Quantifying Political Interactions

Simon Weschle*
Carlos III–Juan March Institute
sweschle@inst.uc3m.es

January 16, 2017

Abstract

This paper introduces a new approach to measure cooperation and conflict in public communication among politicians, non-partisan political actors, and societal actors. I use an extensive trove of machine coded news reports and employ latent factor models to create a spatial representation of public elite interactions, which is used to compute scores that quantify the degree of cooperation and conflict for any pair of actors. This approach has a number of advantages over existing techniques: It captures public interactions in a multitude of venues on a continuous basis, locates political and non-political actors in a common space, and reflects that cooperation is not unidirectional but rather a back and forth. To demonstrate the value of the approach, I apply it to 13 Western European countries from 2001 to 2014 and show the influence of coalition status and policy positions on party cooperation.

Word Count: 3900

*For helpful comments and advice, I thank James Adams, Pablo Fernández-Vázquez, Sebastián Lavezzolo, and Michael Ward.

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 other members of society may shape, among others, whether voters like or dislike them, how citizens perceive policy positions, or the degree to which they feel represented in the political system (e.g. Fenno, 1978; Mansbridge, 2003; Strom, 2008). However, a lack of comparative data has hampered the empirical study of the causes and consequences of public communication involving political elites. Due to advances in automated text analysis methods, this has begun to change recently (for an overview see Lucas et al., 2015). The focus so far has been on examining parliamentary speeches (e.g. Martin and Vanberg, 2008; Proksch and Slapin, 2015; Eggers and Spirling, 2014; Herzog and Benoit, 2015) and politicians' press releases (e.g. Grimmer, 2013; Grimmer, Westwood and Messing, 2015; Sagarzazu and Klüver, 2015). 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 an approach that uses large-scale event data and latent factor scaling techniques to provide a new comparative measure of cooperation and conflict in public elite communication among politicians, non-partisan political actors, and societal actors. Instead of relying on data from a single venue, I use an extensive trove of news reports as my source data. The reports are machine coded to provide tens of thousands of interactions among many hundreds of political and societal actors. I then create a *spatial representation* of these public interactions by estimating latent factor models (Hoff and Ward, 2004; Hoff, 2005, 2015; Minhas, Hoff and Ward, 2016). They place all actors in a low-dimensional "social space," where those who are likely to have a cooperative relationship are placed in the same direction, while those who are in conflict with each other are placed in different directions. From the latent positions, it is possible to compute "cooperation scores" for any pair of actors in the data.

This approach to quantify public communication has a number of advantages over existing techniques. 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. Instead of examining only one-sided communication *by* politicians, it also incorporates non-partisan political and societal actors. The data thus reflect the fact that communication is not a one-way street, but that it involves senders and receivers, and that politicians can be both. It also makes it possible to locate political and societal actors in a *common* space. Importantly, because the latent factor approach takes third order 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 measure can be estimated for a large number of countries over many years.

The new measure of public cooperation for the first time makes it possible to study interactions between politicians, non-partisan political actors (such as the bureaucracy, police, or judiciary), and societal actors (like citizens, unions, companies, or religious organizations) in a comparative manner. It can help answer questions in research areas such as coalition politics, polarization, or democratic representation and satisfaction. In this article, I describe the details of the measure and demonstrate its value by applying it to 13 Western European countries for the years 2001 to 2014. I show that the degree of cooperation and conflict between political parties relates to their policy positions and coalition status.

Measuring and Scaling Public Political Interactions

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 a number of academic research teams to help U.S. policy analysts predict violent as well as 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 news reports from a large number of sources at the international and national level. The first six sentences of each news 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 Ramshaw et al., 2011; Boschee et al., 2015). As an example, consider the following sentence from a Reuters story in 2008 about a proposed new law in 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 a number of variables, three of which are relevant here: The *event source* is Michael Glos, Brigitte Zypries is the *event target*, and the *event 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 for the ICEWS Event Data was 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 duplicates. Validation studies find that the machine coded information triplet 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 data have been made publicly available (Boschee et al., 2015) and provide a large amount of information: For the 13 countries I consider in this paper, it

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

records 250,316 domestic events involving 3422 actors reported by 232 media sources. In the Online Appendix, I provide detailed further information on the ICEWS event data, including technical details on the coding algorithm, descriptive statistics, a list of media sources, and frequency tables of actors.

I further process the 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 roughly 350 event type categories into cooperative and conflictual. For example, "make optimistic comment" or "express intent to settle dispute" are cooperative categories, whereas "accuse of aggression" or "bring lawsuit against" are conflictual ones.[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.

Scaling

The next step is to apply a scaling approach that transforms the large quantity of event data into a measure that can easily be interpreted and is amenable to further quantitative analysis. The key idea is to realize that the interactions arise from a network of relations, and I use a latent factor approach to infer its structure (Hoff and Ward, 2004; Hoff, 2005, 2015; Minhas, Hoff and Ward, 2016). Denote the number of cooperative interactions between i and

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

j by m^+_{ij}, and the number of conflictual interactions with m^-_{ij}. The relationship between the two actors can be summarized by $y_{ij} = y_{ji} = \ln\left(\frac{m^+_{ij}+1}{m^-_{ij}+1}\right)$, which is then modeled as follows:

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

The latent factor approach decomposes the dependent variable into several components. The random effects a_i and a_j capture overall differences in the tone of interactions among actors. The 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. Finally, the remaining variance in y_{ij} is absorbed by $\boldsymbol{u}_i' \boldsymbol{\Lambda} \boldsymbol{u}_j$, the multiplicative effects term that captures latent nodal characteristics. The K-dimensional vector $\boldsymbol{u}_i$ summarizes i's behavior in interactions (net of their overall propensity to interact cooperatively or conflictually), and $\boldsymbol{u}_i$ summarizes the same for j. $\boldsymbol{\Lambda}$ is a $K \times K$ diagonal matrix (see Hoff, 2015).

The multiplicative effects term $\boldsymbol{u}_i' \boldsymbol{\Lambda} \boldsymbol{u}_j$ 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 a latent 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 $\boldsymbol{u}_i' \boldsymbol{\Lambda} \boldsymbol{u}_j$ 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 relations with unions, but conflictual ones with business representatives, they are likely to be 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 $\boldsymbol{\Lambda}$ is simply a vector) and recall that y_{ij} is positive when interactions are cooperative, and negative when conflictual interactions dominate. If $\boldsymbol{\Lambda}$ is positive and $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ are both positive or both negative, then $\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j$ will be positive, corresponding to a cooperative relationship between i and j. If one of $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ is negative and one is positive, so the actors are located in different directions in the space, then $\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j$ is negative.[3] 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 (conditional on their overall level of cooperation) or if they are connected through mutual cooperation with third actors. The latent space thus summarizes the complex network of relations into a low-dimensional representation that is easily interpretable and can be used for quantitative analysis. In particular, the multiplicative effect $\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j$ provides an estimated cooperation score between actors i and j, no matter whether we actually observe interactions between them or not.

Quantified Political Interactions

To show what the cooperation scores $\boldsymbol{u}_i'\boldsymbol{\Lambda}\boldsymbol{u}_j$ derived from the latent factor models look like in practice, Figure 1 plots them for Germany from 2001 to 2009, using one-dimensional latent positions.[4] In this case, I have aggregated all partisan actors in the data so parties are unitary actors.[5] Scores involving two parties are written out, dark gray dots represent cooperation scores between one partisan and one non-partisan actor (both non-partisan political and societal), and light gray is two non-partisan actors. The measure shows considerable face validity. From 2001 to 2004, the party dyad that consistently has the largest cooperation scores is SPD-Grüne, which formed the government at the time. This indicates that the coalition partners interacted mostly cooperatively at the time, and that they also tended to interact in similar ways with third actors (e.g. other parties, unions, protesters). The lowest scores are between the SPD and the main opposition party CDU/CSU, as well as the other opposition parties PDS/Linke and FDP.

[3]Again assuming that $\boldsymbol{\Lambda}$ is positive.

[4] See the Online Appendix for estimation details.

[5] See Online Appendix for a list of parties. Cooperation scores in which political actors are not aggregated are available as well.

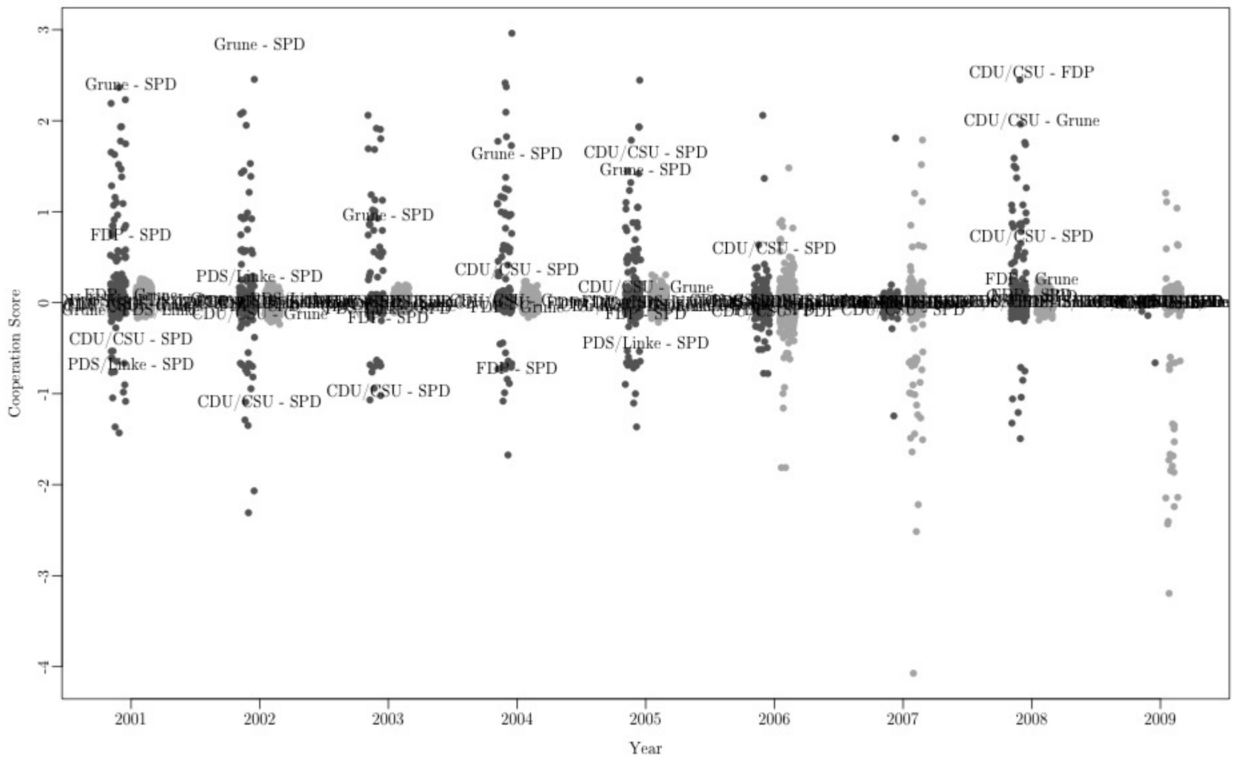

Figure 1: Cooperation Scores, Germany, 2001-2009. Cooperation scores involving two parties written out, scores involving one party in dark gray, and scores only non-partisan actors in light gray.

In 2005, the SPD-Grüne government was replaced by a "grand coalition" between SPD and CDU/CSU, which is reflected by the fact that this becomes the party dyad with the highest scores. Cooperation between the two wanes as 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 cooperation scores for 2009, the year in which the Great Recession dominated the agenda. All party dyads exhibit neutral relationships. Notice that this is also true for the scores of one party and one non-partisan actor, while there is an increase of polarization for non-partisan dyads (for a systematic analysis see Weschle, 2017).

Application: Public Relationships among Political Parties

The approach introduced in this paper relies on news reports as source material. These reports chronicle interactions involving political parties as well as non-partisan societal actors, and the latent factor model locates them in a common space. This means that it is possible to compute and analyze cooperation scores between political parties and societal actors, as well as between political parties. The previous section has shown that in the example of Germany, the cooperation scores between parties make intuitive sense. But what determines them more broadly?

To answer this question, I estimate the cooperation scores for 13 Western European countries annually from 2001 to 2014.[6] When thinking about which party pairs are most likely to directly and indirectly interact in a cooperative manner, two determinants come to mind immediately. The first one is whether parties are in a coalition together. Parties join a governing coalition to achieve a common set of policy goals together. By the same token, one of the major functions of opposition parties is to hold the government accountable and criticize it where necessary. The expectation is therefore that parties are more cooperative with each other when they are in a coalition.

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

To test these conjectures, I estimate a set of models in which the dependent variable are the dyadic cooperation scores involving two parties. I regress this on a dummy variable

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

indicating whether two parties are in a coalition together. To account for potential differences between government-opposition and opposition-opposition dyads, I also add a dummy for the latter. To test the effect of policy, I compute the absolute distance between the left-right positions of the two parties according to the Comparative Manifestos Project (CMP). I also include the absolute distance between the positions on multiculturalism as the second policy dimension in European party competition.[7]

Table 1: Determinants of Party-Party Cooperation Scores. Cooperation scores from 13 Western European countries from 2001 to 2014. 95 percent confidence intervals in parentheses, based on standard errors clustered at the dyad-level.

	(1)	(2)	(3)
Coalition	0.172	0.213	0.251
	(-0.013, 0.357)	(0.019, 0.408)	(0.041, 0.462)
Opposition	-0.030	-0.060	-0.067
	(-0.101, 0.041)	(-0.136, 0.015)	(-0.161, 0.027)
$\lvert\Delta$ CMP Left-Right$\rvert$	0.001	0.001	0.001
	(-0.001, 0.002)	(0.000, 0.003)	(-0.001, 0.003)
$\lvert\Delta$ CMP Multiculturalism$\rvert$	-0.008	-0.008	-0.011
	(-0.015, -0.001)	(-0.015, -0.001)	(-0.022, -0.001)
Country-Year Controls	✓	✓	
Country and Year FE		✓	
Country-Year FE			✓
N	1150	1150	1150
R^2	0.065	0.097	0.226

All models include $\lvert\Delta$ Vote Share$\rvert$ as a control. Country-year controls: Election year, number of parties, population (log), number of events, mean cooperation score. The models take the estimation uncertainty of the cooperation scores into account by deriving the dependent variable separately for all 500 posterior draws of a latent space, running the regressions for each, and then combining the results from these estimations.

Table 1 shows the results of three specifications: a pooled model, one with country and year fixed effects, and one with country-year fixed effects. In all models, I control for the difference in vote shares between the parties. For the first two, I also add a set of controls at the country-year level: whether there was a national parliamentary election in the same year, the number of parties, the population size of the country (logged), as well as the number of events in the ICEWS database for each country-year. I also include the mean cooperation score for each country-year to addresses potential concerns about the comparability of the

[7] While the cooperation scores are available on a yearly basis, CMP positions can only be measured in election years. I therefore assign the most recent available CMP position.

separately estimated latent spaces. It is unnecessary to include these controls in the final specification, as they are absorbed by the country-year fixed effects. All models cluster the standard errors at the dyad level and take the estimation uncertainty of the cooperation scores into account.

The specifications provide evidence supporting both conjectures laid out above. Being in a coalition together is indeed associated with a more cooperative public relationship compared to one party being in government and one being in the opposition, as well as to both parties being in the opposition. Depending on the specification, being in a coalition is associated with an increase in the dependent variable by 0.42 to 0.6 standard deviations compared to the government-opposition baseline. If both parties are in the opposition, the cooperation score is somewhat lower than in the baseline, although the confidence intervals include zero.

Policy distance also has an effect on the degree of cooperation. However, it is *not* the difference in left-right positions that affects the cooperation scores. Instead, second-dimension differentiation matters as parties with a greater difference in their positions towards multiculturalism have lower cooperation scores. A one standard deviation increase in the variable is associated with a decrease in the dependent variable by 0.07-0.12 standard deviations.

Thus, a systematic analysis of the cooperation scores shows that political parties are more likely to have a cooperative relationship if they are in a coalition together, and if they have similar views on second-dimension policies. The application thus demonstrates the face validity of the proposed approach to quantify public interactions. At the same time, it hints at the potential of the measure for the study of political interactions.

Conclusion

In this paper, I have introduced a novel approach to quantify public political interactions. I use machine coded news reports as my source material and create a spatial representation by estimating latent factor network models. They allow me to compute cooperation scores between any dyad of actors, be they partisan political, non-partisan political, or societal, and irrespective of whether any interaction between the pair was actually observed. The measure captures public interactions in a multitude of venues on a continuous basis, locates political and non-political actors in a common space, and reflects that communication is not unidirectional but rather a back and forth. In an application, I have demonstrated the value of the approach for the study of party politics. However, this is only one area of study in which the new data can be useful. In particular, they can also be used to study the public interactions between politicians and the non-partisan state apparatus, and between politicians and societal actors.

In recent years, scholars have started to increasingly make use of novel, large-scale data to address long-standing questions in comparative politics. The quantified political interactions introduced here contribute to this trend. The cooperation scores derived from large-scale machine coded event data provide a wealth of new information to explain and be explained. They open the door for new research on, for example, elite polarization and its effect on mass polarization, coalition dynamics and the prediction of coalition formation and collapses, democratic representation, citizens' perceptions of parties, and democratic satisfaction.

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 Bercovitcch 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.

Lucas, Christopher, Richard Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer and Dustin Tingley. 2015. "Computer Assisted Text Analysis for Comparative Politics." *Political Analysis* 23(2):254–277.

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.

Minhas, Shahryar, Peter D. Hoff and Michael D. Ward. 2016. "Inferential Approaches for Network Analysis: Amen for Latent Factor Models." Unpublished Manuscript.

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.

Ramshaw, Lance, Elizabeth Boschee, Marjorie Freedman, Jessica MacBride, Ralph Weischedel and Alex Zamanian. 2011. SERIF Language Processing–Effective Trainable Language Understanding. In *Handbook of Natural Language Processing and Machine Translation: DARPA Global Autonomous Language Exploitation*, ed. Joseph Olive, Caitlin Christianson and John McCary. New York: Springer.

Sagarzazu, Inaki and Heike Klüver. 2015. "Coalition Governments and Party Competition: Political Communication Strategies of Coalition Parties." *Political Science Research and Methods* tba.

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* tba.