SOCIAL INTERACTIONS AND LEGISLATIVE ACTIVITY

NATHAN CANEN, MATTHEW O. JACKSON, AND FRANCESCO TREBBI

Abstract. We develop a model of social interactions, as well as strategic interactions that depend on such social activity, and use it to measure social complementarities in the legislative process. Our model allows for partisan bias and homophily in the formation of relationships, which then impact legislative output. We use it to show how increased electoral competition can induce increased social behavior and the nonlinear effects of political polarization on legislative activity. We identify and structurally estimate our model using data on social and legislative efforts of members of each of the 105th-110th U.S. Congresses (1997-2009). We find large spillover effects in the form of complementarities between the efforts of politicians, both within and across parties. Although partisanship and preference differences between parties are significant drivers of socializing, our empirical evidence paints a less polarized picture of the informal connections of legislators than typically emerges from legislative votes alone.

Date: July, 2022

Canen: Department of Economics, University of Houston.
E-mail: ncanen@uh.edu.

Jackson: Department of Economics, Stanford University and External Faculty of Santa Fe Institute.
E-mail: jacksonm@stanford.edu.

Trebbi: University of California - Berkeley, Haas School of Business; CEPR and NBER
E-mail: ftrebbi@berkeley.edu.

Kareem Carr and Juan Felipe Riaño provided excellent assistance in the data collection. The authors would like to thank Matilde Bombardini, Gabriel Lopez-Moctezuma, and seminar participants at various institutions for their comments and suggestions. Particularly, we thank Antonio Cabrales and Yves Zenou for fruitful discussion of their model. Jackson gratefully acknowledges support from the NSF under grants SES-1629446 and SES-2018554. Trebbi gratefully acknowledges support by the Canadian Institute For Advanced Research and the Social Sciences and Humanities Research Council of Canada. Previous versions of this work circulated under the title "Endogenous Networks and Legislative Activity" and "Endogenous Network Formation in Congress".

1. Introduction

Deliberative bodies, especially large ones, rely on informal interactions in order to function productively. Individuals form relationships with each other to craft and pass legislation. Because of the salient role of interpersonal ties in the legislative process, its study dates at least to the 1930s (Routt, 1938), but only in the last fifteen years has research on this topic grown in prominence (Lazer, 2011).

The challenge of simultaneously modeling such social behavior and political decision-making is one reason for this delay. Interpersonal ties are not drawn at random – legislators strategically choose how much and whom to socialize with (e.g., with whom to cooperate and collaborate). In turn, benefits are strategic as well – having key allies enables a politician to craft and pass legislation that would otherwise not be possible. Finally, these decisions are made in an environment rife with identity-based (party) affiliation with an immense number of potential connections, making empirical analysis challenging. The construction and analysis of the model presented here addresses these challenges.

Our model simultaneously incorporates such social activity based on individual choices, strategic decisions made which anticipate such socialization, and homophily.[1] We prove statistical identification of the parameters driving each of these features and show how various predictions in the model, including the role of electoral competition in increasing pro-social behavior and the role of political polarization in equilibrium effort levels, are validated in the data. This is done using both reduced-form and structural empirical methods.

Our model generalizes the tractable and powerful framework of Cabrales, Calvó-Armengol, and Zenou (2011) in several important directions. As in Cabrales et al. (2011), our model has two strategic choices: legislators choose both how much socializing to do with other politicians as well as how much effort to exert crafting and passing legislation. Socializing efforts result in formed relationships that increase the success of legislative efforts, and so social and legislative efforts are complements. Importantly, social and legislative efforts are also complementary to those of the other politicians with whom a given politician has ties, both within and outside of his/her party. The two main generalizations in our model are as follows. First, while in Cabrales et al. (2011) relationships form completely at random, our model admits homophily – allowing social ties to form at a different rate within and across groups. Legislators can collaborate with members of their own party at a different rate than with members of the opposition. Second, we allow the returns to social and legislative efforts to be individual-specific. This captures important institutional or time-specific differences across parties, including who holds a majority, as well as individual differences, such as one's home district's electoral competitiveness, all of which can affect the returns to effort. The model is also general enough to accommodate multiple distinct groups.

Our model provides comparative statics which are corroborated by our empirical analysis. In a first result, we show how a district's electoral competitiveness can change a politician's incentives to socialize. An increase in electoral competition is associated with an increase

[1]This model is the first to capture all of these features, which should be useful beyond the application to legislative production. In Section B.2, we compare our model to those in the literature and, in the conclusion, we discuss applications beyond political economy.

in the value of passing a legislative bill. As the latter is rewarded by voters, legislators have a higher return to social connections and legislative effort, which may spill over to their connections. We find robust validation of this prediction in the data. In a second illustration, we show that increases in bipartisanship (i.e., cross-party social interactions) do not imply unambiguous increases in governmental activity and legislative success. A party can benefit from being less exposed to less-engaged types in the opposition: ringfencing their members may induce higher effort levels and bill passage rates. This insight has welfare implications regarding polarization that would not arise without such network spillovers.

We structurally estimate our model employing data on cosponsorship and legislative efforts of members of House of Representatives from the 105th-110th U.S. Congresses. Using the model enables us to identify, quantify and distinguish social effort (forming relationships) and legislative effort (crafting and passing legislation), and the partisan forces as well as complementarities driving them. Our model is particularly apt at capturing social efforts in political systems like the U.S. one where individual legislators maintain a certain degree of individual independence from the party. This is because, even within the context of party caucuses and bipartisan meetings which we model explicitly, politicians maximize their individual utility, and not party-specific objectives. This is obviously an approximation, even for the American case, but one that may be less appropriate for strong party systems like the Westminster one.

A first empirical result is that the complementarities among politicians are significant and stable across our sample period. The estimated social marginal multiplier on legislative effort is between a tenth and a quarter of the direct incentive for legislative effort, with larger values for Democrats.[2] This means that a nontrivial fraction of the incentives for efforts of politicians appears to be driven by what other connected politicians are doing.

A second empirical finding is that the base payoffs from passing legislation differ significantly across parties. The two parties have different base payoffs from passing legislation, both in terms of average and variance (both higher for Democrats). These differences lead to higher levels of social and legislative efforts by Democrats, all else equal. They appear largely attributable to the increased electoral return for Democrats in passing bills, particularly when they are in the minority.

Third, we document the role of electoral incentives in socialization, focusing on heterogeneity among congressmembers. We show that politicians facing tougher electoral challenges also engage in more socialization, as they face a higher return to such effort.

Fourth, we also show that our model that imputes equilibrium socialization effort has better in-sample properties (model fit) than models of social interactions in a legislature based on exogenous or predetermined graphs, including those based solely on alumni connections and committee membership. As such, legislative behavior can be better explained by explicitly modeling the strategic decisions that drive these underlying networks.

[2]As is clear in the analysis that follows, the parameter multiplying the full product of social and legislative efforts ranges from 0.03 to 0.05, which when multiplied by other legislators' efforts of 4 to 6, and social efforts around 1, leads to a multiplier of 0.12 to 0.3. This is compared to direct incentives for legislative activity ranging from 1.0 to 1.4.

Fifth, we find evidence that partisan bias is empirically relevant: a model with biased interactions fits the data significantly better than a model with no bias. However, we also find that social interactions are far from being an exclusively partisan affair. Using cosponsorships in the U.S. Congress as a proxy for socialization levels,[3] we find that intermediate levels of partisanship – including partisan biases in the range of 10 percent – fit the data significantly better than a fully partisan model where 100 percent of interactions are exclusively within party. In fact, it is hard to reconcile the thousands of bipartisan cosponsorships in recent data with a hypothesis of unmitigated polarization between parties. The data are more nuanced than the common narrative of a legislature segregated along party lines that has emerged from recent literature mostly based on post-1980 congressional roll call evidence (Fiorina, 2017, see also Canen et al., 2020, 2022). The stark posturing and divisive language (Gentzkow and Shapiro, 2015; Gentzkow, 2017), and some metrics of formal political activity, miss bipartisan interaction that more informally takes place among legislators – especially with respect to the bulk of less controversial bills that constitute day-to-day law and budget making.

1.1. **Relation to the Literature.** From the theoretical perspective, this paper contributes to a literature that examines peer-influenced behavior when accounting for how networks are formed.[4] Our model allows for homophily, so that people interact more within groups than across groups, and also allows the value of social interaction to differ across groups *and* within groups. This meaningfully generalizes the model of Cabrales et al. (2011) to have group membership impact the value of social interaction and the rate at which that interaction happens within as opposed to across groups. Both factors matter significantly in our empirical application. Introducing a tractable and estimable form of asymmetry in the process of socializing of members of Congress within this framework should be valuable for other applications that involve multiple groups with homophily.[5] Although our model does not allow a politician to explicitly target their social effort choices, its empirical fit and theoretical results provide a foundation for future research to build upon.

This paper also contributes to the growing literature showing that social networks matter in legislative environments. For instance, Fowler (2006) uses a connectedness measure based on cosponsorships to show that more connected members of Congress are able to get more amendments approved and have more success on roll call votes on their sponsored bills.[6] Also

[3]The choice of cosponsorships as a measure of social effort is suggested by Fowler (2006), among others, due to it being an (observable) way of showing support for other Congress members with good predictive power for social-based outcomes (e.g., passing amendments, considered a form of legislative influence). However, unlike their work, we do not use cosponsorships as the de-facto network, but simply as a proxy for the equilibrium social effort that generates congressional networks. We revisit these points in detail below.

[4]See Bramoullé, Djebbari, and Fortin (2009), Mauleon et al. (2010), Goldsmith-Pinkham and Imbens (2013a), Goldsmith-Pinkham and Imbens (2013b), Manski (2013), Jackson (2013), Badev (2017), Jackson (2019), Hsieh and Lee (2016), Mele (2017), Baumann (2017) (and see Jackson (2005), Jackson (2008), Mauleon and Vannetelbosch (2016), Jackson and Zenou (2015) for surveys of the network formation and games on networks literatures).

[5]Homophily in peer group formation is also theoretically explored in Baccara and Yariv (2013), who further explore group stability.

[6]See also Zhang et al. (2008).

using cosponsorship links, Cho and Fowler (2010) show that Congress appears subdivided in multiple dense parts tied together by some intermediaries. These network features correlate with legislative productivity over time (number of important laws passed, as defined by Mayhew, 2005).[7]

The network analysis of legislation is growing, and the literature provides increasing evidence that social relationships matter substantially and are causal in nature. Kirkland (2011) shows a correlation between bill survival and weak ties of the sponsor for eight state legislatures and for the US House of Representatives. Cohen and Malloy (2014) employ identification restrictions aimed at ascertaining causal effects of networks on voting behavior (using the quasi-at-random seating arrangements of Freshman Senators).[8] Rogowski and Sinclair (2012) also use random spatial arrangements to estimate the causal effect of interactions on legislative voting and cosponsorship in the House. In particular, the authors use lottery office assignment affecting certain classes of members of the House of Representatives, not finding a significant affect of office proximity on co-behavior.[9] Harmon et al. (2019) study the role of exogenously shifted social connections within the European Parliament also using seating arrangements.

Importantly, none of those papers model interaction as a choice variable. In an important theoretical contribution, Squintani (2020) studies endogenous legislative networks, and the role of ideological positions on information transmission. Our focus is on legislative production rather than information transmission, and our analysis enables us to see how incentives to socialize differ across parties, relate to legislative productivity, and have changed over time. The role of political networks in connection to special interest politics is studied in Groll and Prummer (2016) and Battaglini and Patacchini (2018). Meanwhile, Battaglini et al. (2020) focus on legislative effectiveness of legislators based on their Bonacich centrality - taking it as exogenous, but employing a Heckman two-step procedure based on alumni

[7]There is other work on cosponsorship. For example, Alemán and Calvo (2013), Koger (2003), and Bratton and Rouse (2011) study the incentives for cosponsoring in different settings (focusing on ideological similarity, tenure, etc.). Beyond their role in social networks, Wilson and Young (1997) study the signaling content of cosponsorships, noting that cosponsorship is a cheap way of signaling to the median voter about one's congressional activity. They identify three different explanations for cosponsorships and their possible signaling impact: (i) bandwagoning (signaling strong support for the bill), (ii) ideology, and (iii) expertise. They find a null to moderate effect of cosponsorship on bill success, as measured as successive progress of the bills through Congress hurdles. Kessler and Krehbiel (1996) instead point out that the timing of cosponsorships would indicate that it is not as much a signaling to voters, as to other politicians (for example, they show that extremists seem to cosponsor earlier). Meanwhile, Anderson et al. (2003) find correlations with legislative productivity (i.e. the bill passing through different stages in Congress) for Congress member who sponsor more bills and use more floor time (albeit at a declining marginal rate).

[8]For a similar approach see also Masket (2008).

[9]They do not interpret these results as an absence of peer effects in Congress, but rather that office proximity - and the exogenous changes in connections caused by it - do not significantly explain congressional behavior. In fact, the most important network effects might be those from endogenously formed connections. However, these are hard to identify in reduced form and cannot be captured in their study. As they conclude, "Paying attention not only to the structure of networks but also to how that structure came to be can help remedy many of the difficulties in providing causal evidence for network effects." (p. 327). Our structural model fills this gap.

networks to correct for network endogeneity in the empirical analysis.[10] In a complementary effort, we present a tightly connected theoretical and empirical structure of strategic interactions and networks and how it contributes to legislators' decisions. This allows us to estimate socialization efforts and how they affect legislation, and enables us to do comparative static exercises based on the estimated model. In subsequent work, Battaglini et al. (2021) study a game of network formation with strategic decisions made on the graph in Congress. The authors focus on the choice of directed links using a setting akin to general equilibrium: while politicians can choose who to link to directly, in equilibrium the solution is characterized by "prices" that clear the market for connections. We differ by modeling strategic interactions within a game-theoretic framework and by proving identification (the focus in the cited paper is on Bayesian estimation).

2. The Model

2.1. **Legislators, Parties, and Partisanship.** The legislature is composed of a set $N = \{1, 2, ..., n\}$ of politicians. For simplicity, we focus on one chamber (e.g. the House), and clearly the model applies to a variety of deliberative bodies, legislatures, committees, and organizations.

The set of politicians N is partitioned into K parties, with a generic party ℓ denoted P_ℓ.

Each party P_ℓ has a level of partisanship $p_\ell \in [0, 1]$, which is a fraction, and can be thought of as a structural form of homophily. In particular, members of party ℓ spend a fraction p_ℓ of their total interaction at exclusively party ℓ events, so only mixing and meeting with own-party events, and the remaining fraction, $1 - p_\ell$, at events in which they mix with members of all parties. This can include party and caucus meetings, joint sessions, fund-raising events, committee works, social gatherings and formal events, etc. For our empirical application, the U.S. House of Representatives, examples of party-specific events are closed sessions called Party Conferences for Republicans and Party Caucuses for Democrats (their respective chairs represent the number 3 position of official party leadership rankings).

Politician i belongs to party $P(i)$, and $p(i)$ denotes the level of partisanship of politician i's party.

In our empirical analysis, there are $K = 2$ parties, $1, 2$, and then we index the n politicians so that the first q of them belong to $P_1 = \{1; ...; q\}$ and the remainder to $P_2 = \{q + 1; ...; n\}$. Let $q \geq n/2$, so that party 1 is the majority party.

However, we describe the equilibrium for the general case where K may differ from 2, and then specialize the solution later to the case of $K = 2$. While our empirical application interprets K as the two parties in the U.S. Congress, K can be thought more broadly in other contexts. This includes the case of multiple parties in other legislatures, or as groups when socialization is biased along geographical or social identity lines.

2.1.1. *Socializing.* Each politician chooses an effort level of socializing (i.e., interacting with other politicians), denoted $s_i \in I\!R_+$. Socializing forms connections with other politicians.

[10] In Appendix, we provide a similar reduced-form identification strategy, exploiting differential characteristics of networks within the same bill across Senate and House. We discuss this in more detail in Section 4.

The network $G = \{g_{i,j}\}_{i,j \in N}$, where g_{ij} denotes the strength of a pair i and j's connection, arises from the social efforts, $\mathbf{s}$, and is:[11]

$$g_{ij}(\mathbf{s}) = s_i s_j m_{ij}(\mathbf{s}),$$

where if $j \in P(i)$ then

$$m_{ij}(\mathbf{s}) = p(i)\frac{p(j)}{\sum_{k \in P(i), k \neq i} p(k)s_k} + (1-p(i))\frac{(1-p(j))}{\sum_{k \neq i}(1-p(k))s_k},$$

and if $j \notin P(i)$ then

$$m_{ij}(\mathbf{s}) = (1-p(i))\frac{(1-p(j))}{\sum_{k \neq i}(1-p(k))s_k}.$$

The term m_{ij} determines the baseline rate in which politicians within/across parties meet each other. The first equation in its definition reflects that politicians meet own-party members in two ways: at their own events and at general events. The second case is for members of opposing parties, who only meet at general events. Politicians are met with the relative frequency with which they are present at events.

However, the actual strength of connections (denoted $G = \{g_{i,j}\}_{i,j \in N}$) also depends on the relative socialization efforts by all politicians: it is not enough for there to be low homophily, because politicians must also want to interact with each other. The strength of such connections is increasing in both politicians' effort levels, since $g_{ij}(\mathbf{s}) = s_i s_j m_{ij}(\mathbf{s})$, and this is the case even conditional on their parties' baseline meeting rates. Note that each entry g_{ij} can be interpreted as a strength of connection because the total amount of connections for i satisfies $\sum_{j \neq i} s_i s_j m_{ij}(\mathbf{s}) = s_i$, so that it is proportional relative to s_i.[12]

Note that when $p_\ell = 0$ for each P_ℓ this simplifies to coincide with the model of Cabrales et al. (2011). When $p_\ell = 1$ for each P_ℓ, instead, each party is completely cut off from the other. Then, within each party again Cabrales et al. (2011) applies.

The specification above captures key features of how politicians socialize in practice, while maintaining tractability. First, members of the same party meet more often and, hence, are more likely to have stronger connections.[13] Second, more social members (those with higher s_i, which we show to be the higher types in equilibrium) are more likely to connect

[11] This is for the case in which $s_j > 0$ for at least two people in each party. If other agents are not putting in social effort, then there is nobody to match with, and some of these equations do not apply (they divide by 0). In those cases the matching is described as follows. If at most one $s_j > 0$, then set $m_{ij} = 0$ for all ij and the entire network equal to 0. If there are at least two people with $s_j > 0$, but also at least one party with $s_j > 0$ for no more than one agent, then set $m_{ij} = g_{i,j} = 0$ for all members of a party that does not have more than one $s_j > 0$, and use the remaining equations specified above for $g_{i,j}$'s for any other combinations.

[12] This follows from the definition of m_{ij} above, since the left hand side of the expression is equal to:

$$\sum_{j \neq i, j \in P(i)} s_i s_j \underbrace{\left(p(i)\frac{p(j)}{\sum_{k \in P(i), k \neq i} p(k)s_k} + (1-p(i))\frac{(1-p(j))}{\sum_{k \neq i}(1-p(k))s_k}\right)}_{m_{ij}\ if\ j \in P(i)} + \sum_{j \neq i, j \notin P(i)} s_i s_j \underbrace{(1-p(i))\frac{(1-p(j))}{\sum_{k \neq i}(1-p(k))s_k}}_{m_{ij}\ if\ j \notin P(i)}$$

$$= s_i\left(p(i) + (1-p(i))\left(\frac{\sum_{j \neq i, j \in P(i)} \sum_{j \neq i, j \notin P(i)}(1-p(j))s_j}{\sum_{k \neq i}(1-p(k))s_k}\right)\right) = s_i\left(p(i) + (1-p(i))\left(\frac{\sum_{j \neq i}(1-p(j))s_j}{\sum_{k \neq i}(1-p(k))s_k}\right)\right) = s_i.$$

[13] As previously described, examples of single party events include fund-raisers and party caucuses.

(and have a stronger connection) to members in their own party.[14] Third, socialization is not deterministic: observable decisions do not fully determine social connections. In this model, two socially active members with equal characteristics will not necessarily have strong connections. As in any activity, we allow for randomness to be a component of social relationships, although driven by efforts and socialization characteristics. In practice, personalities, unobserved characteristics and preferences play a large role in such connections, but chance may also play a role.[15] Furthermore, the events in which legislators interact are often unobserved by researchers and the public.[16]

Our specification with a biased random matching protocol allows us to capture such socialization features on aggregate, while maintaining a tractable framework in which we can characterize equilibria, obtain closed-form theoretical predictions, and test them empirically.[17] Such randomness in behavior is consistent with empirical models whereby realized networks are randomly drawn from a class of possible network structures.

Finally, we note that our model is able to generate a rich class of different network configurations, G, which arise naturally in equilibrium (described below) as a function of all the parameter configurations. Some examples are shown in Appendix B.3. While the $p(i)$s affect the overall meeting rate across parties, individual politicans have different incentives to socialize and legislate, and that can generate very heterogeneous patterns of socialization among and between parties. Some politicians socialize much more with the opposing party than other politicians, for instance, and the parameters α_i and ϕ_i affect not only the incentives to legislate, but also to socialize, and ends up affecting which politicians cross party lines most frequently.

Nevertheless, the current approach does assume that partisanship is exogenous. Another possibility would be to endogenize partisanship. Having three action variables for each party/agent renders the model even more challenging analytically. In this paper, we focus on the two that seem most important to endogenize and, in the empirical part of the paper, we estimate the third. At the end of Appendix B, we discuss some potential approaches to, and challenges with, endogenizing partisanship that could be addressed in future research.

---

[14]We explore these patterns empirically in a later section, showing that these higher types coincide with those in more influential committees, for example.

[15]This is also the case beyond the U.S. Congress. For instance, the alphabetical seating arrangement in the European Parliament changes legislators' likelihood of meeting and connecting with each other, and such chances influence legislative outcomes (Harmon et al., 2019). Similarly, lottery-based seating allocations in Iceland generate durable connections (Saia, 2018; Jo and Lowe, 2020). In Jo and Lowe (2020), such connections are measured using cosponsorships, a measure we use for the U.S. Congress.

[16]These include being at the gym at the same time, going to particular restaurants or attending certain cultural events. A journalistic description for the case of the Senate can be found in Roll Call's piece, https://www.rollcall.com/news/behind_the_doors_of_the_senate_gym-222790-1.html, accessed January 21, 2019. Meanwhile, Norton (2019) provides a thorough analysis of the role of informal gathering spaces (e.g. tea rooms, smoking rooms, pubs) in the political behavior of MP's in the U.K. House of Commons.

[17]As an alternative one could consider a model of links g_{ij}, where politician i identifies a specific partner j deterministically and based on j's other choices. This appears less tractable once one introduces any forms of heterogeneity. Our framework instead allows for closed-form solutions and for a realistic modicum of stochasticity in the formation of certain social ties. We acknowledge that our approach is a first step that empirically accommodates such diversity of strategic interactions.

2.2. **Legislative Effort and Preferences.** The other choice of politicians is their legislative effort $x_i \in I\!R_+$. The benefits from legislative efforts are described by:

$$\alpha_i x_i + \phi_i \sum_{j \neq i} s_i s_j m_{ij}(\mathbf{s}) x_i x_j.$$

As in a large class of models, of which Cabrales et al. (2011) is a salient instance, there is a direct benefit from private effort, with idiosyncratic weight α_i. In addition, there are complementarities in legislative efforts between politicians who have formed connections: the more effort they both expend, the more likely their legislation is to pass. The size of this interaction effect is governed by a legislator-specific parameter, ϕ_i, whose quantification is a relevant goal in the empirical analysis that follows. Note that this interaction effect can vary at the individual level depending on each legislator's incentives. In the next section, we extend the model and show that the electoral competitiveness of i's district and the likelihood of a good electoral or bill-level "shock" can induce different values of ϕ_i. Such values of ϕ_i may be correlated, but not necessarily identical, within a party.

Politicians choose x_i and s_i simultaneously, and both forms of effort are costly for a politician. The cost of legislative effort is given by $\frac{c}{2}x_i^2$, with $c > 0$, and the total cost of socializing is given by $\frac{1}{2}s_i^2$. The parameter c governs the relative cost of legislative effort to social effort.

Taken together, the politician's preferences are the amount of legislation that he or she produces less the costs of legislative and social efforts. This is given by:

$$\tilde{u}_i(x_i, x_{-i}, s_i, s_{-i}) = \alpha_i x_i + \phi_i \sum_{j \neq i} s_i s_j m_{ij}(\mathbf{s}) x_i x_j - \frac{c}{2} x_i^2 - \frac{1}{2} s_i^2. \tag{2.1}$$

If G was exogenous and known, equation (2.1) would collapse to the set-up in Ballester et al. (2006). As a result, equilibrium legislative efforts would be proportional to a weighted measure of the politician's centrality. If that were the case, we could base an empirical specification on this foundation, as done by Acemoglu et al. (2015) in the context of public good provision on a geographical network. However, in this paper, we assume that the politician network results from (strategic) socialization choices. As a result, we must model the choice of social effort that generates the underlying network in a way that can be identified in the data. The current set-up accomplishes this through equilibrium restrictions and additional data in an appropriate way, as we show in the next sections.

2.3. **An Electoral Motive for Socialization.** There are many different ways to justify the preferences in (2.1), as the natural presumption is that politicians care to maximize the legislation that they pass. Here, we provide an example of an electoral motive that extends the model above in a two-period set-up, generating new predictions that are empirically tested below.

Politicians care about being reelected and can affect the probability of being reelected by exerting effort in the legislature and by building connections instrumental to having specific legislation passed (e.g. policy favorable to the politician's constituents), as in Mayhew (1974), for instance.

Each politician anticipates these effects on his/her reelection chances. More specifically, each cycle is composed of a legislative stage, in which politicians choose and undertake legislative activity, followed by an electoral stage where voters choose whether to reelect legislators given their behavior. The latter stage provides the incentives that drive activity in the former. Politicians are career motivated and exert costly efforts with the aim of increasing their chances of being reelected. We normalize the politician's payoff from not being reelected to 0.

In their term, each legislative member can present a policy proposal, which for brevity we refer to as a "bill". The bill consists of a policy goal the legislative member intends to fulfill, for instance passing a statute targeted to his or her constituency, landing a subsidy, or obtaining an earmark beneficial to firms in the home district. We describe below how getting i's policy goal fulfilled maps into an increase in i's chances of being reelected.

The timing of the extended game is as follows: during the first stage (their term) the legislator simultaneously chooses (s_i, x_i) before the election to maximize the probability of being reelected. They cannot guarantee reelection because of bill-level and electoral-level uncertainty, described below. In the second, an election occurs which depends on a politician's electoral approval rate which, in turn, depends on whether it passed the bill and its electoral competitiveness. We take politicians to be myopic in the sense that they only consider reelection for the next cycle, so that existing networks do not influence utilities after the election, and thus voters must be backward-looking.[18]

Suppose a politician's utility is given by:

$$u_i = Pr(reelected) - c_i(x_i) - \frac{1}{2}s_i^2, \tag{2.2}$$

where $c_i(x_i)$ is the relative cost of legislative effort vis-à-vis social effort for politician i.[19]

We assume that the cost of legislative effort is heterogeneous across politicians and given by:

$$c_i(x_i) = \frac{1}{2}cx_i^2 - \alpha_i x_i. \tag{2.3}$$

This cost function is strictly convex, possibly heterogeneous across politicians, and such that those with higher types α_i have lower marginal costs of legislative effort vis-à-vis socialization (the difference in marginal costs between i and j is given by $\alpha_j - \alpha_i$). Those with higher types also have lower absolute costs to legislative effort. Hence, we interpret the distribution of α_i as the distribution of politician types. Since an elected politician can always choose $s_i = x_i = 0$, they always weakly prefer to be reelected.

[18] As there is complete information in this model, politician types are observed and, hence, there is no role for learning. Reelection is backwards looking to incentivize higher effort levels which may improve voter welfare, as voter welfare increases with politician effort (see V_i, below). Meanwhile, politician welfare also increases with effort levels because the costs of effort are quadratic, but the returns to effort are cubic.

[19] Equation (2.2) assumes that politician i has the same relative preferences for reelection and the cost of socialization. While this is unnecessary for theory (an additional parameter on either term would simply rescale other parameters in equation (2.6)), it matters empirically, as the latter would not be separately identified. We pursue the current version, which is what we produce empirically. Alternatively, one can interpret either ϕ_i or $\gamma_{P(i)}$ as scaled relative to that additional parameter.

The choice of x_i, the level of legislative activity exerted by i, affects the support for i's legislation, Y_i, through a function:

$$Y_i = \varepsilon_i x_i \left(\sum_{j \in N} g_{i,j}(\mathbf{s}) x_j \right).$$

Both i's own legislative effort, x_i, and that of his or her connections in the network, $\sum_{j \in N} g_{i,j}(\mathbf{s}) x_j$, matter for the ultimate support received by i's bill. The bill is approved if $Y_i > m$, where $m > 0$ is a generic institutional threshold.[20]

Y_i is stochastic and also depends on a random shock ε_i, assumed to be standard Pareto distributed with scale parameter $m\gamma_{P(i)} > 0$ and i.i.d. across politicians. We allow γ to be party-specific, reflecting the different uncertainty (probability distributions) that parties face to passing a bill, even conditional on effort.[21] We assume that ε_i is realized after the choice of $\mathbf{x}$, the vector of x_j across all politicians $j \in N$. Because ε_i is a shock following the realized legislative support, i must take expectations over its value when choosing (s_i, x_i).[22]

The probability of having the bill approved is thus given by:

$$Pr(Y_i > m) = \Pr\left(\varepsilon_i > \frac{m}{x_i \left(\sum_{j \in N} g_{i,j}(\mathbf{s}) x_j \right)} \right) = \gamma_{P(i)} \left(\sum_{j \in N} g_{i,j}(\mathbf{s}) x_j \right) x_i, \tag{2.4}$$

where we use the distributional assumption on ε.[23] Actual passage of the bill sponsored by i is represented by the indicator function $I_{[Y_i > m]}$.

To get reelected, the politician must have an approval rate in his/her electoral district that is sufficiently large. Similarly to Bartels (1993), the electoral approval rate of i is modeled as a random variable V_i which depends on past electoral performance/the historical competitiveness of the district, $V_{i,0}$, on whether the legislator approved the desired policy, and on a random variable η_i:

$$V_i = \rho V_{i,0} + \zeta_{P(i)} I_{[Y_i > m]} - \eta_i,$$

where η_i is assumed to follow an Exponential distribution, with parameter λ. Since reelection occurs if this $V_i > 0$ (i.e. $\rho V_{i,0} + \zeta_{P(i)} I_{[Y_i > m]} > \eta_i$), η_i can be interpreted as a random variable which captures popularity shocks or uncertain turnout in large elections (e.g. Myerson, 1998).

This set-up allows for approval rates to be persistent, but also to react when a politician is capable of getting a bill approved $I_{[Y_i > m]}$. The parameter $\rho > 0$ measures this persistence

[20]Naturally, m can be function of a simple majority requirement or even supermajority restrictions.

[21]In this formulation, $\gamma_{P(i)}$ is interpretable relative to the institutional threshold m.

[22]Notice that each link between politician i and j is a function of the endogenous social efforts of everybody else, hence the dependency $g_{i,j}(\mathbf{s})$ on $\mathbf{s}$, the vector of s_j efforts across all politicians $j \in N$.

[23]More generally, one can take Y_i to represent the average approval rate of i's multiple bills. In this case, each b is a separate bill by a politician i. The conditions for our model are unchanged, as long as bills are not strategically introduced (i.e., the shocks ε_b are still i.i.d. within i).

in winning margins/past electoral races and implies that electoral competitiveness can incentivize legislative behavior. Meanwhile, the parameter ζ, which could be equal to zero empirically, governs the relative importance of a bill actually passing vis-à-vis legislative effort. Notice that, while there is no direct value to the voters of the politician having more socializing, the value of s_i matters implicitly, being instrumental in getting legislation passed.

We solve the legislator's decision by backwards induction. After observing whether the bill-passed or not, the probability of reelection is:

$$Pr(\rho V_{i,0} + \zeta_{P(i)} I_{(Y_i > m)} > \eta_i) = 1 - e^{-\lambda(\rho V_{i,0} + \zeta_{P(i)} I_{(Y_i > m)})}$$

where we use the distributional assumption on η_i.

As the legislator simultaneously chooses (s_i, x_i) before knowing whether the bill has passed (i.e. before the Pareto distributed shock ε_i is realized), the probability of reelection when choosing (s_i, x_i) is given by:

$$\begin{aligned} Pr(reelected) &= (1 - e^{-\lambda(\rho V_{i,0} + \zeta_{P(i)})}) P(Y_i > m) + (1 - e^{-\lambda \rho V_{i,0}})(1 - P(Y_i > m)) \\ &= 1 - e^{-\lambda \rho V_{i,0}} + P(Y_i > m)(e^{-\lambda \rho V_{i,0}} - e^{-\lambda(\rho V_{i,0} + \zeta_{P(i)})}) \\ &= 1 - e^{-\lambda \rho V_{i,0}} + \gamma_{P(i)}(1 - e^{-\lambda \zeta_{P(i)}}) e^{-\lambda \rho V_{i,0}} \sum_j g_{ij} x_i x_j \end{aligned} \tag{2.5}$$

where the third line substitutes the expression for $P(Y_i > m)$.

Replacing (2.3) and (2.5) and defining $\phi_i \equiv \gamma_{P(i)}(1 - e^{-\lambda \zeta_{P(i)}}) e^{-\lambda \rho V_{i,0}}$ into the utility function yields:

$$u_i(x_i, x_{-i}) = 1 - e^{-\lambda \rho V_{i,0}} + \phi_i \sum_j g_{ij} x_i x_j - (\frac{1}{2} c x_i^2 - \alpha_i x_i) - \frac{1}{2} s_i^2. \tag{2.6}$$

Since the term $1 - e^{-\lambda \rho V_{i,0}}$ does not affect the maximization problem, (2.6) can be rewritten as the specification given in (2.1).

We remark that the return to social effort, ϕ_i, is a function of three components in this set-up. It depends on (i) the likelihood of passing a bill conditional on effort, parameterized by $\gamma_{P(i)}$, (ii) the electoral returns to passing a bill, measured by $(1 - e^{-\lambda \zeta_{P(i)}})$, and (iii) electoral competition in i's district, driven by $\rho V_{i,0}$. The first term captures that, if the politician expects to have better draws of passing a bill given chosen effort levels, (s)he has higher returns from legislative effort. This is because for the same amount of effort, it is more likely to pass the bill and have the electoral reward for it. The second term has a similar interpretation to the first. It maps the higher electoral return from passing a bill $(\zeta_{P(i)})$ to a higher return to effort. Finally, the term $\rho V_{i,0}$ determines the role of electoral competitiveness: those in historically more competitive races (a lower $V_{i,0}$) face higher returns to effort and passing a bill than those in safe districts. While the first two terms depend only on party-level returns, the latter varies within a party. Electoral competition means that each legislator i may face a different return from social effort, even compared to their co-partisans.

We conclude this section by noting our model's timing. The efforts (s_i, x_i) are chosen simultaneously. This is one possible approach. An alternative could assume that social effort, s_i, is chosen before legislative effort, x_i. In this case, politicians will increase their investment in social choices anticipating their effect in then driving legislative choices. Then voters observe a larger network, suggestive of higher future legislative effort, and so forward-looking voters may then reelect legislators based on such networks, although they would have to forecast the set of future legislators across all possible races, as mentioned above. This may be a fruitful possibility for future research.

2.4. **Analysis and Comparative Statics.** We examine the pure strategy Nash equilibria of the game in which all politicians simultaneously choose s_i and x_i to maximize (2.1). We work with the same approximation as in Cabrales et al. (2011): We operate "at the limit", when the number of politicians grows.[24] In particular, we solve for equilibria under the assumption that the term $\sum_{j\neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*$ is the same for all i of the same party and study best responses and comparative statics of the equilibria.

The first order conditions with respect to s_i and x_i that characterize the best response of politician i imply that interior equilibrium levels of (s_i^*, x_i^*) must satisfy:[25]

$$s_i^* = \phi_i \sum_{j\neq i} s_j^* m_{ij}(\mathbf{s}^*) x_i^* x_j^* \tag{2.7}$$

and

$$c x_i^* = \alpha_i + \phi_i \sum_{j\neq i} s_i^* s_j^* m_{ij}(\mathbf{s}^*) x_j^*. \tag{2.8}$$

Substituting (2.8) into (2.7) yields

$$s_i^* = \frac{1}{c}\left(\alpha_i + s_i^* \phi_i \sum_{j\neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*\right) \phi_i \sum_{j\neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*$$

or

$$s_i^* = \frac{\alpha_i \phi_i \sum_{j\neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*}{c - \left(\phi_i \sum_{j\neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*\right)^2}. \tag{2.9}$$

Then, from (2.8) it follows that

$$x_i^* = \frac{\alpha_i}{c - \left(\phi_i \sum_{j\neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*\right)^2}. \tag{2.10}$$

Given these equations, we state some general predictions of our model. In the next section, we develop results that characterize the equilibria of this game and provide conditions for its existence.

PROPOSITION **2.1.** In any equilibrium of the game above:

(1) An increase in ϕ_i increases both equilibrium effort levels s_i^* and x_i^*.
(2) An increase in i's type, α_i, increases both s_i^* and x_i^*.

[24] Alternatively, this could be justified via a continuum of politicians of each type, or by examining an epsilon equilibrium with a large n.

[25] Note that second derivatives are everywhere negative.

(3) A decrease in the relative cost of legislative effort, c, increases both s_i^* and x_i^*.

These results follow from (2.9) and (2.10). In addition, the electoral motive in our model lets us decompose the first result: an increase in ϕ_i can be due to increases in (i) the probability of passing a bill, $\gamma_{P(i)}$, (ii) electoral returns to passing a bill, $\zeta_{P(i)}$, and (iii) an increase in electoral competition (a decrease in $V_{i,0}$). Hence, it follows that:

Corollary **2.1.1.** Politicians who face greater electoral competition (lower $V_{i,0}$, all else held equal) have higher equilibrium effort levels (s_i^*, x_i^*).

The results above apply to settings beyond the legislative one. They show that groups that have greater values from unilateral action (α_i) or socialization (ϕ_i), will generally do more of *both*, as they are complementary.

Meanwhile, increases in partisanship, $p(i)$, are not as easy to sign. They have ambiguous effects on equilibrium outcomes, as they depend on relative group sizes as well as all other parameters. Hence, we illustrate comparative statics in the context of 2 groups via simulations in Section 2.5.

For the results to follow, it is convenient to rewrite (2.7) as

$$\frac{s_i^*}{x_i^*} = \phi_i \sum_{j \neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*. \tag{2.11}$$

Since $\sum_{j \neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*$ is the same for all i of the same party in a "large n" approximation, $\frac{s_i^*}{x_i^*}$ is the same for all agents within a party. Using (2.11) in (2.8) yields:

$$\begin{aligned} c x_i^* &= \alpha_i + s_i^* \phi_i \sum_{j \neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^* \\ &= \alpha_i + \frac{s_i^{*2}}{x_i^*}. \end{aligned}$$

Dividing through by x_i^* implies that

$$c = \frac{\alpha_i}{x_i^*} + \frac{s_i^{*2}}{x_i^{*2}}. \tag{2.12}$$

2.5. **Characterizing Equilibria in a Simple Setting.** To fully characterize equilibria in order to get a closed form solution, we examine the case in which ϕ_i is the same for all the members of a given party. This decreases the number of ϕ_i parameters in the model from n to K and can be borne out from an assumption that either $\rho = 0$, or that electoral persistence/competition varies at the party-level (conditional on politician characteristics). While the latter assumption has been made frequently since at least Snyder (1989), for us, this is simply used to illustrate the properties of our model in a simpler setting. For instance, whether $\rho = 0$ is an empirical question for which we provide reduced form evidence in Section 4.2.1, and structurally test in Section 6.

Since $\frac{s_i^*}{x_i^*}$ is the same for all agents within a party provided that ϕ_i is the same for all agents within a party, (2.12) implies that $\frac{\alpha_i}{x_i^*}$ is the same for all agents within a party. This further implies that:

$$x_i^* = \alpha_i X_{P(i)},$$

for some $X_{P(i)}$. In addition, the fact that $\frac{s_i^*}{x_i^*}$ is the same for all agents within a party, implies that

$$s_i^* = \alpha_i S_{P(i)},$$

in equilibrium for some $S_{P(i)}$.[26]

To get explicit expressions compatible with our empirical analysis of the U.S. Congress, we now specialize the analysis to the case of $K = 2$ parties.

For each party $j = 1, 2$ define

$$\begin{aligned} A_j &= \sum_{i \in P_j} \alpha_i, \\ B_j &= \sum_{i \in P_j} \alpha_i^2. \end{aligned}$$

Thus,

Proposition **2.2.** The (interior) Nash equilibria of the limit game of the simple model are positive solutions to the system given by:

$$x_i^* = \alpha_i X_{P(i)}, \text{ and} \tag{2.13}$$

$$s_i^* = \alpha_i S_{P(i)}, \tag{2.14}$$

where

$$\frac{S_1}{X_1} = \phi_1 \left(\frac{p_1 B_1 X_1}{A_1} + \frac{(1-p_1)^2 B_1 S_1 X_1 + (1-p_1)(1-p_2) B_2 S_2 X_2}{(1-p_1) A_1 S_1 + (1-p_2) A_2 S_2} \right), \tag{2.15}$$

$$\frac{S_2}{X_2} = \phi_2 \left(\frac{p_2 B_2 X_2}{A_2} + \frac{(1-p_2)^2 B_2 S_2 X_2 + (1-p_1)(1-p_2) B_1 S_1 X_1}{(1-p_1) A_1 S_1 + (1-p_2) A_2 S_2} \right), \tag{2.16}$$

$$cX_1^2 = X_1 + S_1^2, \qquad cX_2^2 = X_2 + S_2^2. \tag{2.17}$$

All proofs appear in Appendix A.

If $p_1 = 1$ or $p_2 = 1$, then things reduce to the case of two separate parties with no interaction across them. That is, they are two copies of the model in Cabrales et al. (2011). Similarly, if $p_1 = p_2 = 0$ then there is no impact of party affiliation, and again the model simplifies to that of Cabrales et al. (2011). The novel case is when at least one partisanship level is positive, yet both levels are below 1. This biases the interaction of at least one party, leaving room for interaction across parties. In this case there will be both social mixing across different parties and partisanship in socializing.

[26]In contrast to results in network games with exogenous networks, equilibrium actions in our model are not expressed as only being proportional to a centrality measure of the network (e.g., Katz-Bonacich centrality). This results from the equilibrium interactions between social and legislative efforts. The legislative efforts still have to satisfy a version of the usual characterization on the margin. Therefore, an empirical approach using only centrality measures for estimation instead of our structural equations would only capture one dimension of the model's predictions.

Generally, there are multiple equilibria. This includes multiple interior equilibria, only some of which are stable.[27] In addition, there is always an (unstable) equilibrium in which $s_i = 0$ for all i. In that case, since no other politician provides effort, a given politician's efforts results in no connections and so the best response is also to provide no effort.[28] In our empirical specification below, we will focus on the interior stable equilibrium.

A sufficient condition for existence of an interior equilibrium is as follows.

Proposition 2.3. A sufficient condition for the existence of an interior equilibrium is

$$\frac{2c^{3/2}}{3\sqrt{3}} \geq \max\left[\phi_1, \phi_2\right] \max\left[\frac{B_1}{A_1}, \frac{B_2}{A_2}\right].$$

In this setting with two parties and nontrivial partisanship, there will generally be either two or four interior equilibria (except at a degenerate set of values where the system switches from two to four equilibria).[29]

For the case of more than two parties or groups, a similarly sufficiently large c is required to bound behaviors and ensure equilibrium existence, but as a function of many more parameters. More generally, the number of equilibria can grow exponentially in K, as the complementarities between groups' efforts, lead to a K-dimensional lattice. In practice, one can use the best response iteration described in Appendix B to guarantee existence of equilibria for specific parameter choices in the general set-up.

We now illustrate some of the comparative statics in partisanship, (p_1, p_2), in this simplified context. In contrast to the comparative statics in α_i characterized in Proposition 2.1, which apply quite generally, changes in $p(i)$ produce a more ambiguous result. Hence, they are difficult to sign in a general case. By changing $p(i)$, we change the rate that legislators meet members of their own party relative to those in other parties.[30] The final effect on equilibrium effort levels depends on the distribution of types in each party, the complementarity among them, as well as other parameters of the model. However, we can see how these move via some computations for specific parameter values. This is shown in Figure 1 below.

---

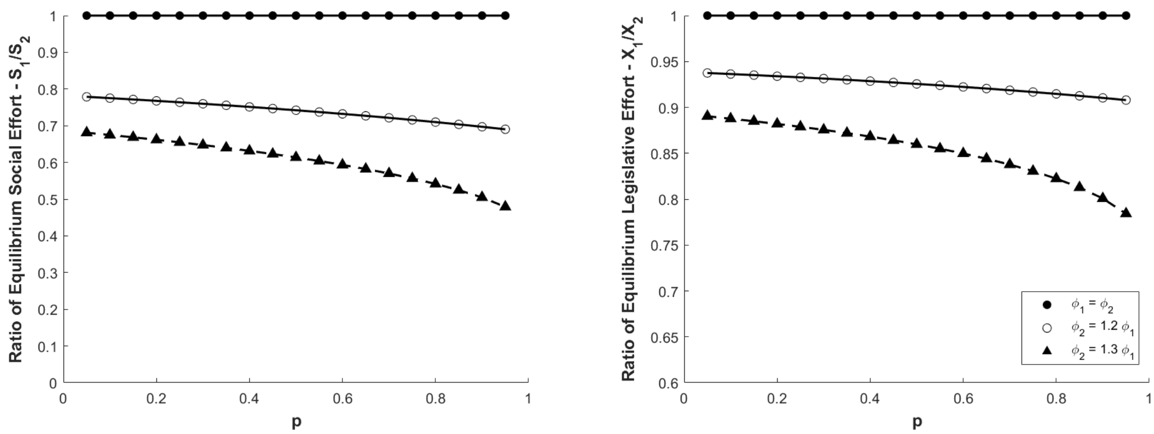
[27]This is in contradiction with Proposition 1 in Cabrales et al. (2011) which claims stability of all interior equilibria. In their model, contrary to the original proof, the largest equilibrium is unstable. In the proof of that proposition the matrix Π cannot be approximated by setting off-diagonal terms to 0. In fact, the eigenvalue can change sign if the off-diagonal terms are included and are on the order of $1/n$. This reverses their conclusion.

[28]Legislative effort, x_i is still provided in the model because there are direct incentives α_i for legislative effort, but no law is passed. Such an outcome is not desirable for politicians or voters as it is Pareto dominated by an interior (stable) equilibrium. Furthermore, this semi-corner equilibrium is unstable, in the sense that, were any politician to deviate to a positive social effort s_j, so would all the other politicians. The semi-corner equilibrium is not observed, given that we observe positive socialization empirically. Such a complete shutdown of socialization effort would be unstable, and so should not be observed for any length of time.

[29]These equilibria correspond to when both parties exert high levels or low levels of social efforts, and then for some parameters there are also two additional equilibria in which one party does medium-high and the other does medium-low socializing.

[30]As a result, increases in partisanship can be interpreted as increases in polarization. However, this "type" of polarization is at the party level, in contrast to changes to individual ideologies, which would be captured instead as changes to individual types, α_i.

Figure 1. Comparative Statics in Partisanship, $p(i)$



(a) Effect of Partisanship in Meeting Rates on Relative Social Effort S_1/S_2

(b) Effect of Partisanship in Meeting Rates on Relative Legislative Effort X_1/X_2

We present numerical simulations of how increases in partisanship affect equilibrium choices across parties. The left panel shows the effects on the ratio of equilibrium social effort, S_1/S_2, while the right one shows the ratio of equilibrium legislative effort, X_1/X_2. For simplicity, parties have identical types $\alpha_i = 1$ for all i, sizes ($n_1 = n_2 = 100$) and costs $c = 2.25$. We set partisanship to be symmetric, $p_1 = p_2 = p$. The only source of heterogeneity is the party-specific returns to social effort for Party 2, ϕ_2, which increases across specifications relative to $\phi_1 = 1$.

To emphasize the main mechanisms in the model, our simulations have two parties that are identical in sizes, types (i.e. $\alpha_i = 1$ for all i) and partisanship $p_1 = p_2 = p$. In the absence of any heterogeneity across parties (i.e. when $\phi_1 = \phi_2$ as well), parties will have the same equilibrium choices of effort and that does not vary with p. This is shown in the line with black circles in Figure 1. After all, a party's own members would be indifferent between interacting among themselves or with the opposition. However, as we increase the returns for social effort for one party relative to the other ($\phi_1 < \phi_2$ in the figure), the effect of partisanship increases. More interaction across parties – a low p – means legislators interact more across parties, pulling their relative equilibrium choices closer together. By contrast, when p is higher, legislators' choices are driven more by interaction with their own types. This amplifies the heterogeneity in equilibrium choices, which was first set by the different returns to social effort in each party. As p increases, effort levels in the party with the highest returns become proportionally larger, as shown by the respective lines moving (nonlinearly) downwards in Figure 1.

2.6. **Brief Overview of the Model's Contributions.** We conclude this section by briefly highlighting the contributions of the model. The model is designed to include five key features simultaneously: (i) strategic decisions on a network - agents choose behaviors and their payoffs from those behaviors depend on their neighbors' behaviors, (ii) strategic socialization - agents have discretion over how they interact, the resulting network depends on such interactions, and agents base those choices over the anticipated payoffs from their (expected) network position accounting for the ensuing behaviors and externalities, (iii) group memberships and homophily - both the payoffs and meeting rates of agents can be biased to privilege group identity, (iv) statistical identification of the three previously discussed different features, and (v) practical estimation of this model for a moderately sized network from a single observation of a network and associated behaviors.[31]

These five features are all essential for our analysis, as we anticipate they would also be in many other applications. This list of desirable features of the model requires some stylizing of the model, hitting an appropriate tradeoff between tractability and richness. The model has to be rich enough to provide a good fit and explain much of the variation in the data along several dimensions (more on this below), and yet tractable enough to solve and estimate.

There are models that combine one or more of the various features mentioned above, but none that allow for all of them. For example, the growing literature on games on networks (e.g., see Jackson and Zenou, 2015 for a survey) addresses peer interactions on networks and how those vary as a function of the network. The linear-quadratic framework here, first explored in Ballester, Calvó-Armengol, and Zenou (2006), has become a standard approach in that literature given its tractability. Here, this is combined with network formation and homophily. For further details and extensive comparisons of our model relative to others across these five features, please see Online Appendix B.2.

3. Data

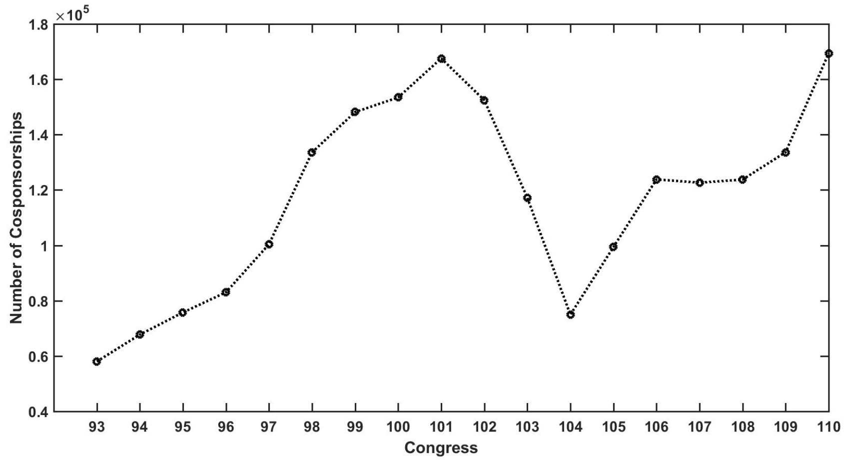
Our empirical application is the U.S. Congress, particularly the U.S. House of Representatives. It is characterized by two parties, Democrats and Republicans. To empirically study our model, we need data on social effort choices, $s^*_{i,\tau}$, legislative effort $x^*_{i,\tau}$, previous electoral competition, $V_{i,0}$, the passing of bills, and individual legislator characteristics across legislators i and Congresses τ. We denote $\mathcal{N}_\tau$ as the members of Congress τ.

First, we use cosponsorship data from Fowler (2006) as a proxy for social effort. This data is compiled from the Library of Congress, covering the 105th to the 110th United States Congress (from 1997 to 2009).[32] This data contains cosponsorship decisions by politician, and within that data, who sponsors and who cosponsors each bill. It also contains information on whether each bill was approved in Congress or not (we focus on passage in the House of Representatives). Figure 2a shows that measures of inter-connectedness of Congress, for example the total number of cosponsorship links in legislative acts across members of the House (Fowler, 2006), have been steadily increasing. Figure 2b then breaks down how cosponsorships vary within and across parties.

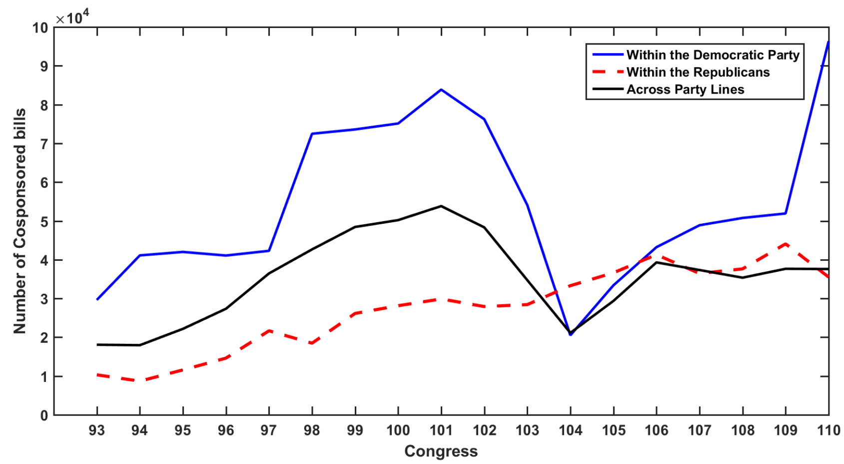
[31]Congressional data allow for a time series, but the agents and their preferences change over time, and so we need a model that can be fit from a snapshot of the network and behaviors.

[32]We restrict the data to Congresses 105th-110th for multiple reasons. First, the 104th Congress (corresponding to the Republican Revolution) provides a structural break in the analysis of Congressional behavior. With multiple changes to Congressional composition and structure during the 104th, it becomes hard to compare the costs and socializing of this specific Congress to others, preceding or following, without having to further delve into the exceptionality of this particular congressional cycle, which is not the aim of this work. Second, the data for floor speeches, which we use to compute our proxy for legislative effort and which we describe below, is only available from the 104th Congress onwards.

Figure 2. Number of Cosponsorships per Congressional Cycle



(A) Total Number of Cosponsorships



(B) Number of Cosponsorships Within and Across Parties

The figure shows the evolution of the total number of (unique) cosponsorships during a congressional cycle (i.e. anytime a politician has cosponsored another in a directed way) over time. The first figure shows the total number of cosponsorships, while the second decomposes it by party.

Per Congressional cycle, our measure for cosponsorships is the log of how many hundred bills each politician cosponsors. This function of cosponsorships acts as an empirical proxy for the social effort for legislator i in Congress τ, $\{s^*_{i,\tau}\}_{i \in N_\tau}$.

We note here that cosponsorship differs from bill sponsorship. Sponsoring a bill refers to the introduction of a bill for consideration (and can be done by multiple legislators drafting the bill, the "sponsors"). These sponsors are the authors of the bill. Instead, cosponsorships refer to the decision of adding one's name as a supporter of the bill (becoming a "cosponsor" of the bill). In contrast to sponsorship of a bill, the decision to cosponsor does not involve any writing of legislation. Instead, cosponsorships serve as a sign of support to that current bill (or potentially, to its authors), without ownership of the legislation itself. Cosponsorships are prevalent in Congress, as can be seen in Table 1, and the presence of cosponsorship across party lines is still quite common, notwithstanding the trends in polarization discussed in Fiorina (2017) or Canen et al. (2020, 2022), as evident from the time series in Figure 2b.

Our use of cosponsorships as a proxy for social effort is supported by extensive arguments in the political science literature. The importance of cosponsorships in the legislative process is clearly laid out early on by Campbell (1982) who noted, with examples, that sponsors spend "significant effort to recruit members as cosponsors" and that "the number and the diversity of cosponsors ... are often cited by legislators during floor and in public discussions as evidence of the bill's support." Although one might worry that signing a name as support is cheap and uninformative, Campbell notes that "to the typical congressman, the decision to cosponsor seems to be neither a rare nor a common occurrence." This is still the case as seen in Figure 2. More recently, Fowler (2006) argued that cosponsorships can be interpreted as a "sign that they [sponsor and cosponsor] have spent time together and established a working relationship," or at least that "it is likely that legislators make their cosponsorship decisions at least in part based on the personal relationships they have with sponsoring legislators." Altogether, this suggests cosponsorships are informative about political connections, social effort, and provide variation that can be exploited in empirical specifications. In Section 6.1, we show that a naïve network based on cosponsorships (as in Fowler, 2006) outperforms measures based on alumni connections or committee memberships. This coincides with findings in Fowler (2006) that cosponsorship based centrality measures outperform alternatives in predicting legislators' numbers of successful amendments, a proxy for legislative influence, and roll call votes. It has also been extensively used in political economy as measures of socialization both within the U.S. and abroad, as discussed in Section 1.1.

We follow Anderson et al. (2003) and use data on floor speeches as a measure of individual legislative effort. To do so, we compile the amount of words that each Congress member used in his/her floor speeches across the duration of one term. Our proxies $\{x^*_{i,\tau}\}_{i\in N_\tau}$ are given by the *Floor Speeches* variable constructed as $\log(1 + Words_{i,\tau}/100)$. We log and rescale this variable to a scale that is comparable to other legislative activities.[33] Data on floor speeches comes from Gentzkow and Shapiro (2015), available on ICPSR.[34] Furthermore, we show that our reduced form results in Section 4 are robust to using a measure based on the number of bill sponsorships (i.e., number of bills introduced) per politician to proxy legislative effort. This also follows Anderson et al. (2003), as the writing and introduction of bills for legislative consideration is costly. However, this measure cannot be our main choice in structural estimation in Section 5, as it is separately accommodated in the model (since politicians may present multiple bills for consideration).

[33]Dividing the number of words by 100, reflects an appropriate scale to compare cosponsorships to these speeches. It is a reasonable scale as House rules explicitly limit one minute speeches, a useful tool for politicians (Schneider, 2015), to 300 words.

[34]As there are changes in the composition of Congress within a term, for instance due to death or resignation among other reasons, we have some observations whose cosponsorship numbers and word counts do not correspond to a full term. To mend this, we scale up values proportionally to the recorded behavior while in Congress. In other words, if a politician leaves halfway through his term, we double the values of these observations.

We denote $\{y_{i,\tau} = I_{(Y_{i,\tau}>m)}\}$ whether each bill was approved or not, where $i \in N_\tau$ and τ is a given Congressional cycle. Our congressional data also includes the individual bill success outcome (i.e. if the bill passes or not) which maps into $\{y_{i,\tau}\}_{i\in N}$. We then use the sponsorship information to link the outcome of the bill to the network characteristics and individual decisions.

We use the winning margins in the previous election as the random variable, $V_{i,0}$, first introduced in Section 2.3. The previous winning margin captures the previous competitiveness of the district. By construction, this measure is available for the vast majority of legislators present in Congress τ (the exceptions being legislators who replace others mid-term due to death or retirements). The data on winning margins comes from the Clerk of the House, who provides these Election Statistics online in the History, Art & Archives of the United States House of Representatives. We accessed them and processed them in December 2020, while also making use of its digitized form available from the MIT Election Lab.

Finally, we use data on a variety of legislator observable characteristics. Our covariates are ideology (measured by DWNominate from VoteView), tenure (how many terms a politician has served in Congress, with data coming from the Library of Congress), and committee memberships.[35]

Data on committee memberships comes from the work of Stewart and Woon (2016). To quantify the value of the committees a politician is in, we use the *Grosewart* measure (Groseclose and Stewart, 1998). Groseclose and Stewart (1998) and Stewart (2012) estimate a cardinal value of how much an assignment to a given committee is valuable to politicians. Such estimates are based upon data on how often politicians accept transfers from one committee to another. The more desirable committees are those that politicians accept to be transferred to often, but rarely accept to be transferred away from. The *Grosewart* measure sums up the values of the committees in which a politician is present. We use the estimates given in Stewart (2012), since they are the updated values for the period we study.[36] Summary statistics for all our variables of interest can be found for reference in Table 1.

[35]We also perform a final, additional, trimming of the data across all Congresses. We drop a set of 19 observations (out of 2636), that have the number of words in Floor Speeches set to 0 in the data of Gentzkow and Shapiro (2015). These observations relate almost exclusively to a politician who either resigned or died during that term (e.g. Representatives Jo Ann Davis in the 110th Congress, Sony Bono in the 105th, or resignations as Representative Bobby Jindal in the 110th). Since the data is zero, the rescaling above does not prove to be adequate, so we drop these observations. We also drop one observation in which politicians that have cosponsorship figures less than 3 bills over a full term, since identification relies on the existence of cosponsorship and most cosponsor in the hundreds, so scaling is also inappropriate. The results do not depend on this cutoff.

[36]Below, we also consider an alternative measure for committee memberships. We construct dummy variables for whether a politician has been assigned to a given committee during that congressional term. We then focus on the main committees for parsimony: Appropriations, Energy and Commerce, Oversight and Government Reform, Rules, Transportation and Infrastructure, and Ways and Means. We also include a variable *Leadership* of whether the politician was the Speaker, the Majority or Minority Leader, or the Majority or Minority Whip.

Table 1. Summary Statistics

	Congress					
	105	**106**	**107**	**108**	**109**	**110**
Cosponsorships						
Mean	185.74	234.57	229.79	226.75	230.74	269.65
Standard Deviation	85.79	102.91	127.03	124.08	119.48	135.90
Floor Speeches (Words)						
Mean	32938.633	36282.23	27906.61	33490.47	33985.21	37416.96
Standard Deviation	38503.19	39234.14	34421.74	42334.30	45922.73	51212.574
Winning Margins (Previous Election)						
Mean	0.301	0.411	0.381	0.400	0.390	0.344
Standard Deviation	0.225	0.283	0.236	0.244	0.238	0.239
Ideology (DWNominate)						
Mean	0.0674	0.0695	0.0865	0.1116	0.1276	0.0784
Standard Deviation	0.4428	0.4549	0.4682	0.4823	0.4966	0.5031
Tenure						
Mean	4.8439	5.1839	5.4498	5.6073	6.0479	6.0584
Standard Deviation	3.9562	3.7690	3.7741	3.9005	4.0137	4.2412
Grosewart						
Mean	0.2725	0.2797	0.2896	0.2352	0.3046	0.3180
Standard Deviation	1.0815	1.1207	1.1224	1.1545	1.1591	1.1654
Approval of House Bills						
Mean	0.1087	0.1246	0.0981	0.1138	0.0957	0.1285
Standard Deviation	0.3758	0.3782	0.3092	0.3439	0.3690	0.3687
Number of Politicians N	442	435	440	439	438	445
Number of Bills	4874	5681	5767	5431	6436	7340

The table presents summary statistics for the variables used in our empirical estimation, across Congresses. Number of words said in floor speeches aggregates the number of words said by a politician across all his speeches in a term. Cosponsorships and number of words are scaled to full term length (i.e. if a politician leaves mid-office and is replaced mid-office; then both him and the replacement have those variables multiplied by 2.). For estimation, we remove the observations (bills and politicians) we do not have or cannot match to identifying numbers, and those with less than 3 cosponsorships (see the Data Section). These are mostly Congressmen who substitute others mid-term. Data used for bills is House bills (H.R.).

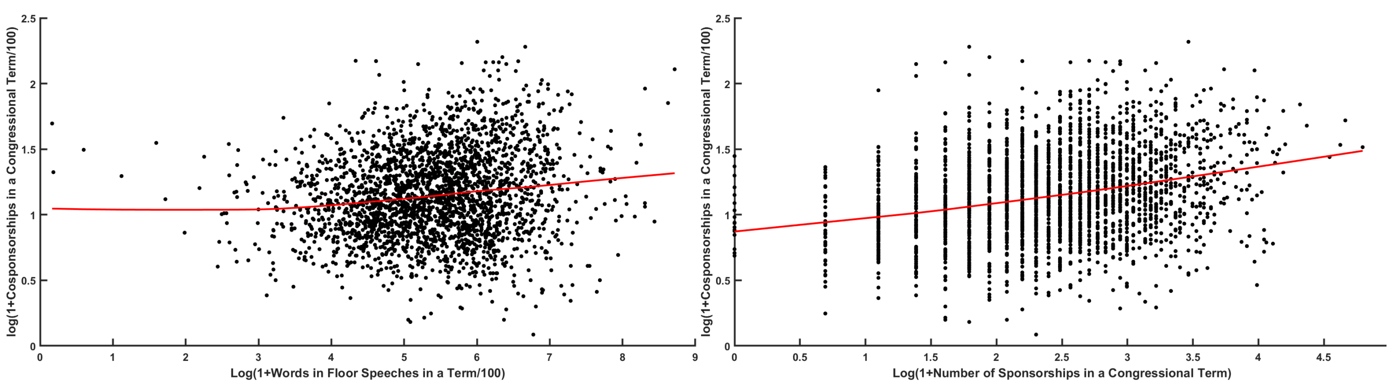
4. Reduced Form Estimation: Validating Assumptions and Model Preliminaries

We use the data to test key assumptions and predictions of the model. We start with the assumptions of complementarity of legislative and social effort, and that social connections help bill approval. We find evidence supporting both assumptions. Then, we show that the model's predictions of reelection incentives and partisanship on equilibrium effort levels are also supported by the data. Altogether, we conclude that our model is consistent with important reduced-form patterns of observed legislative behavior.

4.1. **Validating the Model.**

4.1.1. *Complementarity of Social and Legislative Effort.* In our model, social effort and legislative effort are assumed to be strategic complements. This assumption is consistent with the data, as shown in Figure 3. The proxies of link formation and our main measure of legislative activity (floor speeches) have a statistically significant and positive correlation, whether in raw form (regression coefficient = 0.046, t-stat = 6.51 with robust standard errors) or once we control for party and Congress (regression coefficient = 0.038 and t-stat = 6.38). This is also robust to using the number of bill sponsorships (i.e., bills introduced for consideration) in a congressional term as a measure of legislative effort, as shown in the right-hand side of Figure 3 (with an associated regression coefficient of 0.220, t-stat = 15.1, and a coefficient of 0.222, t-stat = 15.3 when controlling for party and Congress).

Figure 3. Correlation between Raw Measures of Legislative Effort and Social Effort.



(A) Cosponsorships and Floor Speeches (B) Cosponsorships and Bills Sponsored

We show the positive correlation between proxies for socializing (log(1 + number of cosponsorships/100)) and legislative effort (on the left figure, log(1+ number of words in floor speeches/100); on the right, log(1+bills sponsored)). The graphs present the sample used in estimation (see Data section) with variables in raw form, without rescaling. In red, we present a LOWESS (locally weighted scatterplot smoothing) fit, with bandwidth (span) equal to 0.9, fitting the relationship between the variables.

4.1.2. *Social Effort is Positively Correlated with Bill Approval.* The extension in Section 2.3 posits that the role of connections is to increase the likelihood of passing a bill. In Appendix F, we present evidence validating this mechanism: socializing is positively and significantly correlated with bill approval.

The more parsimonious specifications compare bills by the same legislator within the same Congress, but with different numbers of cosponsors. We find that adding one more cosponsor who has 10 links (i.e. increasing the support network for a bill), is associated with an increase of about 1.7 percentage points in the probability of bill approval. As this identification strategy uses within-sponsor and within-Congress variation, it controls for a legislator's average ability to socialize, network, write the bill and produce legislative effort within the same Congress (i.e. where legislators are subject to the same set of potential links and institutional constraints). A further specification addresses the role of unobserved bill quality: bills from the same sponsor might have more cosponsors and be more likely to pass simply because they are better. We compare identical bills (thereby controlling for their quality) introduced in the House of Representatives and in the Senate. Such bills only differ in the networks of cosponsors/supporters, conditional on historical differences between the chambers. Our results again confirm that social effort is associated with increased likelihood of bill approval. See Appendix F for further details.

4.2. **Reduced Form Evidence.** We now assess the model's predictions in a regression framework. We focus on two empirically relevant predictions: the correlations between past electoral competition and social effort, and changes in partisanship meeting rates and equilibrium effort levels.

4.2.1. *Past Winning Margins and Equilibrium Effort.* Corollary 2.1.1 shows that electoral competition, denoted as $V_{i,0}$, has a negative effect on legislative activity in Congress. Intuitively, past winning margins are persistent and reflect the likelihood of being reelected in the future. The legislator invests less in costly social connections, needed to pass bills, when (s)he is in a less competitive district.

In the data, we can test the resulting correlation by regressing our proxy for social effort - a politician's bill (hundreds of) cosponsorships in a Congress - on the winning margins that the legislator faced in previous elections. This is shown in Table 2 below.

We present different specifications that control for a variety of individual characteristics (e.g. measures of ideology, tenure, committee membership, party), Congress and state fixed effects, and check whether the results are driven by outliers in non-competitive races. As we can see, there is a strong negative correlation of past winning margins/electoral competition with cosponsorships. The results are robust across specifications - e.g., whether using different controls - and it is not driven by uncontested races (see the last column). While our outcomes are the log hundreds of cosponsorships (the same used in the next section), our results have the same sign and statistical significance if we use the number of cosponsorships instead.

4.2.2. *Changes of Partisanship and Equilibrium Effort.* Another prediction of the model relates to the role of partisanship, (p_1, p_2), on equilibrium outcomes. While this relationship

Table 2. Evidence on Lower (Past) Winning Margins Being Positively Correlated with Social Effort

	Outcome: Log(1+Cosponsorships in a Congressional Term/100)				
	(1)	(2)	(3)	(4)	(5)
$V_{i,0}$ - Previous Win. Margin	-0.088***	-0.064***	-0.078***	-0.077***	-0.114***
	(0.022)	(0.022)	(0.021)	(0.023)	(0.031)
Ideology Controls	Yes	Yes	Yes	Yes	Yes
Additional Individual Controls	No	Yes	Yes	Yes	No
Congress Fixed Effects	No	No	Yes	Yes	No
State Fixed Effects	No	No	No	Yes	No
N	2580	2580	2580	2580	2424
R^2	0.341	0.366	0.415	0.430	0.344

Notes: Robust standard errors in parentheses. The outcome is the log of (1+the Number of Cosponsorships in a Congressional term/100). Ideology controls are the politician's DW-Nominate score, and its DW-Nominate score squared. Additional individual controls include party fixed effects, tenure and Grosewart score to measure the value of Committee assignments (see the Data section). The last column drops candidate-Congress observations whose previous election was uncontested (i.e. winning margins above 0.9).

can be highly nonlinear, we focus on one empirically salient prediction generated by the model: the possibility that an increase in $p(i)$ can actually increase aggregate equilibrium effort levels, all else constant. A numerical example is shown in Figure 4.[37]

Capturing this possibility is important because of its empirical relevance: many measures suggest that the U.S. Congress has become increasingly polarized (e.g. McCarty et al., 2006; Canen et al., 2022 using roll call voting, Bonica, 2014 using campaign contributions, Gentzkow et al., 2019 using floor speeches), despite increased cross-partisan socialization (e.g. Figure 2).[38]

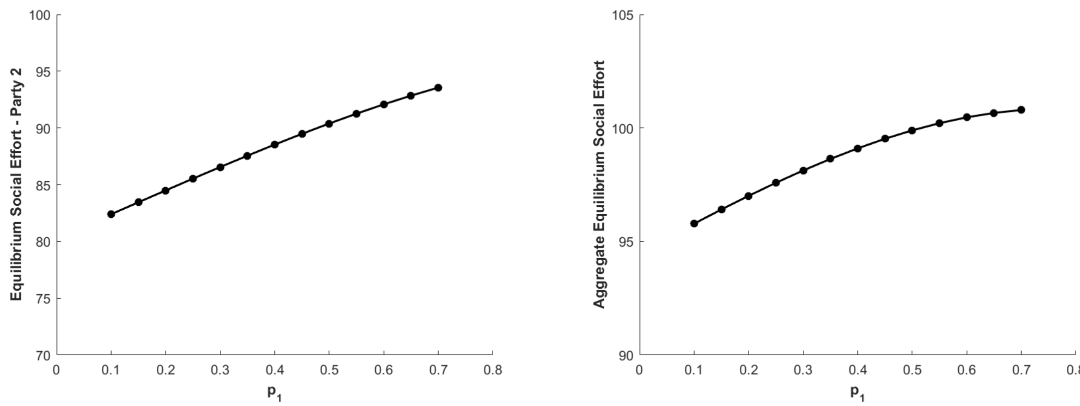
Our model provides a clear explanation for this correlation. When one party has politician types α_i that are sufficiently higher than the other (in Figure 4, $\sum_{i\in P_1}\alpha_i < \sum_{i\in P_2}\alpha_i$), or when their socialization incentives are larger (e.g. $\sum_{i\in P_1}\phi_i < \sum_{i\in P_2}\phi_i$, in line with what is estimated in later sections), then an increase in p_1 directly benefits members of the higher

[37]A formal test of this prediction without estimating the model is complicated due to data limitations. A measure of $p(i)$ should capture polarization in meeting rates across politicians in different parties. However, such a measure: (i) would only vary at the congressional-party level (i.e. it does not vary at the individual level, and we only have 6 Congresses in the data), (ii) is very hard to observe. Even the arguably simpler task of measuring ideologies is already subject to extensive issues of identification (see Canen et al., 2020, 2022 for recent contributions). This further motivates the structural approach pursued in the next section.

[38]For instance, the correlation between cross-party cosponsorships and DW-Nominate's distance across party medians in the first dimension is above 0.8.

type party (Party 2 in this case). The latter meet less often with politicians in the opposition, who are lower types on average (or have lower incentives to exert effort). The opposition generally chooses lower effort levels, do not have as many connections and, as a consequence, do not generate as many spillovers from interactions. This isolation greatly benefits Party 2 members because their high type/highly incentivized politicians meet and connect more often with their co-partisans who are also, on average, higher types. This further increases the gains from both modes of effort, given their strategic complementarity. This dynamic spurs choices for those in Party 2 which are sufficiently higher and increase aggregate effort levels.

FIGURE 4. Numerical Example Where Increases in Partisanship Increases Social Effort



(A) Equilibrium Social Effort, Aggregated across Party 1 Members (Higher Types)

(B) Aggregate Equilibrium Social Effort Across All Legislators

The figure shows a numerical example of how increases in polarization (p_1, p_2), can yield increases in equilibrium effort levels for both parties. To mimic the House of Representatives, our empirical application, we set the number of members in Party 1 to be 217, and those in Party 2 to 218, for a total number of 435 politicians. The example is based on having the distribution of types of members of Party 2 to be sufficiently higher on average than those in Party 1. To do so, we draw $\alpha_i \sim U[0, 0.5]$ if $i \in P_1$, and $\alpha_i \sim U[1, 1.5]$ if $i \in P_2$. For the remaining parameters, we set $c = 0.6, \phi_i \sim U[0, 0.15]$, $p_2 = 0.1$ and vary p_1 across a grid between 0.1 and 0.7. We use the best response dynamics to converge to a (stable) equilibrium. See Appendix B for details. While we present results for social effort, the same results hold for legislative effort as both effort choices are strategic complements.

Polarization itself is not necessarily effort reducing in this model. In fact, polarization may lead to higher effort levels than in the non-polarized equilibrium.

5. Identification and Estimation of the Structural Model

5.1. **Preliminaries to Structural Estimation.** While the previous section is useful to validate the model assumptions and its reduced-form predictions, other components of the model can only be evaluated after estimating its parameters. Some of the model's predictions are borne from unobserved heterogeneity in party types (α_i), in the returns to socialize (ϕ_i) and in polarization of meeting rates. Furthermore, we do not observe the underlying network in the data. Such (possibly nonlinear) predictions require recovering the model's parameters, which can only be done by explicitly using the model. We now pursue this objective by structurally estimating the model.

In Appendix C, we show that if we observed equilibrium outcomes exactly, all $\{\alpha_{i,\tau}\}$, as well as the other empirically relevant parameters of the model (e.g. $c, \lambda\rho, \lambda\zeta_{P(i)}$) are nonparametrically identified solely from the model's equilibrium restrictions (together with a location normalization). Hence, measurement errors and parametrizations are not required for identification. This further implies that our model can accommodate unobserved heterogeneity: $\alpha_{i,\tau}$ is unobserved and we do not impose structure on it except for estimation.

However, equilibrium effort levels s_i^*, x_i^* are not measured exactly and are observed with noise. For instance, bill cosponsorships are end products that miss other forms of socializing (e.g. close-doors meetings, fund-raisers, and so on). Similarly, legislative effort is only partially observed using imperfect proxies (e.g. times the Congress member was present on the floor for speeches, presence in roll call voting, or number of bills written[39]).

To account for this, we assume that the proxies for legislator i's choices in Congress τ are observed with classical measurement error[40]:

$$\tilde{s}_{i,\tau} = s^*_{i,\tau} e^{-\epsilon_{i,\tau}} \tag{5.1}$$

$$\tilde{x}_{i,\tau} = x^*_{i,\tau} e^{-v_{i,\tau}}. \tag{5.2}$$

s_i^* denotes the equilibrium social effort chosen by legislator i, which is hit with independent noise and $\tilde{s}_i$ is observed instead (and similarly for $\tilde{x}_i$). The measurement error in each equation is mean zero conditional on equilibrium choices and independent of other measurement errors (i.e. across individuals and time). We do not need to impose that the measurement errors in both types of effort have the same distribution.

For estimation purposes alone, we use the following parametrization of α_i. Let:

$$\alpha_{i,\tau} = e^{z'_{i,\tau}\beta_{P(i),\tau}} \tag{5.3}$$

where $z_{i,\tau}$ indicates a vector of individual observables (e.g. ideology, tenure, committee membership), and $\beta_{P(i),\tau}$ are party-specific and Congress-specific parameters that we estimate.

This parametrization is useful to gain statistical power. Our model allows $\alpha_{i,\tau}$ to vary at the politician and Congress levels (as politicians' types could change after gaining experience

[39] Both highlighted as important for legislative success in Anderson et al. (2003).

[40] The data is observed for multiple Congresses and we provide identification results for parameters specific to each Congress. This means we allow our parameters to differ across different Congresses and we can construct time-series estimates of the parameters.

etc.). Pooling information across legislators through equation (5.3) increases the amount of information we can explore empirically.

While one might think that (5.3) limits unobserved heterogeneity, adding noise to this specification under reasonable conditions preserves the empirical specification we use (i.e., the additional heterogeneity gets incorporated into the measurement error $\epsilon_{i,\tau}$). Measurement error's independence with respect to individual politician covariates is not a particularly stringent assumption in our context and is needed to accommodate the possibility of a partial mismatch between data and model effort predictions. Covariates capture much of the individual-level heterogeneity in benefits from legislative effort and common behavior: in the model, through α_i, and in the data, as we allow α_i to vary according to ideology, tenure, strength of committee positions held (this is verified in the model fit section). As a result, our i.i.d. assumption is on the (mis)measurement of equilibrium actions, not the actions themselves. It implies that our mismeasurement on average is not worse for certain politicians than others, conditional on their characteristics.[41]

In addition, our specification is not oblivious to common shocks driving social and legislative effort and legislative success. In the simplified set-up, equations (2.7) and (2.8) show how our structural approach in fact operates under the theoretical result of dependence of individual efforts from party and time specific common $X_{P(i)}$ and $S_{P(i)}$ factors through equations (2.13)-(2.14).

The information we employ for structural estimation is the following. First, we use whether a bill is approved or not ($\{y_{i,\tau} = I_{(Y_{i,\tau}>m)}\}$ where $i \in N_\tau$ and τ is a given Congressional cycle). $\{\tilde{s}_{i,\tau}\}_{i,\tau}$ indicates the (log of hundreds of) cosponsorship decisions per politician $i \in N_\tau$. This is our proxy for the equilibrium social effort $\left\{s^*_{i,\tau}\right\}$. The use of logs and rescaling allows us to keep this effort proxy in the same scale as our proxy variable for legislative effort. $\{\tilde{x}_{i,\tau}\}_{i,\tau}$ indicates a vector of observable proxies for legislative effort $\left\{x^*_{i,\tau}\right\}_{i,\tau}$ As already discussed, this is constructed using data on floor speeches (i.e., the log of hundreds of words in speeches per politician during a term). Meanwhile, $V_{i,0}$ are the winning margins in the previous election. Finally, our covariates z_i include tenure, the DW-Nominate ideology measure, Committee membership (either as separate dummy variables or as a Grosewart variable), as described in the Data section.[42]

As we perform our analysis within a Congress, we suppress the notation τ. We assume that a single pure strategy Nash equilibrium is played in each Congress. It is not necessary to identify the full set of equilibria, but instead just to use the implications that we are observing *some* (interior) equilibrium. More precisely, we show that, given the observed data, one can uniquely pin down the equilibrium that is played, as long as only one is played during each Congressional term. We do not impose, however, that the same equilibrium is played across different Congresses, although we focus on stable equilibria. Stability is verified using best response iterations using estimated parameters, as described in Appendix B.

[41]While we make this assumption explicitly in this paper, it is commonly (implicitly) used for inference in most empirical models, whether reduced form (e.g. using instrumental variables as in Battaglini et al., 2020) or structural, for the validity of normal approximations of test statistics and asymptotic distributions.

[42]Sponsorship of bills is already included, as we use the separate bills independently.

5.2. **Identification.** Given $\{y_i, \tilde{s}_i, \tilde{x}_i, V_{i,0}, z_i\}_{i\in N}$, we can identify and then estimate the structural parameters of the model, including $\{c, \lambda\zeta_1, \lambda\zeta_2, \lambda\rho, \beta, \{\alpha_i\}_{i\in N}\}$ for each Congress.[43] The basis for identification are our equilibrium conditions.

Formal identification of our model is demonstrated in Appendix C. The arguments are based on the equilibrium conditions (2.7) and (2.8) and equation (2.4) from our extension. Here we provide a heuristic overview of those arguments.

To obtain our identification results, we must make an additional normalization assumption. This is done to pin down the location of the distribution of α_i in the first Congress in the sample. Without it, we could move the distribution of types $\{\alpha_i\}$ and adjust c accordingly to rationalize the same observed behavior. Such an assumption is standard in fixed effects models, for instance. With it, we can identify the parameters listed above.

Intuitively, $\lambda\rho$ is identified from the variation in the ratio of effort choices s_i^*/x_i^* across members of the same party. This is because electoral competition is the only source of within-party variation in ϕ_i. By affecting ϕ_i within a party, it affects the ratio of different types of effort. The argument still holds when using the observed data $\tilde{s}_i/\tilde{x}_i$.

Meanwhile, $\lambda\zeta_{P(i)}$ is identified by the average probability of passing a bill for politicians in $P(i)$ given their observed effort levels. Finally, we can also identify the ratio $\gamma_1 \sum_j s_i^* s_j^* m_{i\in P_1,j}(s^*)/\gamma_2 \sum_j s_i^* s_j^* m_{i\in P_2,j}(s^*)$. This is identified from the relative choices s_i^*/x_i^* across members of different parties (i.e. given the known costs and effects of social effort, differences in choices across parties are due to different probabilities of passing a bill across parties).

The variation in choices of legislative effort across politicians pins down α_i and c through (2.8). The relative cost of socialization c is pinned down by comparing the relative choices of social to legislative effort across legislators given knowledge of the average types (the normalization). Individual α_i are pinned down by comparing legislative effort across politicians, given c.

We cannot separately identify p_1, p_2, as different combinations of those parameters can induce the same meeting probabilities $m_{ij}(s^*)$, which imply the same equilibrium choices.[44] As $m_{ij}(s^*)$ is unobserved, both party-level $\sum_j s_i^* s_j^* m_{ij}(s^*)$ and $\gamma_{P(i)}$ are unobserved and cannot be separately identified (only the product is identified). Nevertheless, we can construct a grid of $(p_1, p_2) \in [0,1] \times [0,1]$ and find the set of γ_1, γ_2 that are consistent with some meeting probability. This implies we set identify $\gamma_{P(i)}$. As we show below, the estimated sets for these parameters are very tight: their length is less than 0.001. And since, $\phi_i = \gamma_{P(i)}(1 - e^{-\lambda\zeta_{P(i)}})e^{-\lambda\rho V_{i,0}}$, it is also set-identified.

5.3. **Moment Equations.** We now describe our estimation procedure. Our approach is based on rewriting the equilibrium conditions (2.7), (2.8) and (2.4) used for identification

[43] We note that only the products $\lambda\rho$ and $\lambda\zeta_{P(i)}$ are identified, rather than λ separately from ρ and $\zeta_{P(i)}$. This does not hamper our analysis, because those product suffice. This lack of identification is intuitive: both ρ and ζ_1, ζ_2 are scaled relative to reelection shocks. In Section 2.3, each of those parameters measure the effects of some choice on reelection behavior, and so they depend on the scale of those shocks. This can be clearly seen in equation (2.2). The latter is a function solely of the products $\lambda\rho$ and $\lambda\zeta_{P(i)}$, but not of their separate components.

[44] In fact, there is a ridge of (p_1, p_2) that can induce the same meeting probabilities.

as moment conditions. To do so, we use our parameterization for α and the measurement errors specification given in equations (5.1) and (5.3). Appendix D shows that these moment conditions can be rewritten as:

$$\mathbb{E}\left(\log\left(\frac{s_i}{x_i}\right) - \log(\tilde{A}_1\gamma_1) - \log(1-e^{-\lambda\zeta_1}) + \lambda\rho V_{i,0}\right) I_{\{i\in P_1\}} = 0 \tag{5.4}$$

$$\mathbb{E}\left(\log\left(\frac{s_i}{x_i}\right) - \log(\tilde{A}_2\gamma_2) - \log(1-e^{-\lambda\zeta_2}) + \lambda\rho V_{i,0}\right) I_{\{i\in P_2\}} = 0 \tag{5.5}$$

$$\mathbb{E}\left(\log\left(\frac{s_i}{x_i}\right) - \log(\tilde{A}_{P(i)}\gamma_{P(i)}) - \log(1-e^{-\lambda\zeta_{P(i)}}) + \lambda\rho V_{i,0}\right) V_{i,0} = 0 \tag{5.6}$$

$$\mathbb{E}\left(\log(x_i) - z_i'\beta + \log\left(c - (\tilde{A}_{P(i)}\gamma_{P(i)}e^{-\lambda\rho V_{i,0}}(1-e^{-\lambda\zeta_{P(i)}}))^2\right)\right) = 0 \tag{5.7}$$

$$\mathbb{E}z_i\left(\log(x_i) - z_i'\beta + \log\left(c - (\tilde{A}_{P(i)}\gamma_{P(i)}e^{-\lambda\rho V_{i,0}}(1-e^{-\lambda\zeta_{P(i)}}))^2\right)\right) = 0 \tag{5.8}$$

$$\mathbb{E}(\log(P(y_i=1)) - \lambda\rho V_{i,0} + \log(1-e^{-\lambda\zeta_1}) - 2\log(s_i))I_{i\in P_1} = 0 \tag{5.9}$$

$$\mathbb{E}(\log(P(y_i=1)) - \lambda\rho V_{i,0} + \log(1-e^{-\lambda\zeta_2}) - 2\log(s_i))I_{i\in P_2} = 0, \tag{5.10}$$

where $\tilde{A}_{P(i)} \equiv \sum_{j\neq i} s_j^* m_{ij}(s^*)x_j^*$ which is constant within each party. These equations have the added advantage of clearly delineating the identification arguments from the previous section: the first three equations clearly pin down $\tilde{A}_1\gamma_1, \tilde{A}_2\gamma_2, \lambda\rho$. The last two equations pin down $\lambda\zeta_1, \lambda\zeta_2$. The remaining equations pin down c and β. The normalization assumption guarantees that the fourth equation pins down c alone, rather than c together with a constant term in β. Then, we construct a grid of $(p_1, p_2) \in [0,1]\times[0,1]$ to estimate the set of feasible γ_1, γ_2 as each value in the grid pins down a separate value of $\tilde{A}_{P(i)}$.

For our main empirical exercise, we let Party 1 denote the Democratic Party (with its variables denoted by the subscript *Dem*) and Party 2 denote the Republican Party (analogously denoted with a *Rep* subscript). We carry out the estimation process via GMM under the assumption that the mean of $\alpha_i = e^{z_i'\beta}$ with $z_i = 0$ is known (our normalization assumption), which is done by not including a constant in z_i. We estimate the model separately for every Congress.

In what follows, we choose to estimate and analyze the more convenient parameter $\tilde{\zeta}_{P(i)} = \log(1-e^{-\lambda\zeta_{P(i)}})$ rather than $\lambda\zeta_{P(i)}$. $\tilde{\zeta}_{P(i)}$ is a known and increasing function of $\zeta_{P(i)}$, retaining its quantitative interpretation and conclusions. However, our moment conditions are linear in $\tilde{\zeta}_{P(i)}$, but highly nonlinear in $\lambda\zeta_{P(i)}$. This adds significant numerical stability. Furthermore, the former is more robust to underlying parametric assumptions.

Concerning the information of whether a bill passed or not $\{y_{i,\tau}\}_{i\in N}$, the model is agnostic on how many bills a politician proposes. Because a good fraction of members of Congress sponsor multiple bills, however, we work with $L > N$ bills in the actual data. This is easily accommodated in the estimation. Recall that ε are i.i.d. across time and bills. For each politician i, all of i's bills have the same associated network $g_{i,j}$ in the model. They come from the same politician's effort choices (as well as those of his network induced by those choices). The different ε realizations, however, represent different bill qualities or institutional arrangements within politician, meaning that the same politician may have one bill approved and not another. The dimensionality of the problem can be decreased by simply averaging out each bill's success by politician. This is made possible by the fact that equation (2.4) holds for all bills, implying that it must hold for all politicians as well. Hence, we use the average pass rate of bills for politician i as its estimate of the probability of bill approval.

5.4. **Estimation via GMM.** To estimate the model, we replace equations (5.4)-(5.10) by their empirical counterparts and stack them into a vector of the form $\frac{1}{n}\sum_{i=1}^{n} g(\tilde{s}_i, \tilde{x}_i, y_i, z_i; \theta)$. Since all moments have expectations taken over ϵ_i, v_i, which are i.i.d. and mean zero for all politicians, the empirical counterpart replaces the expectation operator by the mean over i.[45]

Furthermore, we average over the approval rates for bills for each politician to get the estimated probability of approval at the politician level. We then minimize the quadratic form of those moments. Since the model is exactly identified, we use the identity matrix as the weighting matrix. Our estimates are obtained using appropriate starting values. We use consistent estimators for all parameters apart from c, and then explore different starting values for c in our algorithm. Further details are given in Appendix D. Standard errors are computed by plugging in our estimates in the asymptotic variance matrix, which is derived analytically.

6. Results

Table 3 presents our parameter estimates, which delineate a series of intuitive relationships emerging from the data. Table 4 shows the distributions of the estimated individual types, α_i, and the individual returns to social effort, ϕ_i, over time and by party. These distributions appear stable across Congresses. They are computed based on the estimates in Table 3.

Splitting the samples by party, we observe important differences in the estimated distributions of Republicans and Democrats. While both parties have similar mean α_i, Democrats have a higher dispersion from Congresses 105-109 (when they are in a minority), while Republicans have tighter distributions and slightly higher types in those periods. This implies different social effort patterns across parties, as Democrats socialize meet higher types more often (by our social meeting function).

From the main results, we can compute the estimated sets for the returns to social effort, ϕ_{Dem} and ϕ_{Rep}. These are presented in the second panel of Table 4. For Democrats, the mean estimates range from $[0.037, 0.047]$, while for Republicans, the range is $[0.031, 0.037]$. Both are slightly increasing over time. To put this into context, note that the marginal utility of an increase in x_i is

$$\alpha_i + \phi_i s_i \sum_{j \neq i} x_j s_j m_{ij}(s) - c x_i.$$

The direct benefit α_i ranges from 1 to 1.4, while the network benefit $\phi_i s_i \sum_{j\neq i} x_j s_j m_{ij}(s)$ ranges from about 0.1 to 0.3. Therefore, the social incentive is somewhere between a tenth to a fourth of the direct incentives, a substantial consideration in politician's choices.

We note that the value for socializing, ϕ_{Dem}, for Democrats is higher than that for Republicans, even as we control for the difference in types, partisan bias, and preferences across

[45]That is, the expectation operator has one observation for each politician, and averages across all politicians.

TABLE 3. Main Results, Specification 1

	Congress					
	105	**106**	**107**	**108**	**109**	**110**
c	0.270	0.277	0.295	0.305	0.296	0.273
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\tilde{\zeta}_{Dem}$	5.890	5.507	5.720	5.554	5.546	3.125
	(0.160)	(0.179)	(0.174)	(0.176)	(0.174)	(0.119)
$\tilde{\zeta}_{Rep}$	3.403	2.695	3.117	2.880	2.935	4.311
	(0.178)	(0.146)	(0.170)	(0.172)	(0.158)	(0.181)
$\lambda\rho$	0.101	0.138	0.015	0.010	0.038	0.023
	(0.072)	(0.057)	(0.068)	(0.073)	(0.080)	(0.070)
Rep	0.089	0.142	0.163	0.109	0.091	-0.092
	(0.047)	(0.043)	(0.051)	(0.048)	(0.051)	(0.090)
Ideology	-0.421	-0.290	-0.386	-0.450	-0.422	-0.364
	(0.049)	(0.064)	(0.058)	(0.126)	(0.088)	(0.047)
Tenure	0.008	0.011	0.009	0.011	0.009	0.003
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)
Grosewart	-0.018	-0.004	-0.023	-0.023	-0.010	-0.015
	(0.010)	(0.011)	(0.010)	(0.032)	(0.012)	(0.009)
Ideology × *Rep*	0.607	0.434	0.487	0.709	0.688	0.768
	(0.092)	(0.092)	(0.098)	(0.145)	(0.115)	(0.121)
Tenure × *Rep*	0.010	0.006	0.007	0.008	0.003	0.007
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)
Grosewart × *Rep*	0.007	-0.008	0.001	-0.002	-0.021	-0.037
	(0.015)	(0.015)	(0.016)	(0.034)	(0.015)	(0.018)
γ_{Dem}	[0.0001,0.0001]	[0.0002,0.0002]	[0.0002,0.0002]	[0.0002,0.0002]	[0.0002,0.0002]	[0.0020,0.0020]
γ_{Rep}	[0.0011,0.0011]	[0.0023,0.0023]	[0.0015,0.0015]	[0.0018,0.0019]	[0.0018,0.0018]	[0.0005,0.0005]
N	424	427	426	431	429	426

Notes: Asymptotic standard errors in parentheses. The table presents the GMM estimates for the parameters of interest, as described in the Estimation section. Standard Errors are computed from estimates of the asymptotic variance for a GMM estimator using its analytical formula. Details are in Appendix D. The estimates of γ_1 and γ_2 are their estimated sets. $\tilde{\zeta}_{P(i)}$ is a known increasing function of $\lambda\zeta_{P(i)}$. *Rep* represents the dummy variable of whether a politician was in the Republican Party. Hence, a variable *Tenure* × *Republican* represents the additional estimate of the *Tenure* variable for the Republican Party, as compared to the Democratic one.

both parties. It is quantitatively large, approximately 17% of the returns to social effort (in the order of 0.006/0.036 across Congresses). Furthermore, it is of a significant magnitude relative to c, which ranges in $[0.27, 0.30]$ and does not change much over time, on average.

The relative cost of legislative effort to social effort, c, is stable over time. As a result, interactions between politicians are more valuable, as there is an increasing return of socializing against a stable cost. This may be consistent with an increase in the complexity of

TABLE 4. Heterogeneity: Differences in the Distributions of α_i and ϕ_i Across Parties

Congress	**105**	**106**	**107**	**108**	**109**	**110**
			Types, α_i			
Democrats:						
Mean α_i	1.218	1.183	1.210	1.256	1.249	1.156
Standard Deviation of α_i	0.091	0.077	0.082	0.100	0.095	0.067
Republicans:						
Mean α_i	1.292	1.345	1.343	1.416	1.360	1.230
Standard Deviation of α_i	0.078	0.076	0.074	0.103	0.081	0.103
			Returns to Social Effort, ϕ_i			
Democrats:						
Mean ϕ_i	0.037	0.043	0.047	0.046	0.045	0.045
Standard Deviation of ϕ_i	0.001	0.002	0.0002	0.0001	0.0004	0.0003
Republicans:						
Mean ϕ_i	0.031	0.033	0.034	0.033	0.034	0.037
Standard Deviation of ϕ_i	0.001	0.001	0.0001	0.0001	0.0003	0.0001

We show the mean and the standard deviation of the (estimated) distributions of α_i and of ϕ_i for each party, highlighting the differences in those distributions. They are computed using the estimates from Table 3.

extant statutes, making it more difficult to approve legislation. This is for example evident from an average number of pages per statute of 3.6 in 1965-66 to 18.8 in 2015-16[46], making interactions between politicians in drafting and drumming up support for legislations on the Capitol more important. Although this would not change the costs to social effort, it would appear to change the returns from it.

The estimates of $\tilde{\zeta}_{Dem}$ and $\tilde{\zeta}_{Rep}$ are also significant and large in magnitude. This indicates that politicians see positive gains from having bills approved. We can see large differences across parties. In particular, the Democrats have the largest ζ between Congresses 105-109, when they are in the minority. In Congress 110, this is reversed, with the minority Republicans having the largest ζ. This intuitively suggests that the returns from getting bills approved for a politician depend on their party's majority status in the House. A high ζ for the minority party indicates that voters reward legislators more when it is harder for them to pass bills. This is consistent, for instance, with voters rewarding their representatives for more difficult outcomes, or voters learning from legislative successes about the quality of their representatives.

[46]Vital Statistics of Congress 2017, Chapter 6, available at www.brookings.edu

Our extension allows us to decompose ϕ_i into its three components and infer the main mechanism which induces the higher returns for Democrats. First, we note that Democrats do not, on average, face more competitive districts than Republicans (i.e. $V_{i,0}$ is similar, on average, across parties). Meanwhile, the evolution of $\gamma_{P(i)}$ closely tracks majority status. Hence, neither mechanism can be driving this systematic difference in legislative behavior across all Congresses. According to our model, the predominant explanation why Democrats socialize more is their higher electoral return to passing bills as a minority ($\tilde{\zeta}_{Dem}$ is significantly larger than $\tilde{\zeta}_{Rep}$, implying the same about ζ_{Rep} and ζ_{Dem}). In Congress 110, this return is still relatively high and its decrease is compensated by its improved likelihood of passing a bill (γ_{Dem}). This could be because Democrat voters may further assign higher value to government legislative activity.

Using the estimated values of $\tilde{\zeta}_{P(i)}$ for both parties, we can also calculate the probability of bill approval for each politician. We show these in Figures 5a for Democrats and 5b for Republicans.

By comparing Figures 5a - 5b with the average bill passage rates in the summary statistics (Table 1), we can see that the model can generate a good match of the mean bill success rates (which we observe). Our structural assumptions allow us to represent the whole distribution of expected probabilities of having a bill approved across different politicians. These indicate some variation over time. Later Congresses (108th and 110th) show a higher predicted approval rate for most politicians. Furthermore, approval rates are highly correlated with majority party status. Democrats have a much higher rate as a majority in Congress 110, while Republicans have higher ones as a majority in 105-109.

We can also discuss the statistical significance of different covariates in explaining direct benefits, α_i. With our baseline specification that uses $Ideology$, $Tenure$, $Grosewart$ for z_i in Table 3, we see that ideology is statistically significant (especially in later Congresses). The estimates suggest that Democrats on the left of the ideological spectrum have higher direct returns of exerting legislative effort. We do not observe such a systematic result among Republicans. Meanwhile, the *Grosewart* variable, capturing the impacts of committee assignments, appears to be noisy.

We also consider another specification where we replace the *Grosewart* variable by dummy variables of committee assignments to each of Congress main committees. This is shown in Table E.1 in Appendix. We can see that the results from our main specification are robust. It is noteworthy that in this alternative specification, the estimate of being in the Rules Committee is positive and significant, suggesting those are higher types.[47] Similarly, those in leadership positions also have higher types, for similar reasons as described for the Rules Committee.

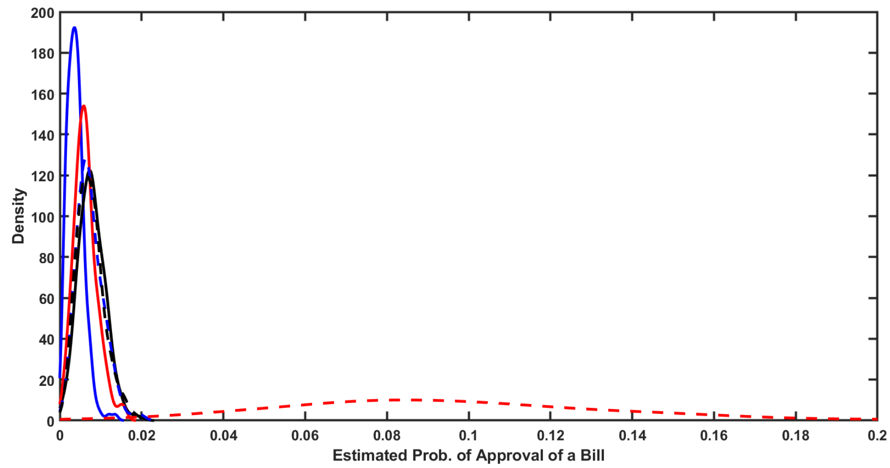
6.1. Fit and Further Discussion.

To conclude the section, we conduct three complementary exercises. Each one illustrates a different dimension of our model's fit (in-sample,

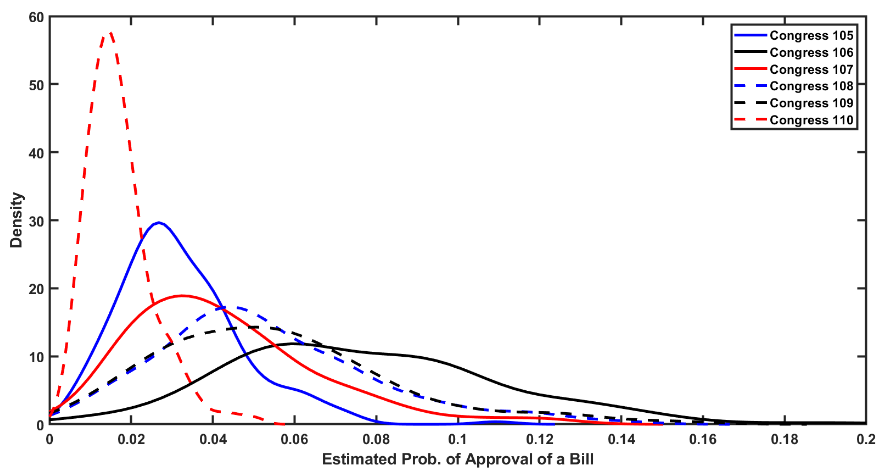
[47] The Rules Committee is the committee in charge of determining the rules that allow each bill to come to the floor, fundamental for the progress of legislation. It seems consistent that politicians in that committee have a higher return for effort in it, even conditional on having the same ideology, party, and tenure.

Figure 5. Estimated Probability of Approval

(a) Democrats



(b) Republicans



out-of-sample, and comparing to alternatives in the literature) and, hence, we draw conclusions from their collective results. In the first exercise, we consider an out-of-sample validation of our approach, by assessing the fit of our model of moments of the socialization patterns not used in estimation. In the second, we look at an in-sample validation comparing the fit of empirical approval rates of bills to the estimated values. In the third, we compare the prediction of legislative behavior according to our model generated network to other salient examples of political networks used in the literature, including cosponsorship networks, alumni-connections networks, and those based on committee memberships.

In this first exercise, we predict the i, j links for each pair of politicians based on what predicted by our theoretical $g_{i,j}(\mathbf{s})$ function based on the estimated parameters for each Congress in our sample and then compare them to the actual cosponsorship pairs. The goal is to show that our estimated network model fits such proxies for social ties, commonly employed in the literature following Fowler (2006), even without explicitly modeling pairwise socialization decisions. Although our analysis employs i's total cosponsorships to proxy for his/her social effort $s^*_{i,\tau}$, the more fine-grained data on pairwise cosponsorship information between i, j politician pairs is not used in estimation.

The correlations between the estimated $g_{i,j}(\mathbf{s})$ and any i, j pairwise cosponsorships are reported in Table 5. The table illustrates correlations for two possible definitions of links based on actual i, j cosponsorship in the data. In the top panel cosponsorships are considered directed from i to j and in the second panel cosponsorships are considered a-directional. In the two cases the correlations with the model-implied $g_{i,j}(\mathbf{s})$ are 0.374 and 0.447 respectively and statistically significant. Similar results hold throughout interior choices of (p_1, p_2). Thus, the model appears able to capture disaggregated socializing proxies not directly targeted in estimation, reassuring on the plausibility of our socialization structure.

We can draw an additional conclusion from this exercise. Results of Fisher's z-transformation tests also suggest that our model with $p_{Dem} > 0, p_{Rep} > 0$ is better at capturing the relationships from the pairwise cosponsorship data than alternative models with either full partisanship (at least one of $p_{Dem} = 1$ or $p_{Rep} = 1$) or without partisanship ($p_{Dem} = p_{Rep} = 0$). These comparisons are possible as different $g_{i,j}(\mathbf{s})$ can be generated using different values for p_{Dem}, p_{Rep}. This results hold at different interior values. In the table, we describe the cases of $p_1 = p_2 = 0.1$ and $p_1 = p_2 = 0.5$.

Although recent political economy research highlights a hollowing out of the moderate middle ground in congressional voting (Fiorina, 2017; McCarty et al., 2006; Canen et al., 2020, 2022), even our model with p_{Dem}, p_{Rep} around 0.1 produces a substantially better fit of the cosponsorship data than a model with complete polarization $p_{Dem} = p_{Rep} = 1$, which is statistically dominated. Also, while the exact point estimates of p_{Dem}, p_{Rep} cannot be pinned down due to their nonlinearity in the model, we believe that the rejection of $p_{Dem} = p_{Rep} = 1$ has to be considered more general. The raw data in Figure 2b display a sufficient degree of cross-party cosponsorship to cast doubt on an hypothesis of "full sorting" among party members in the House. A model with full polarization would predict 0 relationships emerging across different parties, that is obviously not the case. Meanwhile, a model with no polarization would predict a much higher degree of social connection across party lines than observed (it predicts approximately 60% of Democrat connections to other Democrats, compared to around 70% for our model and the data; and fewer connections among Republicans alone, relative to Republicans and Democrats, when compared to our model and the data).

Possibly, reconciling a world of both more polarized legislators and active cosponsorships across party lines (e.g. Figure 2b), may come from noting that, as ideologies may diverge, engagement across party lines becomes more important for getting legislation to the floor and passed. Our model appears to capture such phenomena.

In a second exercise, we fit the empirical success rate y_i for each politician in each Congress to the estimated one, given by equations (5.9)-(5.10).

For this exercise, we focus on politicians who have sponsored sufficient bills (over 10 House Bills sponsored), so to have a modestly reasonable empirical approximation to their average success rate. We then use the estimates of our model presented in Table 3 and those used

Table 5. Model Fit: Correlation of Estimated Network of the Model to the Cosponsorship Networks in the Data

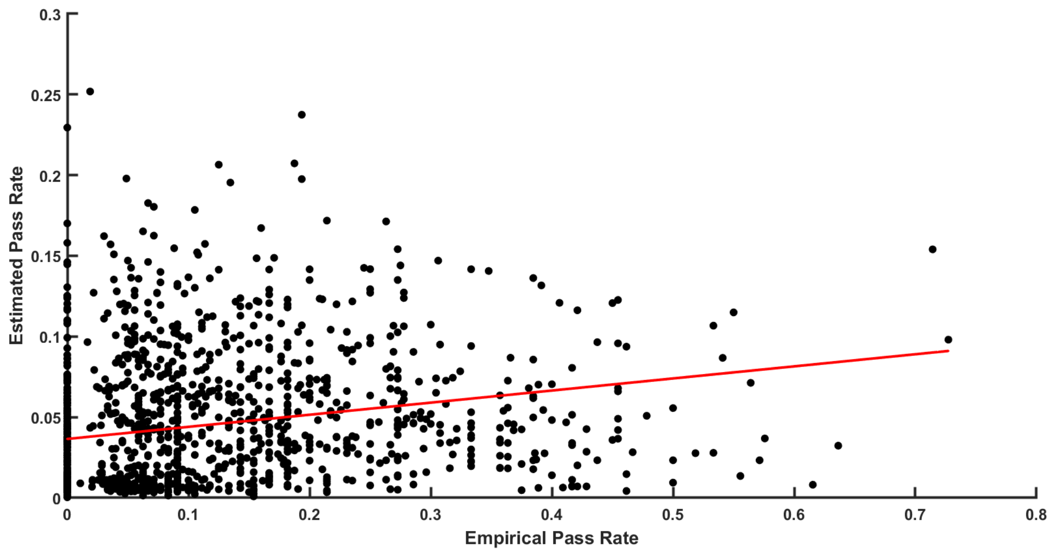
Congress	Correlation	Fisher's z-statistic
Data from Directed Cosponsorships:		
Model: $p_{Dem} > 0, p_{Rep} > 0$	0.374	-
Model: $p_{Dem} = 0, p_{Rep} = 0$	0.325	41.60***
Model: $p_{Dem} = 1$	0.342	27.45***
Data from "Combined" Cosponsorships:		
Model: $p_{Dem} > 0, p_{Rep} > 0$	0.447	-
Model: $p_{Dem} = 0, p_{Rep} = 0$	0.388	52.72***
Model: $p_{Dem} = 1$	0.409	35.12***

We compare the performance of the partisan model (with $p_{Dem} > 0, p_{Rep} > 0$) to the performance of the model without partisanship ($p_{Dem} = p_{Rep} = 0$) and complete partisanship ($p_{Dem} = 1$), in explaining the observed cosponsorships in the data. In the first panel, cosponsorships are measured by the directed number: how many times i cosponsors j. In the second panel, "combined cosponsorships" are measured by the number of times i cosponsors j and j cosponsors i, creating a symmetric undirected graph. To calculate the statistics for each model, we first generate the links using the theoretical definition $g_{ij}(s) = s_i s_j m_{ij}(s)$ under our estimated parameters. Since p_1, p_2 are not identified, we set them at the interior values of $p_1 = p_2 = 0.5$. The correlations with $p_1 = p_2 = 0.1$ are 0.348 and 0.416 respectively, and imply the same results in the statistical tests. We present Fisher's z-transformation statistic, for the test that the correlation of the adjancency matrix of the Model with $p_{Dem} > 0, p_{Rep} > 0$ with the data is equal to the correlation of the alternative model (without partisanship/complete partisanship) with the data. Since our model generates a symmetric adjacency matrix by construction, we consider the correlations of the lower triangular adjacency matrices. *** represents that the null hypothesis of equal correlations can be rejected at 1% significance level, ** at 5%. Note that, when estimating the model, we did not use cosponsorships at the ij level. We aggregate all Congresses in the analysis above.

for Table 5. The results are shown in Figure 6. We find a significant positive correlation of 0.223 between estimated and empirical bill success rates. The relationship displayed in Figure 6 is statistically significant (slope t-stat of 7.69, p-value is 0.000). It is, however, noisy. Thus, we conclude that our model captures only part of the empirically observed relationship. Reassuringly, this correlation is present as we use more precise measures (e.g., those with more than 15 or 30 sponsorships), despite the decreasing sample size.

Finally, we conduct a horse race comparing the in-sample properties of the predicted equilibrium network from our model to exogenous alternative G's in the literature. We compare these networks in terms of their predictive power/model fit of legislative behavior using the outcome from our main specification. To do so, assume that $g_{ij}(\mathbf{s})$ is known, as

FIGURE 6. Model Fit: Estimated and Empirical Approval Rate of Bills



The figure shows the correlation between estimated and empirical approval rates for politicians. We consider politicians with over 10 sponsorships, so that a meaningful estimate of the average approval rate can be obtained. The correlation is 0.223 in this case.

we will feed it from the data. Then, the game defined by (2.1) collapses to that in Ballester et al. (2006), with the Nash equilibrium given by:

$$(I - \phi G)x^* = \alpha, \tag{6.1}$$

where $\alpha = \{\alpha_i\}_{\{i=1,\ldots,n\}}$. Using our measurement error assumption of x^* in (5.1) and the parametrization of α in (5.3), equation (6.1) can be rewritten as:

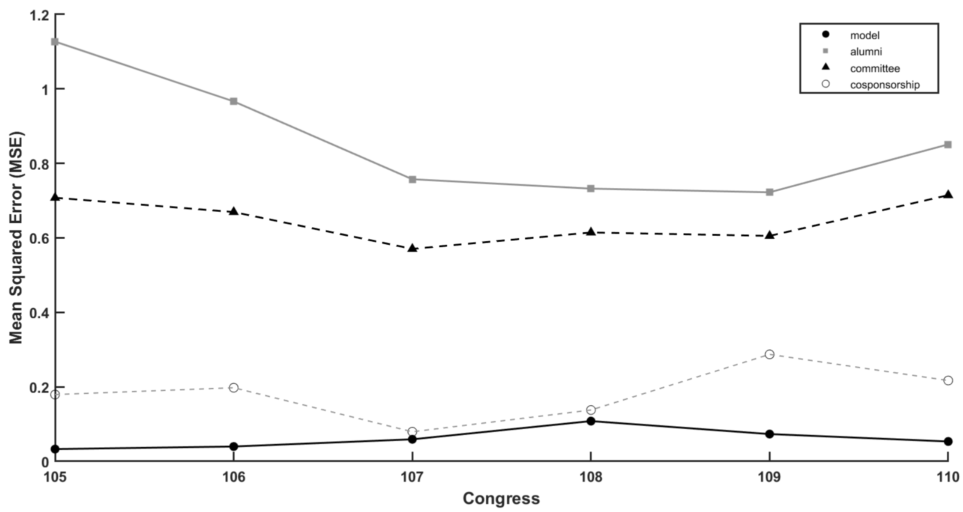
$$\begin{aligned} log(\tilde{x}_i) &= log((I - \phi G)^{-1} e^{z_i'\beta}) + v_i. \\ &\approx log((I + \phi G + \phi^2 G^2 + \phi^3 G^3) e^{z_i'\beta}) + v_i, \end{aligned} \tag{6.2}$$

for small ϕ and G of (approximate) full rank. We can then compare the fit of equation (6.2) across different networks G, as this model does not include any of our own model's structure. We compare the predicted results using our model's output to graphs based on directed cosponsorships (Fowler, 2006), alumni connections (Battaglini and Patacchini, 2018; Battaglini et al., 2020), and committee membership (from the main committees used in our estimation). We use the third-order approximation (6.2), rather than the full equilibrium equation (6.1), due to the sparsity in some of the alternative networks, which would preclude inversion of equation (6.1).[48]

[48]For illustration purposes, we present figures of these networks in Congress 110 in Figure E.1 in Appendix. From those graphs, one can check that the alumni and committee-based networks are much sparser than cosponsorship-based networks. For example, approximately 270 legislators are isolated in the alumni network, with an average degree below 2, while no legislator is isolated in the cosponsorship case. The data for the committee and cosponsorship based networks is the same as used in estimation in the previous section. The committee network has edges defined by whether legislators sat on the same committees. For alumni

To estimate ϕ, β in (6.2) across models, we first look at the subgraph of non-isolated nodes for each case. We estimate those parameters by Non-Linear Least Squares using the covariates of Ideology, Tenure, and Grosewart, used in our main specification of Table 3. The outcome is our proxy for legislative effort (floor speeches), which is positive and unbounded. Then, we use the estimated parameters to fit the Mean Squared Error (MSE) across all observations for each alternative network G. We note that this is a fair comparison across models - while we do use cosponsorships to inform our model, we only use aggregate information about one's cosponsorships. Instead, the directed cosponsorship network in the alternative uses information about pairwise decisions to completely determine G. Furthermore, covariates are essential in this specification so that one can generate heterogeneity in α_i.[49] The results for this exercise are shown in Figure 7.

Figure 7. Comparison of Model Fit Across Our Model Network and Competitors



The figure compares the performance in terms of mean squared error (MSE) of different legislative networks used in the literature (our estimated model, one based only on directed cosponsorships, one based on university alumni networks, and one based on committee memberships) in fitting legislative behavior (our $\tilde{x}_i$). To do so, we estimate the Nash equilibrium of the game in which the network G is given, which collapses to the problem in Ballester et al. (2006).

Our model fits legislative behavior substantially better (lower MSE) than the sparser alumni and committee networks. This is because it can capture correlated behavior across the hundreds of nodes that those alternative models assume are isolated, but that in reality are highly correlated. While alumni and committee networks can fit the behavior of very central and influential nodes, they miss out on the majority of legislators who do not have

networks, we scrape the Congressional bioguide webpage and use fuzzy matching based on the politician's university and date of graduation to generate the network. A link on this network exists if politicians went to the same university within 8 years of one another (as in Battaglini and Patacchini, 2018).

[49]Monte Carlo simulations have found that ϕ, β can be recovered reliably in this set-up. Nevertheless, these simulations also illustrate the empirical limits of a model based on (6.2). For example, discrete covariates generate identification problems due to the lack of variation in the support of those variables. For this reason, we only keep the covariates from our main specification in this exercise.

such easily codified connections. In addition, our model shows slight improvements over a pure cosponsorship network under the current approximation. This could be because observed cosponsorships are not as well related to underlying relationships, or because of the chosen approximation.[50] For the first explanation, our model may be using cosponsorship information more efficiently, as it is able to explore how cosponsorships correlate with other legislative behavior through equilibrium restrictions. Regarding the second explanation, the cosponsorship network is denser and could benefit from higher-order terms, although we are unable to compare all four models in that set-up.

Altogether, this last exercise further illustrates benefits of simultaneously modeling socialization choices and strategic interactions in Congress, complementing the model's previously discussed in-sample and out-of-sample performance.

7. Conclusions

We have developed and estimated a structural model of legislative activity in which endogenous, partisan social interactions play an important role in promoting bill passage. We estimate that social effort matters significantly for legislative activity. Such results are also validated in separate regressions exploring alternative identification strategies.

By endogenizing both legislative and social efforts, we accommodate complementarities in actions that appear to be strong. In particular, we find that complementarities among politicians are quantitatively substantial (on the order of 0.1 to 0.25 of the direct incentives), and are fairly stable across our sample period. Overall, we show how the process of informal social interaction among legislators paints a less extreme, although still partisan, picture of the internal operation of a legislature.

From the methodological perspective, our tractable model of biased socialization may be fruitful for further investigation of the behavior of other political agencies and in more general environments where both endogenous socialization and homophily are relevant features, even when there are multiple groups. One salient example beyond political economy is the study of social interactions and educational choices in a labor market. In such contexts, socialization is homophilous, as biased interactions may be due to wealth, race, culture or nationality; but is also driven by potential labor market outcomes (e.g., "networking" for future jobs). Our approach provides a tractable theoretical framework that can be taken to the data, and would be complementary to other frameworks (e.g., Albornoz, Cabrales, and Hauk, 2019; Bolte, Immorlica, and Jackson, 2020). Future work could further extend this model to targeted socialization, although that is likely to generate more complex identification and data requirements.

[50]The cosponsorship networks performs more comparably to our model when using a fourth-order approximation (with sometimes lower and other times higher in-sample MSE).

References

Acemoglu, D., C. García-Jimeno, and J. A. Robinson (2015). State capacity and economic development: A network approach. *American Economic Review 105*(8), 2364–2409.

Albornoz, F., A. Cabrales, and E. Hauk (2019). Occupational choice with endogenous spillovers. *The Economic Journal 129*(621), 1953–1970.

Alemán, E. and E. Calvo (2013). Explaining policy ties in presidential congresses: A network analysis of bill initiation data. *Political Studies 61*(2), 356–377.

Anderson, W. D., J. M. Box-Steffensmeier, and V. Sinclair-Chapman (2003). The keys to legislative success in the us house of representatives. *Legislative Studies Quarterly 28*(3), 357–386.

Baccara, M. and L. Yariv (2013). Homophily in peer groups. *American Economic Journal: Microeconomics 5*(3), 69–96.

Badev, A. (2017). Discrete games in endogenous networks: Theory and policy. *Mimeo, University of Pennsylvania*.

Badev, A. (2020). Nash equilibria on (un) stable networks. *arXiv preprint arXiv:1901.00373*.

Bala, V. and S. Goyal (2000). A noncooperative model of network formation. *Econometrica 68*(5), 1181–1229.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). Who's who in networks. wanted: The key player. *Econometrica 74*(5), 1403–1417.

Banerjee, A. V., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2018). Changes in social network structure in response to exposure to formal credit markets. *SSRN paper 3245656*.

Bartels, L. M. (1993). Messages received: the political impact of media exposure. *American Political Science Review 87*(02), 267–285.

Battaglini, M. and E. Patacchini (2018). Influencing connected legislators. *Journal of Political Economy 126*(6), 2277–2322.

Battaglini, M., E. Patacchini, and E. Rainone (2021). Endogenous social connections with unobserved networks. Technical report, Mimeo, Cornell University.

Battaglini, M., V. L. Sciabolazza, and E. Patacchini (2020). Effectiveness of connected legislators. *American Journal of Political Science 64*(4), 739–756.

Baumann, L. (2017). A model of weighted network formation. *SSRN Working Paper*.

Bolte, L., N. Immorlica, and M. O. Jackson (2020). The role of referrals in immobility, inequality, and inefficiency in labor markets. *SSRN working paper https://ssrn.com/abstract=3512293*.

Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science 58*(2), 367–386.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009). Identification of peer effects through social networks. *Journal of Econometrics 150*(1), 41–55.

Bratton, K. A. and S. M. Rouse (2011). Networks in the legislative arena: How group dynamics affect cosponsorship. *Legislative Studies Quarterly 36*(3), 423–460.

Cabrales, A., A. Calvó-Armengol, and Y. Zenou (2011). Social interactions and spillovers. *Games and Economic Behavior 72*(2), 339–360.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Campbell, J. E. (1982). Cosponsoring legislation in the us congress. *Legislative Studies Quarterly*, 415–422.

Canen, N., C. Kendall, and F. Trebbi (2020). Unbundling polarization. *Econometrica 88*(3), 1197–1233.

Canen, N., C. Kendall, and F. Trebbi (2022). Political parties as drivers of us polarization: 1927-2018. Technical report, Working Paper, University of California-Berkeley.

Chandrasekhar, A. G. and M. O. Jackson (2016). A network formation model based on subgraphs. *SSRN Working Paper*.

Cho, W. K. T. and J. H. Fowler (2010). Legislative success in a small world: Social network analysis and the dynamics of congressional legislation. *The Journal of Politics 72*(01), 124–135.

Cohen, L. and C. J. Malloy (2014). Friends in high places. *American Economic Journal: Economic Policy 6*(3), 63–91.

Currarini, S., M. O. Jackson, and P. Pin (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica 77*(4), 1003–1045.

Currarini, S., M. O. Jackson, and P. Pin (2010). Identifying the roles of race-based choice and chance in high school friendship network formation. *Proceedings of the National Academy of Sciences 107*(11), 4857–4861.

Currarini, S. and M. Morelli (2000). Network formation with sequential demands. *Review of Economic Design 5*, 229 – 250.

Dutta, B. and S. Mutuswami (1997). Stable networks. *Journal of Economic Theory 76*, 322 – 44.

Fiorina, M. P. (2017). The political parties have sorted. *Hoover Institution Essay on Contemporary Politics, Series No. 3*.

Fowler, J. H. (2006). Connecting the congress: A study of cosponsorship networks. *Political Analysis 14*(4), 456–487.

Gentzkow, M. (2017). Polarization in 2016. *Essay, Stanford University*.

Gentzkow, M. and J. Shapiro (2015). Congressional record for 104th-110th congresses: Text and phrase counts. *ICPSR33501-v4*.

Gentzkow, M., J. M. Shapiro, and M. Taddy (2019). Measuring group differences in high-dimensional choices: Method and application to congressional speech. *Econometrica 87*(4), 1307–1340.

Goldsmith-Pinkham, P. and G. Imbens (2013a). Rejoinder: Social networks and the identification of peer effects. *Journal of Business and Economic Statistics 31:3*, 279–281.

Goldsmith-Pinkham, P. and G. Imbens (2013b). Social networks and the identification of peer effects. *Journal of Business and Economic Statistics 31:3*, 253–264.

Graham, B. S. (2017). An econometric model of network formation with degree heterogeneity. *Econometrica 85*(4), 1033–1063.

Groll, T. and A. Prummer (2016). Whom to lobby? targeting in political networks. Working Paper 808, London.

Groseclose, T. and C. Stewart, III (1998). The value of committee seats in the house, 1947-91. *American Journal of Political Science 42*(2), 453–474.

Harmon, N., R. Fisman, and E. Kamenica (2019). Peer effects in legislative voting. *American Economic Journal: Applied Economics 11*(4), 156–80.

Herings, P. J.-J., A. Mauleon, and V. Vannetelbosch (2009). Farsightedly stable networks. *Games and Economic Behavior 67*(2), 526–541.

Hiller, T. (2017). Peer effects in endogenous networks. *Games and Economic Behavior 105*, 349–367.

Hsieh, C.-S., M. D. König, and X. Liu (2020). A structural model for the coevolution of networks and behavior. *Working Paper*.

Hsieh, C.-S. and L. F. Lee (2016). A social interactions model with endogenous friendship formation and selectivity. *Journal of Applied Econometrics 31*(2), 301–319.

Hsieh, C.-S., L.-F. Lee, and V. Boucher (2019). Specification and estimation of network formation and network interaction models with the exponential probability distribution. *Available at SSRN 2967867*.

Jackson, M. O. (2005). A survey of models of network formation: Stability and efficiency. *Group Formation in Economics; Networks, Clubs and Coalitions, ed. G Demange, M Wooders. Cambridge, UK: Cambridge Univ. Press*.

Jackson, M. O. (2008). *Social and Economic Networks.* Princeton, NJ, USA: Princeton University Press.

Jackson, M. O. (2013). Comment: Unraveling peers and peer effects (comments on goldsmith-pinkham and imbens' "social networks and the identification of peer effects"). *Journal of Business and Economic Statistics 31:3*, 270–273, DOI: 10.1080/07350015.2013.794095.

Jackson, M. O. (2019). The friendship paradox and systematic biases in perceptions and social norms. *Journal of Political Economy 127*(2), 777–818.

Jackson, M. O. and A. Watts (2002). The evolution of social and economic networks. *Journal of Economic Theory 106*(2), 265–295.

Jackson, M. O. and A. Wolinsky (1996). A strategic model of social and economic networks. *Journal of Economic Theory 71*(1), 44–74.

Jackson, M. O. and Y. Zenou (2015). Games on networks. Volume 4 of *Handbook of Game Theory with Economic Applications*, pp. 95–163. Elsevier.

Jo, D. and M. Lowe (2020). Social interactions of lawmakers and the partisan divide: A natural experiment in iceland.

Kessler, D. and K. Krehbiel (1996). Dynamics of cosponsorship. *American Political Science Review 90*(03), 555–566.

Kirkland, J. H. (2011). The relational determinants of legislative outcomes: Strong and weak ties between legislators. *The Journal of Politics 73*(03), 887–898.

Koger, G. (2003). Position taking and cosponsorship in the us house. *Legislative Studies Quarterly 28*(2), 225–246.

König, M., C. J. Tessone, and Y. Zenou (2009). A dynamic model of network formation with strategic interactions.

Lazer, D. (2011). Networks in political science: Back to the future. *PS: Political Science & Politics 44*(01), 61–68.

Leung, M. P. (2015). Two-step estimation of network-formation models with incomplete information. *Journal of Econometrics 188*(1), 182–195.

Leung, M. P. (2019). A weak law for moments of pairwise stable networks. *Journal of Econometrics*.

Manski, C. F. (2013). Comment: Social networks and the identification of peer effects. *Journal of Business and Economic Statistics 31:3*, 273–275.

Masket, S. E. (2008). Where you sit is where you stand: The impact of seating proximity on legislative cue-taking. *Quarterly Journal of Political Science 3*, 301–311.

Mauleon, A., H. Song, and V. Vannetelbosch (2010). Networks of free trade agreements among heterogeneous countries. *Journal of Public Economic Theory 12*(3), 471–500.

Mauleon, A. and V. Vannetelbosch (2016). Network formation games. *The Oxford handbook of the economics of networks, edited by Bramoullé, Yann and Galeotti, Andrea and Rogers, Brian*.

Mayhew, D. R. (1974). *Congress: The electoral connection.* Yale University Press.

Mayhew, D. R. (2005). *Divided we govern: Party control, lawmaking and investigations, 1946-2002.* Yale University Press.

McCarty, N., K. Poole, and H. Rosenthal (2006). *Polarized America.* MIT Press.

Mele, A. (2017). A structural model of segregation in social networks. *Econometrica 85 (3)*, 825–850.

Mele, A. (2018). A structural model of homophily and clustering in social networks. *mimeo, Johns Hopkins University*.

Myerson, R. B. (1998). Population uncertainty and poisson games. *International Journal of Game Theory 27*(3), 375–392.

Norton, P. (2019). Power behind the scenes: the importance of informal space in legislatures. *Parliamentary Affairs 72*(2), 245–266.

de Paula, Á., I. Rasul, and P. Souza (2019). Recovering social networks from panel data: identification, simulations and an application.

de Paula, Á., S. Richards-Shubik, and E. Tamer (2018). Identifying preferences in networks with bounded degree. *Econometrica 86*(1), 263–288.

Rogowski, J. C. and B. Sinclair (2012). Estimating the causal effects of social interaction with endogenous networks. *Political Analysis 20*(3), 316–328.

Routt, G. C. (1938). Interpersonal relationships and the legislative process. *The Annals of the American Academy of Political and Social Science 195*(1), 129–136.

Saia, A. (2018). Random interactions in the chamber: Legislators' behavior and political distance. *Journal of Public Economics 164*, 225–240.

Schneider, J. (2015). One-minute speeches: Current house practices. *CRS Report RL30135.*

Sheng, S. (2020). A structural econometric analysis of network formation games through subnetworks. *Econometrica 88*(5), 1829–1858.

Snyder, J. M. (1989). Election goals and the allocation of campaign resources. *Econometrica: Journal of the Econometric Society*, 637–660.

Squintani, F. (2020). Information transmission in networks of political decision makers. *Working Paper*.

Stewart, III, C. (2012). The value of committee assignments in congress since 1994. *MIT Political Science Department Research Paper 2012-7. Available at SSRN: https://ssrn.com/abstract=2035632* (2), 453–474.

Stewart, III, C. and J. Woon (2016). Congressional committee assignments, 103rd to 114th congresses, 1993–2017: House, august 12, 2016. *Data Archive*.

Wilson, R. K. and C. D. Young (1997). Cosponsorship in the u. s. congress. *Legislative Studies Quarterly 22*(1), 25–43.

Zhang, Y., A. Friend, A. L. Traud, M. A. Porter, J. H. Fowler, and P. J. Mucha (2008). Community structure in congressional cosponsorship networks. *Physica A: Statistical Mechanics and its Applications 387*(7), 1705–1712.

ONLINE APPENDIX

Social and Legislative Activity in Congress

Nathan Canen, Matthew O. Jackson and Francesco Trebbi

Appendix A. Proofs

Proposition 2.2: The limit equilibrium is defined by equations (2.15)-(2.17).

Proof of Proposition 2.2. Recall that we have from equations (2.12) and (2.11), from the First Order Conditions, that:

$$c = \frac{\alpha_i}{x_i^*} + \frac{s_i^{*2}}{x_i^{*2}}, \tag{A.1}$$

and

$$\frac{s_i^*}{x_i^*} = \phi_i \sum_{j \neq i} s_j^* m_{ij}(s^*) x_j^*. \tag{A.2}$$

We also use that:

$$x_i^* = \alpha_i X_{P(i)} \tag{A.3}$$

$$s_i^* = \alpha_i S_{P(i)}, \tag{A.4}$$

for some $X_{P(i)}, S_{P(i)}$, which comes from the fact that $\frac{s_i^*}{x_i^*}$ and $\frac{x_i^*}{\alpha_i}$ are the same for all agents within a party. Let $P(i) \in \{1, 2\}$ be arbitrary.

Using (A.3) in (2.12) implies:

$$\begin{aligned} c &= \frac{\alpha_i}{x_i^*} + \frac{s_i^{*2}}{x_i^{*2}} \\ &= \frac{\alpha_i}{\alpha_i X_{P(i)}} + \frac{\alpha_i^2 S_{P(i)}^2}{\alpha_i^2 X_{P(i)}^2} \\ &= \frac{1}{X_{P(i)}} + \frac{S_{P(i)}^2}{X_{P(i)}^2}. \end{aligned}$$

Multiplying both sides by $X_{P(i)}^2$ yields:

$$cX_{P(i)}^2 = X_{P(i)} + S_{P(i)}^2, \tag{A.5}$$

which is (2.17).

Let us now substitute (A.3) in (2.11):

$$\begin{aligned}
\frac{\alpha_i S_{P(i)}}{\alpha_i X_{P(i)}} &= \phi_{P(i)} \sum_{j\neq i} \alpha_j S_{P(j)} m_{ij}(s^*) \alpha_j X_{P(j)} \\
\frac{S_{P(i)}}{X_{P(i)}} &= \phi_{P(i)} \sum_{j\neq i} \alpha_j^2 X_{P(j)} S_{P(j)} m_{ij}(s^*) \\
&= \phi_{P(i)} \sum_{j\neq i} \alpha_j^2 X_{P(j)} S_{P(j)} \left(p(i) \frac{p(j)}{\sum_{k\in P(i), k\neq i} p(k) s_k^*} + (1-p(i)) \frac{(1-p(j))}{\sum_{k\neq i}(1-p(k))s_k^*} \right) I_{\{j\in P(i)\}} \\
&\quad + \phi_{P(i)} \sum_{j\neq i} \alpha_j^2 X_{P(j)} S_{P(j)} \left((1-p(i)) \frac{(1-p(j))}{\sum_{k\neq i}(1-p(k))s_k^*} \right) I_{\{j\notin P(i)\}}.
\end{aligned}$$

Note that for the first two terms, $p(i) = p(j)$ because they are only summed when $j \in P(i)$. For the last, $p(i) \neq p(j)$ as it is summed when $j \notin P(i)$.

Rewriting the above with this implies:

$$\begin{aligned}
\frac{S_{P(i)}}{X_{P(i)}} &= \phi_{P(i)} \sum_{j\neq i} \alpha_j^2 X_{P(i)} S_{P(i)} \left(p(i) \frac{p(i)}{\sum_{k\in P(i), k\neq i} p(i) s_k^*} + (1-p(i)) \frac{(1-p(i))}{\sum_{k\neq i}(1-p(k))s_k^*} \right) I_{\{j\in P(i)\}} \\
&\quad + \phi_{P(i)} \sum_{j\neq i} \alpha_j^2 X_{P(j)} S_{P(j)} \left((1-p(i)) \frac{(1-p(j))}{\sum_{k\neq i}(1-p(k))s_k^*} \right) I_{\{j\notin P(i)\}}.
\end{aligned}$$

Using that $s_k^* = \alpha_k S_{P(k)}$ leads to:

$$\begin{aligned}
\frac{S_{P(i)}}{X_{P(i)}} &= \phi_{P(i)} \sum_{j\neq i} \alpha_j^2 X_{P(i)} S_{P(i)} \left(\frac{p(i)^2}{p(i) \sum_{k\in P(i), k\neq i} \alpha_k S_{P(k)}} + \frac{(1-p(i))^2}{\sum_{k\neq i}(1-p(k))\alpha_k S_{P(k)}} \right) I_{\{j\in P(i)\}} \\
&\quad + \phi_{P(i)} \sum_{j\neq i} \alpha_j^2 X_{P(j)} S_{P(j)} \left(\frac{(1-p(i))(1-p(j))}{\sum_{k\neq i}(1-p(k))\alpha_k S_{P(k)}} \right) I_{\{j\notin P(i)\}}.
\end{aligned}$$

Let us focus on the case of $P(i) = 1$, as the other case is symmetric.

$$\begin{aligned}
\frac{S_1}{X_1} &= \phi_1 \sum_{j\neq i} \alpha_j^2 X_1 S_1 \left(\frac{p_1}{\sum_{k\in P(i), k\neq i} \alpha_k S_1} + \frac{(1-p_1)^2}{\sum_{k\neq i}(1-p(k))\alpha_k S_{P(k)}} \right) I_{\{j\in P(i)\}} \\
&\quad + \phi_1 \sum_{j\neq i} \alpha_j^2 X_2 S_2 \left(\frac{(1-p_1)(1-p_2)}{\sum_{k\neq i}(1-p(k))\alpha_k S_{P(k)}} \right) I_{\{j\notin P(i)\}}.
\end{aligned}$$

Finally, we use that:

$$\begin{aligned}
\sum_{k\neq i}(1-p(k))\alpha_k S_{P(k)} &= \sum_{k\neq i, k\in P(i)}(1-p(k))\alpha_k S_{P(k)} + \sum_{k\neq i, k\notin P(i)}(1-p(k))\alpha_k S_{P(k)} \\
&= \sum_{k\neq i, k\in P(i)}(1-p_1)\alpha_k S_1 + \sum_{k\neq i, k\notin P(i)}(1-p_2)\alpha_k S_2 \\
&= (1-p_1)S_1 \sum_{k\neq i, k\in P(i)}\alpha_k + (1-p_2)S_2 \sum_{k\neq i, k\notin P(i)}\alpha_k \\
&= (1-p_1)S_1A_1 + (1-p_2)S_2A_2.
\end{aligned}$$

To finalize the calculations, we use the simplification above for the denominators of the second and third terms.

Note that only α_j is now a function of the summand j itself, in the main expression. We also note that we can now use the indicators of $j \in P(i)$ for the first two terms, and $j \notin P(i)$ of the last term, within sums. These observations lead to the final equation:

$$\begin{aligned}
\frac{S_1}{X_1} &= \phi_1 X_1 S_1 \sum_{j\neq i}\alpha_j^2 \left(\frac{p_1}{S_1\sum_{k\in P(i),k\neq i}\alpha_k} + \frac{(1-p_1)^2}{(1-p_1)S_1A_1+(1-p_2)S_2A_2}\right) I_{\{j\in P(i)\}} \\
&\quad + X_2S_2\phi_1 \sum_{j\neq i}\alpha_j^2 \left(\frac{(1-p_1)(1-p_2)}{(1-p_1)S_1A_1+(1-p_2)S_2A_2}\right) I_{\{j\notin P(i)\}} \\
&= \phi_1 X_1 S_1 \sum_{j\neq i, j\in P(i)}\alpha_j^2 \left(\frac{p_1}{S_1A_1} + \frac{(1-p_1)^2}{(1-p_1)S_1A_1+(1-p_2)S_2A_2}\right) \\
&\quad + X_2S_2\phi_1 \sum_{j\neq i, j\notin P(i)}\alpha_j^2 \left(\frac{(1-p_1)(1-p_2)}{(1-p_1)S_1A_1+(1-p_2)S_2A_2}\right) \\
&= \phi_1 X_1 S_1 B_1 \left(\frac{p_1}{S_1A_1} + \frac{(1-p_1)^2}{(1-p_1)S_1A_1+(1-p_2)S_2A_2}\right) + X_2S_2\phi_1 B_2 \left(\frac{(1-p_1)(1-p_2)}{(1-p_1)S_1A_1+(1-p_2)S_2A_2}\right) \\
&= \phi_1 \left(\frac{X_1B_1p_1}{A_1} + \frac{X_1S_1B_1(1-p_1)^2}{(1-p_1)S_1A_1+(1-p_2)S_2A_2} + \frac{X_2S_2B_2(1-p_1)(1-p_2)}{(1-p_1)S_1A_1+(1-p_2)S_2A_2}\right) \\
&= \phi_1 \left(\frac{p_1X_1B_1}{A_1} + \frac{(1-p_1)^2B_1S_1X_1 + (1-p_1)(1-p_2)B_2X_2S_2}{(1-p_1)A_1S_1+(1-p_2)A_2S_2}\right).
\end{aligned}$$

□

Proof of Proposition 2.3. Recall that an interior equilibrium is a solution to (2.15) to (2.17).

So, rewriting these:

(A.6) $$S_1 = X_1\phi_1 \left(\frac{p_1B_1X_1}{A_1} + \frac{(1-p_1)^2B_1S_1X_1 + (1-p_1)(1-p_2)B_2S_2X_2}{(1-p_1)A_1S_1+(1-p_2)A_2S_2}\right).$$

(A.7) $$S_2 = X_2\phi_2 \left(\frac{p_2B_2X_2}{A_2} + \frac{(1-p_2)^2B_2S_2X_2 + (1-p_1)(1-p_2)B_1S_1X_1}{(1-p_1)A_1S_1+(1-p_2)A_2S_2}\right).$$

$$cX_1^2 = X_1 + S_1^2, \qquad cX_2^2 = X_2 + S_2^2. \tag{A.8}$$

Substituting (A.6) into (A.8) leads to

$$cX_1^2 = X_1 + X_1^2\phi_1^2 \left(\frac{p_1 B_1 X_1}{A_1} + \frac{(1-p_1)^2 B_1 S_1 X_1 + (1-p_1)(1-p_2) B_2 S_2 X_2}{(1-p_1)A_1 S_1 + (1-p_2)A_2 S_2} \right)^2.$$

or

$$cX_1 = 1 + X_1\phi_1^2 \left(\frac{p_1 B_1 X_1}{A_1} + \frac{(1-p_1)^2 B_1 S_1 X_1 + (1-p_1)(1-p_2) B_2 S_2 X_2}{(1-p_1)A_1 S_1 + (1-p_2)A_2 S_2} \right)^2. \tag{A.9}$$

There is a similar expression for S_2, X_2. Note that the right hand side of (A.9) lies above the left hand side as we approach $X_1 = 0$ (same for X_2). To have an interior solution, we need the right hand side to sometimes fall at or below the left hand side for positive X_1.

Suppose that the equilibrium (when it exists) is such that $X_1 \geq X_2$, and the other case is analogous just reversing subscripts everywhere. Then the right hand side is less than what we get by replacing X_2 by X_1, and so we want

$$cX_1 \geq 1 + X_1^3\phi_1^2 \left(\frac{p_1 B_1}{A_1} + \frac{(1-p_1)^2 B_1 S_1 + (1-p_1)(1-p_2) B_2 S_2}{(1-p_1)A_1 S_1 + (1-p_2)A_2 S_2} \right)^2. \tag{A.10}$$

for some interior X_1. Rewriting

$$cX_1 \geq 1 + X_1^3\phi_1^2 \left(\frac{p_1 B_1}{A_1} + \frac{(1-p_1)^2 B_1 + (1-p_1)(1-p_2) B_2 \frac{S_2}{S_1}}{(1-p_1)A_1 + (1-p_2)A_2 \frac{S_2}{S_1}} \right)^2. \tag{A.11}$$

The right hand side is maximized either at $\frac{S_2}{S_1} = 0$ or $\frac{S_2}{S_1} = \infty$, and so it is sufficient to have

$$cX_1 \geq 1 + X_1^3\phi_1^2 \left(p_1 \frac{B_1}{A_1} + (1-p_1) \max\left[\frac{B_1}{A_1}, \frac{B_2}{A_2} \right] \right)^2. \tag{A.12}$$

Let

$$D_1 = p_1 \frac{B_1}{A_1} + (1-p_1) \max\left[\frac{B_1}{A_1}, \frac{B_2}{A_2} \right]$$

Then (A.12) can be rewritten as

$$cX_1 \geq 1 + X_1^3\phi_1^2 D_1^2. \tag{A.13}$$

for some positive X_1. Note that

$$D_1 \leq D = \max\left[\frac{B_1}{A_1}, \frac{B_2}{A_2} \right]$$

So, it is sufficient to have

$$cX_1 \geq 1 + X_1^3\phi_1^2 D^2. \tag{A.14}$$

for some positive X_1.

It is necessary and sufficient to check that the left hand side and right hand side are tangent at the point at which the slope of the right hand side is c. This happens at $X_1 = \sqrt{\frac{c}{3\phi_1^2 D^2}}$

and then the corresponding sufficient condition becomes:

$$c\left(\frac{c}{3\phi_1^2 D^2}\right)^{1/2} \geq 1 + \left(\frac{c}{3\phi_1^2 D^2}\right)^{3/2} \phi_1^2 D^2, \tag{A.15}$$

or

$$\frac{2c^{3/2}}{3\sqrt{3}} \geq \phi_1 D. \tag{A.16}$$

Having this hold also for the other case, leads to the claimed expression. □

Appendix B. Additional Aspects of the Theory

B.1. **Best Response Dynamics.** Best response dynamics are described as follows. Consider starting at some vectors s^0, x^0. Then the best response dynamics are described by:

$$s_i^t = x_i^{t-1}\phi_i \sum_{j\neq i} m_{ij}(s^{t-1}) s_j^{t-1} x_j^{t-1}, \tag{B.1}$$

and

$$x_i^t = \frac{\alpha_i}{c} + s_i^{t-1}\frac{\phi_i}{c}\sum_{j\neq i} m_{ij}(s^{t-1}) s_j^{t-1} x_j^{t-1}. \tag{B.2}$$

It follows that if $s^0 = \mathbf{0}$, then $m_{ij}(s^{t-1}) = 0$ for all ij (recall Footnote 11) and we get immediate convergence to $s_i^t = 0, x_i^t = \frac{\alpha_i}{c}$ for all t. Otherwise, s^t, x^t will be positive for all t.

To see how these best response dynamics work for a special case, let us consider the situation in which there is some S^0, X^0 such that $s_i^0 = \alpha_i S^0$ and $x_i^0 = \alpha_i X^0$ (which has to eventually hold at any limit point) - i.e. when we can use Proposition 2.2.[51]

In that case, working with the limiting or continuum case, in which the matching function is symmetric within a party, and presuming that $S_k^{t-1} > 0$ for each party (which happens after the first period if some $s_j^0 > 0$ and otherwise the solution is already described above), we end up with the following dynamics. For party k (letting k' denote the other party):

$$S_k^t = X_k^{t-1}\phi_k\left(m_{kk}(S^{t-1})B_k S_k^{t-1} X_k^{t-1} + m_{kk'}(S^{t-1})B_{k'} S_{k'}^{t-1} X_{k'}^{t-1}\right), \tag{B.3}$$

and

$$X_k^t = \frac{1}{c} + S_k^{t-1}\frac{\phi_k}{c}\left(m_{kk}(S^{t-1})B_k S_k^{t-1} X_k^{t-1} + m_{kk'}(S^{t-1})B_{k'} S_{k'}^{t-1} X_{k'}^{t-1}\right). \tag{B.4}$$

where

$$m_{kk}(S^{t-1}) = \frac{p_k}{S_k A_k} + \frac{(1-p_k)^2}{(1-p_1)S_1 A_1 + (1-p_2)S_2 A_2},$$

and

$$m_{kk'}(S^{t-1}) = \frac{(1-p_1)(1-p_2)}{(1-p_1)S_1 A_1 + (1-p_2)S_2 A_2}.$$

[51]This is also useful in determining the instability of equilibria.

B.2. **Discussion of the Model.** The extensive literature on network formation, starting from its early incarnation in Jackson and Wolinsky (1996); Dutta and Mutuswami (1997); Bala and Goyal (2000); Currarini and Morelli (2000); Jackson and Watts (2002); Jackson (2005); Herings, Mauleon, and Vannetelbosch (2009), provides insight into how networks form, when inefficient networks form, and how that depends on the setting. More recently, the literature has also begun to develop models that incorporate some heterogeneity and are still tractable enough to allow for fitting the models to data, as in Leung (2015); Sheng (2020); Chandrasekhar and Jackson (2016); Mele (2017); Graham (2017); de Paula, Richards-Shubik, and Tamer (2018); Leung (2019); and some of that literature also allows for homophily, such as Currarini, Jackson, and Pin (2009, 2010); Banerjee, Chandrasekhar, Duflo, and Jackson (2018); Mele (2018). The models that are tractable enough to fit to data require a structure that limits the multiplicity of stable (equilibrium) networks, and such that those can be estimated with a practical number of calculations.

We only have a handful of such estimable models that involve non-trivial interaction effects; and generally those models are stylized in some way. For instance, as shown in Sheng (2020), the choice of the specific model can be important since models with indirect network effects (utility from friends-of-friends) lead to (i) a lack of identification (multiple configurations of parameters leading to the same outcomes) and, (ii) computational intractability with as few as 20 players, due to a curse of dimensionality. To make progress, she proposes a model with endogenous links that have "dependence [that] has a particular structure such that conditional on some network heterogeneity and individual heterogeneity, the links become independent." An alternative approach is that of Mele (2017). He proposes an empirical model of network formation that allows for homophily in network formation. Again, he shows that there is a curse of dimensionality in using standard estimation methods unless some strong asymptotic independence conditions are satisfied. Other approaches are to have certain subgraphs generate value and then model the formation of those subgraphs directly (Chandrasekhar and Jackson, 2016), or to have payoffs based on combinations of individual characteristics, geography, or assortativity (e.g., Currarini et al., 2009; Leung, 2015; Graham, 2017; Leung, 2019).

Here we want a model in which the value to a given pairing depends on their subsequent mutual (legislative) efforts, and so we need a model in which expected values of links can be calculated conditional upon future efforts, and those efforts can also be characterized as a function of the pairings. Using random meeting probabilities to derive link formation does exactly this by reducing the dimension of choices while allowing for rich interdependencies, homophily, and still yielding a clean characterization of both types of efforts. The model we work with is the only one we have found in the literature that fits all of these criteria, and which are needed for this application.

In summary, one has to be judicious in modeling network formation to obtain a formulation that also allows for homophily, as well as choices that affect network positions, and remains both well-identified and estimable. Meanwhile, existing empirical models of games on networks that are well-identified (e.g., de Paula et al., 2019, advancing the work of Bramoullé et al., 2009) do not allow for endogenous networks - they assume that the network is fixed and exogenous, and require a different data set-up than ours.[52]

Thus one can see why models that incorporate both behavior and network formation are few: Cabrales, Calvó-Armengol, and Zenou (2011); König, Tessone, and Zenou (2009); Goldsmith-Pinkham and Imbens (2013b); Hiller (2017); Badev (2017, 2020); Hsieh, Lee, and Boucher (2019); Hsieh, König, and Liu (2020). These models necessarily sacrifice some richness in order to incorporate both network formation and endogenous behaviors and to allow for an interaction between them. Nonetheless, they can still be quite rich and, as we show here, can still fit data well. Of this class, in order to work with a tractable model that we can extend to have yet a third dimension of group identity and homophily, and still take to (static) data, we build upon the model of Cabrales, Calvó-Armengol, and Zenou (2011).[53] This introduces another dimension to the estimation, of group interaction rates, and thus requires that the model be tractable enough to still solve with a third dimension of endogeneity. Finally, a close look at the fit of the data provides support for our modeling choices.[54] In Section 6.1, we show that the predicted links from the model are highly correlated with measures that use disaggregate data between Congress members (e.g. Fowler, 2006), even if we do not use the latter in estimation. Second, strategic outcomes on this network, such as the probability of bill approval, are fit well within and across parties/politicians. Third, homophily is quantitatively important: a model with homophily fits the data significantly better than one without it. Fourth, our model is shown to outperform alternative approaches to the characterization of G in terms of in-sample mean squared error.

B.3. **Examples of Equilibrium G Across Parameter Configurations.** In our model, $\{\alpha_i, c, \phi_i, p_1, p_2\}$ parametrize the incentives for social and legislative effort. The network G is then generated by agents' strategic decisions, taking those incentives into account. Here, we showcase the rich class of equilibrium networks G that can arise in as the parameters vary.

The parameter values are purposefully kept similar to those in Figure 1 for comparability, although we set $n_1 = n_2 = 2$ to visualize G (a 4 x 4 matrix). For convenience, we let politicians 1 and 2 be in Party 1, and politicians 3 and 4 in Party 2, so that the first two rows/columns of G below represent connections among Party 1 members. We set $c = 2.25$ and we keep ϕ_i constant within parties. For Party 1, we set $\phi_1 = 1$. We set $\alpha_i = 1$ for one

[52]For example, to recover an unobserved exogenous network as they set out to, de Paula et al. (2019) assumes (i) an exogenous network that is sufficiently sparse, (ii) the network does not change over time, (iii) a panel data structure, with large enough time-series dimension, and (iv) a linear in means model. Our set-up and data structure do not have any of these 4 properties, as alluded to previously. Furthermore, the estimation of this set-up must involve shrinkage estimators and their resulting bias due to the size of the parameter space (N^2 parameters to recover from just the network itself).

[53]Recent alternatives that accommodate all three dimensions include Badev (2020), Hsieh et al. (2019) and Hsieh et al. (2020). While they constitute important advances to the literature, their solutions are not applicable to our problem. For instance, all of them assume observable networks and unidimensional actions (in the case of Badev, 2020, actions are binary). In terms of estimation, we complement their Bayesian methods with a frequentist approach which shows identification of our parameters and provides a less computationally intensive estimation procedure.

[54]Most notably, we see the value of (i) a (biased) random socialization protocol, (ii) choices made on effort levels, and (iii) mean-zero i.i.d. measurement errors on our observed proxies.

member in each party (politicians 1 and 3). Our examples below vary the remaining five parameters $\{p_1, p_2, \phi_2, \alpha_2, \alpha_4\}$.

Example 1: Complete Bipartisanship.
Parameters: $p_1 = p_2 = 0, \alpha_2 = \alpha_4 = 1, \phi_2 = 1$.

$$G = \begin{bmatrix} 0 & 0.063 & 0.063 & 0.063 \\ 0.063 & 0 & 0.063 & 0.063 \\ 0.063 & 0.063 & 0 & 0.063 \\ 0.063 & 0.063 & 0.063 & 0 \end{bmatrix}$$

There are no partisan biases in relative rates of meeting potential partners, and all members end up equally connected. This is because all politicians are identical, have high enough types for social interaction to occur, there is no homophily to bias social interactions ($p_1 = p_2 = 0$), and there are no differential incentives to socialize by party ($\phi_1 = \phi_2$). Here, all politicians exert the same social efforts s_i.

Example 2: Bias in Mixing
Parameters: $p_1 = p_2 = 0.5, \alpha_2 = \alpha_4 = 1, \phi_2 = 1$.

$$G = \begin{bmatrix} 0 & 0.094 & 0.031 & 0.031 \\ 0.094 & 0 & 0.031 & 0.031 \\ 0.031 & 0.031 & 0 & 0.094 \\ 0.031 & 0.031 & 0.094 & 0 \end{bmatrix}$$

Relative to Example 1, the introduction of partisanship (structural homophily $p_1 = p_2 = 0.5$) biases social interactions along party lines, despite politicians having identical types (α_is) and identical party-level incentives to socialize (ϕs). In this example, politicians all exert the same social efforts s_i.

Example 3: Full Partisanship
Parameters: $p_1 = 1, p_2 = 0.5, \alpha_2 = \alpha_4 = 1, \phi_2 = 1$.

$$G = \begin{bmatrix} 0 & 0.125 & 0 & 0 \\ 0.125 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.125 \\ 0 & 0 & 0.125 & 0 \end{bmatrix}$$

Party 1 is fully partisan, so its members can never meet those in Party 2. This induces full sorting along party lines. Even though party 2 would be willing to mix with party 1, they do not manage to, given that party 1 does not mix.

Example 4: Nuanced Biased Socialization
Parameters: $p_1 = p_2 = 0.5, \alpha_2 = \alpha_4 = 0.5, \phi_2 = 1$.

$$G = \begin{bmatrix} 0 & 0.048 & 0.032 & 0.016 \\ 0.048 & 0 & 0.016 & 0.008 \\ 0.032 & 0.016 & 0 & 0.048 \\ 0.016 & 0.008 & 0.048 & 0 \end{bmatrix}$$

We revisit Example 2, but decrease the type of one politician in each party ($\alpha_2 < 1, \alpha_4 < 1$). As a result, socialization is still biased along parties, but in a heterogeneous way. Politicians socialize more often within parties, but the higher type politicians (politicians 1 and 3) have greater incentives to legislate, and that induces them to socialize more, and they are more likely to meet across party lines than the low type ones.

Example 5: Nuanced Biased Socialization II
Parameters: $p_1 = p_2 = 0.5, \alpha_2 = \alpha_4 = 0.5, \phi_2 = 1.3$.

$$G = \begin{bmatrix} 0 & 0.046 & 0.039 & 0.019 \\ 0.046 & 0 & 0.019 & 0.010 \\ 0.039 & 0.019 & 0 & 0.075 \\ 0.019 & 0.010 & 0.075 & 0 \end{bmatrix}$$

We now increase ϕ_2 relative to Example 4. This increases the incentives to socialize for politicians in Party 2, yielding stronger equilibrium connections among them. However, politicians in Party 1 understand this and increase their social efforts as well. This allows them meet those Party 2 members more often (since the latter's externalities are now higher). Compared to Example 4, this yields stronger connections between Politician 1 and those in the opposing party, but weaker connections within Party 1 *despite homophily*.

Example 6: Reversal of Partisanship Despite Homophily
Parameters: $p_1 = 0.3, p_2 = 0.5, \alpha_2 = 0.5, \alpha_4 = 1, \phi_2 = 1.3$.

$$G = \begin{bmatrix} 0 & 0.045 & 0.051 & 0.051 \\ 0.045 & 0 & 0.026 & 0.026 \\ 0.051 & 0.026 & 0 & 0.160 \\ 0.051 & 0.026 & 0.160 & 0 \end{bmatrix}$$

We now begin with the parameters in Example 5, and then decrease p_1 and increase α_4. The high type politician in Party 1 now has stronger connections with opposing party politicians than with 1's own party member, *despite Party 2's strong homophily*. Politicians in Party 1 internalize the stronger types and incentives to socialize in Party 2, and choose effort that is large enough to overcome such homophily.

B.4. **Endogenous Partisanship.** A natural extension of our model would be to endogenize the p_i's. We comment here on potential directions and issues that arise.

First, it is easy to see that if one simply endogenized the p_i's within the current model without introducing any costs of affecting p_i, then the solutions would be corner solutions. If a group can choose its p_i without having any costs of selecting p_i, then (generically in the parameters) one of the two groups would want to be entirely partisan, since one of the two groups would find interacting with itself more beneficial than interacting across the aisle. Such a corner solution is clearly of little interest, and is incompatible with our empirical estimates.

More generally, there are interactions, both within and across parties, that happen naturally due to committee membership among other things and would be difficult to prevent, and others that might be costly to encourage. This suggests that there would minimum and maximum levels of partisanship that could be attained and also that one would need to model a nonlinear cost of partisanship. Once one provided a nonlinear cost to capture the high cost of going to either extreme of $p_i = 0$ or $p_i = 1$, one would end up with an interior equilibrium. A challenge would be that this could be dependent upon the cost formulation, and so one would need to work with a flexible enough cost function to allow the model to fit the data.

Having three endogenous choices for each of the two parties - partisanship, social effort, and legislative effort - would then end up producing a model for which analytic characterizations of the equilibrium would no longer be possible, and for which the multiplicity of equilibria would more difficult to ascertain. There would be two approaches. One would be to work entirely with numerical simulations. Since the interest in endogenizing partisanship levels would presumably be to understand how they interact with other variables and change incentives, this would require a very rich and complex set of simulations, especially as they would be sensitive to the choice of the cost function.

Another approach, and perhaps the most fruitful, would be to fix one of the other effort variables and return to a model in which there are just two different action variables that agents/parties are making. Given the importance of partisanship on the endogeneity of the network, a starting point might be to fix the x_i's and then work with the other variables. This could be an interesting approach for further research. We chose to work with endogenizing the network and legislative effort, holding partisanship constant, as these seem to be the first-order questions, but understanding partisanship is also a very interesting topic.

Appendix C. Formal Arguments for Identification

Recall that in our extension in Section 2.3, preferences are given by:

$$\tilde{u}_i(x_i, x_{-i}) = \alpha_i x_i + \phi_i \sum_j g_{ij} x_i x_j - \frac{1}{2} c x_i^2 - \frac{1}{2} s_i^2, \tag{C.1}$$

where α_i is now interpreted as the heterogeneous marginal cost of legislative effort for i and $\phi_i \equiv \gamma_{P(i)} e^{-\lambda\rho V_{i,0}}(1 - e^{-\lambda\zeta_{P(i)}})$, where $\zeta_{P(i)}$ was the electoral return to passing a bill (in the reelection equation), $\gamma_{P(i)}$ was the scale parameter in the shock for passing the bill, λ was the parameter from the exponentially distributed reelection shock, and $V_{i,0}$ is the winning margin for i in the previous election.

We now prove (point) identification of the following parameters from this model: $\{\{\alpha_i\}_{i=1}^n, \lambda\rho, \{\lambda\zeta_{P(i)}\}_{P(i)=1,2}, c, \tilde{A}_1\gamma_1, \tilde{A}_2\gamma_2\}$. To prove identification, we make use of the equilibrium conditions of s_i^* and x_i^* derived from the first order conditions. These are given in equations (2.7) and (2.8) in the main text. We also use equation (2.4) on the probability of passing a bill. It will also prove useful to work with the equation combining (2.8) into (2.7):

$$x_i^* = \frac{1}{c}\left(\alpha_i + \frac{s_i^{*2}}{x_i^*}\right) \tag{C.2}$$

Finally, recall that we impose a normalization to pin down the location of the distribution of α_i in the first Congress in the sample. Below, we simply assume that there is a legislator 0 with α_0 known, although in the empirical specifications, we simply omit the constant from z_i (which implies that we know α_i for an i with $z_i = 0$) since we parametrize α_i. Note that the arguments below do not rely on having measurement errors or on the parametrization of α_i. We drop the notation τ as our identification arguments are valid within each Congress.[55]

Dividing both sides of (2.7) by x_i^* for an arbitrary politician i yields:

$$\frac{s_i^*}{x_i^*} = \gamma_{P(i)}(1 - e^{-\lambda\zeta_{P(i)}}) e^{-\lambda\rho V_{i,0}} \sum_{j\neq i} s_j^* m_{ij}(s^*) x_j^*. \tag{C.3}$$

Now, dividing (C.3) by its analogue for a politician j in the same party as i for whom $V_{j,0} \neq V_{i,0}$ yields:

$$\frac{s_i^*/x_i^*}{s_j^*/x_j^*} = e^{-\lambda\rho(V_{i,0} - V_{j,0})}, \tag{C.4}$$

where we have used that $P(i) = P(j)$ and that $\tilde{A}_{P(i)} = \sum_{j\neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*$ is constant across politicians. It follows that we identify the product $\lambda\rho$.[56]

[55] The normalization assumption can be imposed only in one Congress, as we can rely on the overlap of politicians across Congresses to maintain the assumption in later periods.

[56] For completeness, when s_i^*, x_i^* are observed with measurement error as in (5.1), we find that:

$$\frac{s_i^*/x_i^*}{s_j^*/x_j^*} = \frac{s_i/x_i}{s_j/x_j} e^{(\epsilon_i - \epsilon_j) + (v_j - v_i)}. \tag{C.5}$$

Since the measurement errors are independent and mean 0, we can apply a log operator and then the expectation operator on both sides of (C.5). Hence, $\lambda\rho$ is still identified.

We can identify $\zeta_{P(i)}$ across parties by rewriting (2.4) using (2.7):

$$\begin{aligned} P(y_i = 1) &= \gamma_{P(i)} \sum_{j \neq i} g_{ij}(s^*) x_i^* x_j^* \\ &= \gamma_{P(i)} s_i^* \sum_{j \neq i} s_j^* m_{ij}(s^*) x_i^* x_j^* \\ &= \frac{\gamma_{P(i)}}{\phi_i} s_i^{*2} \\ &= \frac{1}{e^{-\lambda\rho V_{i,0}}(1 - e^{-\lambda\zeta_{P(i)}})} s_i^{*2}. \end{aligned} \tag{C.6}$$

where the third line uses (2.7) and the last line uses the definition of ϕ_i. The only unknown in the last line is $\lambda\zeta_{P(i)}$. When accounting for measurement error, it suffices to note that $\log(P(y_i = 1)) = \log(\frac{1}{1-e^{-\lambda\zeta_{P(i)}}} e^{\lambda\rho V_{i,0}} s_i^2) + 2\epsilon_i$, where ϵ_i is mean 0. Hence, $\lambda\zeta_{P(i)}$ is identified for both parties by the average probability of passing a bill for politicians in $P(i)$ given their observed effort levels.

Now, let us return to (C.3). The product $\tilde{A}_{P(i)}\gamma_{P(i)}$ is the only unknown on the right hand side, so it is identified for the arbitrary party $P(i)$. As a result, the ratio $\tilde{A}_1\gamma_1/\tilde{A}_2\gamma_2$ can be identified. The intuition is easily seen by dividing (C.3) for i and k for different parties.

$$\frac{s_i^*/x_i^*}{s_k^*/x_k^*} = \frac{\tilde{A}_{P(i)}\gamma_{P(i)}(1 - e^{-\lambda\zeta_{P(i)}})}{\tilde{A}_{P(k)}\gamma_{P(k)}(1 - e^{-\lambda\zeta_{P(k)}})} e^{-\lambda\rho(V_{i,0} - V_{k,0})}, \tag{C.7}$$

so that this ratio is identified by the systematic variation in relative choices of social and legislative effort across members of opposite parties.

We now proceed with identification of α_i for all i. To do so, we rewrite (C.2) as:

$$x_i^* = \frac{\alpha_i}{c - \left(\phi_i \sum_{j \neq i} s_j^* m_{ij}(\mathbf{s}^*) x_j^*\right)^2} \tag{C.8}$$

Taking logs and using (5.1) implies[57]:

$$\begin{aligned} \log(x_i) + v_i &= \log(\alpha_i) - \log(c - \tilde{A}_{P(i)}^2 \phi_i^2) \\ \log(x_i) &= \log(\alpha_i) - \log(c - \tilde{A}_{P(i)}^2 \phi_i^2) - v_i \\ &= \log(\alpha_i) - \log\left(c - (\tilde{A}_{P(i)}\gamma_{P(i)} e^{-\lambda\rho V_{i,0}}(1 - e^{-\lambda\zeta_{P(i)}}))^2\right) - v_i. \end{aligned} \tag{C.9}$$

Recall that the term $((\tilde{A}_{P(i)}\gamma_{P(i)})e^{-\lambda\rho V_{i,0}}(1 - e^{-\lambda\zeta_{P(i)}}))^2$ has already been identified, as each of its 3 components are identified. Hence, (C.9) has only 2 unknowns: c and α_i. Since this equation is valid for every i, it is also valid for the normalizer politician "0" whose type α_0 is known by assumption. Hence, c is pinned down by equation (C.9) for the normalizer, as c is the only unknown in that case and $\log(\cdot)$ is a strictly monotonic function. With measurement error, we also use that v_i is mean 0 and i.i.d.

Once c is pinned down, then α_i is identified for every i from equation (C.9) under the analogous argument, as α_i is the only unknown.

[57]In the absence of measurement error, simply replace $v_i = 0$ below and the same arguments stand.

Appendix D. Rewriting the Model in terms of Moment conditions over i

In this Section, we provide the derivation for transforming the model's equilibrium outcomes to the moment equations described in Section 5. Let us begin with (C.3):

$$\frac{s_i}{x_i} e^{\epsilon_i - v_i} = \tilde{A}_{P(i)} \gamma_{P(i)} (1 - e^{-\lambda \zeta_{P(i)}}) e^{-\lambda \rho V_{i,0}}.$$

We can rewrite this expression as:

$$\log \left(\frac{s_i}{x_i} \right) = \log(\tilde{A}_{P(i)} \gamma_{P(i)}) + \log(1 - e^{-\lambda \zeta_{P(i)}}) - \lambda \rho V_{i,0} + (v_i - \varepsilon_i).$$

Applying expectations over the measurement errors (which are mean zero) on both sides of the expression yields the first set of equations above. They are analogous to moment conditions which coincide with the OLS estimator with party-specific intercept parameters.[58]

For the second set of equations, we use the parametrization (5.3) in (C.9) to obtain:

$$\log(x_i) = z_i' \beta - \log \left(c - (\tilde{A}_{P(i)} \gamma_{P(i)} e^{-\lambda \rho V_{i,0}} (1 - e^{-\lambda \zeta_{P(i)}}))^2 \right) - v_i.$$

Exploiting the orthogonality conditions on v_i yields the second set of moment conditions. We note that the location normalization is important here: otherwise, c could not be separately identified from the constant in $z_i'\beta$. In fact, the cost of legislative effort c is pinned down by the average legislative behavior of politicians, conditional on individual characteristics and electoral returns. But it could be increased if all types are similarly increased.

For the final equation, we rewrite (C.6) using (5.1):

$$P(y_i = 1) = \frac{1}{e^{-\lambda \rho V_{i,0}} (1 - e^{-\lambda \zeta_{P(i)}})} s_i^2 e^{2\epsilon_i},$$

which implies that:

$$\log(P(y_i = 1)) = \lambda \rho V_{i,0} - \log(1 - e^{-\lambda \zeta_{P(i)}}) + 2 \log(s_i) + 2\epsilon_i.$$

D.1. Details on Estimation.

We now provide further details on how the estimation procedure was implemented, including the starting values for the numerical solution to the GMM optimizer and numerical details on the computation of standard errors.

D.1.1. *OLS and plug-in Approach as Starting Values for Optimization.*

For the starting values for GMM optimization, we use simple closed form estimates for most parameters of interest, borne out of the separability of the moment equations. We then use different starting points for the remaining parameter, c.

[58] We note that $\log(1 - e^{-\lambda \zeta_{P(i)}})$ can be split from the term $\log(\tilde{A}_{P(i)} \gamma_{P(i)})$ since it is identified from another equation.

More precisely, recall that the estimating equations are the empirical counterparts to equations (5.4)-(5.10) and are given by:

$$\frac{1}{n}\sum_{i=1}^{n}\left(\log\left(\frac{s_i}{x_i}\right)-\log(\tilde{A}_1\gamma_1)-\tilde{\zeta}_1+\lambda\rho V_{i,0}\right)I_{\{i\in P_1\}} = 0 \tag{D.1}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\log\left(\frac{s_i}{x_i}\right)-\log(\tilde{A}_2\gamma_2)-\tilde{\zeta}_2+\lambda\rho V_{i,0}\right)I_{\{i\in P_2\}} = 0 \tag{D.2}$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(\log\left(\frac{s_i}{x_i}\right)-\log(\tilde{A}_{P(i)}\gamma_{P(i)})-\tilde{\zeta}_{P(i)}+\lambda\rho V_{i,0}\right)V_{i,0} = 0 \tag{D.3}$$

$$\frac{1}{n}\sum_{i=1}^{n}(\log(x_i)-z_i'\beta+\log\left(c-(\tilde{A}_{P(i)}\gamma_{P(i)}e^{-\lambda\rho V_i}e^{\tilde{\zeta}_{P(i)}})^2\right)) = 0 \tag{D.4}$$

$$\frac{1}{n}\sum_{i=1}^{n}z_i(\log(x_i)-z_i'\beta+\log\left(c-(\tilde{A}_{P(i)}\gamma_{P(i)}e^{-\lambda\rho V_i}e^{\tilde{\zeta}_{P(i)}})^2\right)) = 0 \tag{D.5}$$

$$\frac{1}{n}\sum_{i=1}^{n}(\log(P(y_i=1))-\lambda\rho V_{i,0}+\tilde{\zeta}_1-2\log(s_i))I_{i\in P_1} = 0 \tag{D.6}$$

$$\frac{1}{n}\sum_{i=1}^{n}(\log(P(y_i=1))-\lambda\rho V_{i,0}+\tilde{\zeta}_2-2\log(s_i))I_{i\in P_2} = 0, \tag{D.7}$$

Careful inspection of equations (D.1) - (D.7) shows how to come up with appropriate starting points.

First, it is immediate that the only parameters in (D.1)-(D.3) are $\log(\tilde{A}_1\gamma_1), \log(\tilde{A}_2\gamma_2)$ and $\lambda\rho$. Furthermore, those three equations are exactly the moment conditions implied by OLS estimation of $\log\left(\frac{s_i}{x_i}\right)$ on $I_{i\in P_1}, I_{i\in P_2}, V_{i,0}$. The OLS coefficients of this regression set equations (D.1) - (D.3) to exactly 0. Hence, we use these OLS estimates as starting values for $\log(A_1\gamma_1), \log(A_2\gamma_2), \lambda\rho$. We use an analogous argument on equations (D.6) and (D.7). In this second "regression", we use the OLS estimates of the outcome $\log(P(y_i = 1)) - (\lambda\rho)^{start}V_{i,0} - 2\log(s_i)$ on $I_{i\in P_1}, I_{i\in P_2}$, where $(\lambda\rho)^{start}$ are the starting values for $\lambda\rho$. This second regression results in estimates for $(\tilde{\zeta}_1, \tilde{\zeta}_2) = (\log(1-e^{-\lambda\zeta_1}), \log(1-e^{-\lambda\zeta_2}))$ which set (D.6)-(D.7) to 0, which we use as starting values. Finally, equations (D.4)-(D.5) also come from a separate OLS regression where the outcome is $\log(x_i)$, the independent variables are a constant and z_i. The OLS coefficients on z_i is the starting value for β and set equations (D.5) to 0. The normalization assumption - not including a constant in z_i - guarantees that only one c satisifies (D.4). Hence, the outlined procedure delivers us starting values (and consistent estimators) for all parameters except c.

Our GMM estimator are the set of parameters that minimize the GMM objective function given moments (D.1-(D.7), given starting values for all parameters except c as described above, and across different starting values for c.

D.1.2. *Computation of Standard Errors.* To compute the standard errors for our GMM estimates, we use a consistent estimator based on its asymptotic value. Given the model is exactly identified, we can use the identity matrix as a weighting matrix.

As it is well known, the asymptotic variance matrix (of $\sqrt{n}$ times) our parameters of interest is then given by $(\Gamma'\Omega^{-1}\Gamma)^{-1}$, where $\Gamma = \mathbb{E}\frac{\partial g(\tilde{s}_i, \tilde{x}_i, \theta)}{\partial \theta'}$ and $\Omega = \mathbb{E}(g(\tilde{s}_i, \tilde{x}_i, y_i, z_i, \theta)g(\tilde{s}_i, \tilde{x}_i, y_i, z_i, \theta)')$.

We compute Γ analytically, by taking derivatives of each moment equation in relation to each parameter. We then replace the expectation by its empirical counterpart (the mean across all politicians).

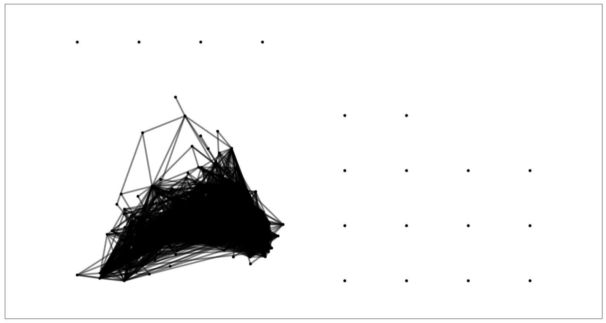
D.1.3. *Finite Sample Corrections for the Standard Errors.* In finite samples, Ω can be close to singular. This appears to be the case in some of the specifications in our paper. To improve the finite sample performance, we implement the correction used in Cameron et al. (2011). This involves increasing the standard errors in Ω by adding a small perturbation to its eigenvalues. This perturbation is sufficient to remove singularity.

Such a procedure uses the spectral decomposition of $\Omega = D\Lambda D'$, where Λ is a diagonal matrix of eigenvalues. We then add a small $\delta_\Omega > 0$ to the diagonal of $\hat{\Lambda}$, therefore increasing the eigenvalues of $\hat{\Omega}$. Since this procedure increases standard errors, the new standard errors are still valid for our parameters.

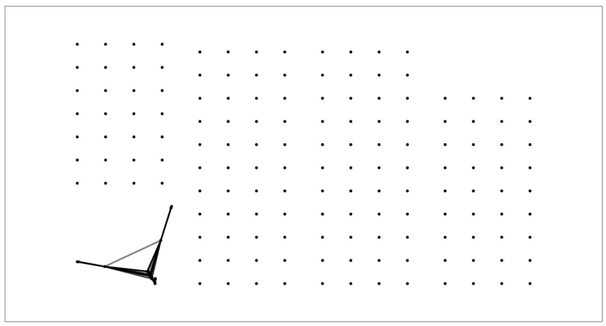
In practice, we pick $\delta_\Omega = 0.00001$, and use it on the eigenvalues that are smaller than 10^{-7}. This is typically 1 or 2 of the eigenvalues of our estimated Ω.

Appendix E. Additional Tables and Figures

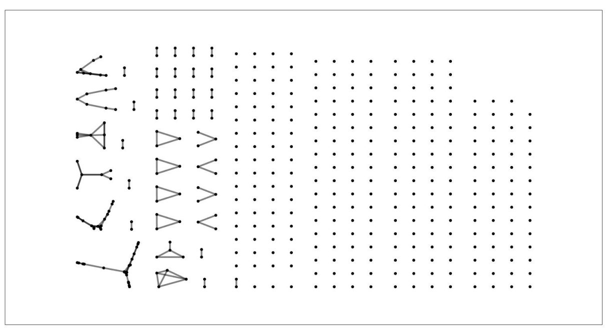
Figure E.1. Examples of Alternative Congress 110 Networks



(a) More than 3 Directed Cosponsorships



(b) Committee Network



(c) Alumni Network

We show illustrations of alternative networks used in the literature that we use to compare our model against. A link in the committee network exists if two legislators sit in one of the 7 main committees together (see the Data section). A link in the alumni network exists if two legislators attended the same university within 8 years of one another. While in the empirical specification of equation (6.2) the cosponsorship network is taken as the amount of directed cosponsorships, we illustrate it here by plotting the upper triangular matrix of directed cosponsorships, with a link formed if a legislator cosponsors more than 3 bills by another.

Table E.1. Main Results, Specification 2

	Congress					
	105	106	107	108	109	110
c	0.263	0.277	0.292	0.291	0.293	0.271
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\tilde{\zeta}_{Dem}$	5.890	5.507	5.720	5.554	5.546	3.125
	(0.160)	(0.179)	(0.174)	(0.176)	(0.174)	(0.119)
$\tilde{\zeta}_{Rep}$	3.404	2.695	3.117	2.880	2.935	4.311
	(0.178)	(0.146)	(0.170)	(0.172)	(0.158)	(0.181)
$\lambda\rho$	0.101	0.138	0.015	0.010	0.038	0.022
	(0.072)	(0.056)	(0.068)	(0.073)	(0.080)	(0.070)
Rep	0.078	0.143	0.212	0.098	0.108	0.047
	(0.053)	(0.049)	(0.059)	(0.062)	(0.057)	(0.084)
Ideology	-0.387	-0.293	-0.368	-0.400	-0.382	-0.322
	(0.053)	(0.067)	(0.065)	(0.105)	(0.112)	(0.048)
Tenure	0.007	0.009	0.007	0.010	0.009	0.004
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Appropriations	-0.007	0.035	-0.041	-0.076	-0.046	-0.062
	(0.042)	(0.033)	(0.038)	(0.040)	(0.044)	(0.032)
Energy and Commerce	-0.024	-0.011	-0.065	-0.143	0.016	-0.003
	(0.039)	(0.039)	(0.046)	(0.135)	(0.046)	(0.026)
Oversight	0.014	0.069	0.029	0.022	0.018	0.023
	(0.034)	(0.038)	(0.058)	(0.053)	(0.051)	(0.033)
Rules	0.122	0.128	0.098	0.160	0.210	0.149
	(0.023)	(0.025)	(0.035)	(0.041)	(0.057)	(0.020)
Leadership	0.124	0.110	0.140	0.123	0.280	0.109
	(0.040)	(0.034)	(0.066)	(0.108)	(0.033)	(0.117)
Transportation	-0.052	-0.020	0.011	-0.123	0.001	-0.002
	(0.037)	(0.047)	(0.030)	(0.106)	(0.044)	(0.025)
WaysAndMeans	-0.087	-0.018	-0.015	-0.033	0.010	-0.023
	(0.043)	(0.036)	(0.038)	(0.049)	(0.056)	(0.037)
Ideology × *Rep*	0.541	0.419	0.409	0.622	0.610	0.591
	(0.094)	(0.095)	(0.104)	(0.131)	(0.136)	(0.120)
Tenure × *Rep*	0.010	0.008	0.006	0.008	0.004	0.005
	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)
Appropriations × *Rep*	0.036	-0.044	-0.067	-0.002	-0.016	-0.127
	(0.052)	(0.042)	(0.052)	(0.056)	(0.061)	(0.075)
Energy and Commerce × *Rep*	-0.023	-0.003	0.010	0.052	-0.064	-0.084
	(0.053)	(0.052)	(0.060)	(0.139)	(0.061)	(0.064)
Oversight × *Rep*	0.062	0.046	0.003	-0.007	0.048	-0.005
	(0.047)	(0.045)	(0.066)	(0.065)	(0.063)	(0.055)
Rules × *Rep*	0.021	-0.003	0.008	-0.050	-0.005	0.058
	(0.045)	(0.036)	(0.044)	(0.053)	(0.069)	(0.037)
Leadership × *Rep*	0.000	-0.111	-0.059	0.006	-0.207	0.059
	(0.045)	(0.057)	(0.094)	(0.128)	(0.078)	(0.146)
Transportation × *Rep*	0.015	-0.012	-0.051	0.114	-0.015	-0.062
	(0.049)	(0.056)	(0.043)	(0.113)	(0.057)	(0.048)
WaysAndMeans × *Rep*	0.035	-0.011	0.016	0.008	-0.077	-0.157
	(0.056)	(0.055)	(0.057)	(0.062)	(0.072)	(0.075)
γ_{Dem}	[0.0001,0.0001]	[0.0002,0.0002]	[0.0002,0.0002]	[0.0002,0.0002]	[0.0002,0.0002]	[0.0020,0.0020]
γ_{Rep}	[0.0011,0.0011]	[0.0023,0.0023]	[0.0015,0.0015]	[0.0018,0.0019]	[0.0018,0.0018]	[0.0005,0.0005]
N	424	427	426	431	429	426

Notes: Standard errors in parentheses. The table presents the results from the GMM estimation under the second specification. That is, we replace the Grosewart measure by dummy variables for the most important committees. The variable Leadership represents a dummy of whether the politician was the Speaker, the Majority or Minority Leader, or the Majority or Minority Whip. *Rep* is a dummy variable for belonging to the Republican Party. The estimates of γ_{Dem} and γ_{Rep} are their estimated sets. Standard errors are estimated as discussed in Appendix D. All other notes follow those in Table 3.

Appendix F. Additional Reduced Form Evidence on The Effect of Cosponsors on Bill Passage

We begin by looking at specifications which show the correlation between cosponsors of a bill and whether the bill is approved or not. In our model, cosponsorships can only help bill approval through extending the (endogenously formed) network

F.1. **Data.** We use data from the 93rd (1973-1975) to the 110th Congress (2007-2009), originally from the Library of Congress, and used in Fowler (2006). The data includes all bills (both House and Senate) in these periods, with data for the politicians in each Congress (such as tenure, party, ideology measure), the cosponsoring decisions for each bill in each Congress and Senate (i.e. who sponsored and cosponsored each one) and the outcomes for each (passed house, passed Senate, was vetoed or not, and so forth).

With this data, it is possible to construct network variables such as: the number of cosponsors for each bill, average number of cosponsors for a politician's own bills, a network graph using cosponsorship decisions as links. The focus is on House bills. Table F.1 presents the summary statistics.

Table F.1. Summary Statistics for Appendix F

Variable	Obs	Mean	Std. Dev.	Min	Max
Pass	139021	0.077	0.267	0	1
Party	138986	60.32% Democrat			
		39.37 % Republican			
Ideology	137426	-0.069	0.388	-.757	1.685
Tenure	138986	5.974	4.096	1	27
Number of cosponsors	139021	10.311	27.084	0	406
Avg. cosponsors of cosp.	139021	6.239	8.55	0	175

F.2. **Empirical Specifications.** A first approach to this problem is to test whether networks do impact bill approval in Congress. To do so, we can check whether the number of cosponsors of a bill and the extended network of those cosponsors are positively correlated with passing rates in Congress (as in our structural model). To do so, consider the following regression:

$$pass_{i,k} = \beta_1 cosponsors_{i,k} + \beta_2 average\ cosponsors\ of\ cosponsors_{i,k} + X_i'\gamma + \varepsilon_{i,k} \tag{F.1}$$

where *cosponsors* represents the number of cosponsors of bill k (proposed by sponsor i); and *average cosponsors of cosponsors* represents the average number of cosponsors that cosponsors of this bill have (in their own bills). The latter captures the influence, or additional order network effects of those agents. X_i represent a series of politician level controls, such as the sponsor's ideology, tenure, party.

Equation (F.1) implies that having additional cosponsors (captured by β_1) and those cosponsors being more influential/with larger networks (β_2) are associated with the approval of legislation.

One may expect the OLS estimates of (F.1) to be inconsistent. First, it is possible that certain sponsors/politicians are more politically able and/or have better bills, and so would attract more cosponsors and better networks. In our model, higher types/returns α_i socialize more and have larger and more influential networks, and hence would be observed to cosponsor more on average.

To control for that, consider the fixed effects regression:

$$pass_{i,k} = \alpha_i + \beta_1 cosponsors_{i,k} + \beta_2 average\ cosponsors\ of\ cosponsors_{i,k} + \varepsilon_{i,k} \tag{F.2}$$

where α_i is a fixed effect for the politician who sponsors the bill. This effect captures the above problem, and would use the following variation: different bills by the same sponsor can have different number of cosponsors/extended network. The differences in their outcomes in Congress would then be attributed to the different (observed proxies for) networks.

A threat to identification in (F.2) is that we are not controlling for bill quality. The same sponsor can have some bills which are better than others, which by themselves attract more cosponsors. To deal with this issue, one can increase the set of controls, for instance focus on the specific characteristics of the Senate sponsor of the House bill.

This is done using the following specification:

$$pass_{i,j,k} = \alpha_i + \gamma_j + \beta_1 cosponsors_{i,j,k} + \beta_2 average\ cosponsors\ of\ cosponsors_{i,j,k} + \varepsilon_{i,j,k} \tag{F.3}$$

where α_i, γ_j represents a fixed effect for the House sponsor (i) and Senate sponsor (j) pair. The bills studied here are those present in both chambers.

Our preferred specification further controls for bill type. Although the above intuitively should do so, there is still a threat that part of the bill quality is not being captured by having the same sponsors in both chambers.

For that reason, consider the within bill variation model:

$$pass_{i,j,k,h} = \delta_k + \beta_1 cosponsors_{i,j,k,h} + \beta_2 average\ cosponsors\ of\ cosponsors_{i,j,k,h} + \varepsilon_{i,j,k,h} \tag{F.4}$$

$$pass_{i,j,k,s} = \delta_k + \beta_1 cosponsors_{i,j,k,s} + \beta_2 average\ cosponsors\ of\ cosponsors_{i,j,k,s} + \varepsilon_{i,j,k,s} \tag{F.5}$$

In this version, we are using variation in outcomes for the identical bills across chambers (h for House, s for Senate). We posit that the same bill, if it faces different results in separate chambers, must have that due to differential (networks) supporting it. It cannot be coming from bill quality, as it is the same bill in both scenarios. It cannot be coming from different politician abilities, as these are spanned by δ_k. The difference in outcomes is due to network effects.

Identification in (F.4)-(F.5) is due to the availability of bills that switch status across chambers.

We also use the definitions of identical bills in the Senate, as defined by the Library of Congress. This is done by checking for identical bills in the Senate (under related bills) for all house bills in Congresses 93-110. Table F.2 shows that there are bills that switch status across chambers, which is key to our identification. These constitute around 20% of the sample.

Table F.2. Bills with "Switching" Outcomes

	(1) Outcome $\Delta_{h-s} Billpass$	(2) Frequency	(3) Percent
Panel A: All identical bills			
	Pass Senate, Not Pass in House	1,073	8.30
	Same Outcome in Both	10,473	81.02
	Pass House, Not Pass in Senate	1,380	10.68
N: 12926			
Panel B: $Cosponsors > 0$ in both			
	Pass Senate, Not Pass in House	524	8.13
	Same Outcome in Both	5120	79.43
	Pass House, Not Pass in Senate	802	12.44
N: 6446			

Panel A: All bills with paired observations. Panel B: Only those with number of cosponsors bigger than zero in both the House and Senate.

F.3. **Results.** Table F.3 presents the results across our various specifications (F.1), (F.2), (F.3) and (F.4)-(F.5).

As can be seen, the estimates of β_1 and β_2 are positive and very significant across specifications (Linear, Linear with controls, House Sponsor Fixed Effects, House and Senate Sponsor Fixed Effects and within bill variation). The number of cosponsors is positively associated with the approval of bills, as is their influence along the congressional network.

TABLE F.3. (Appendix) Main Results

	(1) Linear	(2) Linear w/ Controls	(3) House Sp. FE	(4) House & Sen. Sp. FE	(5) Within bills
Cosponsors	0.000589***	0.000554***	0.000594***	0.000584***	0.00107***
	(0.0000540)	(0.0000547)	(0.0000527)	(0.000107)	(0.0000713)
Average	0.00275***	0.00140***	0.00227***	0.00162***	0.00104***
Cosp. of cosp.	(0.000214)	(0.000209)	(0.000208)	(0.000543)	(0.000369)
Constant	0.0536***	-0.0742***	0.0566***	0.102***	0.0854***
	(0.00250)	(0.00772)	(0.00135)	(0.0252)	(0.00447)
N	137703	137426	137703	12852	12926
R^2	0.015	0.035	0.010	0.044	0.014

Standard errors in parentheses, clustered at the House Sponsor level (first 4 columns) and Senate Sponsor (Column (5), due to lack of data to cluster at the House sponsor). Individual controls in Column (2) include Tenure, Party, Ideology and Congress.$^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$. The first column is the OLS regression, the second puts controls (described above), the third is fixed effects at the House Sponsor level, the fourth has fixed effects of both House and Senate sponsor. Column (5) is the specification with within bill variation. N for Column (5) is the number of bills we have pairs of observations. It is larger than (4) because it does not use information on the id code of the sponsor in the House.

The estimate of β_1 is between 0.0003 and 0.0005. This represents that an additional cosponsor correlates with a (directly) increased probability of approval by 0.05%. This is a small, but non negligible amount, as bills usually have many cosponsors. The coefficient for β_2 amplifies this effect, and is estimated to be around 3 times as large as β_1 (in Columns (1)-(4)). This implies that adding a cosponsor who has on average 10 cosponsors on his own bill, is associated with an average increase of $0.000541 + 10 \times 0.00162 = 0.0167$, or a 1.67 point increase in the percentage probability of approval.

Table F.4 allows for heterogeneity in the effects for the House and the Senate, for the specification of (F.4)-(F.5). The results confirm the positive and significant effects in the House, and shows that the influence term β_2 is really important in the House, although not so much from the Senate, which presents noisy estimates.

Our results indicate that it is seemingly advantageous to have additional cosponsors. In the context of the structural model, this means there are gains in having larger networks and more connections. We should hence, observe denser networks in Congress. This seems to be the case in our structural model. It also seems to hold in evidence in Fowler (2006) and Cho and Fowler (2010). This suggests that models with sparse equilibrium interconnections would not provide a good fit for Congressional activity.

Table F.4. Effect Heterogeneity

	(1) pass	(2) pass	(3) pass	(4) pass
House Outcome (Indicator)	0.0386***	0.00273	0.0529***	-0.0323*
	(0.00512)	(0.00804)	(0.00651)	(0.0182)
Cosponsors	0.00111***	0.00102***	0.000971***	0.000813***
	(0.000112)	(0.000136)	(0.000124)	(0.000152)
Average cosp. of cosp.	0.00204***	0.000850*	0.00102	-0.00127
	(0.000374)	(0.000433)	(0.000696)	(0.000874)
Interaction: House × cosponsors		0.000126		0.000226
		(0.000165)		(0.000207)
Interaction: House × avg. cosp. of cosp.		0.00333***		0.00478***
		(0.000599)		(0.00103)
Constant	0.0544***	0.0717***	0.0543***	0.103***
	(0.00543)	(0.00620)	(0.0133)	(0.0168)
N	12926	12926	6446	6446
R^2	0.021	0.025	0.021	0.026

Standard errors in parentheses, clustered at the senate sponsor level. Tests reject the hypothesis that the coefficients of the interactions are the same as those without. Columns (1) and (2) focus on all bills with paired observations. Columns (3) and (4) only on bills with positive number of cosponsors in both the House and the Senate. N is the number of bills (each bill has 2 observations). $^{*}p < .1$, $^{**}p < .05$, $^{***}p < .01$