Cosponsorship in the U.S. Senate: A Multilevel Two-Mode Approach to Detecting Subtle Social Predictors of Legislative Support

Abstract

Relational data and relational approaches to data analysis are quickly becoming more common tools for political scientists examining behaviors like legislative cosponsorship, international conflict, and interest group donations. Many of these relational behaviors of interest are best conceptualized as bi-partite networks, where actors are connected to one another through some other means of affiliation. Legislators are connected to one another through bills. Countries may be connected to one another through IGO membership. Groups may be connected to one another through common donations to candidates. This second network mode presents a statistical challenge to the traditional tools for network analysis like exponential random graph models. In this research we offer an alternative approach based on multilevel modeling that can simultaneously capture actor, dyad, and alternative mode level effects in bi-partite networks. Our method is flexible and easily extended to a wide variety of multi-mode network problems. We illustrate the method with an extensive examination of cosponsorship in the 108th U.S. Senate.
1 Introduction

Legislative scholarship has begun to take seriously the notion of statistical (and theoretical) interdependence in legislatures with the adoption of social networks as a useful lense through which to examine legislative behaviors. Social network theory and analysis provide powerful insights regarding the development of complex systems, the paths of influence in such systems, and the flow of information across relationships. Social network scholars wishing to study legislatures as these complex systems of interactions have focused the majority of their empirical analysis on the cosponsorship choices by legislators. The inherently relational nature of a cosponsorship decision makes it an obvious candidate for social network analysis and this particular legislative choice is free of many of the agenda setting mechanisms that influence roll call votes (another commonly used measure of legislative choices). Unfortunately, the majority of work studying cosponsorship has either obscured relational determinants of cosponsorship by studying cosponsorship choices as though they were independently distributed, or obscured bill-specific determinants of cosponsorship by generating a network of connections between legislators directly. Cosponsorship choices connect legislators through bills, providing complementary sources of variance. In this paper, we present a statistical model capable of accounting for both sources of variance simultaneously, thus overcoming the inherent specification problems that have plagued previous analyses.

Legislators frequently rely on one another as sources of information (Kingdon, 1973; Matthews and Stimson, 1975). They use each other to convey signals to one another (Kessler and Krehbiel, 1996). They use bills to signal to their constituents (Wilson and Young, 1997). This implies that aggregate legislative behaviors are both conditional on the nature of legislation about which a signal is being sent, and conditional on the behavior of other legislators.\(^1\) This relational, inter-actor dependence is best conceptualized as a social network

\(^1\)The other cosponsors on a bill help convey a certain type of signal to the chamber at large. There is also recent evidence that the choice to cosponsor legislation together is a function of how well-connected each
of legislative interactions. Unfortunately, wrapped up in this network of legislative signaling and information exchange is a second level/source of variance. Cosponsorship measures the connected or networked behavior of legislators through legislation. Legislators may make a choice to cosponsor a bill because of who else is on the bill, because they genuinely wish the bill to pass, or because they wish to signal support for legislation that is unlikely to pass (Koger, 2003). In most studies of legislative networks, scholars collapse this second level variance, constructing a single mode, or unipartite network where legislators are directly connected to each other. This collapse brings with it the assumption that legislative choices about co-support are only influenced by the behavior of other legislators. The notion that legislative behavior is only dependent on the behavior of other legislators, or conversely only dependent on the nature of the current piece of legislation, is in either case an untenable oversimplification. Here, we present a statistical model of relational interactions capable of accounting for variance in legislative choices at the individual, dyadic, and bill level. Using multilevel models with mixed effects, we demonstrate that there are bill-level, dyad-level, and individual-level motivations for cosponsorship choices.

The traditional use of multilevel models has been to account for unobserved heterogeneity in observations due to clustering, or a lack of exchangeability in observations. While on its surface, the handling of relational data may not seem like a problem of clustering, we demonstrate that choices about cosponsorship relationships can be thought of as choices by individuals, nested within dyads, nested within bills. The multilevel framework then allows us to decompose sources of variance at each level, thus capturing legislator-to-legislator influences while allowing for bill-level heterogeneity. While we implement this framework for cosponsorship choices, the flexibility of the multilevel model structure makes it an appropriate choice for modeling affiliation networks of many kinds.

---

(legislator in the dyad is (Kirkland, 2011).)
2 Cosponsorship and Social Network Analysis

By the regulations of Congress, each proposed piece of legislation may have only a single official sponsor in each chamber. In reality, several members of Congress (MCs) may be involved in crafting and pushing a bill from the earliest stages, but only one can be designated as sponsor. Cosponsorship has been permitted in the Senate since the 1930s and in the House since the 1960s, identifying those others involved or simply wishing to be on the record as supporting a measure. Nonetheless, for many years, cosponsorship was all but ignored by Congressional scholars, who tended to regard it as a legislative behavior without a clear procedural purpose until Campbell (1982) tried to explain cosponsorship in terms of individual ideology and electoral considerations using an ordinary least squares (OLS) regression.

Various studies have followed this line of investigation, seeking the reasons for cosponsorship and debating the relative merits of electoral motivations versus institutional signaling mechanisms (Kessler and Krehbiel, 1996; Wilson and Young, 1997; Koger, 2003). Explaining cosponsorship typically has meant taking as the dependent variable the number of bills cosponsored rather than the degree to which certain dyads are more prone to cosponsorship than others. A problem in this line of inquiry arises, however, if we believe that two legislators’ choices about what to cosponsor are conditional on one another. If, for example, Al Franken takes cues about what to cosponsor in the U.S. Senate from Thad Cochran, then their choices regarding cosponsorship are interdependent. Even if direct cue-taking and advice-seeking were insufficient motivators for interdependence, the legislative norm of reciprocity will also generate a lack of independent decision-making amongst legislators. Bernhard and Sulkin (2010) indicate that violations of the norm of reciprocity are met with group level consequences. Actors who fail to honor their commitments to support one another through cosponsorship experience decreases in the size of their coalitions and more limited legislative
success. This interdependence on individual pieces of legislation will also characterize their total counts of cosponsorship, thus making traditional regression approaches (which assume observations to be independent of one another) an inappropriate way to analyze cosponsorship choices. Indeed, a recognition of interdependence in cosponsorship choices and an interest in the dyadic patterns of cosponsorship implies that the more appropriate analytic approach is to treat cosponsorship as a social network.

Recognition that cosponsorship patterns in Congress may fruitfully be analyzed as a social network is very recent. Fowler (2006a,b) and Kirkland (2011) have sought to determine the influences of cosponsorship patterns on legislative outcomes, both concluding that actors’ positions and choices in the cosponsorship network play critical roles in their ability to pass legislation. Additionally, Tam Cho and Fowler (2010) have demonstrated that the aggregate cosponsorship network is predictive of the number of important pieces of legislation Congress will pass. Other efforts have focused less on the consequences of the cosponsorship network and more on the motivation for and description of the cosponsorship network. Faust and Skvoretz (2002) uses comparative techniques to compare the network of connections between legislators to other kinds of social networks.\footnote{Faust and Skvoretz’s comparative techniques indicate that the network of cosponsorships between Senators most resembles the network of grooming amongst cows.} More sophisticated measurement techniques employed by Zhang et al. (2008) and Waugh et al. (2009) demonstrate that cosponsorship patterns reveal the same polarization that characterizes floor voting, but that that polarization seems to precede the polarization of roll calls. Modeling approaches using the exponential random graph model by Bratton and Rouse (2011) reveal that cosponsorship patterns amongst state legislators are highly clustered activities, a result that reinforces the interdependence of the cosponsorship action. Finally, Kirkland and Gross (2012) show that the cosponsorship network is highly responsive to changes in Congressional approval.

While the recognition of interdependence in cosponsorship decisions (and legislative
choices more generally) is critical, all of these studies treat cosponsorship decisions as though they link legislators directly to one another. In fact, cosponsorships represent connections between legislators through specific pieces of legislation. To our knowledge, no previous model of cosponsorship has managed to simultaneously capture the dyadic (or “relational”) patterns in cosponsorship as well as bill-to-bill heterogeneity in cosponsorship choices. Thus, we intend to extend the work of these scholars by building a statistical framework which can accommodate dyadic, relational choices that vary depending on the nature of the bill. Thus, our model represents an option for the statistical analysis of affiliation (or “bipartite”) networks, where actors are connected to one another through a common category or cluster.

3 Exploratory Analysis

To demonstrate the utility of our approach to studying cosponsorship, we will use data from all bills and amendments proposed by members of the U.S. Senate during the 2003–2004 (108th) session and receiving at least one endorsement via cosponsorship, but no more than 98. (For convenience, we refer simply to these as bills, but unless otherwise noted, we mean the term to include bills and amendments, but not non-binding resolutions.) For each of these 2166 bills, the sponsor and all co-sponsors have been recorded, yielding 99 dyadic observations per bill, or 214,434 observations in total. We also collected information about individual senators and their relations. The main data on proposed legislation during the 108th Congress is maintained in the Thomas database of the Library of Congress, and were compiled and cleaned by Fowler (2006a). These include sponsorship and cosponsorship information, committees of jurisdiction, and legislative outcomes. State socioeconomic figures, as well as personal information on profession, veteran status, and the religious affiliation of senators were obtained from Barone et al. (2003).
Cosponsorship is at least in part a relational act, an expression of support by a legislator for the proposal of a colleague, and patterns in cosponsorship may be thought of as characterizing a social network linking MCs. The most common approach to a two-mode network (here, legislators and bills), in which the relations of interest are on one mode (here, the legislators), is to “collapse” the network and restrict attention to the network among these individuals. Why throw away all the information contained in the other mode (here, bills)? As complicated as the dependence structures arising in a typical one-mode network tend to be, the additional complications in bipartite graphs make modeling and inference trickier still. In particular, visualization of two-mode networks tends to be difficult and can hide the patterns of greatest interest. Thus, it is always a good idea in such cases to consider a collapsed version of the underlying network, at least as an exploratory step. With this in mind, we begin by representing the patterns of cosponsorship by aggregating all acts of cosponsorship for each of the \( P_{ij} = 9900 \) directed dyads and collecting them in an adjacency matrix, from which we may generate network graphs.

Creating such an adjacency matrix \( P \) requires some care, as the act of aggregation introduces what amounts to a measurement problem. We may think of individual acts of cosponsorship as being generated by a non-observable propensity of support manifesting itself only through such acts. Each element \( P_{ij} \) should indicate the propensity of senator \( i \) to cosponsor legislation proposed by senator \( j \). This will be a \( 100 \times 100 \) matrix with 9900 non-trivial entries, and structural zeroes on the diagonal. In related work on coauthorship and scientific collaboration networks, typically \( P_{ij} = 1 \) if \( i \) and \( j \) have ever collaborated and zero otherwise (Newman, 2001a). This works for sparse networks without strict boundaries, but not for the U.S. Congress, especially the Senate, where the cosponsorship network is much denser than in other collaboration networks (67% of the 9900 direct dyads in our data provide).

\(^3\)A bipartite network is one in which nodes/actors are connected to one another through a node of another type. For example, legislators are connected to one another through bills. It is this additional level of heterogeneity we wish to explore.
set are non-empty). On the other hand, a simple count of how many times \( i \) cosponsors \( j \) may also be inappropriate, since it exaggerates the support of senators who sponsor a lot of legislation. Instead, we take \( n_j \), the number of bills sponsored by \( j \), as the number of opportunities for each other senator to show support for \( j \); if these events were independent, we might think of \( n_{ij} \), the number of times \( i \) cosponsors a bill by \( j \), as being distributed \( \text{Bin}(n_{ij}, P_{ij}) \). However, the more cosponsors appearing on a bill, the less informative the act of cosponsorship is. To take this into account, Fowler (2006a) adopts a measure employed by Newman (2001a,b) in his study of scientific coauthorship. Each act of cosponsorship is scored fractionally, in inverse proportion to the number of cosponsors on the bill. The resulting measure can still exaggerate ties to senators sponsoring a great deal of legislation; for exploratory purposes, we adapt it slightly, dividing Newman and Fowler’s weighted raw score by the total possible score any colleague of \( j \) could accrue if appearing as cosponsor on all of \( j \)’s bills.

Specifically, taking the lowest level observation to be \( Y_{ij(k)} = 1 \) if \( i \) cosponsors \( j \)’s \( k \)th bill, and 0 otherwise, we let \( c_{j(k)} \) = the number of cosponsors on senator \( j \)’s \( k \)th bill. Then the relative weighted propensity to cosponsor of \( i \) for \( j \) shall be defined as:

\[
WPC_{ij} = \frac{\sum_{k=1}^{n_j} \frac{Y_{ij(k)}}{c_{j(k)}}}{\sum_{k=1}^{n_j} \frac{1}{c_{j(k)}}}
\]  

The denominator represents the total raw score possible if \( i \) were to appear as cosponsor on every single piece of legislation proposed by \( j \). While the normalization of \( WPC \) is perhaps not optimal, it is satisfactory for exploratory purposes.

Most visualizations of social networks work with dichotomous ties, i.e., treat all extant ties alike. To visualize the network, we dichotomize, including only ties above a threshold
strength $t$. To make this less arbitrary, we take “snapshots” as we raise $t$, noting which sorts of relationships seem to persist as weaker connections are removed. The choice of threshold determines what sort of picture we get, from extremely dense at $t = 0$ to completely disconnected when $t$ is large. In Figure 2, nodes are placed by “spring embedding,” so that nodes sharing many neighbors tend to be near one another (Borgatti et al., 2002).

Three likely drivers of cosponsorship patterns become apparent in distinct visualizations. These include one entirely unsurprising factor (party), one factor (representing the same state) that, while not unexpected, seems even stronger than may have been anticipated, and one factor (committee leadership pairs) that might easily have been overlooked (Moody et al., 2005).

In Figure 2, with threshold set to 0.02, the approximate mean of all $W_{PC_{ij}}$, cosponsorship by party dominates the visible pattern. Notice that Zell Miller, the conservative southern Democrat is situated far to one side, deep in Republican territory, and that the New England Republicans widely recognized as most liberal at the time, Chafee (RI), Collins (ME), and
Figure 2: Weighted Propensity to Cosponsor (Threshold = 0.02) (This and subsequent figures, through Figure 4, were made with Netdraw, from the software suite UCINET. Labels are of the form name/party/state. Arrowheads indicate the direction of support. Democrats appear as circles (in blue if color is available), Republicans as a (red) squares; the independent former Republican Jeffords is a (yellow) triangle.

Snowe (ME) are surrounded by Democrats. Other Republicans—Coleman (MN), Voinovich (OH), and Bennett (UT), Specter (PA) and Lugar (IN)—hover along the periphery of the cloud of Democrats. The density of the network is also clearly greater on the Democratic side, as the proliferation of edges darken the neighborhood around Senators Clinton (NY), Kennedy (MA), and Harkin (IA).

In order to further explore potential drivers of cosponsorship, we raise the threshold for the visual display to 0.25. Thus, the remaining connections in the graph are stronger than
those eliminated from the party-based graphic. To display regional patterns, we place nodes for the senators roughly according to the spatial layout of the states they represent on the map. In Figure 3, same-state pairs abound, with little remaining of the purely partisan clustering seen at lower thresholds. Other than the state links, there is not much by way of regional alignment at this level. In other words, when we move to a higher threshold for “counting” cosponsorship as symbolizing a connection, same-state connections persist while same-party connections are less prominent.

We do, however, see a number of seemingly random cross-party ties stretched across the map. It turns out that these transcontinental streaks are far from haphazard. An interesting committee based pattern under lies these extremely strong, cross-state, cross-party ties. In examining the pairs in question, we rearrange them so that the pattern becomes more obvious.4

4 Jeffords had only recently left the GOP, and Chafee, despite his reputation as a centrist, would be voted
Just as we arranged nodes geographically to make that pattern stand out, we can do something similar for committee leaders. In Figure 4, Republican Chairs have been situated directly above their Democrat counterparts (Ranking Minority Members, or RMM) on the same committee. The remaining sixty are mostly placed near their state partners along the edge of the figure. At the selected WPC threshold of 0.25, with only around a hundred remaining among 9900 possible directed ties, thirteen of twenty potential ties from RMM to Chair persist (only two of which are reciprocated, it is worth noting). Thus, these remaining cross-party, cross-state ties are largely representative of Democratic Ranking Members cosponsoring with the Republican Chairs of their committees.

out of office on a wave of anti-Bush sentiment in the 2006 elections.
Why Incorporate Bill Level Heterogeneity?

Much of the earlier literature explaining cosponsorship patterns adopts a non-network perspective and thus focuses on bill specific reasons for cosponsorship choices (Kessler and Krehbiel, 1996; Wilson and Young, 1997; Koger, 2003). That is, in this line of reasoning, legislators are motivated to cosponsor bills because of the nature of the bills themselves. The ideological placement of a legislator, a bill, and a legislator’s constituents is key to determining when a legislator will sign onto a piece of legislation. More recent literature adopting a network approach to examining cosponsorship has instead focused on relational reasons for cosponsorship (Bratton and Rouse, 2011; Cranmer and Desmarais, 2011). That is, legislators cosponsor bills because of the nature of their relationship with the bill’s sponsor. Two legislators being from the same party, serving on the same committees, and being of the same race strongly predicts when legislators will cosponsor each other’s proposals, an explanation our exploratory research strongly supports. However, it seems extremely likely that there are both dyadic and bill-specific motivators for cosponsorship patterns, so that any model addressing only one set of explanations is necessarily incomplete. Relational and bill-specific determinants of cosponsorship choices should be incorporated into a single model.\footnote{Harward and Moffett (2010) also use a multilevel model to estimate cosponsorship choices including bill and sponsor specific variables, but fail to incorporate dyadic intercepts. Thus, their model misses the interdependence in choices that characterizes social network choices.}

So far, the scholarship on cosponsorship that has managed to incorporate both levels of variance has done so through differential weighting schemes, as in Fowler (2006a) and our own exploratory analysis above. A model, on the other hand, would both provide additional flexibility and impose fewer assumptions on the data than a weighting scheme. Whereas a weighting scheme developed by researchers and then imposed on the data assumes that the scholars properly understand how the actors in a network weight their acts, a model
that incorporates bill-level variance would allow scholars to evaluate the varying import of such actions empirically rather than assume they were correctly understood \textit{a priori}. Such a model would also allow scholars to incorporate both bill-level and relational literatures into a single framework. It is to the construction of such a model that we now turn our attention.

4 A Multilevel Model of Individual Acts of Cosponsorship

4.1 The Importance of Addressing Interdependence

We next develop a model to handle both relational interdependence and bill-level heterogeneity. We begin by addressing the critical step of accounting for interdependence in relational data. Statisticians, quantitative psychologists, and sociologists have long been aware of a fundamental problem in the study of relational data, of network analysis in particular. Analogous to problems encountered in temporal (time series) and in spatial data, the basic linear modeling assumption of no error autocorrelation (or, equivalently, independence among observations) fails spectacularly, but in systematic ways that may potentially be modeled. And, as with these other examples where correct inferences are threatened by autocorrelation, the interdependence among network observations do not always represent nuisance but may indeed be a primary object of inference. A basic strategy common in all cases is to appropriately specify conditional probabilities for separable components in order to allow factoring of the likelihood function associated with such models. To oversimplify a bit, the patterns of dependence typically observed in time series are relatively predictable compared to those encountered in social networks. In many applications, it is reasonable to assert (and straightforward to test) an ARMA($d$) process, in which observations more than $d$ time steps apart are conditionally independent given the intervening observations. The
identification of conditional independence conditions is arguably more challenging when it comes to spatial analysis, but spatial autocorrelation models take advantage of similar logic, whereby distant observations are taken to be independent conditional on more proximate ones. Special problems arise when attempting to use this sort of approach on social (or other) network data. For one thing, no natural distances exist in social space corresponding to those in time or two/three-dimensional Euclidean space; suitable metrics generally lack the triangle inequality \( d(x, z) < d(x, y) + d(y, z) \)\(^6\) that constrains the types of conditional independence possible in time and space. Additionally, the measurement issues in temporal and spatial autocorrelation are far less problematic than those in social networks; while relative proximities in these other contexts are trivial to measure, such proximities are typically themselves objects of study in social networks. Not only are individuals (nodes in a network) interdependent, dyads (pairs of nodes), triads, and higher-order structures may all interact in complex ways.

Early breakthroughs that would lead to current state-of-the-art approaches utilized an approach based on log-linear modeling (Fienberg and Wasserman 1981) and eventually improved upon the resulting technique by applying an important result from spatial interdependence theory (Besag 1974) to accommodate the non-independence of dyads (Frank and Strauss 1986, Strauss and Ikeda 1990). Prominent modeling frameworks to have emerged in political science during the explosion of network analytic modeling over the past several years include the ERGM/p* approach (Wasserman and Pattison, 1996), latent space models (Hoff et al., 2002; Hoff, 2003a,b, 2004), and actor-oriented stochastic models (Snijders, 1996). None of these explicitly deals with bipartite graphical structure, but Hoff’s latent space modeling bears resemblance to what we do here. Our setup is even more directly inspired by the multilevel social relations model of Snijders and Kenny (1999), adapted for

\(^6\)That is, the sum of social distances between Senators A and B and between B and C is not necessarily larger than the social distance between Senators A and C.
dichotomous observations on a two-mode network, as well as Gill and Swartz (2001, 2004).

4.2 The Special Challenges of Bipartite Networks and Social Acts

A particular type of network structure arises in what has often been referred to as affiliation networks. Classic examples include individual socializers connected to one another via the gatherings they have attended (as in analysis of the “Southern Women” data of Davis et al. (1941)), businesspeople via shared Boards of Directors (Robins and Alexander, 2004), or legislators via shared committees (Porter et al., 2005). In this arrangement, all links among actors are indirect in the sense that they are operationalized via links to a different type of entity (celebrations, boards, committees, etc.) Defined precisely, such networks correspond to bipartite graphs, consisting of two distinct sets of nodes (vertices), \( L, B \), with each entity appearing in one or the other set, and a set of ties (edges) connecting pairs of these nodes, one from each set. In our case, \( L \) contains all legislators, while \( B \) contains all bills or amendments. Then an edge list might look like:

\[
\{a_7, b_1\}, \{a_{42}, b_1\}, \{a_{100}, b_1\}, \ldots, \\
\vdots \\
\ldots \{a_{87}, b_{2166}\}, \{a_{99}, b_{2166}\},
\]

with all \( a_i \in L \) and \( b_k \in B \), and each pair indicating a legislator and a piece of legislation he or she has cosponsored.

If this were treated as a standard (symmetric) affiliation network, we would infer relationships based on shared connections to bills, but we are not simply interested in connecting cosponsors through their shared bills; rather, in the current research, we wish to consider a directed bipartite graph, in which cosponsors are connected to sponsors via bills.

The state of the art in bipartite network analysis is still rather underdeveloped. Most
work in the area seems to have had a focus on global network properties and visualization. For example, Galois lattices (Freeman and White, 1993) and correspondence analysis (Roberts, 2000) have been suggested for simultaneously visualizing how people are related through organizations and organizations through people. In other research, blockmodels have been extended to the two-mode situation (Doreian et al., 2004) and new measures of centrality for two-mode networks have been proposed (Sinclair, 2004).

But our primary interest is neither the large-scale (global) properties nor the small-scale (individual) properties treated by these scholars. Our research questions instead concern dyads: What factors make certain members of Congress likely to support other members via cosponsorship?

The data we consider are, upon careful reflection, not measurements of persistent pairings, but of acts linking individuals at particular moments. In a way, the modeling situation shares common ground with network applications in neuroscience, where interest lies in predicting excitation between neurons. As with neurons firing, the observations consist of fleeting events in continuous time. While we face less of a computational challenge from the size of our data set (fewer senators than neurons and fewer bills than firings), this means we must learn more with less. Indeed, this is what makes the partial pooling of multilevel modeling so useful.

The Basic Model

Our exploration of the data in the previous section suggests possible elements to include in specifying a model; but no less important is what this exercise demonstrates in terms of the richness of the data set. In particular, the data visualization approach allows us to examine the relationship of interest, between Senators, though at the cost of oversimplifying the data-generating process. The different weighting schemes incorporated at the exploratory stage represent an attempt to mitigate the damage done by treating the network as unipartite. At
the modeling stage, we would like to directly incorporate bill-level information likely to have an impact on legislators’ decision-making. That is, the assumption that each opportunity to cosponsor a colleagues’ proposal is exchangeable with any other is incorrect. If we wish to learn, for example, whether the apparent tendency of a committee’s Ranking Member to exhibit a special inclination to support a bill sponsored by the Chair only applies to legislation passing through that committee, we must model observations at the bill-level rather than simply the dyad-level. Similarly, if wish to attribute more meaning to individual cosponsors when they are among just a few on a bill, we may incorporate this information as a covariate rather than through a weighting scheme.

Most analyses that have treated cosponsorship behavior patterns as constituting a social network (or providing clues about an underlying support network) have counted incidents of cosponsorship from each \( i \) to \( j \) and taken these counts, or some function of them, as indication of tie strength, as we did in our exploratory analysis (Burkett, 1997; Burkett and Skvoretz, Burkett and Skvoretz; Faust and Skvoretz, 2002; Fowler, 2006a,b). However, the implicit assumption that the probability \( p_{ij(k)} \) of one senator cosponsoring legislation by another is constant over bills (i.e. \( p_{ij(k)} = p_{ij}, \forall k \)) is too strong. To see why a multilevel model may adequately handle the data structure at hand, one should consider the nature of the dependence structure. The standard example for multilevel (hierarchical) modeling comes from educational testing (Raudenbush and Bryk, 2002), and finds a rough analogue in the cosponsorship data. Just as a student may take a number of distinct exams, with all test scores grouped by student, so are there a number of opportunities for the link from Senator \( i \) to \( j \) to manifest itself. Just as students are clustered hierarchically in schools nested in districts, each instance of cosponsorship (or lack thereof), \( Y_{ij(k)} \), belongs to a directed dyad \( (i,j) \), which in turn is nested together in the undirected dyad or pair \( \{(i,j),(j,i)\} \). Furthermore, as students may have an instructor and an exam grader drawn from the same teacher pool, each observation in our data set has cross-classified effects from cosponsor and
sponsor, drawn from the same pool of senators.

When we do not address the various types of clustering of observations inherent in the data structure, inferences will go wrong in two main ways: first, the degrees of freedom will be vastly overestimated, yielding estimated standard errors that are far too small\(^7\), and, second, estimates may be biased. The greatest advantage of multilevel modeling for us is that, by acknowledging the dependence structure, we are able to essentially take weighted regressions based on the amount of variance at each level of clustering. This “partial pooling” (Gelman and Hill, 2007) improves the accuracy of our estimates, as well as giving much more realistic standard errors than we would by simply ignoring the structure and pooling results.

If we consider the joint probability of all dyadic relations, taken as a socioarray \((n_j \text{ observations for each dyad } i_j)\), the interdependence just discussed translates into a lack of independence among the components of our data array. Recalling that \(Y_{ij(k)} = 1\) if Senator \(i\) cosponsors Senator \(j\)’s \(k\)th bill, and is 0 otherwise,

\[
Pr(Y) \neq \prod_{i,j,k} Pr(Y_{ij(k)}) ,
\]

This failure to factorize amounts to correlation among error terms, invalidating a crucial assumption of OLS and standard generalized linear models (GLM), including logistic regression. However, if we can model the systematic components of this dependence, the remaining disturbance terms will no longer be correlated (Hoff, 2003b). In forming such a model, we proceed on the assumption that the dependence among the observations is not utterly arbitrary and intractable but follows patterns common to networks and relational data generally.

Consider the basic building block of the model:

\(^7\text{This is not only due to the obvious lack of independence among observations on common dyads, but among those with common sender, receiver, and higher order interaction.}\)
\[ g(Y_{ij(k)}) = \mu + \epsilon_{ij(k)}, \]

where \(g(\cdot)\) is some suitable link function, and the grand mean \(\mu\) is the unconditional probability of cosponsorship, the so-called density of the network, the proportion of observations where cosponsorship has indeed taken place. In the spirit of ANOVA-type models, \(\mu\) serves as a baseline from which particular senators and pairs of senators deviate. This parameter will become the fixed intercept of our full model. As it stands, the error term \(\epsilon_{ij(k)}\) will have some hopelessly complicated and interdependent distribution, involving the sponsor, potential cosponsor and particular bill. The goal is to decompose the potential sources of variance to the point where any remaining error can be reasonably assumed to have some nice distribution.

The general model is as follows. If \(Y_{ij(k)}\) equals one if legislator \(i\) cosponsors legislator \(j\)’s \(k\)th bill or amendment, or zero if not, then:

\[
\begin{align*}
\text{logit} (Y_{ij(k)}) &= \mu + \alpha' S + \beta' R + \delta' X_{ij} \\
&= \mu + \left(\alpha' S_i + a_i\right)_{\text{sender}} + \left(\beta' R_j + b_j\right)_{\text{receiver}} \\
&\quad + \delta' X_{ij} + d_{ij}^{\text{directed}} + \rho_{(ij)}^{\text{undirected}} + c_j + \epsilon_{ij(k)} \ ,
\end{align*}
\]

where \(\alpha\) and \(\beta\) are vectors of shared,\(^8\) constant coefficients on attributes \(S_i\) and \(R_j\) of the potential cosponsor (sender) and sponsor (receiver), respectively, \(\delta\) are shared, constant coefficients on relational variables measured on dyads, while \(a_i, b_j, c_j, d_{ij}, \) and \(\rho_{(ij)}\) are varying intercepts assumed drawn from distributions specified below. Sender effects \(a_i\) capture expansiveness, accounting for variation particular to potential cosponsors (the tendency of

\(^8\)By shared, we mean shared across clusters, or common to all observations.
legislators to cosponsor a lot or a little), while receiver effects \( b_j \) capture popularity, and account for the tendency of some legislators to attract more explicit support than others when they sponsor legislation. Directed dyads are nested (in pairs) within the corresponding undirected dyads, so that the directed dyadic effects \( d_{ij} \) represent random departures from mutual (reciprocated) effects \( \rho_{ij} \), which themselves are attributed to the dyad and unexplained by included covariates. Finally, \( c_j \) is an additional random intercept associated with committee \( j \), accounts for bill cluster. This can be expected to allow for variation in cosponsorship activity based on committee dynamics, and the bill context. As always, we need to assume a distribution for these varying intercepts, treating them as if committees were drawn from a hypothetical universe of possible committees. We take \( c_j \sim N(0, \sigma_c^2) \), similar to our handling of actor effects \( a_i \) and \( b_j \). The final disturbance term, \( \epsilon_{ij(k)} \) represents any remaining error associated with \( i \)'s choice of whether to cosponsor \( j \)'s kth proposal. In the second rendering of the equation, labeled (3), variables are grouped by unit of observation, with superscripts labeling these units to assist interpretation.

The actor-specific random effects are assumed to be distributed as

\[
\begin{bmatrix}
a_i \\
b_j
\end{bmatrix}
\sim MVN \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix} \right),
\]

\[
\begin{bmatrix}
\epsilon_{ij(k)} \\
\epsilon_{ji(k)}
\end{bmatrix}
\sim MVN \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\epsilon}^2 & \rho \sigma_{\epsilon}^2 \\ \rho \sigma_{\epsilon}^2 & \sigma_{\epsilon}^2 \end{bmatrix} \right),
\]

with \( \rho \) a reciprocity (mutuality) factor (Hoff, 2003b) and we do not assume actor effects to be uncorrelated, since (for example) sender and receiver effects may well be correlated for the same actor — senators’ relative sponsoring activity may covary with their level of cosponsorship activity. Thus, calling the entire random portion of the model
\[ \eta_{ij(k)} = a_i + b_j + d_{ij} + \rho_{ij} + c_j + \epsilon_{ij(k)}, \] 

(4)

the notable corresponding first and second moments will be

\[
\begin{align*}
E(\eta_{ji(k)}) &= 0 , \\
E(\eta_{ij(k)}^2) &= Var(\eta_{ij(k)}) + \mu(\eta_{ij(k)}) \\
&= Var(a_i) + Var(b_j) + Var(\epsilon_{ij(k)}) = \sigma_a^2 + \sigma_b^2 + \sigma_{\epsilon}^2 , \\
E(\eta_{ij(k)}\eta_{ji(k)}) &= Cov(\eta_{ij}, \eta_{ji}) + E(\eta_{ij(k)})E(\eta_{ji(k)}) = 2\sigma_{ab} + r\sigma_{\epsilon} , \\
E(\eta_{ij(k)}\eta_{im(k)}) &= \sigma_a^2 , \\
E(\eta_{ij(k)}\eta_{mj(k)}) &= \sigma_b^2 , \\
E(\eta_{ij(k)}\eta_{mi(k)}) &= \sigma_{ab} = E(\eta_{ij(k)}\eta_{mj(k)}) .
\end{align*}
\]

Finally, \( E(\eta_{ij(k)}\eta_{mn(k)}) = 0 \), where all four indices are distinct.

Beyond these, all other expected products are assumed zero.

As is common in multi-level modeling, we postulate Gaussian distributions for the varying intercepts (Hoff, 2003b; Snijders and Kenny, 1999; Gill and Swartz, 2001). This is more credible if we explicitly include the most plausible covariates, so that the remaining randomness comes from the sum of many smaller causes, and is thus (by the Central Limit Theorem) apt to follow a Gaussian distribution.

Variables

We are particularly interested in the role of certain structural (relational) predictors. In our data, these are all functions of individual-level attributes. We will outline a set of dyadic variables below based on the social networks concepts of propinquity and homophily.
Note that the distinction between homophily and propinquity is not always clear cut (as for example with geographical variables Same Region and Same State, which may be thought of as either). Propinquity refers to the opportunities for two actors to interact, while homophily refers to traits two actors have in common (McPherson et al., 2001).

**Proximity/Opportunities for Interaction (Propinquity)**  
Increased opportunities for interaction between senators and their staffs will be associated with higher propensity to cosponsor. In particular, we examine whether (a) office location of senators, and (b) number of shared committee assignments tend to be predictive of cosponsorship. Senators serve on anywhere from two to six committees. Since the greatest opportunity for interaction among senators and staff exists during committee business, committee membership will be of interest. Moreover, the nature of the relationship between shared committee memberships and propensity to cosponsor is likely to depend upon whether the bill or amendment in question is under the jurisdiction of the committee two legislators share in common. That is, a pure propinquity effect, whereby simple increased contact between legislators makes them more likely to cosponsor each other’s proposals, would be supported if simply serving on the same committee increases cosponsorship *even on bills not passing through the shared committee*. The tendency to cosponsor bills passing through one’s own committee is likely to have more to do with interest in the legislation and the strategic importance of signals from committee members.

*Offices on Same Floor* = 1 if offices are located on the same floor of same Senate office building, 0 otherwise.

*Both on Committee of Jurisdiction* = 1 if $i$ and $j$ both serve on the committee of jurisdiction for bill $k_j$, 0 otherwise.

*Both on Other Committee* = 1 if both serve on at least one committee other than that
with jurisdiction over bill \( k_j \), 0 otherwise.

\textbf{Tenure Difference} = \( j \)'s number of years serving as Senator office minus \( i \)'s number of years serving.

\textbf{Common Traits (Homophily)} Common traits based on shared identity and personal history will be associated with higher propensity to cosponsor. Specifically, we attempt to verify this with (a) common religious affiliation, (b) former profession, (c) gender (both female), and (d) veteran status. We also examine the extent to which (e) same state and (f) same region—but not state—drive cosponsorship and compare these to the other principal driving forces, party and ideology. Shared religious affiliation, professional background, and veteran status (Barone et al., 2003) are examined as aspects of personal identity that may similarly play a role, if not overtly, then perhaps indirectly by way of common experience and exposure to similar points of view. A few religious denominations are well-enough represented in the Senate to be considered as covariates. We have considered three: Presbyterian (twelve), Catholic (twenty-three), and Jewish (eleven). Professional background was included in the form of an indicator for what was by far the most common type: legal (fifty-four). An additional indicator for veteran status (thirty-seven), including those in the Reserves, was also considered. Race and ethnicity would also make sense here, but the homogeneity of the 2003-04 Senate ruled this out.

\textbf{Ideological Distance} Absolute difference in point estimates of position in one-dimensional space (in the interval \([-1.0, +1.0]\), yielding values of no more than two in theory, and 1.65 as scaled in the dataset), according to NOMINATE-DW scores for the session under consideration.\(^9\) We use the primary component of Poole’s DW-Nominate scores

\(^9\)Inclusion of this variable renders superfluous the variable \textit{Same Party} in the contemporary Senate, as it implicitly contains the information present in the simple partisan indicator. Estimates of other coefficients were hardly affected by the choice of this variable instead of \textit{Same Party}. During the 108th Congress (January 2003 through December 2004), the United States Senate was nearly evenly split by party. Of the one hundred
for the 2003-04 Senate (108th) to measure relative ideology (Poole and Rosenthal, 1997). These are calculated based on roll-call votes and scaled on a single dimension, typically interpreted as liberal-conservative. A number of researchers have pointed out that one dimension is enough to explain most variation in roll-call voting during most periods in American history, and certainly during the past decade or two (Poole and Rosenthal, 1997).

**Same Region** = 1 if both represent states from the same geographical region (but not the same state), 0 otherwise. A number of issues addressed in legislation are of particular regional importance, and senators might additionally feel certain cultural affinity with colleagues from their particular region, so we examine whether common region may be associated with increased cosponsorship. To examine this, we group the states into the Northeast, Southeast, Midwest, Northwest, West, and the non-contiguous Pacific (AK, HI).

**Same State** = 1 if both represent the same state, 0 otherwise. Since they are elected “at large”, senators from the same state represent exactly the same electorate. Moreover, we might expect that having only two senators from each states invites, if not necessarily a special bond, then at least a common sense of obligation on state-specific concerns.

**Difference in Urbanization, Difference in Poverty Levels** are additional measures of dissimilarity between the constituencies that senators serve, the absolute difference in each respective measure for pairs of senators.

**Both Lawyers, Both Veterans** are indicators included for certain prevalent types of shared senators, fifty-one were Republicans, -eight were Democrats, and one—Jim Jeffords—a former Republican turned Independent. In the figures of dichotomized support propensities, Jeffords is shown in yellow, but for modeling purposes, he will be treated as a non-Republican, since he chose to align with Democrats on party line votes.
Figure 5: Nearly half the 7123 bills and amendments considered by the Senate during the 108th Congress contained no cosponsors at all. Of those with at least one cosponsor, mean = 7.40.

Both Female = 1 if both female, 0 otherwise. Fourteen women served as senators during the 108th Congress. Shared identity and concerns may give rise to higher cosponsorship rates among women (Swers, 2005).

Both Jewish, Both Catholic, Both Presbyterian are indicators for shared religious affiliations.

5 Results and Discussion

We fit two sets of models, all according to the setup presented in Equation (2/3). Unlike the exploratory analysis, we consider only bills, not amendments, but as before, just those having at least one cosponsor, but fewer than ninety-nine. Eliminating amendments will help reduce estimation issues arising from yet another source of clustering, as amendments are
naturally connected to the bills they amend. (Future work should model the interdependence of bills and amendments as well.) This gives us 2166 bills, yielding \(2166 \times 99 = 214,434\) observations in all, one for each possible cosponsor on every bill.

We begin by comparing the full multilevel logit model to a conventional logit model, noting differences in the point estimates and measures of uncertainty. The results are summarized in Figure 6, with inner and outer bars corresponding to one and two estimated standard errors, respectively (around 68% and 95% confidence, assuming normality). Afterwards, we fit the multilevel logit model separately for each of four subsets of observations, based on the distinct type of directed dyad by party (both Republican, both Democrat, Republican cosponsor of Democrat-sponsored bill, and Democrat cosponsor of Republican-sponsored bill). Since there may be a somewhat distinct process going on in each of these four situations, we may thus compare the resulting estimates rather than assuming a homogeneous process; we do so via predicted probability estimates, with the results summarized in Figure 7. The results have been computed using the \texttt{lme4} package (Bates et al., 2007), called from within the \texttt{arm} package (Gelman et al., 2011), in the R statistical programming environment (R Development Core Team, 2011).

**Multilevel vs. Standard Logit**

In the side-by-side comparison between the multilevel coefficient estimates and those in the standard logit, one notices some modest shrinkage toward zero in the former, a typical consequence of the partial pooling among clusters of observations. For the most part, coefficients that are clearly distinguishable from zero (with the anticipated sign) in the standard logit are also distinguishable in the multilevel version. These include *Both on Committees of Jurisdiction*, *Both on Other Committee*, *Both Female*, *Same Region*, *Same State*, *Ideological Distance*, *Number of Cosponsors on Bill*. Others are indistinguishable from zero in either
approach: *Both Lawyers, Both Presbyterian, Tenure Difference, Difference in Urbanization, Difference in Poverty Levels*. However, the coefficients on *Both Jewish, Both Catholic, and Both Veterans*, all clearly positively associated with propensity to cosponsor in the standard logit, are not statistically distinguishable from zero once the various types of interdependence are taken into consideration via the multilevel model, though point estimates remain positive. One coefficient, *Offices on the Same Floor*, which appears distinguishable from zero (in the “wrong” direction) according to the basic logit version ($p < .001$), is less obviously so according to the multilevel estimates ($p = .017$). In only one case does a result from the multilevel model potentially lead us astray where the basic logit get it right; a confidence interval (at any conventional level) around the coefficient for *Difference in Urbanization* includes zero under the standard GLM model but not using the multilevel estimates. We expect that even as we have come much closer to correct estimates of standard errors, we may still be underestimating a bit due to a lack of inclusion of higher order interdependence such as that due to triadic clustering. Thus, our suspicion is that coefficients on *Difference in Urbanization* and *Offices on the Same Floor* are effectively zero.

Figure 6 provides little insight into the substantive importance of each covariate, both because of the difficulty in directly interpreting coefficients on their original logodds scale and the lack of standardized units across covariates. For example, looking at the 95% confidence interval on the coefficient for *Number of Cosponsors*: (0.093, 0.095), we may mistakenly perceive this as a minor effect. Since the standard deviation on number of cosponsors per bill is around 10.7, the logodds of an individual act of cosponsorship may be expected to jump by somewhere between .99 and 1.01 with a standard deviation increase in number of cosponsors, corresponding to odds increasing by a factor of 2.7, or an increase in predicted probability of up to .25.¹⁰ That is, all else equal, an extra 11 cosponsors on a bill may mean

¹⁰Gelman and Hill (2007) recommend dividing logodds by four to get a rough idea of predicted probability, as this is the approximate slope of the logistic curve at its steepest.
up to an additional 0.25 probability that a given senator has cosponsored that bill.\textsuperscript{11}

It is also useful to keep in mind the variability observed in cluster-specific errors (or the random intercept distributions). We are not interested in particular instances of cluster-specific variation, as the main purpose in employing the multilevel approach is to control for the various levels of interdependent errors rather than to estimate the individual errors themselves. However, it helps us maintain a sense of perspective in discussing the relative contribution of various predictors if we understand the amount of variation due simply to the heterogeneity across dyads, individual senators, and committees, beyond what is explained by our included covariates. Looking at Table 1, standard deviations of the clustered errors range from around 0.22 to 0.74 on the logodds scale, or from around .05 to .18 as upper bounds on the probability scale. For instance, the distribution of idiosyncratic effects for sponsors (variation in how much individual senators attract cosponsorship, beyond what is predictable from covariates) has a standard deviation of 0.221; if one could look at the extra boost and penalty associated with specific individuals, the standard deviation of these probabilities would be no more than around 0.05. Variability based on individuals’ tendency to cosponsor a lot or a little seems to be a greater; standard deviation of 0.744 for \textit{Sender}_i is between three and four times as great as that for \textit{Receiver}_j. In other words, there is a great deal more variability in how often Senators cosponsor others than there is in how often Senators are themselves cosponsored. Bills from some committees will be more likely to receive a lot of cosponsors than those from others, and this is reflected in the standard deviation of 0.418 in logodds for an upper bound of around 0.10 on the probability scale.

\begin{footnote}
\textsuperscript{11} We have included this variable as a proxy for attractiveness or valence of a particular bill, but note that the data do not permit us to distinguish between (1) a high-valence bill drawing much support on its own perceived merits, and (2) a cascade effect, whereby individual senators sign on as they notice their colleagues do so.
\end{footnote}
Figure 6: Coefficients from a Standard Logistic Regression and from a Mixed Effects Logistic Regression Predicting Cosponsorship Patterns
Table 1: Estimated Variation in Cluster-Specific Effects (Error Components)

<table>
<thead>
<tr>
<th>Error Component</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bill Committee</td>
<td>0.418</td>
</tr>
<tr>
<td>Dyad (_{ij})</td>
<td>0.465</td>
</tr>
<tr>
<td>Directed Dyad (_{ij})</td>
<td>0.418</td>
</tr>
<tr>
<td>Receiver (_j)</td>
<td>0.221</td>
</tr>
<tr>
<td>Sender (_i)</td>
<td>0.744</td>
</tr>
</tbody>
</table>

Separate Models by Partisan Pattern

There is a fair amount of anecdotal evidence to suggest that sponsorship serves quite different purposes for members of the minority party than for those in the majority, and it would not be all that surprising to find that patterns of cosponsorship differ by party combinations of dyads. There are only four possible arrangements of directed dyad by party (DD, RR, DR, and RD). We run separate versions of our model using only observations where potential connections would have occurred between the combinations listed. Thus, our Democrat-Democrat model contains all the potential dyadic connections between Democrats on every bill in the chamber. Thus, the dependent variable is coded 1 when a Democrat cosponsors a bill introduced by a fellow Democrat, and 0 when a Democrat fails to cosponsor a bill sponsored by a fellow Democrat.

Figure 7 plots the change in predicted probability for several of the relevant relational variables in our models along with their upper and lower 95% confidence bounds.\(^{12}\) The changes in predicted probability are calculated using posterior simulation in which all other variables in the model are fixed at their modes for categorical data and means for continuous data. For dichotomous independent variables, the reported change occurs when the value switches from 0 to 1. For continuous variables, the reported change is a result of a change in the independent variable from its 75 percentile to its 25 percentile, a large but not atypical level of variation in the data. As evident in the figure, the change in probability of dyadic

\(^{12}\)Full results from these models may be provided upon request.
cosponsorship may vary by dyadic type. A decrease in ideological distance seems to correspond to increased propensity to cosponsor in any party combination, although the effect is not distinguishable from noise in the case of Democrats (the minority party) cosponsoring bills proposed by Republicans (the majority party). Among Democrats, a large drop in ideological distance corresponds to a large jump in predicted probability of cosponsorship (somewhere from +0.10 to +0.30, controlling for all other included covariates). Among Republicans, the predicted impact may have been more modest, as the simulations yield predicted spikes of somewhere between +0.05 and +0.18 for a similar change in standardized ideological distance (one-dimensional DW-NOMINATE scores). By contrast, a large drop in the difference in seniority (tenure in the Senate) between sponsor and potential cosponsor makes no statistically discernible difference on predicted cosponsorship, and the magnitude is low in any case (from −0.02 to +0.02 for a large change in seniority difference).

The figure also indicates that the geographic motivations for cosponsorship first observed in our exploratory analysis retain their predictive power in these subset models. Indeed, a move from out-of-state dyads to in-state dyads produces the largest change in the predicted probability of cosponsorship of any covariate included in the model, enormous jumps in predicted probability of anywhere from around +0.20 to +0.50 when Democrats are the potential cosponsors, and possibly a bit lower in the case of Republicans as potential cosponsors. Coupled with the positive effect associated with two actors being from the same region, but not the same state, this indicates a powerful geographic element to cosponsorship choices in the Senate, a pattern evident in our initial visualizations. Along with the negligible impact of differences in seniority, the relevance of both potential partners being female or both potential partners being Jewish seem to be barely discernible in any dyadic configuration, and on the order of a +0.04 at most in simulations, so substantively small in any case.

Finally, while two legislators both being on the committee of jurisdiction results in a predicted probability of cosponsorship by a relatively large margin, at a level distinguishable
from noise, the possible propinquity effect of simply serving on committees together even when these committees are not charged with handling the bill in question, is not apparent. In the case of intra-party propinquity via shared non-germane committees, statistically significant but substantively small effects appear, but even such modest effects are absent in the case of cross-party dyads.

6 Conclusions

Political networks are frequently two-mode, or bipartite networks where connections between the actors of interest occur through some secondary set of actors/objects. One prominent example of this phenomenon is the study of cosponsorship networks where legislators are connected to one another through their common cosponsorship of legislation. The current approach to dealing with bipartite political networks is to simply ignore the second mode as a potential source of variance to be explained. By imposing oversimplified weighting schemes or simply collapsing the second mode of variance entirely, scholars of social networks in political science have ignored important sources of prediction and explanation. We have provided an improved method for drawing statistical inferences in bipartite networks by developing and extending methods of multilevel modeling to incorporate theoretically interesting patterns of interdependence between observations, while allowing those patterns to vary according to the context provided along the second mode (the means through which individuals connect). Thus our model can incorporate mode-level, dyad-level, and node-level variance in the same statistical model while also accounting for some forms of interdependence that may bias coefficient estimates.

While we have demonstrated this technique using a detailed examination of cosponsorship, the method could be easily extended to other types of affiliation networks, including citizens connected to one another through participation in social groups, countries connected to
Figure 7: The change in predicted probability for different party dyad combinations. Dummy variable plots show the change in predicted probability that results from being coded 1 instead of being coded 0. The continuous variables (difference in seniority and difference in ideology) show the change in predicted probability as the difference decreases from the 75% to the 25% difference.
one another through membership in IGOs, and candidates connected to one another through donations from the same donors. In each of these settings, characteristics of the second mode of the network (groups, IGOs, and donors) are sources of variance that can be utilized in explaining phenomena of interest. Not only is it important to account for these sources of variance as potentially confounding elements in an empirical model; a whole host of theoretical arguments assert that social connections are a function of both the actors making the connections to one another and the means by which they make those connections. Our model allows scholars to disentangle both sets of potential explanations.

Future research might take advantage of the flexibility of this model to develop stronger methods of network comparisons. We have allowed for variance to be decomposed from bills, committees, dyads, and individuals in a legislature. This variance decomposition could be extended to many more levels. For example, scholars could consider cosponsorship patterns in many legislatures (rather than just the Senate) using our model and incorporating random intercepts and slopes for systematic differences between chambers. Additionally, rather than comparing the same behavior in many contexts, the model is capable of comparing many behaviors from the same actors. It would be trivial to extend the model in a way that would allow examination of trade and conflict networks between countries in the same model with varying slopes and intercepts for connection type (commerce or conflict). Thus, the model could distinguish between the common and distinct causes of different behaviors from the same set of actors.

Finally, while our model is exceptionally flexible and allows for network comparisons and estimates in ways other models cannot, it does have some limitations. The multilevel model can only incorporate dyadic or triadic types of interdependence. Complex networks with interdependence extending out beyond three actors will still result in biased estimates due to violations of regression assumptions. Thus, future work on bipartite networks (and network comparisons more generally) would do well to incorporate the exponential random
graph model’s ability to deal with complex forms of interdependence with our model’s ability to capture multiple modes as sources of variance.
References


