Mapping Political Communities: A Statistical Analysis of Lobbying Networks in Legislative Politics^{*†}

In Song Kim[‡]

Dmitriy Kunisky[§]

September 18, 2019

Abstract

Relationships between special interest groups and politicians play a significant role in policymaking. Yet empirical studies of interest group politics have been limited by the difficulty of observing these ties directly. In this paper, we propose a new methodology for inferring political actors' latent memberships in communities of legislative activity that drive their observable interactions. Unlike existing methods, the proposed *Bipartite Link Community Model* (biLCM) (1) applies to two groups of actors, (2) takes into account that actors may be members of more than one community, and (3) allows a pair of actors to interact in more than one way. We apply this method to characterize legislative communities in the 113th U.S. Congress. To overcome the difficulty of directly observing lobbying interactions, we construct an original dataset that connects the politicians who sponsor congressional bills with the interest groups that lobby on those bills based on more than two million textual descriptions of lobbying activities. The proposed methodology provides a quantitative measurement of community memberships ranging from narrow targeted interactions according to industry interests and jurisdictional committee membership to broad multifaceted connections across multiple policy domains.

Keywords: Network analysis, bipartite network, lobbying, ideal point estimation, scaling, stochastic block model, link community model, community detection

Word Count: 10,576 (abstract: 186)

^{*}Funding: This work was supported by the National Science Foundation (SES-1264090, SES-1725235 to I.K.).

[†]Acknowledgements: We thank Pablo Barberá, J. Lawrence Broz, Devin Caughey, Bruce Desmarais, Marina Duque, Jeff Harden, Lisa McKay, Nolan McCarty, Cristopher Moore, Michael Peress, Yunkyu Sohn, Erik Voeten, and Hye Young You for helpful comments. We also thank seminar participants at NYU-Abu Dhabi, Yale University, Duke University, Princeton University, University of California, San Diego, and University of Maryland. Conference participants at the American Political Science Association, the Asian Political Methodology annual meeting, and Political Methodology annual meeting provided thoughtful insights.

[‡]Associate Professor, Department of Political Science, Massachusetts Institute of Technology, Cambridge, MA, 02139. Email: insong@mit.EDU, URL: http://web.mit.edu/insong/www/

[§]Ph.D. Student, Department of Mathematics, Courant Institute of Mathematical Sciences, New York University, New York, NY, 10012. Email: kunisky@cims.nyu.edu, URL: http://www.kunisky.com/

1 Introduction

Special interest groups engage in lobbying to promote their political objectives (e.g., Wright, 1996; Grossman and Helpman, 2001).¹ The nature of the connection between politicians and interest groups, however, is much debated by political scientists. One popular view holds that interest groups take part in such costly political activity in order to buy votes or otherwise persuade policymakers, and in return policymakers gain a means of informing their legislative choices (Bauer, Dexter, and Pool, 1972; Potters and Van Winden, 1992; Austen-Smith and Wright, 1992; Wright, 1996). Others interpret lobbying as a "legislative subsidy" through which interest groups help allied politicians pursue common objectives (Hall and Deardorff, 2006). To assess these possibilities, researchers study patterns in how interest groups and politicians are connected to each other.

Early empirical works, such as Wright (1990) and Heinz et al. (1993), investigated interest group behavior by analyzing interviews with lobbyists, interest groups, and politicians. Though these inquiries have yielded many insights, they are limited in scale and treat only a few policy domains in isolation. More recently, political *networks* have been proposed as a broader framework for understanding how lobbying works (e.g., Heaney and Strickland, 2016). Still, important questions remain: how do committees and other congressional power structures guide interest groups' behavior? How do prominent legislators and interest groups allocate their attention among their many interests? Despite the significance of such behaviors and of the policy contexts that give rise to them, large-scale empirical studies of lobbying networks have been limited by the difficulty of making direct quantitative observations of the ties between interest groups and politicians.

In this paper, we address this challenge by proposing a new methodology for inferring political actors' latent memberships in communities of legislative activity that drive their observable

¹We use the term *special interest group* to refer to any political actor who has a specific policy objective, including firms, trade associations, labor unions, business associations, and professional associations. See https://lobbyingdisclosure.house.gov/amended_lda_guide.html for the definition of lobbying.

interactions. The proposed *Bipartite Link Community Model* (biLCM) departs from existing community detection models in the literature (see Fienberg and Wasserman, 1981; Airoldi et al., 2008, for example). First, it explicitly models mixed memberships through link communities.² That is, our model allows actors not only to belong to multiple political communities (so-called "mixed memberships"), but also to interact with another actor in multiple policy domains ("mixed interactions"). Second, our model uses a bipartite setting with two types of actors. We derive an Expectation-Maximization (EM) algorithm (Dempster, Laird, and Rubin, 1977) to estimate the biLCM, allowing researchers to apply this method in other bipartite networks with mixed memberships and mixed interactions.

Because we cannot directly observe lobbying interactions, we build a dataset that identifies a type of political connection between interest groups and members of Congress by combining two types of political behavior related to a bill: (1) which politician sponsored the bill and (2) which interest groups lobbied the bill. Our dataset is distilled from a larger database of lobbying data, which we build by applying natural language processing techniques on mandatory reports filed by lobbyists to identify the interest groups that lobbied on 108,086 congressional bills introduced between the 106th and 114th Congress. Although lobbying on a single bill does not necessarily imply political ties to its sponsor, *recurring* instances of lobbying that involve the same interest group and sponsor on numerous bills do reliably indicate close political relationships. Therefore, analyzing how often various politicians and interest groups interact in this way lets us infer the structure of political networks in the legislative process.

We apply the biLCM methodology to identify legislative communities in the 113th Congress.³ This yields quantitative measurements of participation in legislative issue domains for each interest group and politician. We find that "specialist" and "generalist" interest groups coexist in

² "Links" is a generic term from network theory for what we call "interactions" elsewhere.

 $^{^{3}}$ We focus on a single recent session of Congress to restrict our attention to a specific modeling task. In Appendix B.3 and Figure F.3, we present some preliminary results comparing biLCM results across several sessions.

U.S. legislative politics. Namely, we find that some interest groups engage in targeted lobbying of members of committees that have jurisdiction over their narrow interests. Meanwhile, other interest groups, particularly those that represent the varied interests of many members (such as the CHAMBER OF COMMERCE), engage in *aggregate* lobbying by targeting politicians who are members of broad "power committees" such as the HOUSE COMMITTEE ON WAYS AND MEANS and the SENATE COMMITTEE ON APPROPRIATIONS (Fenno, 1973). In this way, the proposed methodology improves our understanding of the organization of lobbying in the U.S. Congress.

Second, we find that alliances between disparate interest groups seldom govern interest groups' lobbying patterns. While one might expect alliances, such as those based on ideology, to lead interest groups to expand the scope of their lobbying, we find it rare for a nominally domainspecific interest group to lobby on unrelated legislation. In other words, the effect of alliances on lobbying is not so strong that it overwhelms the specificity of interests. This suggests that the network structure of lobbying is different from other types of political interactions. For instance, groups make campaign contributions based on ideology rather than shared interests (Bonica, 2013; Desmarais, La Raja, and Kowal, 2015). Similarly, civilians and journalists follow ideologically compatible politicians on social media (Barberá, 2014; Bond and Messing, 2015). In contrast, we find that interest groups lobby largely on their own interests. Moreover, the occasional crossinterest alliances we do observe concern social issues, consistent with Hojnacki (1997), who found that civic groups tend to participate in interest group alliances rather than lobbying alone.

Finally, by identifying connections that deviate most from the predictions of the biLCM, we are able to distinguish genuinely significant political connections from routine alignments of lobbying and legislative activity, such as energy firms lobbying on energy bills written by members of an energy committee. The special connections we identify in this way often coincide with campaign contributions, geographic ties between politicians and interest groups, or disruptive and industrychanging legislation. Our analyses also provide technical guidance for further research on political networks. To the best of our knowledge, ours is the first statistical study of lobbying networks in legislative politics that examines both politicians and interest groups using micro data. More broadly, the proposed community detection approach is a novel empirical framework for political scientists to systematically identify overlapping political communities and their members, which will help leverage the rich political network data that have become increasingly available (Keck and Sikkink, 1998; Hoff, Raftery, and Handcock, 2002; Hafner-Burton, Kahler, and Montgomery, 2009; Clark and Lauderdale, 2010; Ward, Stovel, and Sacks, 2011).

Our network data, all pairwise measurements of interactions between politicians and interest groups, the estimated community memberships of political actors, ancillary results from other models, and the visualization tools used in preparing this paper will be made publicly available at https://www.lobbyview.org. The proposed methods for network analysis will be implemented in the open-source software package "polnet: Statistical Analysis of Political Networks."

2 The Bipartite Link Community Model

In this section, we propose a statistical model of bipartite networks of interactions between two types of actors. We begin by motivating the methodology in the context of interest groups' interactions with politicians. In doing so, we describe two key features of our method: mixed memberships and mixed interactions. We then formally introduce the *Bipartite Link Community Model* (biLCM). Finally, we derive an EM algorithm to estimate the model parameters, giving a scalable implementation of the proposed method.

2.1 Motivation

Political networks have proved crucial to our understanding of how actors behave in various policymaking domains (Baumgartner and Leech, 1998; Box-Steffensmeier, Christenson, and Hitt, 2013; Box-Steffensmeier and Christenson, 2014; Hadden, 2015). Specifically, political networks are often organized into distinct *communities*⁴ within which actors typically interact. Researchers have found that members of Congress form partisan communities in cosponsorship networks (Zhang et al., 2008), while larger congressional units such as committees and subcommittees form more complex communities according to subject matter (Porter et al., 2005). Interest groups and politicians also form joint communities in networks based on interactions such as campaign contributions (Desmarais, La Raja, and Kowal, 2015). Our proposed methodology is motivated by the presence of *legislation communities*—collections of interest groups and legislators that are interested in the same bills. Our goal is to model the structure of a particular lobbying network by identifying the most important legislation communities and their members.⁵

Models based on community structure have been widely studied in the network science literature. Perhaps the most common such model is the *stochastic block model* (SBM) of Fienberg and Wasserman (1981), which was adapted by Larremore, Clauset, and Jacobs (2014) into the *bipartite stochastic block model* (biSBM), a similar model specific to bipartite networks. In these models, each political actor belongs to one of k communities, and interactions depend only on actors' respective community memberships. Often, the network structure is *assortative*, meaning that actors interact more when they share a community, and less when they do not.

In addition, our model incorporates two patterns characteristic of complex political interactions that have been identified empirically in the literature (Gray and Lowery, 2000; Heaney, 2004): (1) mixed memberships and (2) mixed interactions. *Mixed memberships* occur when political actors participate in multiple communities. This can occur when actors have heterogenous interests across various policy domains or represent the interests of diverse constituents. For instance, an oil energy firm might be interested in both restrictions on utility providers and automotive pollu-

 $^{^{4}}$ We use the term "community" in the technical sense common in network science: a community is a densely connected subset of actors in a network.

⁵We show below that the presence of community structure is also corroborated by latent space models, which do not assume *a priori* that community structure exists in the lobbying network.

tion regulation. Similarly, most members of Congress are interested in numerous policy domains (Lauderdale and Clark, 2014) and sit on several committees that have distinct legislative jurisdictions. To account for this, our model allows political actors to have simultaneous memberships of varying strengths in *all* legislation communities.

One prominent model in the SBM family that incorporates this enhancement is the *mixed-membership stochastic block model* (mmSBM) (Airoldi et al., 2008). In this model, each actor has a probability distribution over communities. When two actors interact, they "roll the dice" to choose which community they will belong to for this interaction; how many times they interact depends only on that choice of communities. Panels (a) and (b) of Figure 1 illustrate how the biSBM differs from the mmSBM: the single-membership biSBM (panel (a)) assumes that each actor belongs to only one community (indicated by color), and members in the same community tend to interact more frequently with each other than with actors in other communities. In contrast, the mmSBM (panel (b)) allows an actor to act in different communities for different interactions.⁶

Although the mmSBM has some desirable features, it also has the property that an actor cannot be part of more than one community while interacting with a given counterpart. For modeling interactions between interest groups and politicians, this is a limitation, because we expect to observe *mixed interactions*: pairs of political actors who interact with each other in multiple communities. For example, the CHAMBER OF COMMERCE interacted repeatedly with the same legislator, Senator Barbara Boxer (D-CA), by lobbying three diverse bills: 113th S. 601 "Water Resources Development Act of 2013," 113th S. 462 "United States-Israel Strategic Partnership Act of 2013," and 113th S. 809 "Genetically Engineered Food Right-to-Know Act." Similarly, the SPECIALTY EQUIPMENT MARKET ASSOCIATION lobbied on 113th S. 983 "Keep the IRS Off Your Health Care Act of 2013," 113th S. 2635 "21st Century Endangered Species Transparency Act,"

⁶For the sake of comparison, we illustrate a hypothetical adjustment of the original mmSBM to an assortative community structure in a bipartite network, with interactions having integer magnitudes.



Figure 1: Schematic Comparison of Community Models. We illustrate three generative network models: the biSBM, the mmSBM, and the biLCM. The first two models are illustrated in the *assortative* case, where each politician community interacts strongly with just one interest group community and vice-versa. Colors indicate community memberships and, in the mixed membership cases, interaction types. (a) In the single membership model, political actors sharing a community interact more frequently. (b) The mmSBM allows for mixed memberships, but all interactions of a single pair of actors are constrained to be of a single type. (c) The biLCM allows for mixed membership as well as mixed interaction types. This captures political interactions in which actors with diverse memberships interact repeatedly for different political reasons.

and 113th S. 725 "Small Business Taxpayer Bill of Rights Act of 2013." All three bills, which address very different substantive issues, were sponsored by Senator John Cornyn (R-TX). In both cases, a model in the mmSBM family would bizarrely insist that a single common community membership must account for every interaction involving the two actors.

To overcome this limitation, instead of modeling each interaction as a *mixture* of possible interactions in different communities, we model each interaction as a *sum* of independent interactions in *all* possible communities. Panels (b) and (c) of Figure 1 illustrate this difference. Under our model, illustrated in panel (c), interactions between a single pair of actors can belong to multiple communities (see also Equation (1) below). In this way, we arrive at a model similar to the *link community model* (LCM) of Ball, Karrer, and Newman (2011) that we adapt specifically to the bipartite setting. The most similar prior work we are aware of is that of Li, Zhang, and Zhang (2015), who treat the link community detection task in bipartite graphs as an optimization problem that they solve with an ad hoc genetic algorithm. Unlike that work, our model provides an underlying probabilistic generative model and therefore a statistical interpretation.

In summary, motivated by examples from interest group lobbying, we posit that political actors are organized into legislation communities, that actors are often members of multiple communities, and that two actors may interact in more than one community. As we will see, this allows us to accurately describe the activity of both powerful actors with broad interests and less prominent actors with narrower interests within the same model.

2.2 The Model

We now give a general mathematical description of the biLCM. Suppose that we have two disjoint groups of political actors, U and V, which we index by $i \in U = \{1, \ldots, m\}$ and $j \in V = \{1, \ldots, n\}$. We denote the number of interactions between the pair i, j by $A_{i,j}$, and organize these numbers as the entries of the *interaction matrix* $\mathbf{A} \in \mathbf{R}^{m \times n}$. This matrix may be viewed as the adjacency matrix of a bipartite graph with weighted edges, where actors in the groups U and V lie on opposite sides of the partition.

We suppose that interactions occur in k "link communities" (a generic term for "legislation communities" from the presentation above), which we index by $z \in \{1, \ldots, k\}$. Each actor iand j has a vector of parameters $\alpha_{i,z}$ and $\beta_{j,z}$, respectively, which represents their involvement in community z. The number of total interactions between i and j in community z is modeled as Poisson with a mean proportional to $\alpha_{i,z}\beta_{j,z}$. To impose parameter identification, we assume that for each fixed z, $\sum_{i=1}^{m} \alpha_{i,z} = \sum_{j=1}^{n} \beta_{j,z} = 1$, and we introduce another parameter κ_z to capture the overall level of activity in community z, so that the number of interactions between i and j in community z has mean $\kappa_z \alpha_{i,z}\beta_{j,z}$. We assume that these Poisson variables are independent, thus the joint distribution of the interaction matrix is

$$\mathbf{P}(\mathbf{A} \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}) = \prod_{i=1}^{m} \prod_{j=1}^{n} \operatorname{Poisson} \left(A_{i,j} \mid \sum_{z=1}^{k} \kappa_{z} \alpha_{i,z} \beta_{j,z} \right).$$
(1)

2.3 Estimation

We derive an EM algorithm for this model in Appendix B.1. The algorithm alternates expectation and maximization update steps until the log-likelihood of the model converges. The update equations produced by our derivation are given below, including ancillary optimization parameters $q_{i,j}(1), \ldots, q_{i,j}(k)$. The first equation is the expectation step for the ancillary parameters, and the last three equations are maximization steps for the model parameters.

$$q_{i,j}(z) = \frac{\kappa_z \alpha_{i,z} \beta_{j,z}}{\sum_{z=1}^k \kappa_z \alpha_{i,z} \beta_{j,z}},\tag{2}$$

$$\kappa_{z} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} A_{i,j} q_{i,j}(z)}{\sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,z} \beta_{j,z}},$$
(3)

$$\alpha_{i,z} = \frac{\sum_{j=1}^{n} A_{i,j} q_{i,j}(z)}{\sum_{i=1}^{m} \sum_{j=1}^{n} A_{i,j} q_{i,j}(z)},$$
(4)

$$\beta_{j,z} = \frac{\sum_{i=1}^{m} A_{i,j} q_{i,j}(z)}{\sum_{i=1}^{m} \sum_{j=1}^{n} A_{i,j} q_{i,j}(z)}.$$
(5)

In practice, because the EM algorithm is not guaranteed to converge to the maximum likelihood parameters, we perform 50 randomly initialized runs and choose the parameters that attain the highest likelihood value.

3 Empirical Findings

In this section, we use the biLCM to better understand the structure of a network of interest groups and legislators. We begin by introducing an original database of lobbying data. We then apply the biLCM to a specific dataset concerning the 113th Congress and present the estimated community memberships of interest groups and legislators. Our analysis illustrates a distinction between "specialist" and "generalist" groups and identifies the legislators with whom each kind is connected. We also discuss the role of alliances among interest groups and a class of anomalous political connections identified by the biLCM. Finally, we discuss the benefits of this methodology

over other network models.

3.1 The Lobbying Network Database

All lobbying organizations are required by law to file quarterly reports describing the issues on which lobbyists have engaged in political activities, including lobbied bills. However, organizations need not identify their political contacts. This is unfortunate, since almost 90% of lobbying reports indicate that at least one member of Congress or member of their staff was contacted.

Because we cannot directly observe lobbying interactions, we construct an original lobbying database that indirectly captures the connections between interest groups and politicians. Our database is built from the universe of reports filed between 1999 and 2017.⁷ We use a suite of automated systems to (1) identify lobbied bills, (2) identify the session of Congress that those bills were introduced in, and (3) identify each bill's sponsor. We briefly outline this process below.

First, we analyze more than two million "specific lobbying issues" found in the lobbying reports, which refer to lobbied bills by either number or title.⁸ In practice, bills are often referenced by alternative names or subtitles, so bill references must be algorithmically identified. Second, we identify the session of Congress that each bill belongs to, which the report often does not mention explicitly. This empirical challenge has prevented prior works from analyzing lobbying accurately. To overcome this difficulty, we mine several signals from the report text to predict the session number, including phrases like "Act of [YEAR]," other bills identified in the same report, and similarity between bill texts and the report text mentioning them.⁹ Next, we record

⁷Compliance with disclosure requirements is closely monitored, strictly enforced, and annually audited by the Government Accountability Office (GAO). According to the 2014 GAO audit report, 90% of organizations filed reports as required and 93% could provide documentation related to expenses.

⁸Each report may describe several issues lobbied, resulting in a set of 2,396,693 corpora to parse.

⁹When bills are mentioned only by number, it is difficult to determine which bill is being described, as bills are numbered starting at 1 in every new congressional session. This presents a substantial challenge; for instance, the OpenSecrets.org database from the Center for Responsive Politics, which is often used in academic research, erroneously states that INTEL lobbied "Value Our Time Elections Act" instead of "National STEM Education Tax Incentive or Teachers Act of 2011." Both bills are numbered H.R. 289. Such examples call into question the validity of previous findings based on these data. See http://www.opensecrets.org/lobby/billsum.php?id=hr289-113.



Figure 2: **Descriptive Statistics of Lobbying.** The left panel presents the numbers of bills introduced, lobbied, and voted on in each session from the 106th through the 114th Congress. Note that the increase in the number of bills lobbied beginning with the 110th Congress is likely a function of the digitization of lobbying records and does not necessarily reflect an actual increase in lobbying. The right panel shows how many interest groups lobbied on each bill introduced in the 113th Congress. The distribution is highly skewed: the median is three while the maximum is 946 (on 113th H.R. 1, "Tax Reform Act of 2014"), and 23% of lobbied bills are lobbied by only one interest group.

the politicians sponsoring each bill. Finally, we identify all the actors of the lobbying sector (lobbyists, their lobbying firms, the interest groups that hire them, and so forth) that are involved in lobbying on each bill. We repeat this process for each of 1,111,859 lobbying reports, which in total link 20,092 special interest groups with 1,164 (current and former) members of Congress.

The left panel of Figure 2 shows that about 12,000 bills are introduced in each Congress and the majority of them are lobbied by at least one interest group. Very few of these bills make it to a vote as written, however. Instead, they tend to be merged into larger bills following a complex process of amendment and compromise. Importantly, we observe interactions *before* this noisy process takes place, and thus we witness the lobbying network at a particularly granular level. There is ample evidence that lobbyists help to draft bills or even write entire bills on behalf of legislators (Nourse and Schacter, 2002; Hertel-Fernandez, 2014). Indeed, as the right panel shows, most bills are lobbied by very few interest groups, suggesting that each individual instance of lobbying tends to reflect narrow interests. A typical example: on "A bill to exempt the aging process of distilled spirits from the production period for purposes of capitalization of interest costs" (113th S. 1457), sponsored by Senator Mitch McConnell (R-KY), the DISTILLED SPIRITS COUNCIL was the only interest group to lobby.¹⁰ This bill has never been voted on in the Senate (and therefore would never appear in roll call data), but has been reintroduced twice by McConnell in subsequent sessions of Congress.

3.2 Findings with the Bipartite Link Community Model

We now apply the biLCM to lobbying in the 113th Congress. To produce a suitable dataset, we take $A_{i,j}$ to count the number of bills interest group *i* lobbied that politician *j* sponsored (see Appendix A.2 for further details on the data used in this section). We then estimate the parameters of the biLCM using k = 8 legislation communities.¹¹

We name the legislation communities that result by examining the interest groups that interact in them most frequently. Some involve straightforward collections of industry- or issue-specific groups, such as "Healthcare," "Veterans' Affairs," "Technology & Telecommunications," "Energy," and "Finance & Insurance." Others are broader: one, which we name "Universities & Research," involves a mix of universities and aerospace and defense research firms (CRAY, GENERAL DYNAM-ICS). Another, named "Civil Society," includes leading civil rights associations (ACLU) as well as lobby organizations for the rights of workers (AFL-CIO), senior citizens (AARP), women (NATIONAL WOMENS LAW CENTER), the disabled (EASTER SEALS), and others. The last, named "Retail & Transportation," includes agricultural groups (NATIONAL CORN GROWERS), manufacturing associations (NATIONAL ASSOCIATION OF MANUFACTURERS), automotive firms (NISSAN), retailers (HOME DEPOT), and fuel manufacturers (BRITISH PETROLEUM).

To illustrate the information the biLCM provides, in the table of Figure 3 we present legislation community memberships for six interest groups that illustrate the range of lobbying behaviors

 $^{^{10}}$ See Kim (2017) for examples of similar patterns in trade bills.

¹¹Our choice is guided by the community structure we observe in the latent space model below. Researchers may specify the number of communities based on the granularity of community structure they wish to observe.

Interest Group	Distribution	Example Bills
Microsoft		Startup Act 3.0 Law Enforcement Access to Data Stored Abroad Act Satellite Television Access and Viewer Rights Act
McAfee, Inc.		Cybersecurity Enhancement Act Department of Defense Appropriations Act Energy Efficient Government Technology Act
Arch Coal		Climate Protection Act Caring for Coal Miners Act Energy Consumers Relief Act
British Petroleum (BP)		Keep American Natural Gas Here Act RFS Reform Act Corn Ethanol Mandate Elimination Act
Philips North America	11.1 1	Protect Medical Innovation Act Medicare DMEPOS Market Pricing Program Act Renewable Energy Parity Act
Chamber of Commerce		Water Resources Development Act United States-Israel Strategic Partnership Act Genetically Engineered Food Right-to-Know Act



Figure 3: Legislation Communities in the biLCM. In the table at the top, we give examples of interest groups and their legislation community memberships, with examples of bills that they lobbied. Memberships are plotted as histograms, with eight bars for the eight communities in the biLCM. In the bottom plot, we show mixed interactions between actors in the "Healthcare" and "Civil Society" communities, by plotting a portion of the lobbying network. Links indicate the number of interactions belonging to these communities under the biLCM, and color indicates their distribution between the two communities.

the biLCM identifies in the network. Firms such as MICROSOFT and ARCH COAL lobby on a single topic, thus they interact with politicians in a single legislation community ("Technology & Telecommunications" and "Energy," respectively) with probability at least 85%. Firms with broader interests, such as MCAFEE, INC. and BRITISH PETROLEUM, lobby primarily in two or three arenas. For instance, MCAFEE, INC., a computer security firm, lobbied on 113th S. 1429 "Department of Defense Appropriations Act" (as did defense firms such as GENERAL DYNAM-ICS and NORTHROP GRUMMAN) and 113th H.R. 756 "Cybersecurity Enhancement Act" (as did technology firms including MICROSOFT and GOOGLE). Accordingly, MCAFEE, INC. interacts in the "Technology & Telecommunications" community with probability 31% and in the "Universities & Research" community with probability 52%. Large holding companies such as PHILIPS NORTH AMERICA or organizations such as the CHAMBER OF COMMERCE lobby most categories of legislation, interacting in any single community with probability no more than 25%.¹²

To demonstrate the network structure we find, we plot part of the lobbying network in the bottom panel of Figure 3. We show actors involved in the "Healthcare" and "Civil Society" communities as well as the classification of their interactions by the biLCM. We show interactions occurring in one community (vertical links on the left and right) in lighter colors and mixed interactions (diagonal links and vertical links in the center) in darker colors. As the figure shows, many repeated interactions are characterized by legislation community mixtures. Thus the mixed interactions modeled by the biLCM are crucial to accurately describing the lobbying network.

Before proceeding, we describe a simple way to quantify how many different legislation communities an actor actively participates in. From the model definition, the mean total number of times interest group *i* lobbies in community *z* equals $\kappa_z \alpha_{i,z}$, and likewise $\kappa_z \beta_{j,z}$ for politician *j*. A natural choice is then to form probability distributions $p_{i,z} = \frac{\kappa_z \alpha_{i,z}}{\sum_{z=1}^k \kappa_z \alpha_{i,z}}$ and $q_{i,z} = \frac{\kappa_z \beta_{j,z}}{\sum_{z=1}^k \kappa_z \beta_{i,z}}$, and

¹²We present analogous examples for politicians in Table F.1, finding that the legislators with the most legislation community memberships are often senior politicians or party leaders.



Figure 4: Committee Membership of Politicians by Community. We present histograms of the top 10 committee memberships for legislators in different legislation communities found by the biLCM. Bars are divided by political party (red for Republican and blue for Democrat), and the horizontal bars at the top show the overall party distribution of each community. (For reference, 52% of members of the 113th Congress were Republican.)

then consider the entropies $H_i = H(p_{i,1}, \ldots, p_{i,k})$ and $H_j = H(q_{j,1}, \ldots, q_{j,k})$, where $H(c_1, \ldots, c_k) = -\sum c_z \log_2 c_z$. In computing entropies, we take logarithms base 2 for the sake of interpretability: an entropy of H may then be interpreted as, roughly speaking, an actor typically participating in 2^H legislation communities. Simply put, actors with higher entropy values are members of more legislation communities than actors with lower entropy values.

Committees Associated with Specific Communities We study how typical congressional committee memberships differ between politicians that interact in a small number of legislation communities versus those that interact in many legislation communities. In Figure 4, we draw histograms of committee memberships for three groups of legislators: legislators who interact with probability at least 40% in the "Finance & Insurance" community, legislators who interact with probability at least 40% in the "Veterans' Affairs" community, and legislators with entropy at least 2, i.e., those who actively participate in at least four different communities.

The first two groups of politicians have committee memberships that are concentrated in committees relevant to the subject matter of their community, such as the HOUSE COMMITTEE ON FINANCIAL SERVICES and the HOUSE COMMITTEE ON ARMED SERVICES. In contrast, the high-entropy politicians have committee memberships on so-called "power committees," like the HOUSE COMMITTEE ON WAYS AND MEANS and the SENATE COMMITTEE ON APPROPRIATIONS. These committees hold the "power of the purse," controlling various financial aspects of governance (Fenno, 1966). These results are also compatible with Wright (1990), who, based on interviews and surveys, found that lobbying affected voting especially strongly in the HOUSE COMMITTEE ON WAYS AND MEANS compared to domain-specific committees.

Our results suggest that several other congressional committees enjoy a similar popularity in the lobbying sector. Specifically, we find that the SENATE COMMITTEE ON COMMERCE, SCIENCE, AND TRANSPORTATION, the SENATE COMMITTEE ON HEALTH, EDUCATION, LABOR, AND PENSIONS, and the SENATE SPECIAL COMMITTEE ON AGING are populated by legislators with high entropy values, meaning they are members of many legislation communities. The latter two may be seen as having some financial control, via issues such as minimum wage and Social Security. These committees also reached their current breadth of responsibility more recently than the other power committees. This analysis therefore suggests that the biLCM may be used to keep the notion of power committees up to date with the evolving landscape of Congress. To do this, we leverage the lobbying sector's knowledge of which committees hold the greatest and broadest power, and thus make this choice in a statistically principled and purely empirical manner.

The Role of Interest Group Alliances Analyzing interest group alliances¹³ in lobbying using the available data is a subtle matter, as the data do not distinguish between lobbying for or against a given piece of legislation. This limitation notwithstanding, we may assess one potential effect on lobbying activity. Namely, if such alliances had a strong role in determining the legislative action of interest groups, we would expect to observe interest groups lobbying *outside* of their nominal interest areas together with allies. That is, alliances would induce a *broadening* of interests.

 $^{^{13}}$ We consider alliances arising from any commonalities other than shared interests that might unite interest groups, such as ideology, geography, or personal connections.

We occasionally observe this dynamic in the lobbying marketplace. For instance, the SIERRA CLUB, an environmentalist organization typically lobbying bills like 113th H.R. 3826 "Protecting States, Opening National Parks Act," also lobbied 113th H.R. 3206 "Global Sexual and Reproductive Health Act of 2013" together with PLANNED PARENTHOOD. On the same issue, the OUTDOOR INDUSTRY ASSOCIATION, an outdoor recreation trade organization, typically lobbying bills like 113th H.R. 5204 "Federal Lands Recreation Enhancement Modernization Act of 2014," also lobbied 113th H.R. 1389 "Military Access to Reproductive Care and Health (MARCH) for Military Women Act," again together with PLANNED PARENTHOOD. Accordingly, the SIERRA CLUB and the OUTDOOR INDUSTRY ASSOCIATION participate in the "Civil Society" legislation community, with probability 33% and 19% respectively.

Examples like these suggest that the biLCM captures some of the broadening effect that alliances, ideological alliances among them, may have on interests. Moreover, we expect this to often be observed in the "Civil Society" community (indeed, that is the only community that appears not to admit a description in terms of a specific industry or issue). On the other hand, that community is only the fourth most active out of eight (assessing this by the κ_z model parameters), and thus accounts for only a modest fraction of lobbying activity. Therefore, we find evidence that interests remain the primary determinant of the scope of lobbying activity, and interest group alliances play at most a secondary role. Moreover, the role of alliances appears to be more pronounced in lobbying on social and civic issues, confirming a thesis of Hojnacki (1997).

Anomalous Connections Another analysis that emerges from the biLCM is to examine the results that deviate the most from the model's predictions. We list these connections in Table 1. As we described earlier, our analysis of the results of the biLCM finds that the legislation communities in the lobbying network are organized around actors' interests: industry-specific groups will lobby politicians with power over that industry; groups with a variety of interests may lobby either

Legislator	Party	State	Interest Group	Observed	Estimated
Tom Marino	R	PA	CVS	10	0.55
Mike Conaway	R	ΤХ	IHEARTMEDIA INC.	7	0.17
Michael Enzi	R	WY	EBAY	11	1.08
Gary Peters	D	MI	MEMA	5	0.09
Fred Upton	R	MI	DIRECTV	7	0.31
Steve Scalise	R	LA	DIRECTV	7	0.32
John Barrasso	R	WY	IHEARTMEDIA INC.	6	0.19
Jeanne Shaheen	D	NH	Novo Nordisk	12	1.61
Jim Costa	D	CA	Met. Water District So. CA	4	0.05
Peter DeFazio	D	OR	TREA SENIOR CITIZENS LEAGUE	6	0.24

Table 1: Anomalous Interactions Identified by the biLCM. We list the 10 pairs of legislators and interest groups whose number of observed interactions deviates the most from the predictions of the biLCM. Probabilities are computed directly from Equation (1) using the estimated maximum likelihood parameters.

a variety of politicians or specific politicians with broad power. The connections in Table 1, in which the numbers of observed interactions *exceed* those predicted by the biLCM, represent pairs of actors whose interactions must then be driven by a mechanism other than shared interests.

Indeed, investigating these interactions suggests a range of interesting explanations. In some cases, actors are connected by geography, such as Senator Jeanne Shaheen (D-NH) and NOVO NORDISK (a Danish firm whose U.S. operations are headquartered in New Hampshire), and Representative Jim Costa (D-CA) and the METROPOLITAN WATER DISTRICT OF SOUTHERN CALIFORNIA. In other cases, campaign contributions show a close connection. For instance, IHEARTMEDIA, INC. donated actively to Representative Mike Conaway (R-TX) and Senator John Barrasso (R-WY) and also interacted with the legislators on legislation. Finally, sometimes the interest group and politician connect over individual legislation that is important to a relevant industry, such as Representative Tom Marino (R-PA) and CVS (over regulation of opioid distribution), and Senator Michael Enzi (R-WY) and EBAY (over internet sales tax). In these cases, the bills in question appear especially closely tied to their sponsors, for instance often appearing in electoral materials or receiving other publicity.

We thus find that the biLCM is a useful means of revealing "exceptional" interactions between

politicians and interest groups that go beyond classification according to interest structure. In this way, the biLCM finds politician-interest group connections with a range of specific substantive explanations. The biLCM thus may be a useful tool for directing future work on the lobbying network that can investigate theoretical explanations for these connections.

Significance of Communities How do we know that the community structure identified by the biLCM arises from the structured interactions of interest groups and legislators rather than merely their individual propensities to lobby or sponsor? To provide an answer, we perform a permutation test on the lobbying network to obtain a null distribution of random networks with similar activity distribution of individual actors (see Appendix B.2 for technical details). We then compute entropies for the biLCM legislation community memberships in these random networks and compare them to the entropies from the model of the actual lobbying network. Since the biLCM can capture both single and multiple community memberships, this is a useful test of the statistical significance of the community structure we have been discussing: if the lobbying network has strong community structure compared to a "typical" similar network, then the biLCM should find fewer community memberships for actors in the lobbying network than for actors in a generic network. Indeed, we find that this is the case: interest groups in the null model have much higher numbers of legislation community memberships than do interest groups in model of the actual lobbying network.¹⁴

On the whole, we find that the biLCM successfully describes the multiple community memberships of political actors and how interactions among actors occur in those communities. The proposed methodology allows researchers to quantitatively estimate participation in various communities for each political actor. Furthermore, it allows researchers to relate interest groups with

¹⁴On the other hand, legislators in the null model have similar numbers of legislation community memberships to legislators in our model of the actual lobbying network. This agrees with our previous observation that varied interests are more typical for politicians than for interest groups: while many interest groups are focused on specific industries, most politicians are involved in a variety of policy domains.

specialized and general interests to politicians with narrow and broad political roles. In sum, the proposed methods not only capture the community structure that we often cannot directly observe in political networks, but also provide a useful guide to the varied interests and complex preference aggregation that drives lobbying in legislative politics.

3.3 Comparison with Other Models

We now compare the biLCM to two other types of model, latent space models and stochastic block models. We adapt these existing models to our setting, and show that the biLCM is able to distinguish key features of the lobbying network that the other models fail to identify.

3.3.1 Latent Space Models

In this section, we compare the results of the **biLCM** with an approach based on a *latent space* model (LSM), a common modeling choice for networks. Of the many variants, our model is formally similar to the "Wordfish" model of Slapin and Proksch (2008) for text analysis, which estimates ideal points of a single type of political actor based on text data. In contrast, we work in the bipartite setting because we have two types of actors. Our model is also related to that of Barberá (2014), which estimates ideal points of politicians and social media users based on the binary interaction of Twitter follows. Our model differs because our interaction measurements have an associated magnitude.

The Model We model the interaction between interest group i and legislator j as a function of the Euclidean distance between the two actors' latent positions in d-dimensional space, $\theta_i \in \mathbf{R}^d$ and $\psi_j \in \mathbf{R}^d$ respectively, and assume that interactions have independent Poisson distributions.

To account for the differences in actors' baseline propensities to sponsor or lobby (see Figure F.1), we include interest group- and legislator-specific "popularity" terms, α_i and β_j respectively (Krivitsky et al., 2009). We then take the mean of the interaction $A_{i,j}$ to be $\exp(\alpha_i + \beta_j - \|\boldsymbol{\theta}_i - \boldsymbol{\psi}_j\|_2^2)$. To implement inference, it is more convenient to reparametrize in terms of $\tilde{\alpha}_i = \alpha_i - \|\boldsymbol{\theta}_i\|_2^2$ and $\tilde{\beta}_j = \beta_j - \|\boldsymbol{\psi}_j\|_2^2$. Then, to impose parameter identification and improve the numerical behavior of sampling, we impose hierarchical priors on the latent space positions and modified popularity factors. The posterior distribution under this model is then given by:

$$\mathbf{P}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{\psi} \mid \mathbf{A}) \propto \prod_{i=1}^{m} \prod_{j=1}^{n} \operatorname{Poisson}\left(A_{i,j} \mid \exp(\tilde{\alpha}_{i} + \tilde{\beta}_{j} + 2\boldsymbol{\theta}_{i}^{\top}\boldsymbol{\psi}_{j})\right) \times \prod_{i=1}^{m} \mathcal{N}\left(\tilde{\alpha}_{i} \mid 0, \sigma_{I}^{2}\right) \mathcal{N}\left(\boldsymbol{\theta}_{i} \mid \mathbf{0}, \operatorname{diag}(\boldsymbol{\tau})\right) \times \prod_{j=1}^{n} \mathcal{N}\left(\tilde{\beta}_{j} \mid \nu, \sigma_{P}^{2}\right) \mathcal{N}\left(\boldsymbol{\psi}_{j} \mid \mathbf{0}, \operatorname{diag}(\boldsymbol{\tau})\right)$$
(6)

Computation We estimate our model with the Hamiltonian Monte Carlo No-U-Turn Sampler and Automatic Differentiation Variational Inference implementations in the Stan software package (Carpenter et al., 2017). In Appendix C, we include code describing our model and discuss how parameter identification is achieved, the choice of dimensionality, the concentration of the posterior distribution, and sampling diagnostics.

Comparison We now apply the LSM to the lobbying network dataset for the 113th Congress. For the two-dimensional LSM, Figure 5 presents the posterior means of the estimated latent spatial positions, θ_i and ψ_j , for all interest groups *i* and politicians *j*, plotted against the DW-NOMINATE ideology dimension for politicians and sized according to popularity factor.

Figure 5 shows that the LSM finds some of the same structure as the biLCM in the lobbying network, identifying communities (shaded regions) of specific issue areas and industries, which arise as geometric clusters. In these communities, interest groups share an industry affiliation (such as "Telecommunications") or an interest in an issue area (such as "Veterans' Affairs"), and politicians sit on committees involved with the policies that affect those groups (see Figure F.5 in the Appendix). While often the dimensions of latent space models encode ideological information in political network models, we find that, even when tested formally with an F-test of regression models involving ideological and other covariates, the inferred LSM coordinates depend strongly



popularity factors inferred from the 113th Congress dataset.^a We indicate several significant clusters corresponding to specific industries Estimated LSM Positions: Full 113th Congress. This figure presents the two-dimensional latent space positions and and issue areas. Interest groups are represented with black dots while politicians are represented with colored dots according to their ideological score on the DW-NOMINATE scale. The size of each dot is proportional to the exponential of the actor's popularity factor $(\exp(\alpha_i)$ or $\exp(\beta_i))$. We annotate the clusters with some representative members. Figure 5:

^aOne actor with outlying mean latent positions (right of the region shown), DIRECTTV, INC., is omitted for the sake of visual clarity

only on committee memberships of politicians and industry affiliations of interest groups (see Appendix C.5). Thus the LSM appears to detect no structure in the lobbying network beyond that identified by the biLCM.

Not all regions in the LSM's spatial representation are characterized by domain-specific political interests, however. We find that a region in the center of the latent space (marked by the dashed oval in Figure 5) is populated by politicians who are lobbied by groups with diverse interests. While in a direct examination of the LSM this region appears to be unstructured and difficult to analyze, we find that it contains precisely those actors of high entropy we identify with the biLCM, who are involved in many different legislation communities. We elaborate on this in Figure 6: the left panel repeats the latent space plot from Figure 5 but now divides the latent space into small hexagonal parcels and shades them according to the mean legislation community membership entropy of actors within them. We observe that actors with more memberships (darker hexagons) are clustered near the center of the latent space. We give a schematic visual representation of this in the right panel, representing each actor with a pie chart as in Figure 3. We find that actors in the LSM center region tend to have memberships in many different political communities. We highlight three examples of such actors that we identified in Figure 3: the CHAMBER OF COMMERCE, BRITISH PETROLEUM, and MCAFEE, INC. In addition, the actors at the edges of the latent space are precisely those that have focused community membership in the biLCM.

The LSM's failure to meaningfully describe the actors in the center region is related to its confounding *strong ties* with *stochastic equivalence*, as studied by Minhas, Hoff, and Ward (2019). Simply put, two actors in the LSM may be nearby either because they themselves interact, or because they interact with others in a similar way. Like the AMEN model of that work, the biLCM mitigates this effect: repeated interactions in the biLCM occur between actors that share one dominant community membership; two actors with similar mixed memberships interact less often.

More nuanced results in the LSM are captured by the biLCM as well. For instance, in the LSM,



Figure 6: Latent Space Position vs. Legislation Community Distribution. The left panel divides the LSM latent space into hexagonal regions (keeping only those containing at least one actor). We shade each region by the average entropy of legislation community memberships of the actors in that region: the darker the color, the more community memberships the actors have. The right panel plots example legislation community distributions at their corresponding latent positions. Three examples discussed earlier are highlighted and labeled for reference.

the geometry of the clusters is consistent with common intuitions about relationships among industries. For example, the "Technology" cluster is adjacent to both the "Telecommunications" and "Energy" clusters, while the "Universities" cluster is adjacent to both the "Industrial Research & Defense" and "Healthcare" clusters. In Figure 6, actors near these border regions tend to have multiple legislation community memberships in biLCM communities. In summary, the results of the biLCM improve on those of the LSM: the biLCM captures the same single community memberships that the LSM does, but also clarifies the regions between communities and in the center of the latent space that represent varied interests of interest groups and politicians.

3.3.2 Stochastic Block Models

In this section, we compare the results of the biLCM to the biSBM of Larremore, Clauset, and Jacobs (2014). Unlike the LSM, the biSBM explicitly models interactions as a function of community memberships; however, it lacks the mixed memberships and mixed interactions of the biLCM. The Model As before, we model $A_{i,j}$ as having independent Poisson distributions whose means now depend exclusively on the community memberships of interest group i and politician j, denoted $x_i \in \{1, \ldots, k\}$ and $y_j \in \{1, \ldots, \ell\}$ respectively. Interactions between communities are described by a matrix $\mathbf{B} \in \mathbf{R}^{k \times \ell}$. We again include popularity factors α_i and β_j for each interest group and politician, respectively.¹⁵ The joint distribution of the interaction matrix is then

$$\mathbf{P}(\mathbf{A} \mid \mathbf{B}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{i=1}^{m} \prod_{j=1}^{n} \operatorname{Poisson}(A_{i,j} \mid \alpha_{i}\beta_{j}B_{x_{i},y_{j}}).$$
(7)

The model described by Equation (7) is called the *Degree-Corrected Bipartite Stochastic Block Model* (dc-biSBM) due to the presence of popularity factors.

Computation We perform maximum likelihood estimation for the dc-biSBM using the EM algorithm implementation provided by Larremore, Clauset, and Jacobs (2014). As with the biLCM, this algorithm is not guaranteed to converge to the parameters maximizing the likelihood, so we take the parameters attaining the highest likelihood value out of 50 randomly initialized runs.

Comparison The dc-biSBM suffers from the same drawbacks as the LSM, identifying single community memberships in the lobbying network but not the multiple community memberships that reflect the actors' actual varied interests. We illustrate this by comparing the LSM results from the previous section with the dc-biSBM membership results (with $k = \ell = 8$ for comparison with the biLCM).

The left panel of Figure 7 overlays the community memberships identified by the dc-biSBM, marked by the same colors as in the biLCM,¹⁶ on the LSM latent space. We also include the grey community boundaries from Figure 5. The figure shows that the LSM and dc-biSBM community structures are remarkably well-aligned.¹⁷ Specifically, the boundaries that we drew in the LSM

¹⁵See Appendix D.1 for a discussion of parameter identification.

 $^{^{16}}$ As justified by the alignment between the dc-biSBM and the biLCM that is discussed in Appendix E.

¹⁷We combine the dc-biSBM groupings of interest groups and legislators into eight joint communities when we present our findings in this section, as justified by the observation of assortativity in Appendix D.2.



Figure 7: Latent Space Position vs. Community Membership. In the left panel, we plot latent positions from the LSM marked with dc-biSBM communities. Each marker designates one community of politicians and one community of interest groups. The gray regions and the dashed oval are the ex post LSM community boundaries from Figure 5. On the right, we plot regions where pairs of dc-biSBM communities overlap in their latent positions. These lie near the central LSM region, which we mark in each plot for reference. In each plot we include only the actors from the relevant pair of communities, and highlight illustrative examples of interest groups.

almost always correspond to distinct political communities identified by the dc-biSBM.

The dc-biSBM provides some further insights about community structure that were not legible in the LSM. First, the dc-biSBM uncovers two distinct communities near the center of the latent space, which we named "Civil Society" and "Retail & Transportation" in the biLCM. These communities also lie near the center region of the LSM, which we found to contain the higher-entropy actors from the biLCM.

Investigating more carefully, the dc-biSBM gives a sharper suggestion of the importance of mixed membership modeling. The right panels of Figure 7 illustrate this point. We zoom in on different parts of the center region of the LSM, focusing on three pairs of adjacent communities. We find that the communities in each pair overlap in the LSM, and actors in their intersections indeed

tend to lobby on bills related to the subject matters of both communities. We again find our three examples of actors with varied interests, the CHAMBER OF COMMERCE, BRITISH PETROLEUM, and MCAFEE, INC., lying in these intersections; under the dc-biSBM, such actors belong to only one of the several communities in which they interact.

These findings confirm that, while the dc-biSBM is able to discern more structure than the LSM, its assignment of each actor to a single community is limited because many actors participate simultaneously in multiple communities in the lobbying network. The dc-biSBM thus furnishes yet more evidence that mixed membership modeling is required in order to account fully for the organization of the lobbying network.

4 Concluding Remarks

Special interest groups influence the U.S. legislative process by lobbying for and against bills. Yet observable connections between interest groups and legislators have proved elusive, since groups need not reveal the individual politicians they contact. We therefore assemble a new *lobbying database* that connects interest groups to legislators via the bills that were lobbied, information the groups must disclose. From this, we distill a network of interactions between interest groups and politicians in the 113th Congress.

While this network can be partially modeled with both latent space and stochastic block models, its key properties are only observable through our novel *bipartite link community model*. With our model, we find that lobbying interactions occur in domain-specific communities organized around industries and political issues such as energy and veterans' affairs. Interest groups with concerns in these domains lobby politicians who sit on committees with relevant jurisdiction, such as the HOUSE COMMITTEE ON ENERGY AND COMMERCE and the HOUSE COMMITTEE ON VETERANS' AFFAIRS. Furthermore, politicians who serve on power committees with broad responsibilities are members of multiple communities and interact with aggregate interest groups that represent heterogeneous interests. Our findings also suggest that alliances rarely broaden interest groups' activity, and then only on social issues. It is individual interests, rather than the interests of allies, that determine the purview of most lobbying.

Our work provides a novel and general methodology for analyzing complex political networks that have more than one type of actor who may interact with others in more than one way. Bipartite networks, formed by interactions between actors in two distinct political groups, are prevalent in politics: they are formed by interest groups and politicians, politicians and pieces of legislation, voters with different party affiliations, and developing and developed countries, to name a few examples. We believe applying the methodology in this paper to other political settings will yield new substantive insights and will suggest further refinements to models suitable for political networks, both to heterogeneous networks beyond the bipartite case and to more complex types of political interactions.

References

- Airoldi, Edoardo M., David M. Blei, Stephen E. Feinberg, and Eric P. Xing. 2008. "Mixed Membership Stochastic Blockmodels." Journal of Machine Learning Research 9 (Sep): 1981– 2014.
- Austen-Smith, David, and John R Wright. 1992. "Competitive Lobbying for a Legislator's Vote." Social Choice and Welfare 9 (3): 229–257.
- Ball, Brian, Brian Karrer, and M.E.J. Newman. 2011. "An efficient and principled method for detecting communities in networks." *Physical Review E* 84.
- Barberá, Pablo. 2014. "Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data." *Political Analysis* 23 (1).
- Bauer, Raymond, Lewis Anthony Dexter, and Ithiel de Sola Pool. 1972. American Business and Public Policy: The Politics of Foreign Trade. New York, NY: Aldine.
- Baumgartner, Frank R, and Beth L Leech. 1998. Basic Interests: The Importance of Groups in Politics and in Political science. Princeton, NJ: Princeton University Press.
- Bond, Robert, and Solomon Messing. 2015. "Quantifying Social Medias Political Space: Estimating Ideology from Publicly Revealed Preferences on Facebook." American Political Science Review 109 (1): 62–78.
- Bonica, Adam. 2013. "Mapping the Ideological Marketplace." American Journal of Political Science 58 (2): 367–386.
- Box-Steffensmeier, Janet M., and Dino P. Christenson. 2014. "The Evolution and Formation of Amicus Curiae Networks." *Social Networks* 36: 82–96.
- Box-Steffensmeier, Janet M., Dino P. Christenson, and Matthew P. Hitt. 2013. "Quality over Quantity: Amici Influence and Judicial Decision Making." *American Political Science Review* 107 (3): 446–460.
- Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan: A Probabilistic Programming Language." Journal of Statistical Software 76 (1).
- Clark, Tom S., and Benjamin Lauderdale. 2010. "Locating Supreme Court Opinions in Doctrine Space." American Journal of Political Science 54 (4): 871–890.
- Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. 1977. "Maximum Likelihood from Incomplete Data Via the EM Algorithm (with Discussion)." Journal of the Royal Statistical Society, Series B, Methodological 39 (1): 1–37.
- Desmarais, Bruce A, Raymond J La Raja, and Michael S Kowal. 2015. "The Fates of Challengers in US House Elections: The Role of Extended Party Networks in Supporting Candidates and Shaping Electoral Outcomes." *American Journal of Political Science* 59 (1): 194–211.
- Fenno, Richard F. 1966. The power of the purse: Appropriations politics in Congress. Boston, MA: Little and Brown.
- Fenno, Richard F. 1973. Congressmen in Committees. Boston, MA: Little and Brown.

- Fienberg, Stephen E., and Stanley S. Wasserman. 1981. "Categorical data analysis of single sociometric relations." Sociological methodology 12: 156–192.
- Gray, Virginia, and David Lowery. 2000. The Population Ecology of Interest Representation: Lobbying Communities in the American States. Ann Arbor, MI: University of Michigan Press.
- Grossman, Gene M., and Elhanan Helpman. 2001. Special Interest Politics. Cambridge, MA: MIT Press.
- Hadden, Jennifer. 2015. Networks in Contention. Cambridge, UK: Cambridge University Press.
- Hafner-Burton, Emilie M., Miles Kahler, and Alexander H. Montgomery. 2009. "Network Analysis for International Relations." *International Organization* 63 (3): 559–592.
- Hall, Richard L., and Alan V. Deardorff. 2006. "Lobbying as legislative subsidy." American Political Science Review 100 (1): 69–84.
- Heaney, Michael T. 2004. "Outside the Issue Niche: The Multidimensionality of Interest Group Identity." American Politics Research 32 (6): 611–651.
- Heaney, Michael T., and James M. Strickland. 2016. "A Network Approach to Interest Group Politics." In *The Oxford Handbook of Political Networks*. Oxford, UK: Oxford University Press.
- Heinz, John P., Edward O. Laumann, Robert L. Nelson, and Robert H. Salisbury. 1993. The hollow core: Private interests in national policy making. Cambridge, MA: Harvard University Press.
- Hertel-Fernandez, Alexander. 2014. "Who passes businesss model bills? Policy capacity and corporate influence in US state politics." *Perspectives on Politics* 12 (3): 582–602.
- Hoff, Peter D, Adrian E Raftery, and Mark S Handcock. 2002. "Latent Space Approaches to Social Network Analysis." Journal of the american Statistical association 97 (460): 1090–1098.
- Hojnacki, Marie. 1997. "Interest Groups' Decisions to Join Alliances or Work Alone." American Journal of Political Science 41 (1): 61–87.
- Keck, Margaret E., and Kathryn Sikkink. 1998. Activists Beyond Borders: Advocacy Networks in International Politics. Ithaca, NY: Cornell University Press.
- Kim, In Song. 2017. "Political Cleavages within Industry: Firm-level Lobbying for Trade Liberalization." American Political Science Review 111 (1): 1–20.
- Krivitsky, Pavel N., Mark S. Handcock, Adrian E. Raftery, and Peter D. Hoff. 2009. "Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models." *Social Networks* 31: 204–213.
- Larremore, Daniel B., Aaron Clauset, and Abigail Z. Jacobs. 2014. "Efficiently Inferring Community Structure in Bipartite Networks." *Physical Review E* 90.
- Lauderdale, Benjamin E., and Tom S. Clark. 2014. "Scaling Politically Meaningful Dimensions Using Texts and Votes." American Journal of Political Science 58 (3): 754–771.
- Li, Zhenping, Shihua Zhang, and Xiangsun Zhang. 2015. "Mathematical Model and Algorithm for Link Community Detection in Bipartite Networks." 5: 421-434.

- Minhas, Shahryar, Peter D. Hoff, and Michael D. Ward. 2019. "Inferential Approaches for Network Analysis: AMEN for Latent Factor Models." *Political Analysis* 27 (2): 208–222.
- Nourse, Victoria, and Jane S. Schacter. 2002. "The Politics of Legislative Drafting: A Congressional Case Study." New York University Law Review 77: 575-623.
- Porter, Mason A, Peter J Mucha, Mark EJ Newman, and Casey M Warmbrand. 2005. "A Network Analysis of Committees in the US House of Representatives." Proceedings of the National Academy of Sciences 102 (20): 7057–7062.
- Potters, Jan, and Frans Van Winden. 1992. "Lobbying and Asymmetric Information." *Public choice* 74 (3): 269–292.
- Slapin, Jonathan B., and Sven-Oliver Proksch. 2008. "A Scaling Model for Estimating Time-series Party Positions from Texts." American Journal of Political Science 52 (3): 705–722.
- Ward, Michael D., Katherine Stovel, and Audrey Sacks. 2011. "Network Analysis and Political Science." Annual Review of Political Science 14: 245–264.
- Wright, John R. 1990. "Contributions, Lobbying, and Committee Voting in the US House of Representatives." American Political Science Review 84 (02): 417–438.
- Wright, John R. 1996. Interest Groups and Congress. Boston, MA: Allyn and Bacon.
- Zhang, Yan, Andrew J. Friend, Amanda L. Traud, Mason A. Porter, James H. Fowler, and Peter J. Mucha. 2008. "Community Structure in Congressional Cosponsorship Networks." *Physica A: Statistical Mechanics and its Applications* 387 (7): 1705–1712.