

Quantifying Political Relationships

Simon Weschle*
Syracuse University

Conditionally accepted at *American Political Science Review*

June 9, 2018

Abstract

In this article, I introduce a method that uses large-scale event data and latent factor network models to provide a new comparative measure of cooperation and conflict in public relationships among politicians, non-partisan political actors, and societal actors. The approach has a number of advantages over existing techniques: It captures public relationships in a multitude of venues on a continuous basis, incorporates both partisan and non-partisan actors, allows quantifying the relationship between any pair of actors, reflects that cooperation is not unidirectional but rather a back and forth, and can be applied to a large number of countries over time. I apply the method to 13 Western European countries from 2001 to 2014 and demonstrate that party relationships are determined by coalition status as well as policy differences. The measure is publicly available and can be incorporated into standard research designs.

*Simon Weschle is an Assistant Professor, Department of Political Science, Maxwell School of Citizenship and Public Affairs, Syracuse University, 100 Eggers Hall, Syracuse, NY 13244 (sweschl@maxwell.syr.edu). For helpful comments and advice, I thank the APSR editor Kenneth Benoit, four anonymous reviewers, James Adams, Ben Barber, Pablo Fernández-Vázquez, Sebastián Lavezzi, Michael Ward, and Christopher Wlezien.

There is a large discrepancy between the importance attributed to the public communication of political actors and the (lack of) empirical analysis devoted to it. The way political elites interact with each other as well as with members of society may shape whether voters like or dislike them, how citizens perceive policy positions, or the degree to which they feel represented in the political system (Fenno, 1978; Mansbridge, 2003; Strom, 2008). However, a lack of data has hampered the empirical study of the causes and consequences of public communication involving political elites. This has begun to change recently due to advances in automated text analysis methods. The focus so far has been on examining parliamentary speeches (e.g. Martin and Vanberg, 2008; Eggers and Spirling, 2014; Herzog and Benoit, 2015; Proksch and Slapin, 2015; Lauderdale and Herzog, 2016) and politicians' press releases (e.g. Grimmer, 2013; Grimmer, Westwood and Messing, 2015; Sagarzazu and Klüver, 2017). These approaches have yielded valuable insights, but have focused exclusively on the unidirectional communication *by* politicians *to* constituents in a single venue, and typically in a single country.

In this article, I introduce a method that uses large-scale event data and latent factor network models to provide a new comparative measure of the public relationships among politicians, non-partisan political actors, and societal actors. Instead of relying on source data from a single setting, I use an extensive collection of machine-coded news reports that provide tens of thousands of interactions from many venues among hundreds of politicians, non-partisan political actors (e.g. bureaucracy, police, or judiciary), and societal actors (e.g. citizens, unions, companies, religious groups). I then infer the underlying networks that give rise to these interactions by estimating latent factor network models (Hoff and Ward, 2004; Hoff, 2005, 2015). They place all actors in a low-dimensional “social space”, where those who have a cooperative relationship are placed in the same direction, while those who have a conflictual relationship are placed in different directions. From the latent positions, it is possible to compute scores that reflect the relationship of any pair of actors in the data.

This approach to quantify political relationships has a number of advantages over existing techniques. First, it is not restricted to a single venue, but captures interactions in many of them: press releases, parliamentary speeches, interviews, campaign events, and so on. Second, instead of examining only one-sided communication *by* politicians, it also incorporates information on communication from and to non-partisan political and societal actors. It also makes it possible to locate political and societal actors in a *common* space. Importantly, because the latent factor model takes third-order network dependencies (e.g. a friend of a friend is a friend) into account, the relation between two actors can be inferred *even if* no direct interaction between them is reported. Finally, the approach can be estimated for a large number of countries over many years.

The new measure opens up the possibility to study the causes and consequences of cooperative or conflictual relationships between partisan political, non-partisan political, and societal actors in a comparative manner. The relationship scores are publicly available and can be used by researchers to answer questions in areas such as coalition politics, polarization, or democratic representation and satisfaction. In this article, I describe the details of the approach and demonstrate its value by applying it to 13 Western European countries from 2001 to 2014.

Approach to Quantify Political Relationships

Data

The event data of interactions among political and societal elites I use stem from the ICEWS project (Boschee et al., 2015). ICEWS is an early warning system that was designed in conjunction with several academic research teams to help U.S. policy analysts predict violent and non-violent political crises (for a detailed introduction see O'Brien, 2013). A central idea of the project is that the tone of interactions between the socio-political actors of a country

can help predict such crises. It has therefore developed an extensive event collection that documents the activities of countries' political and societal elites as comprehensively as possible.

The ICEWS project takes a large collection of news reports and machine-codes them into dyadic events reporting the event source, target, and type. The source material stems from the media repositories of the *Open Source Center* and *Factiva*, which collect international and national news reports from a large number of publications. The first six sentences of each report are coded by BBN ACCENT, a natural language analysis system. It employs a number of linguistic models that were trained using a sample corpus to extract structured information from text (for details see Boschee et al., 2015). Consider the following sentence from a 2008 report from Germany: "Economics Minister Michael Glos, of the government's conservative CDU/CSU coalition partner, attacked a draft proposal from Justice Minister Brigitte Zypries."¹ This is machine-coded into an event described by three variables: The *source* is Michael Glos, Brigitte Zypries is the *target*, and the *type* is "criticize or denounce." For the type, the roughly 350 category coding scheme developed by the Conflict and Mediation Event Observation (CAMEO) project is used (Gerner, Schrodt and Yilmaz, 2009).

The goal of the ICEWS event data is to chronicle the activities of countries' main socio-political actors. Events are therefore extensively screened and filtered to exclude historical events, those unrelated to socio-political activities (e.g. sports or entertainment), and exact duplicates. Validation studies find that the machine-coded information triplets were judged to be correct in around 75 percent of cases (Boschee, Natarjan and Weischedel, 2013; Boschee et al., 2015), exceeding the performance that is typically achieved by human coders (King and Lowe, 2003). The event collection has been made publicly available (Boschee et al., 2015). In the Online Appendix, I provide extensive further information on the data, including technical

¹ <http://reuters.com/article/2008/03/05/autoshow-porsche-idINL0575794620080305>, accessed January 5, 2017.

details on the coding algorithm, descriptive statistics, a list of media sources, frequency tables of actors, and a discussion of potential objections and limitations.

I further process the event data in two ways. First, I hand-code every actor as belonging into one of three categories: partisan political, non-partisan political (e.g. bureaucracy, police, judiciary), and societal (e.g. citizens, corporations, unions, religious groups). For the first category, I code actors' partisan affiliations, taking party switches into account. I also assign partisan affiliations to institutional actors (e.g. head of state, ministry of defense, ruling party) when they can be clearly inferred. Second, I dichotomize the event type categories into cooperative and conflictual. For example, "make optimistic comment" is a cooperative category, whereas "bring lawsuit against" is a conflictual one.² This dichotomization further increases coding accuracy. If, say, the example sentence above had been misclassified as "use unconventional mass violence" rather than as "criticize or denounce", it would still correctly be considered a conflictual interaction.

Relationship Scores

The next step is to summarize this wealth of information on the relations among political and societal actors in a way that can be easily interpreted and is amenable to further quantitative analysis. The key insight is that the interactions arise from *networks* of relations among socio-political actors. My approach for a measure of political relationships is to infer the structure of those networks using a latent factor approach (Hoff and Ward, 2004; Hoff, 2005, 2015)

Denote the number of cooperative interactions between i and j by m_{ij}^+ , and the number of conflictual interactions with m_{ij}^- . The interactions between the two actors can be summarized

² See the Online Appendix for the full list of cooperative and conflictual categories.

by $y_{ij} = y_{ji} = \ln \left(\frac{m_{ij}^+ + 1}{m_{ij}^- + 1} \right)$.³ These are then aggregated into an $n \times n$ sociomatrix \mathbf{Y} , the cells of which are modeled as follows:

$$\begin{aligned} y_{ij} &= \alpha + a_i + a_j + \epsilon_{ij} + \mathbf{u}'_i \boldsymbol{\Lambda} \mathbf{u}_j \\ a_1, \dots, a_n &\sim \text{i.i.d. } \mathcal{N}(0, \sigma_a^2) \\ \{\epsilon_{ij}\} &\sim \text{i.i.d. } \mathcal{N}(0, \sigma_e^2) \end{aligned} \tag{1}$$

The latent factor approach decomposes the dependent variable into several components. The intercept is designated by α . Overall differences in the tone of interactions by actors are captured by the random effects a_i and a_j . These coefficients are larger for actors whose interactions are more cooperative in general. The random effects term ϵ_{ij} captures the correlation of actions between a dyadic pair of actors (e.g. reciprocity). Finally, the remaining variance in y_{ij} is absorbed by $\mathbf{u}'_i \boldsymbol{\Lambda} \mathbf{u}_j$, the multiplicative effects term that captures latent nodal characteristics.⁴

The basic idea of the approach is to represent the network by giving actors positions an unobserved latent space. The K -dimensional vector \mathbf{u}_i characterizes i 's position in that network, and \mathbf{u}_j gives the same for j . $\boldsymbol{\Lambda}$ is a $K \times K$ diagonal matrix of (positive) scaling constants (see Hoff, 2015).⁵ The multiplicative effects term $\mathbf{u}'_i \boldsymbol{\Lambda} \mathbf{u}_j$ then represents the nature of the relationship between i and j . It places two actors that tend to be cooperative with each other, or that interact in a similar way with third actors, in the same direction in the latent social space. Actors that are likely to be in conflict with each other, or that interact in different ways with third actors, are placed in opposing directions. The term $\mathbf{u}'_i \boldsymbol{\Lambda} \mathbf{u}_j$ does not only reflect direct interactions between two actors, it also takes into account their relation through others. For example, if two politicians both have cooperative interactions

³ The ratio is logged to make it linear, so e.g. $\ln(7/10)$ and $\ln(10/7)$ have the same distance to neutral.

⁴ The terms in $\mathbf{u}'_i \boldsymbol{\Lambda} \mathbf{u}_j$ are estimated using an eigenvalue decomposition. See Online Appendix for details.

⁵ The values of $\boldsymbol{\Lambda}$ are negative in the (unlikely) case that nodes with similar characteristics are more likely to have conflictual interactions. See Online Appendix for details.

with unions, but conflictual ones with business representatives, they likely have a cooperative relationship with each other so they are placed in the same direction, even if they did not interact directly with each other.

To understand the multiplicative effects term intuitively, suppose $K = 1$, so \mathbf{u}_i and Λ are simply scalars u_i and λ . Recall that y_{ij} is positive when interactions are cooperative, and negative when conflictual interactions dominate. If λ is positive and u_i and u_j are both positive or both negative, then $u_i \lambda u_j$ will be positive, corresponding to a cooperative relationship between i and j . If one of u_i and u_j is negative and one is positive, so the actors are located in different directions in the space, then $u_i \lambda u_j$ is negative. Since this is not only done for i and j , but also for i and k , j and k , and so on, two actors will be located in the same direction if they have mostly cooperative interactions with each other (net of their overall level of cooperation) or if they are connected through mutual cooperation with third actors. The latent factors thus summarize the complex network of relations into a low-dimensional representation that is easily interpretable and can be used for quantitative analysis. In particular, $\mathbf{u}'_i \Lambda \mathbf{u}_j$ provides an estimated relationship score between i and j based on their positions in the latent communication network.

Focusing on the overall relationship $\mathbf{u}'_i \Lambda \mathbf{u}_j$ rather than simply on y_{ij} , the balance of the dyadic direct interactions, has a number of advantages. Equation (1) filters out actor-specific idiosyncrasies. Some actors tend to be more conflictual or more cooperative in general, no matter who they are interacting with, which would contaminate a simple count or ratio measure. The latent factor approach also partitions out the correlation of actions between a dyadic pair of actors, such as reciprocity. Furthermore, relationships are not only determined by direct interactions, but through interactions with others as well. Using $\mathbf{u}'_i \Lambda \mathbf{u}_j$ has the advantage of taking these third-party communications into account. This makes it possible to analyze the relationship between two actors even if no direct interaction between them is observed. For example, political representation does not require direct communication, so

we can infer how well a party i represents a societal actor j , no matter whether they directly interact or not (see Weschle, 2017). On the practical side, this advantage is particularly relevant if the data are only a partial observation of the network. While the ICEWS event data is the largest and most detailed source there is, it nevertheless can only capture a sample of the universe of interactions.

Applications

I demonstrate the value of my approach to quantify political relationships by applying it to 13 Western European countries for the years 2001 to 2014.⁶ There are a total of 250,316 domestic events involving 3,422 actors reported by 232 media sources. I estimate a separate one-dimensional ($K = 1$) latent factor model for each country-year, which allows me to compute yearly relationships scores for any dyadic pair of actors that are present in the data.⁷ The applications here are focused on the relationships among political parties, but the data can be used to analyze non-partisan political and societal actors as well.

Descriptive Example: Germany

To show what the relationships scores derived from the latent factor models look like in practice, Figure 1 plots them for Germany from 2001 to 2009. In this case I have aggregated all partisan actors in the data, so parties are unitary actors.⁸ Scores involving two parties are displayed by black dots, dark gray dots represent cooperation scores between one partisan and one non-partisan actor, and light gray is two non-partisan actors.

⁶ Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal, Spain, United Kingdom.

⁷ See Online Appendix for details on the Bayesian estimation.

⁸ See Online Appendix for a list of parties. Relationships scores in which political actors are not aggregated are available as well.

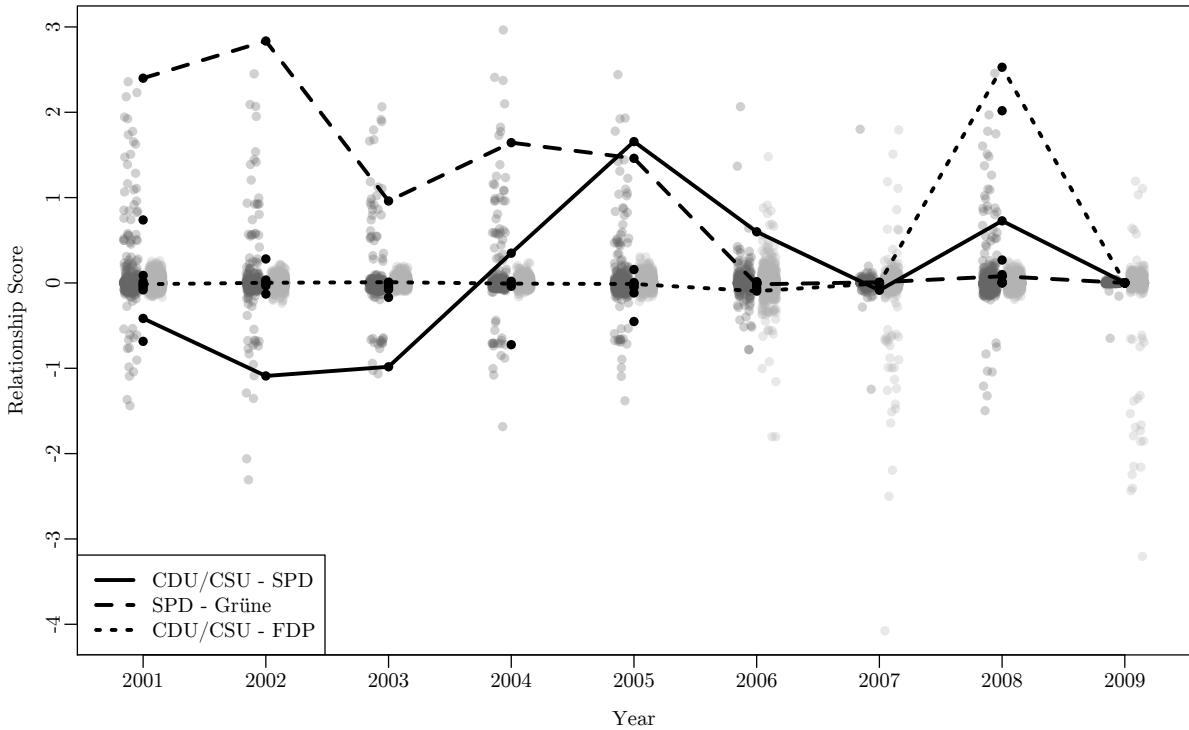


Figure 1: Relationship Scores, Germany, 2001-2009. Black: two parties; Dark gray: one party, one non-partisan actor; Light gray: two non-partisan actors.

To highlight the face validity of the scores, I focus on three party-dyads over time. SPD and Grüne formed a coalition government until 2005. Their relationships scores during these years are consistently large, indicating a positive relationship both in direct interactions as well as through third actors. After 2005, their scores go towards the neutral point. The most conflictual party-dyad relation during the SPD-Grüne coalition is between the SPD and the largest opposition party, the CDU/CSU. This changes once these two parties form a “grand coalition” in 2005, when they become the party dyad with the highest scores. Their relationship deteriorates over time, and the CDU/CSU becomes friendlier with the FDP, foreshadowing the coalition they form after the 2009 elections. However, this is not reflected in the scores for 2009, the year in which the Great Recession dominated the agenda and all party dyads exhibit neutral relationships. Notice that this is also true for the scores involving

one party and one non-partisan actor, while relations for non-partisan dyads become more polarized (see Weschle, 2017).

Public Relationships among Political Parties

Having shown the face validity of the measure, I now demonstrate how it makes a systematic comparative analysis of public political relationships possible. Specifically, I focus on the subset of relations between political parties. What determines whether they have cooperative or conflictual relationships?

Figure 1 suggests that coalition status matters. We would expect parties in government to work together to achieve common goals, and to have similar interactions with other actors in society. Opposition parties, in contrast, are supposed to hold the government accountable, which means direct criticism and cooperation with societal actors opposed to the government. The expectation is therefore that parties have a more cooperative relationship when they are in a coalition.

Second, given the central role that policy plays in structuring political competition in European countries, one can expect that it affects parties' public relationships as well. If two parties advocate similar policies, they should interact in a cooperative manner with each other, and interact similarly with third actors. For example, two economically conservative parties are both likely to have cooperative relations with business and conflictual ones with unions.

To test these conjectures, I estimate a set of models in which the dyadic relationship scores involving two parties are the dependent variable. I regress this on a dummy variable indicating whether two parties are in a coalition together, as well as a dummy for opposition-opposition dyads (making government-opposition the baseline). Regarding the effect of policy, I use the Comparative Manifestos Project (CMP) data and compute the absolute distance between the positions of two parties. In addition to the left-right difference,

I also include a measure of parties' difference on issues of nationalism and multiculturalism as a second policy dimension in European party competition.⁹

Table 1: Determinants of Party Relationship Scores. 95 percent confidence intervals in parentheses.

	(1)	(2)	(3)
Coalition	0.127 (-0.018, 0.272)	0.152 (0.005, 0.299)	0.179 (0.021, 0.336)
Opposition	-0.041 (-0.097, 0.015)	-0.061 (-0.133, 0.011)	-0.065 (-0.154, 0.024)
$ \Delta \text{CMP Left-Right} $	0.001 (-0.001, 0.002)	0.001 (-0.001, 0.003)	0.001 (-0.001, 0.003)
$ \Delta \text{CMP Nationalism/Multiculturalism} $	-0.006 (-0.013, 0.001)	-0.008 (-0.015, -0.001)	-0.011 (-0.022, -0.001)
Control $ \Delta \text{Vote Shares} $	✓	✓	✓
Country-Year Controls	✓	✓	
Country and Year FE		✓	
Country-Year FE			✓
N	1150	1150	1150
R^2	0.060	0.090	0.216

Table 1 shows the results of three specifications. The pooled model (1) and the one with country and year fixed effects (2) include the mean relationship score for each country-year to addresses potential concerns about the comparability of the separately estimated latent spaces. In addition, I add a set of controls at the country-year level (election year, number of parties, log population, number of events). Specification (3) includes country-year fixed effects and thus only uses variation from within each latent space, which ensures comparability. All three specifications control for the difference in vote shares between the parties. I take the uncertainty of the estimated relationship scores into account by deriving the dependent variable separately for all posterior draws of a latent factor model, running the regression and simulating coefficient draws for each (standard errors clustered at the

⁹ Nationalism and multiculturalism position: $(\text{per601} + \text{per607}) - (\text{per602} + \text{per608})$. Since CMP positions can only be measured in election years, I use the most recent available CMP position.

dyad-level), combining the results from these estimations, and calculating the confidence intervals.

The specifications provide evidence supporting both conjectures. Depending on the specification, a coalition partnership is associated with an increase in the relationship score by 0.30-0.42 standard deviations, compared to the government-opposition baseline. If both parties are in the opposition, the relationship score is somewhat lower than in the baseline, although the confidence intervals include zero.

Policy distance also has an effect on what kind of relationship parties have. Interestingly, it is *not* the difference in left-right positions that affects the relationship scores. Instead, parties with a greater difference in their positions towards nationalism/multiculturalism have lower scores. A one standard deviation increase in this independent variable is associated with a decrease in the dependent variable by 0.06-0.10 standard deviations. The public relationships of political parties are therefore not as much driven by whether they are left-wing or right-wing, but instead by the cleavage between “mainstream” and other parties about nationalism and multiculturalism. Thus, a systematic analysis shows that the public relationships among political parties in Western Europe follow predictable patterns.

Using Quantified Political Relationships in Comparative Politics

The public relationships of political actors are important, but a lack of comparative data has largely prevented their empirical study. Using a combination of machine-coded news reports and latent factor models, I have introduced a way to quantify public relationships among political and societal actors. The scores for 13 Western European countries from 2001 to 2014 involving almost 3500 actors are publicly available and can be incorporated by researchers into their analyses in a straightforward way.

Of course, there are some limitations. First, the relationship scores are based on media reports, which tend to focus on high-level political and social actors. The measure does not reflect the actions of e.g. backbencher MPs or societal actors at the local level. A second limitation is that the scores are based on public interactions, which may be different from private ones. Which interactions happen in public may also differ between countries depending on their laws, culture, or institutions. It is therefore important to address potential comparability issues in any analysis, e.g. through country or country-year fixed effects. Third, the measure in its current form treats relationships as symmetric. It is possible to extend the latent factor approach to differentiate between sender and receiver behavior.¹⁰ Finally, I treat all interactions as equal, whereas it is likely that some of them are more important for a relationship than others. However, this concern should be alleviated by the quantity of interactions I am able to use.

Keeping these limitations in mind, the quantified political relationships data opens the door for a host of new research. For example, it provides new data to study elite polarization, and the relationship scores between politicians and societal actors should be of interest for students of political representation. The potential of the new data is perhaps even stronger as an explanatory variable. Public relationships are likely to have an effect on voters, for example on how they perceive parties' policy positions, how polarized they are, or what their level of satisfaction with democracy is.

In the past years, scholars have increasingly made use of novel, large-scale data to address long-standing questions in comparative politics. The approach and data introduced in this article contributes to this trend and makes it possible to re-visit long-standing questions on the causes and consequences of political relationships.

¹⁰ See Online Appendix.

References

- Boschee, Elizabeth, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz and Michael D. Ward. 2015. "ICEWS Coded Event Data." <http://dx.doi.org/10.7910/DVN/28075>.
- Boschee, Elizabeth, Premkumar Natarjan and Ralph Weischedel. 2013. Automatic Extraction of Events from Open Source Text for Predictive Forecasting. In *Handbook of Computational Approaches to Counterterrorism*, ed. V.S. Subrahmanian. New York: Springer.
- Eggers, Andrew C and Arthur Spirling. 2014. "Ministerial Responsiveness in Westminster Systems: Institutional Choices and House of Commons Debate, 1832–1915." *American Journal of Political Science* 58(4):873–887.
- Fenno, Richard F., Jr. 1978. *Home Style: House Members in their Districts*. Boston: Little, Brown.
- Gerner, Deborah J., Philip A. Schrodt and Omur Yilmaz. 2009. Conflict and Mediation Event Observations (CAMEO): An Event Data Framework for a Post-Cold War World. In *International Conflict Mediation: New Approaches and Findings*, ed. Jacob Bercovitch and Scott Sigmund Gartner. New York: Routledge.
- Grimmer, Justin. 2013. *Representational Style in Congress. What Legislators Say and Why It Matters*. Cambridge: Cambridge University Press.
- Grimmer, Justin, Sean J. Westwood and Solomon Messing. 2015. *The Impression of Influence: Legislator Communication, Representation, and Democratic Accountability*. Princeton: Princeton University Press.
- Herzog, Alexander and Kenneth Benoit. 2015. "The Most Unkindest Cuts: Speaker Selection and Expressed Government Dissent during Economic Crisis." *Journal of Politics* 77(4):1157–1175.
- Hoff, Peter D. 2005. "Bilinear Mixed Effects Models for Dyadic Data." *Journal of the American Statistical Association* 100(469):286–295.
- Hoff, Peter D. 2015. "Dyadic Data Analysis with amen." <http://arxiv.org/abs/1506.08237>.
- Hoff, Peter D. and Michael D. Ward. 2004. "Modeling Dependencies in International Relations Networks." *Political Analysis* 12(2):160–175.
- King, Gary and Will Lowe. 2003. "An Automated Information Extraction Tool for International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design." *International Organization* 57:617–642.

- Lauderdale, Benjamin E. and Alexander Herzog. 2016. “Measuring Political Positions from Legislative Speech.” *Political Analysis* 24(3):374–394.
- Mansbridge, Jane. 2003. “Rethinking Representation.” *American Political Science Review* 97(4):515–528.
- Martin, Lanny W. and Georg Vanberg. 2008. “Coalition Government and Political Communication.” *Political Research Quarterly* 61(3):502–516.
- O’Brien, Sean P. 2013. A Multi-Method Approach for Near Real Time Conflict and Crisis Early Warning. In *Handbook of Computational Approaches to Counterterrorism*, ed. V.S. Subrahmanian. New York: Springer.
- Proksch, Sven-Oliver and Jonathan B. Slapin. 2015. *The Politics of Institutional Debate*. Cambridge: Cambridge University Press.
- Sagarzazu, Inaki and Heike Klüver. 2017. “Coalition Governments and Party Competition: Political Communication Strategies of Coalition Parties.” *Political Science Research and Methods* 5(2):333–349.
- Strom, Kaare. 2008. “Communication and the Life Cycle of Parliamentary Democracy.” *Political Research Quarterly* 61(3):537–542.
- Weschle, Simon. 2017. “The Impact of Economic Crises on Political Representation in Public Communication: Evidence from the Eurozone.” *British Journal of Political Science* <https://doi.org/10.1017/S0007123417000023>.