control, restricting the access of the minority's proposals to the floor as in Cox and McCubbins (1993, 2005)). Here, we argue that our results are consistent with such explanations, but also present a more nuanced picture of congressional activity, one discussed by Jenkins et al. (2014).⁶⁶

The first key point is that the types of votes we observe in the data are very heterogeneous. Roll calls include final passage votes on certain bills, but also other votes, such as votes on amendments and motions.⁶⁷ To categorize votes, we merge data from David Rohde's PIPC Roll Call - University of Oklahoma dataset⁶⁸, which provides coarse types for each roll call vote. Broadly speaking, votes can be broken into final passage votes and other (amendments, motions, etc.) In Table C.1, we present the share of Democratic proposals on final passage votes that pin party leaders against one another.

Across all final passage votes (1345 roll called bills out of a full sample of 5424, or 25%), the share of majority proposals according to our proxy (see the Data Section for details) is 49%, which is slightly similar to that in the full sample.⁶⁹ If we further condition our sample to final passage votes in which party leaders vote in opposite directions, we find a clear dominance of the majority party as suggested by the previously cited congressional literature. In this subsample of 394 roll calls, 369 (94%) are proposed by the majority according to our proxy. Therefore, within this sample of salient votes which are often the focus of the literature (Krehbiel (2000)), we do indeed find that

⁶⁶We must also clarify a key distinction between our parameter γ , and the *observed* share of proposals by a party. In the theory, γ is the unconditional probability that the proposing party is the majority (Democrats in our sample). Because our model allows for missing mass, the unconditional share (γ) will differ from the conditional share (observed proposals), regardless of how we assign the proposing party. The fact that this distinction matters is seen in our estimates: Republicans have a larger missing mass than Democrats, which means that the Democrats' unconditional probability of being a proposer ($\gamma \approx 43\%$) is smaller than the observed share of Democrat proposals, 47% (described in Section 5). Thus, γ must be estimated jointly in the second stage of estimation, and cannot simply be calibrated to the observed shares of proposals. While this conceptual distinction is important, it makes only a small quantitative difference given our estimates of the missing mass for each party.

⁶⁷As mentioned in the main text, we treat each roll call vote as a different draw of q in the theory (although possibly the same or close to that of a previous draw). Amendments and motions also change status quo policies, and are subject to the same considerations as final passage votes.

⁶⁸Available online at <http://www.ou.edu/carlabertcenter/research/pipc-votes>.

⁶⁹As we show in a robustness check in Appendix D.1, differences between final passage votes and other votes are not driving our baseline results for the sources of polarization. Comparing Table D.2 to Table 1, we find that party discipline (the sum of $y_D^{max} + y_R^{max}$) is 10-20% larger in the final passage subsample, while ideological distances are slightly smaller - but still close to the baseline estimates.

TABLE C.1. The Sources of Majority Proposals

	Full Sample	Sample Final Passage Votes with Leaders Voting in Opposite Directions
Share of Votes proposed by Democrats	47%	94%
Number of Roll Call Votes	5424	394

Notes: To understand the sources of our result for γ in Table 4, we split our sample of all roll calls votes into 2 subsamples: (i) final passage votes, and (ii) final passage votes with leaders voting in opposite directions. Our model predicts that the bills in (ii) are those that are between the party medians.

the majority maintains agenda setting control. Clearly then, it must be the remaining bills, those with bipartisan support and/or amendments and motions, that are proposed by the minority party. In fact, [Jenkins et al. \(2014\)](#) provide evidence of exactly this.⁷⁰

[Jenkins et al. \(2014\)](#) use data on the sponsorship of bills and amendments to proxy for the party that is responsible for a vote. In doing so, they find the minority to be very active exactly on amendments and other motions. [Jenkins et al. \(2014\)](#) suggest that the minority's focus on these types of votes is strategic. Their aim with these proposals is to gain leverage over the majority through the (over)use of the scarce resource of *time* in Congress. As the authors write, “*In short, we argue that the minority derives leverage by delaying (or threatening to delay) legislative action, which consumes very scarce, very valuable plenary time. To avoid this loss of plenary time, the majority is willing to offer the minority opportunities to have their own proposals considered on the House and Senate floors, respectively. We argue that, with the goal of recapturing majority status in mind, the minority uses these proposal opportunities to schedule roll calls that help separate them from the majority and force majority-party members into casting difficult votes... Specifically, the minority party seems to use its leverage to get floor votes on many amendment proposals that fail, but with the support of the minority party and the opposition of the majority party. This pattern shows up in both chambers, and stands in contrast to voting on the final passage of bills, which shows very little evidence of any kind of minority proposal.*” (p.16).

Summarizing, the reason our estimate of γ is low is because many different types of votes are present in our dataset, and the minority party can be responsible for many of the non-salient votes.

⁷⁰[Jenkins et al. \(2014\)](#) write that the literature has often ignored the minority party and how they act, which is why the existence of such activity by the minority might be surprising. In the literature on ideal point estimation, it is also the case that there is little role for the minority. Most often, models simply pool all roll call votes together and neglect party effects. However, in our model, the minority party can be quite active. It may propose its own bills (as we observe in the data) and can affect the majority's policy proposals indirectly through party discipline.

Given that sponsorship data provides an alternative proxy for the proposing party (compared to ours), one may be inclined to replace our proxy (based on our theory) with sponsorship data. However, a good reason not to do so exists: in the sponsorship data, only one sponsor is recorded, so any bipartisan bill will immediately be erroneously assigned to one party or the other. In Appendix D, we rerun our estimation on the subset of votes for which our proxy and the sponsorship proxy agree, finding that our results are largely unchanged and therefore not likely affected by errors in our proxy.

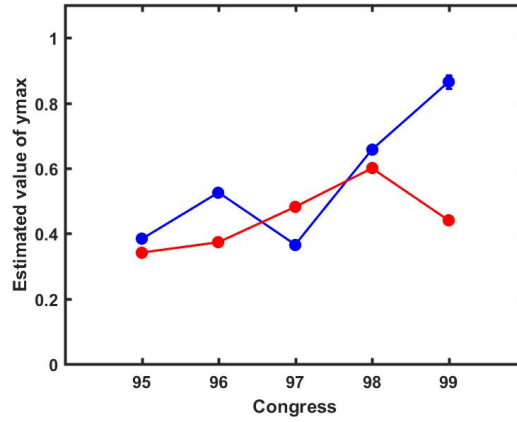
APPENDIX D. (NOT FOR PUBLICATION) ROBUSTNESS CHECKS

We present estimates of four variations of our model to assess their robustness along several dimensions. The first step results are summarized in Figure D.1 and Table D.1. We see that the level and trends in party discipline, as well as the share of perceived polarization due to party discipline are broadly consistent across specifications. We discuss each specification in more detail in the following subsections and then end with a discussion of salient bills.

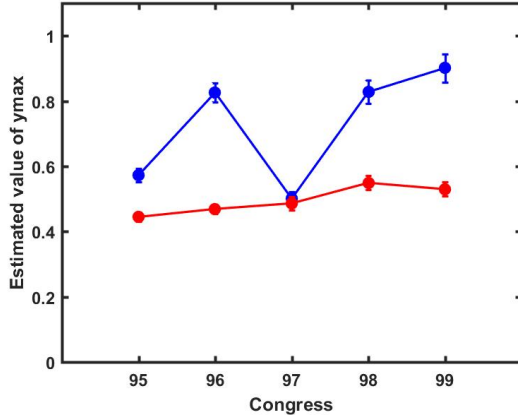
TABLE D.1. Decomposition of Polarization

	Congress				
	95	96	97	98	99
Implications of Table 1 for Polarization					
<i>A: Polarization due to ideology ($\theta_R^m - \theta_D^m$)</i>					
Baseline	1.395	1.473	1.554	1.615	1.698
Only Final Passage Votes	1.135	1.308	1.355	1.401	1.441
Proposer Proxies Coincide (Model & Sponsorship)	1.346	1.423	1.490	1.569	1.645
No Whipping on Lopsided Votes	1.615	1.713	1.796	1.895	2.031
Without Votes that Split Northern/Southern Democrats	1.170	1.306	1.335	1.419	1.471
<i>B: Polarization due to whipping ($y_R^{max} + y_D^{max}$)</i>					
Baseline	0.725	0.899	0.848	1.258	1.305
Only Final Passage Votes	1.018	1.296	0.988	1.378	1.431
Proposer Proxies Coincide (Model & Sponsorship)	0.941	1.126	1.132	1.490	1.563
No Whipping on Lopsided Votes	0.583	0.828	0.814	1.085	1.134
Without Votes that Split Northern/Southern Democrats	0.843	1.143	1.165	1.548	1.559
<i>C: Share of Perceived Ideological Polarization due to whipping ($B/(A+B)$)</i>					
Baseline	0.342	0.379	0.353	0.438	0.435
Only Final Passage Votes	0.473	0.498	0.422	0.496	0.498
Proposer Proxies Coincide (Model & Sponsorship)	0.411	0.442	0.432	0.487	0.487
No Whipping on Lopsided Votes	0.265	0.326	0.312	0.364	0.358
Without Votes that Split Northern/Southern Democrats	0.419	0.467	0.466	0.522	0.515

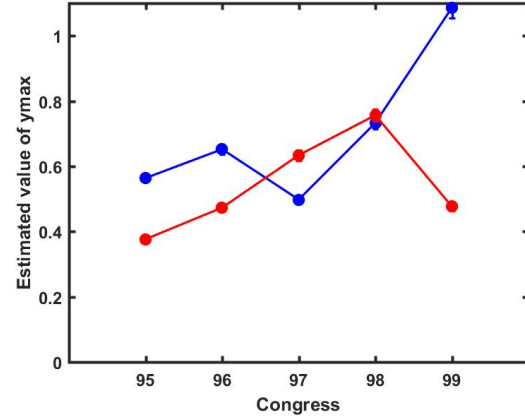
Notes: Decomposition of perceived polarization (polarization in ideologies from a misspecified model that ignores party discipline) into that due to ideological polarization and that due to party discipline, by Congress. In addition to the baseline results from our main model, we present the results of four alternative specifications. The alternative specifications: (i) final passage votes only (i.e. not including roll call votes on motions/amendments, among others), (ii) bills for which our proxy for bill proposer coincides with that in [Jenkins et al. \(2014\)](#) (the sponsor of that vote), (iii) assuming no whipping on lopsided bills, with lopsided bills defined as in [Snyder and Groseclose \(2000\)](#), (iv) dropping votes that split Northern/Southern Democrats, which DWNominate suggests is the source of a second dimension of ideology.

FIGURE D.1. Time Series of y^{max} Parameter Across Specifications


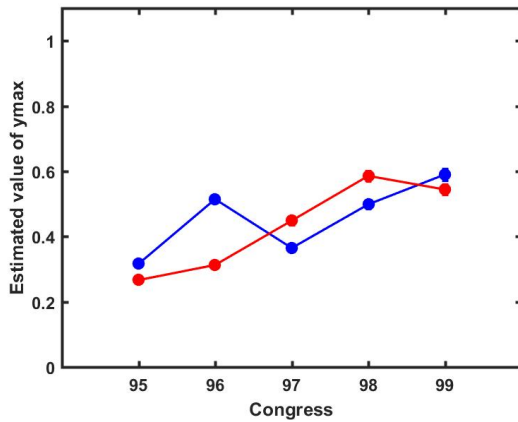
(A) Baseline (Reproduced from Figure 9)



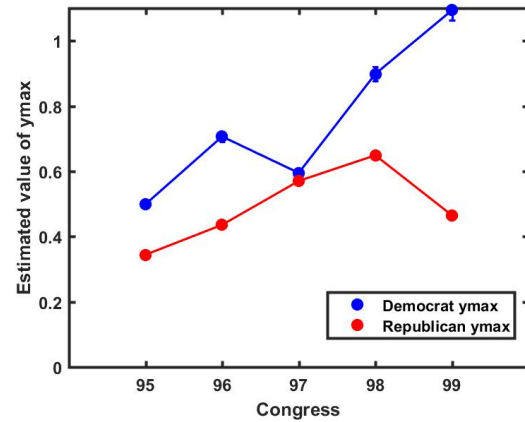
(B) Only Final Passage Votes



(C) Votes Where Our Proxy and Sponsorship Proxy for Proposer Agree



(D) No Whipping on Lopsided Bills



(E) Votes that Split Northern/Southern Democrats Dropped

Notes: Time series estimates of party discipline per party, y^{max} , for four variations of our model (main model reproduced for convenience). The alternative specifications are (i) final passage votes only (i.e. not including roll call votes on motions/amendments, among others), (ii) bills for which our proxy for bill proposer coincides with that in [Jenkins et al. \(2014\)](#) (the sponsor of that vote), (iii) assuming no whipping on lopsided bills, with lopsided bills defined as in [Snyder and Groseclose \(2000\)](#), (iv) dropping votes that split Northern/Southern Democrats, which DWNominate suggests is the source of a second dimension of ideology.

D.1. Robustness - Only Final Passage Votes. It's possible that not all bills are whipped to the same extent. In particular, salient bills might receive more attention. Here, we re-estimate our model on final passage bills to assess this possibility. To identify final passage bills, we merge our data with that on vote types from David Rohde's PIPC Roll Call dataset. We re-estimate our model on they 23% of roll calls that are final passage votes. The results in Table D.2 are similar to those in the main specification, although the party discipline estimates are slightly larger as one would expect if these votes receive greater attention.

TABLE D.2. First Step Estimates - Only Final Passage Votes

Parameter	Congress				
	95	96	97	98	99
Party Discipline y^{max} , Democrats	0.573 (0.010)	0.826 (0.015)	0.500 (0.011)	0.828 (0.018)	0.902 (0.022)
Party Discipline y^{max} , Republicans	0.445 (0.007)	0.469 (0.007)	0.487 (0.010)	0.549 (0.011)	0.530 (0.011)
Standard Deviation of Aggregate Shock σ_η			0.886 (1.228)		
Party Median - Democrats, θ_D^m	-1.198 (0.105)	-1.203 (0.098)	-1.183 (0.093)	-1.188 (0.100)	-1.189 (0.116)
Party Median - Republicans, θ_R^m	-0.064 (0.105)	0.105 (0.109)	0.172 (0.108)	0.214 (0.135)	0.252 (0.098)

$N: 711$
 $T: 192$ Whip Counted bills, 1243 Roll Called bills

Notes: Estimates of the first step parameters for the subset of votes that are final passage votes according to Rohde's PIPC-University of Oklahoma dataset. Asymptotic standard errors are in parentheses. Non time-varying parameters are centered in the table and apply to all five Congresses.

D.2. Robustness - Votes where our Proxy for Proposing Party Coincides with the Bill/Amendment/Motion

Sponsor. Our theoretical model allows us to identify the proposing party and the whipping directions of each party from leadership votes. But, as discussed in Appendix C, another potential proxy is that of the bill/amendment/motion).⁷¹ Here, we re-estimate both stages of our model on the subset of data where these two proxies coincide to assess the robustness of our proxy.

Our sponsorship data comes from two sources. Sponsorship data on final passage bills comes from the Congressional Bills Project. For other types of votes (amendments and motions), we scrape the Voteview website (this website generally lacks sponsorship data for final passage votes).

After merging the sponsorship data, we find that 3882 out of 5424 roll call votes (72%) are assigned the same party proposer regardless of proxy, which is a reassuring first check that our proxy is reasonable. The first and second step estimates for this subsample of votes are given in Tables D.3 - D.4.

Party discipline estimates are similar in level and trend to our main estimates, although slightly larger. Aggregate party discipline ($y_D^{max} + y_R^{max}$) lies between 0.9-1.4 in this subsample (relative to 0.7-1.3 in the main estimates) and it explains between 40-50% of perceived polarization (relative to 33-43% in the main estimates).

For the agenda setting results, the missing mass and distributions of the status-quo's are similar to our main estimates, with only two main differences. First, the mean of the status quo's (μ_q) is closer to the Democrat median in the subsample. Second, the probability that the Democrats are selected as the proposer, γ is now 0.56, higher than in the baseline. Overall, the results confirm that our proxy is reasonable. It coincides with that of the sponsorship data on the majority of votes and our results are not driven by the choice of proxy. The results also shore up our discussion on the origins of $\gamma < 1/2$ in Appendix C - the estimate of γ is close to one-half, even in a subsample for which we are fairly certain of the proposing party.

⁷¹We prefer our proxy for two main reasons. First, relying on the sponsor of the bill means the assignment of the proposer depends on the decision of a single party member, one who may not be influential and may not have coordinated with party leadership. As Jenkins et al. (2014) write, [such proxy] is admittedly noisy: *"This has the potential to be a noisy measure, however, since the views of one minority-party member may regularly deviate from those of his co-partisans."* (p.12). Second, our proxy applies to all votes, including those that are highly bipartisan. For bipartisan votes, may have sponsors from both parties but the sponsorship data (e.g. Voteview and the Congressional Bills Project) report only a single proposer. The choice of sponsor to report is not clear in our reading and so we expect this proxy to be much noisier than ours (which is based on a theoretical prediction about vote shares). Also note that bipartisan votes make up a majority of the votes in which the two proxies are in conflict. In 79% of such votes the party leaders vote in the same direction.

TABLE D.3. First Step Estimates - Model and Sponsorship Proxies Coincide

Parameter	Congress				
	95	96	97	98	99
Party Discipline y^{max} , Democrats	0.564 (0.006)	0.652 (0.008)	0.498 (0.007)	0.733 (0.009)	1.087 (0.017)
Party Discipline y^{max} , Republicans	0.377 (0.004)	0.473 (0.006)	0.634 (0.008)	0.757 (0.009)	0.477 (0.006)
Standard Deviation of Aggregate Shock σ_η			0.786 (0.935)		
Party Median - Democrats, θ_D^m	-1.385 (0.084)	-1.385 (0.070)	-1.385 (0.068)	-1.424 (0.085)	-1.434 (0.149)
Party Median - Republicans, θ_R^m	-0.039 (0.057)	0.038 (0.117)	0.105 (0.147)	0.145 (0.135)	0.212 (0.072)

N: 711

T: 177 Whip Counted bills, 3882 Roll Called bills

Notes: Estimates of the first step parameters in the subset of votes where our model-based proxy for the proposing party coincides with the proxy that comes from the sponsor of the bill/amendment/motion. Asymptotic standard errors are in parentheses. Non time-varying parameters are centered in the table and apply to all five Congresses.

TABLE D.4. Second Step Estimates - Model and Sponsorship Proxies Coincide

	95	96	Congress 97	98	99
Probability Democrat is Proposer, γ			0.563 (0.0004)		
Status Quo Distribution (Mean), μ_q	-0.998 (0.139)	-0.774 (0.121)	-0.222 (0.126)	-0.251 (0.224)	-0.443 (0.125)
Status Quo Distribution (Standard Deviation), σ_q	1.975 (0.150)	1.769 (0.128)	1.835 (0.162)	1.082 (0.228)	1.066 (0.101)

Notes: Estimates of the second step parameters in the subset of votes where our model-based proxy for the proposing party coincides with the proxy that comes from the sponsor of the bill/amendment/motion. Asymptotic standard errors, accounting for estimation error from the first step, in parentheses. Standard errors are computed by drawing 100 samples from the asymptotic distribution of first step estimates, recomputing the second step estimates, and using the Law of Total Variance.

TABLE D.5. Missing Mass: Robustness Check Compared to Main Results

	95	96	Congress 97	98	99
Democrats					
<i>Baseline</i>	0.048	0.057	0.064	0.110	0.071
<i>Robustness (Proxies Coincide)</i>	0.138	0.210	0.141	0.247	0.222
Republicans					
<i>Baseline</i>	0.109	0.180	-	-	-
<i>Robustness (Proxies Coincide)</i>	0.151	0.245	-	-	-

Notes: Mass of status quo policies ('missing mass') that are not pursued by the party at all. Robustness results are for the subset of votes where our model-based proxy for the proposing party coincides with the proxy that comes from the sponsor of the bill/amendment/motion.

D.3. Robustness - The Snyder-Groseclose Procedure (No Whipping on Lopsided Bills). One may be concerned that our results are driven by the assumption that all votes are whipped. We showed in Section D.1 that our results are similar when estimated on final passage bills only (which may be more likely to be whipped). Here, we take a second approach inspired by the work of Snyder and Groseclose (2000).

Snyder and Groseclose (2000) studied the identification of ideal points under the assumption that parties do not discipline lopsided votes. We re-estimate our model using an adaptation of their approach.⁷² Snyder and Groseclose (2000) define a lopsided vote as one in which 70% of the House votes Yes when the majority controls more than 62% of seats and 65% Yes votes otherwise. We adopt this definition and assume that there is no party discipline ($y_D^{max} = y_R^{max} = 0$) exercised on these votes. The first step estimates are shown in Table D.6 below.

Relative to our main estimates, when we assume no whipping on lopsided votes, our results are only mostly unchanged, with the party discipline parameters being slightly smaller. As shown in Table D.1, in this alternative specification, party discipline accounts for 27-34% of perceived polarization, about a 7 p.p. decrease from the baseline.

⁷²Although conceptually similar, our exact procedure differs from that of Snyder and Groseclose (2000). In their work, close votes do not impact the estimation of ideology positions because they estimate ideologies in a first stage using only lopsided bills. In our case, identification of ideal points comes from all votes, including close votes. This difference is made possible by the fact that we explicitly model party discipline.

TABLE D.6. First Step Estimates - Snyder-Groseclose Procedure

Parameter	Congress				
	95	96	97	98	99
Party Discipline y^{max} , Democrats	0.316 (0.004)	0.515 (0.006)	0.365 (0.006)	0.499 (0.007)	0.590 (0.009)
Party Discipline y^{max} , Republicans	0.267 (0.003)	0.313 (0.004)	0.449 (0.007)	0.586 (0.008)	0.544 (0.008)
Standard Deviation of Aggregate Shock σ_η			0.957 (1.104)		
Party Median - Democrats, θ_D^m	-1.666 (0.070)	-1.669 (0.048)	-1.667 (0.045)	-1.696 (0.097)	-1.722 (0.089)
Party Median - Republicans, θ_R^m	-0.051 (0.051)	0.044 (0.149)	0.129 (0.156)	0.200 (0.147)	0.309 (0.061)

N: 711

T: 315 Whip Counted bills, 5424 Roll Called bills

Notes: Estimates of the first step parameters, for the specification in which we assume no whipping on lopsided bills (as defined in [Snyder and Groseclose \(2000\)](#)). Asymptotic standard errors are in parentheses. Non time-varying parameters are centered in the table and apply to all five Congresses.

D.4. Robustness to Removing Votes that Split Northern and Southern Democrats. For tractability, we have assumed a single ideological dimension. But, DWNominate estimates suggest a second dimension that reflects conflict over race and civil rights helps to explain voting patterns. Poole (2007) writes *“The second dimension captured the conflict over race and civil rights. With the passage of the 1964 Civil Rights Act, the 1965 Voting Rights Act, and the 1967 Open Housing Act, this second dimension slowly declined in importance and is now almost totally absent. Race related issues - affirmative action, welfare, Medicaid, subsidized housing, etc. - are now questions of redistribution. Voting on race related issues now largely takes place along the liberal-conservative dimension and the old split in the Democratic Party between North and South has largely disappeared.”* (p. 437)

Our sample starts in the mid-1970’s, when the second dimension has begun to decline in importance. But, it is still plausible that, by not accounting for this second dimension, our results could be biased. To check for this possibility, we re-estimate our model excluding 961 votes that were most likely to reflect a second ideological dimension - those that split the Northern and Southern Democrats (as identified by variable *v17* in David Rohde’s dataset, PIPC - University of Oklahoma).⁷³

Table D.7 presents the results which are similar to those of our main specification. In the subsample, Democratic party discipline is slightly larger, which is intuitive: the presence of bills that split Democrats show up as less effective party control.

⁷³Only 90 of 5424 votes are classified as civil rights issues so excluding only these trivially does not change our results.

TABLE D.7. First Step Estimates - North-South Split Votes Dropped

Parameter	Congress				
	95	96	97	98	99
Party Discipline y^{max} , Democrats	0.499 (0.005)	0.707 (0.008)	0.595 (0.007)	0.899 (0.011)	1.094 (0.014)
Party Discipline y^{max} , Republicans	0.344 (0.003)	0.436 (0.004)	0.570 (0.007)	0.649 (0.007)	0.465 (0.005)
Standard Deviation of Aggregate Shock σ_η			0.852 (0.853)		
Party Median - Democrats, θ_D^m	-1.122 (0.100)	-1.143 (0.060)	-1.116 (0.060)	-1.151 (0.060)	-1.156 (0.160)
Party Median - Republicans, θ_R^m	0.048 (0.093)	0.163 (0.119)	0.219 (0.066)	0.269 (0.143)	0.315 (0.093)
N : 711					
T : 244 Whip Counted bills, 4463 Roll Called bills					

Notes: Estimates of the first step parameters in the sample of votes that excludes those that split Northern and Southern Democrats (per $v17$ in Rohde's PIPC Roll Call - University of Oklahoma dataset). Asymptotic standard errors are in parentheses. Non time-varying parameters are centered in the table and apply to all five Congresses.

D.5. Salient Bills. It is plausible that “salient” bills attract more attention and party discipline, an effect not accounted for in the model. Here, we discuss reasons we think salience is not the primary driver of party discipline. We consider three different definitions of “salient”.

First, we consider Congressional Quarterly’s “Key Votes” - the roughly 16 votes per year they identify as important for policy and legislative opposition. These key votes are the standard way important votes are identified in the political science literature (Shull and Vanderleeuw (1987)). After identifying the Key Votes in our dataset, we find that 32.5% (51/157)⁷⁴ were whip counted (a number similar to Evans (2018), who doesn’t present the results for our timeframe). Thus, the salience of a vote is neither necessary nor sufficient for it to be whip counted, meaning that votes are whip counted (and whipped) for other reasons.

Second, we have shown in Section D.1 that our results are similar when we focus on the subset of final passage bill votes. To the extent that final passage votes are more salient, this again suggests that the salience of a vote is not a key driver of whether or not it is whipped.

Third, one might think that salient votes are those whose outcomes are too close to call. They might, for example, attract the most media attention. But, as shown in our robustness check that applies Snyder and Groseclose (2000)’s procedure (Section D.3), our results are robust to assuming only close bills are whipped, suggesting that discipline is also active on lop-sided votes, at least to some extent.

Summarizing, although it seems likely that the salience of a bill causes it to be more likely to be whip counted and subject to discipline, we don’t find that salience on its own is driving these activities.

⁷⁴We could not merge one CQ Key Vote to a roll call vote in Voteview.

APPENDIX E. (NOT FOR PUBLICATION) PROPERTIES OF OUR ESTIMATORS: MONTE CARLO
SIMULATIONS AND ASYMPTOTICS

We demonstrate good finite-sample properties of both steps of our estimation process via Monte Carlo Simulations, and discuss their asymptotic properties.

E.1. Monte Carlo Simulations.

E.1.1. *Monte Carlos for the First Step.* For the first step estimate, we simulate whip count and roll call votes under true values that correspond to our main first step estimates in Table 1.⁷⁵ In each of $R = 100$ runs, we simulate votes for the number of whip counts and roll call votes we have in the data, and then estimate the parameters. In Table E.1 and Figure E.1, we report averages of the estimated parameters over the $R = 100$ runs. It is clear from the result that the average estimates are very close to the true values for the DGP.

TABLE E.1. Monte Carlo Simulations - First Step Estimation

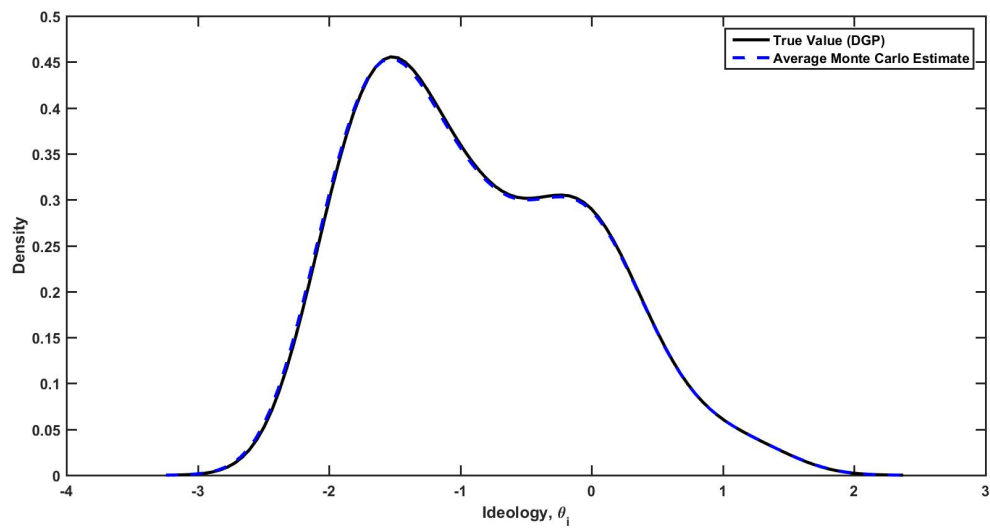
Parameter	True Values					Average Monte Carlo Estimates				
	Congress					Congress				
	95	96	97	98	99	95	96	97	98	99
Party Discipline y^{max} , Democrats	0.383	0.526	0.366	0.658	0.865	0.381 (0.035)	0.518 (0.060)	0.364 (0.019)	0.655 (0.021)	0.863 (0.023)
Party Discipline y^{max} , Republicans	0.342	0.373	0.482	0.600	0.440	0.340 (0.019)	0.376 (0.016)	0.481 (0.012)	0.600 (0.012)	0.440 (0.011)
Standard Deviation of Aggregate Shock σ_η			0.859					0.909 (0.177)		

N : 435
 T : 315 Whip Counted bills, 5424 Roll Called bills

Notes: Average estimates of the first step parameters from 100 Monte Carlo simulations that assume our estimates from Table 1 are the true values. Standard deviations of the estimates are reported in parentheses.

⁷⁵The simulations differ slightly from the true data-generating process in two respects. First, we hold constant the 435 members of Congress, rather than have members move in and out. Second, in order to obtain an estimate for σ , we don't use the estimated $MV_{2,t}$ for bills with both whip counts and roll calls. Instead, we use the estimated $MV_{1,t}$ plus a draw from a Normal distribution with standard deviation equal to our estimated σ .

FIGURE E.1. Monte Carlo Estimates for Ideologies



Notes: We compare the kernel densities of ideologies, θ_i : true values for our Monte Carlo simulations (in black) and our average estimates across simulations (dashed blue). The average standard deviation in θ_i is 0.049.

E.1.2. *Monte Carlos for the Second Step.* As with the first step, we simulate a dataset of the same size as our actual dataset under the assumption that our second step estimates are the true values. We report the average estimates across $R = 100$ runs.⁷⁶ Our estimator demonstrates good finite-sample properties with the average estimates of all parameters being within one standard deviation of the true values.

⁷⁶Some estimates fail to converge (the convergence rate is over 80%), so we run $R = 150$ simulations and report results over the first 100 convergent runs.

TABLE E.2. Monte Carlo Simulations - Second Step Estimation

	True Values					Average Monte Carlo Estimates				
	Congress					Congress				
	95	96	97	98	99	95	96	97	98	99
Probability Democrat is Proposer, γ			0.427					0.438 (0.018)		
Status Quo Distribution (Mean), μ_q	-0.188	-0.227	-0.237	0.045	-0.125	-0.190 (0.122)	-0.142 (0.121)	-0.112 (0.129)	0.068 (0.106)	-0.059 (0.131)
Status Quo Distribution (Standard Deviation), σ_q	2.222	1.816	1.937	1.354	1.252	2.156 (0.143)	1.811 (0.126)	1.973 (0.145)	1.461 (0.143)	1.346 (0.139)
Truncation Points - Democrats										
$q_{l,D}$	-1.585,-1.528	-1.624,-1.572	-1.768,-1.666	-1.507,-1.470	-1.657,-1.617	-1.846,-1.774 (0.207),(0.209)	-1.801,-1.725 (0.194),(0.188)	-1.930,-1.790 (0.186),(0.184)	-1.815,-1.733 (0.195),(0.188)	-1.909,-1.821 (0.230),(0.216)
$q_{r,D}$	-1.220,-0.601	-1.253,-1.018	-1.285,-1.081	-0.906,-0.476	-1.233,-0.915	-0.858,-0.308 (0.307),(0.217)	-0.926,-0.704 (0.291),(0.264)	-1.008,-0.793 (0.247),(0.229)	-0.892,-0.388 (0.395),(0.280)	-1.019,-0.703 (0.260),(0.211)
Truncation Points - Republicans										
$q_{l,R}$	-0.609,-0.365	-0.657,-0.589	-0.886, N/A	-0.531, N/A	-0.508,N/A	-0.757,-0.525 (0.213),(0.247)	-0.692,-0.607 (0.224),(0.241)	-0.769,N/A (0.283),(.)	-0.548,N/A (0.225),(.)	-0.502, N/A (0.255),(.)
$q_{r,R}$	0.242,0.260	0.237, 0.251	N/A, 0.415	N/A, 0.395	N/A, 0.418	0.411,0.439 (0.198),(0.195)	0.435,0.463 (0.204),(0.203)	N/A, 0.591 (.), (0.203)	N/A, 0.512 (.), (0.191)	N/A, 0.603 (.), (0.187)

Notes: Average estimates of the second step parameters from 100 Monte Carlo simulations that assume our estimates from Table 4 are the true values. Standard deviations of the estimates are reported in parentheses.

E.2. Large-Sample (Asymptotic) Properties. Our first stage estimator is a panel-data probit with fixed effects in both the cross-sectional and time-series dimensions. The cross-sectional dimension, N , is the number of politicians (in the hundreds) and the time-series dimension T is the number of votes (in the thousands). We take asymptotics as both N and T increase (N/T going to a constant), since an increased sample size (e.g. additional Congresses) would include both new politicians and votes.

In this nonlinear MLE set-up with large N and large T , the main concern for consistency is the nuisance parameter problem (Neyman and Scott (1948)), due to the fact that the number of parameters increases with N, T . Nevertheless, as shown in Fernández-Val and Weidner (2016), the bias from the nuisance parameter problem in such a set-up is of the order of $o(1/N) + o(1/T)$. This bias is negligible for large N and large T as Fernández-Val and Weidner (2016) show via Monte Carlo simulations of a probit with fixed effects at both the individual and time levels, and is confirmed in our simulations above.⁷⁷ Hence, we obtain consistency of our estimator and can invoke asymptotic normality under standard regularity conditions.

Our second step estimator follows a more standard asymptotic analysis, noting that some of the parameters in the second step likelihood have been previously estimated. This fact means that our inference must account for previous estimation errors. The asymptotic analysis of this type of 2-step MLE is addressed, for example, in Wooldridge (2010). Given the consistency of our first step estimates, and the smooth behavior of our second step likelihood in these parameters, Wooldridge (2010) shows that consistency for the latter parameters follows from standard regularity conditions. Due to the analytical complexity of calculating the standard errors for the second step (which include the standard errors from the first step), we use a resampling procedure that incorporates the estimation errors from the first stage, as described in the notes of Table 4.

⁷⁷To clarify, this argument also holds for the estimator of the ideologies, θ_i because ideologies are identified by the share of Yes votes across roll calls given y^{max} (obtained from changes in No/Yes votes across whip counts to roll calls). Hence, even though for certain Congresses there are few whip counted proposals (e.g. 28 in Congress 97), information about many bills is used for ideologies, allowing us to invoke the large sample asymptotics for θ_i and the other fixed effects.

APPENDIX F. (NOT FOR PUBLICATION) ADDITIONAL TABLES AND FIGURES

TABLE F.1. Number of Whips per Party

Whips	Congress				
	95	96	97	98	99
Democrats (appointed)	14	14	20	26	41
Democrats (elected)	21	23	23	23	23
Republicans (appointed)	16	17	23	22	25

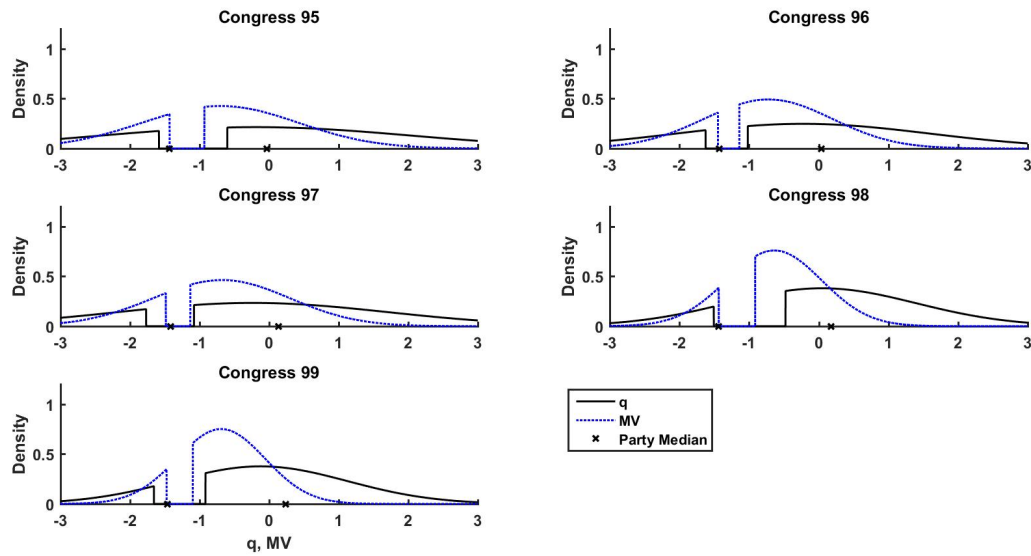
Notes: The table presents the number of whips per Party over the different Congresses. Data is from Meinke (2008). Both party leaderships appointed whips, however, the Democrats also elected a number of whips. Between the 95th and 106th Congresses, the Democrats also elected assistant/zone whips independently of the party leaders (Meinke (2008)).

TABLE F.2. Summary Statistics on Bill Selection

	Congress				
	95	96	97*	98*	99*
A: Total Number of Bills Whip Counted	131	58	28	50	48
B: Number of Bills Whip Counted, but not Roll Called	50	16	8	15	13
C: Total Number of Bills Roll Called	1540	1276	812	906	890

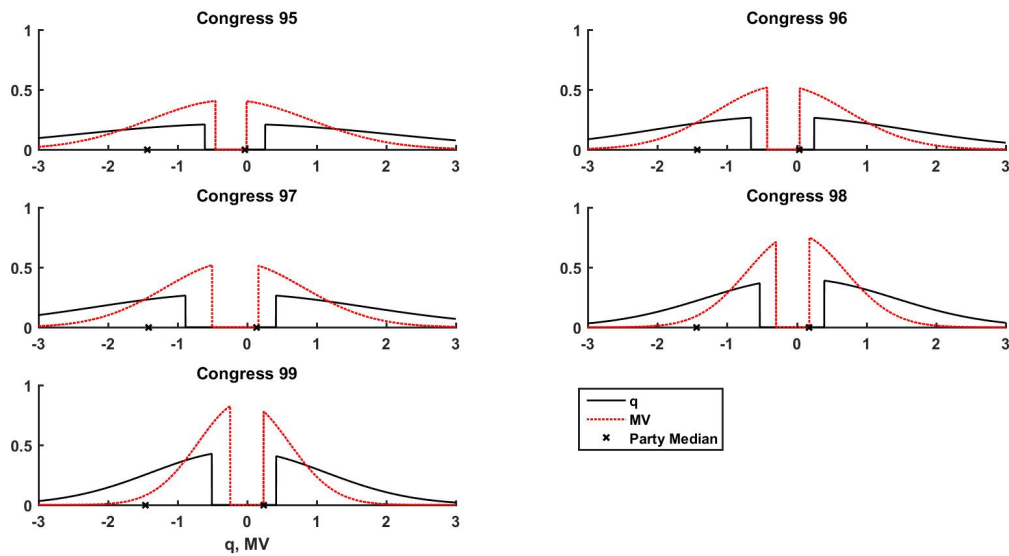
Notes: Number of bills whip counted, whip counted but not roll called, and roll called over Congresses 95-99. *We do not have data for Republican Whip Counts for Congresses 97-99 (see Section 5).

FIGURE F.1. Marginal Voter Distributions: Democrats



Notes: Optimal marginal voters (voters indifferent between status quo and optimal alternative) for Democrats as proposer (solid lines), with the status quo distribution (dashed lines) for reference.

FIGURE F.2. Marginal Voter Distributions: Republicans



Notes: Optimal marginal voters (voters indifferent between status quo and optimal alternative) for Republicans as proposer (solid lines), with the status quo distribution (dashed lines) for reference.

TABLE F.3. Likelihood Ratio Test for Constant y_{max}

Model	Estimated y_{max}	Log-Likelihood
Time Varying y_{max}	See Table 2	-7.940×10^5
Constant y_{max}	Dem: 0.523, Rep: 0.439	-8.441×10^5
p-value for LR test, with 8 degrees of freedom: 0.00		

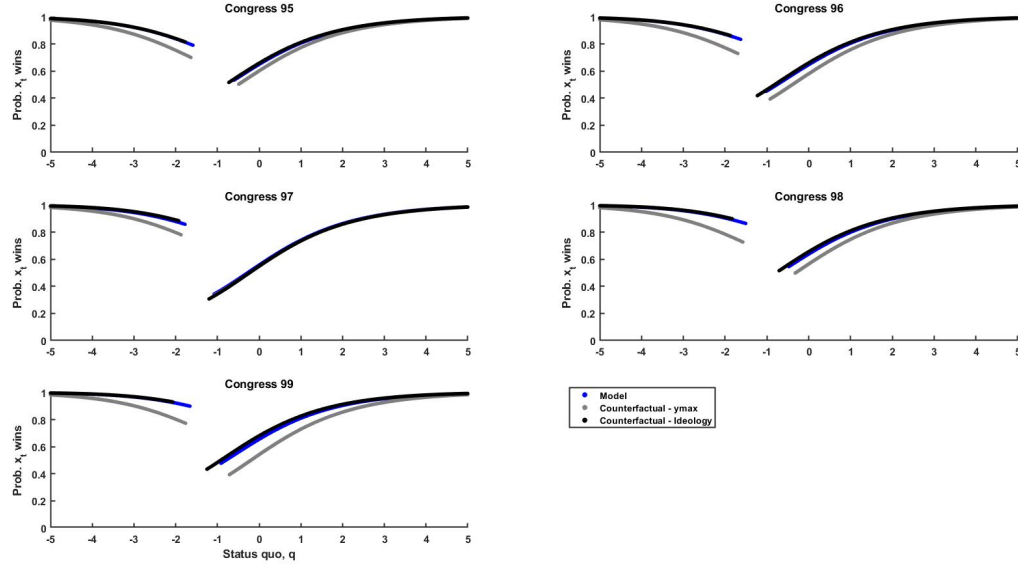
Notes: We test whether the whipping parameter, y^{max} , is constant across all Congresses in our sample. To do so, we fit a restricted version of our model in which each party's y^{max} is the same throughout all Congresses. We compare the restricted model to our original model, and reject the hypothesis of a constant y^{max} with a Likelihood Ratio test.

TABLE F.4. Counterfactual with polarized ideologies: Decomposition

	Congress				
	95	96	97	98	99
A: Polarization due to ideology ($\theta_R^m - \theta_D^m$)	1.758	1.923	1.978	2.244	2.351
B: Polarization due to whipping ($y_R^{max} + y_D^{max}$)	0.725	0.899	0.848	1.258	1.305
C: Share of Polarization due to whipping (B/(A+B))	0.292	0.319	0.300	0.359	0.357

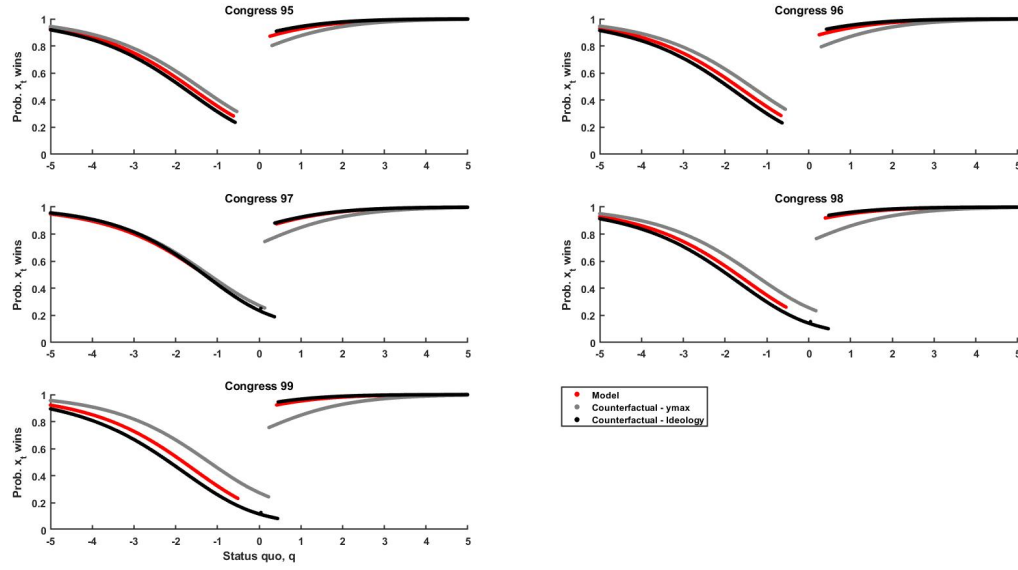
Notes: Decomposition of polarization into its ideological and party discipline components under the counterfactual in which we assume ideologies are farther apart than they actually are (we subtract $y_D^{max}/2$ from Democratic ideal points and add $y_R^{max}/2$ to Republican ideal points). Under the counterfactual, party discipline accounts for around 30% of polarization, compared to 40% in the main model (See Table 2).

FIGURE F.3. Probability of Bill Approval for the Democrats, Main Model and Counterfactuals



Notes: Probability of bill approval for Democratic proposals as a function of the status quo policy. We report the probability resulting from estimates of our main model and under two counterfactuals: (i) set $y^{max} = 0$ for both parties, and (ii) set the ideologies to more polarized values (new ideology equals $\theta_i + y_R^{max}/2$ for Republicans and $\theta_i - y_D^{max}/2$ for Democrats).

FIGURE F.4. Probability of Bill Approval for the Republicans, Main Model and Counterfactuals



Notes: Probability of bill approval for Republican proposals as a function of the status quo policy. We report the probability resulting from estimates of our main model and under two counterfactuals: (i) set $y^{max} = 0$ for both parties, and (ii) set the ideologies to more polarized values (new ideology equals $\theta_i + y_R^{max}/2$ for Republicans and $\theta_i - y_D^{max}/2$ for Democrats).

APPENDIX G. (NOT FOR PUBLICATION) DISCUSSION OF OPTIMAL CLASSIFICATION

This Appendix compares our parametric approach to estimation of legislators' ideal point with alternative statistical approaches. The political science literature on the estimation of ideal points $\{\theta_i\}$ in legislatures is vast, and characterized by several different econometric approaches, typically all within random utility environments. These approaches range from Bayesian, such as [Clinton et al. \(2004\)](#), to parametric ones based on Maximum Likelihood Estimation ([Poole and Rosenthal \(1997\)](#); [Heckman and Snyder \(1997\)](#)), and to nonparametric approaches based on the Maximum Score Estimator (MSE, [Manski \(1975\)](#), [Manski \(1988\)](#)) applied to this specific context (the Optimal Classification approach introduced in [Poole \(2000\)](#)).

Across all of these estimation techniques, an assumption crucial for consistency of the estimators is that party discipline is absent and that members of the legislature legislators “vote sincerely for the alternative that is closest to their ideal point” ([Poole \(2000\)](#)). This assumption has been recognized as problematic and worthy of attention in all the literature cited (e.g. [Clinton et al. \(2004\)](#), [Snyder and Groseclose \(2000\)](#)). In this article, we relax this assumption, instead modeling party discipline explicitly.

Absent an identification strategy designed to address the issue of party discipline, the relative sensitivity of extant approaches to a violation of this assumption on vote choices has been subject of ample discussion. For example, as reported by [Spirling and McLean \(2006\)](#), [Rosenthal and Voeten \(2004\)](#) argue that Optimal Classification (OC) “is preferable to parametric methods for studying many legislatures ... because the nature of party discipline, near-perfect spatial voting, and parliamentary institutions that provides [sic] incentives for strategic behavior lead to severe violations of the error assumptions underlying parametric methods.” In index models, relative to parametric approaches like MLE that assume independence of the random utility shocks, MSE does not rely on distributional assumptions or independence of covariates from the preference shocks.

However, MSE relies on the median error being zero conditional on covariates ([Wooldridge \(2010\)](#)). This means MSE still requires a strict exogeneity assumption, akin to conditional zero mean error in OLS or MLE, which is violated if party discipline is omitted from the vote decision equation. As with MSE, OC cannot achieve consistency in estimation without such an assumption.

Further, while MSE might weaken parametric assumptions, it is characterized by poor statistical properties (e.g. cube-root convergence, non-Normal asymptotic distributions, larger confidence intervals, may display convergence issues due to a discrete objective function compared to concave one, etc.).

Rosenthal and Voeten (2004) attempt to use OC, in the National Assembly of the French Fourth Republic, in a context where party discipline is present. However, Spirling and McLean (2006) show that OC fails to deliver meaningful rank orderings for the modern House of Commons in the UK.

APPENDIX H. (NOT FOR PUBLICATION) ADDITIONAL DISCUSSION OF SOME KEY BILLS USED IN THE COUNTERFACTUALS

H.O.1. *The National Energy Act of 1978*. In April 1977, Jimmy Carter introduced a National Energy Plan he considered to be the defining issue of his presidency (Richardson and Nordhaus (1995)). For reference, oil imports were at around 50% of U.S. oil consumption, during the Cold War and post Oil Crisis when oil self-sufficiency was considered extremely important and within a crisis in the energy sector. This crisis extended beyond oil, with failures in regulation and institutions being blamed for imbalances and crises in that sector, including the supply shortages of 1976-77 (see Richardson and Nordhaus (1995)).

Carter's proposal was extensive, redesigning taxes, regulation and incentives in the energy sector. The proposal included a tax of 5 cents per gallon of gasoline and fuel efficiency standards on automobiles, buildings and home appliances; tax credits to renewable energy; the creation of the Department of Energy; the phasing out of major subsidies and distortions on oil prices; the deregulation of the natural gas sector and the a redesign of the utility rate structure (with peak-load pricing replacing subsidies). With such measures, Carter aimed to curb U.S. dependence on oil. As could be expected, the bill was subject to controversy and political battles.

We have data on H.R. 8444, the National Energy Act of 1978 that passed the House. Its text, as passed in the House, was incorporated in H.R. 5146 in the Senate, with certain other provisions inserted in H.R. 4018, H.R. 5289 and H.R. 5037 (Richardson and Nordhaus (1995)).

H.O.2. *The Panama Canal Treaty*. According to Skidmore (1993); “*The Panama Canal treaties proved among the most contentious pieces of legislation in American history*”. As Skidmore (1993) describes, the legislation was seen as a battleground for the future of U.S. foreign policy. While President Carter pushed foreign policy geared at a step back from military influence, conservatives mobilized to influence public opinion. The negotiations provided two treaties: one leading to the gradual transfer of the Panama Canal from U.S. control to Panama, and the other providing U.S. defense rights and neutral status for the Canal after 2000 (Skidmore (1993)). Conservative groups spent significant resources to sway public opinion, mostly against the treaties (Smith III and Hogan

(1987)). The Carter administration focused on political deals to arrive at the numbers needed for passage.