TITLE: An empirical examination of echo chambers in US climate policy networks

SUPPLEMENTARY INFORMATION:

Constructing a Data Set and Sampling

No comprehensive dataset of elite United States climate policy actors exists, so one had to be constructed for the purposes of this study. A dataset was assembled from three publicly available sources. We began with a list of all the policy actors who participated in climate-related hearings in the United States Congress during the two sessions of Congress prior to when we entered the field. These were the 109th and 110th sessions, which were held from January 2005 to January 2009 *(6)*. To this dataset, we added all of the registered lobbyists who were named in the March 2009 "Climate Change Lobby" inventory compiled by the Center for Public Integrity *(52),* which lists all non-governmental actors who lobbied the US government on the issue of climate change. Finally, we cross-referenced this list with a roster of all attendees from the United States who participated in the international climate change negotiations (COP-15) in Copenhagen in December 2009 *(53)*. By drawing from these varied sources that span the previous sessions of the Congress, we were able to assemble a dataset that measured sustained engagement in the climate policy network over the five years leading up to our period of data collection. Although it would have been ideal to include a list of the speakers who participated in the 111th session of the US Congress as well, since data collection was taking place in the middle of this particular session, a complete list of hearings and participants was not available.

Together, these sources yielded a universe of 1,549 different political actors engaged in climate change politics in the United States in the five years prior to our entering the field to collect data in summer 2010. As such, they spanned the period during which the leadership in the

Congress changed from Democrat to Republican, and during which the presidency went from George W. Bush to Barack Obama. To restate, by including this time span, we are able to measure *sustained* engagement in the policy network.

Next, actors in this dataset were ranked according to how often they were included within hearings, international negotiations, and on lobbyist registries. Testimonies were weighted, such that multiple appearances before Congress indicated greater participation. The top 100 actors—that is, the 100 most active participants in the US climate policy network during the sampling period—were selected to make up our sample following the 'events based' approach to determining network boundaries *(54)*.

As we were interested in the position of the policy actor and not the individual holding a specific job, in cases where people changed jobs or left specific organizations prior to our data collection, his or her replacement was contacted to participate in the study. Although scientists are policy actors engaged in this network, they differ from other policy actors in that they represent their own research when providing Congressional testimony. As such, these scientists were surveyed directly. Following our Human Subjects Protocols from the Institutional Review Boards at Columbia University (# IRB-AAAG2612) and University of Maryland (IRB Protocol #10-0751), these scientists are not named directly in the findings of our study. Instead, they are listed by the name of the Universities where they work.

The top 100 policy actors were contacted to participate in the survey. It is worth noting that these most-engaged actors differed somewhat from the entire sample of actors in our dataset. In particular, the sample of most engaged policy actors undersamples governmental and business actors, oversamples scientists, non-governmental organizations, and Democratic congresspeople, and matches rates for Republican congresspeople. Although the top 100 political actors represents a small proportion of the possible total, it is this core of political elites that have the most influence over the policy process.

Network analysis routinely deals with these kinds of 'boundary' problems *(55)*. Fortunately, we do not wish to generalize our findings to the entire population of actors. Rather, our goal is to understand the dynamics of the most central actors in the climate policy network in the United States at the federal level. We recognize that the processes at work in this sample are unlikely to resemble the dynamics in the periphery *(56)*.

It is important to note that our data set, like those of others working from a policy network perspective *(41,43,57)*, includes actors with all types of organizational affiliations, including both state and non-state actors. We operationalize non-state actors as those who are not employed by the government; non-state actors include business and trade union representatives, members of the environmental and climate policy teams at NGOs, and university scientists. State actors include both climate-focused staff people from Congressional offices, and employees working in the Administrative branches of federal agencies. The authors coded all respondents into one of these categories based on where they worked during the sampling period. Not only does this sampling method provide a range of actor types, but it notably also provided us with actors who represent the full range of ideological positions on the issue of climate change in the United States.

Data Collected

All of the data for this paper were collected in Summer 2010. This time period was particularly interesting for studying climate policy because the American Clean Energy and Security Act, sponsored by Representatives Henry Waxman and Edward Markey, had been passed by the

House of Representatives in summer 2009. During summer 2010, while we were in the field surveying respondents, a similar bill, originally drafted by Senators John McCain and Joseph Lieberman and revised by Senator Barbara Boxer, was working its way through the Senate. If passed, this bill would have been the first case of federal climate legislation passing through the US Congress. Policymakers, business representatives, environmental activists, and climate scientists *(50-51)* were thus primed to speak about climate politics during this period.

Data collection for this project was conducted in accordance with Columbia University (IRB Protocol # IRB-AAAG2612) and University of Maryland (IRB Protocol #10-0751) policies on Human Subjects research. Data were collected through in-person meetings in Washington, D.C. and New York City. In these in-person settings, actors were interviewed and administered a written survey. The survey data are the focus of this paper. For policy actors who were not available to meet or who were located outside of these two areas, surveys were conducted through the online service *SurveyMonkey*. In total, survey data were collected from 64 of the top 100 policy actors, which represents a 64% response rate. This response rate is consistent with other studies of communities of political elites *(58-59)*.

Although respondents included actors from across the political spectrum and from all types of organizational affiliations, we observed some obvious differences between the respondents and the non-respondents in our study. Specifically, the offices of Representatives and Senators in the US Congress had a much lower response rate than the other types of political actors (33% versus an overall response rate of around 65% for all the other political actors). We hypothesize that this low rate is a product of the time period in which we collected data. As mentioned previously, climate legislation was working its way through the Senate. The climate bill was front and center in the media and on Capitol Hill, and drew a lot of attention from the public. The Offices of Congresspeople who were highly engaged in this issue—and thus had been captured in our sampling method—were less likely to make themselves available to speak about their opinions regarding climate politics and the policy network. Coupled with the fact that it was summer, right before the recess, it makes sense that Congresspeople would be underrepresented in the dataset.

Survey Instrument

The survey itself was comprised of three types of questions. Attitudinal questions asked participants to indicate on a scale of 1 to 5—where 1 indicated strong disagreement and 5 indicated strong agreement—their positions on such statements as *Human activities are an important driver of current global climate change* (survey Question 1, item 2: referred to in the paper as Anthropogenic), *Emissions trading (cap and trade) is the best option for reducing US GHG emissions* (survey Question 5-1, item 2: referred to as Cap and Trade), and *There should be an international binding commitment on all nations to reduce GHG emissions* (survey Question 6-3, item 1: referred to as Binding).

Distributions of responses for these three items are presented in figure S1. The survey instrument was developed as part of an international collaborative project on Comparing Climate Change Policy Networks (COMPON). Thus, it includes questions that aim to answer international comparative questions. The US team of this international project, which was led by Dana R. Fisher (co-PI on the NSF grant), was able to include questions in Section 5 of the survey that were aimed at issues of importance to domestic climate politics in the US. As the US Congress was specifically considering a so-called “cap-and-trade” legislation during the time when we collected data, we selected this question to be one of the three questions in our analysis.

Given the propensity of the US Congress to debate the science of the issue, we also included the anthropogenic question that was developed in collaboration with the international COMPON team. Finally, we included a question about perspectives on the international regime and the role that developing countries should play. This issue has been the subject of debate in the US since the original Kyoto Protocol was being drafted and was one of the main drivers behind the Byrd-Hagel Resolution (Senate Resolution 98) in the US Senate in 1997, which has been seen as a vote on climate change legislation by academics and policymakers alike (for a full discussion, see *60-61*). Survey questions also ask specifically which actions these actors have taken in the climate policy sphere, including lobbying, policy advising, media and publicity outreach, technical reporting, activism and mobilization, and coalition-building.

The survey also asked characteristic questions about these policy actors. For those actors representing businesses, think tanks, NGOs, and other organizations, these questions asked about staff size generally, and specifically about staff positions that deal directly with creating, assimilating, or reporting climate science or climate policy. These questions also include inquiries into the size of the organization's operational budget as a proxy for the overall size of the organization.

Finally, and most importantly for the purposes of the present research, this survey asked four network questions. Each of the 100 policy actors in our sample was listed in alphabetical order, and each respondent was presented with four iterations of this list. Participants were then asked to indicate, in order: those actors whom they perceived to be most influential in climate politics, actors whom they identified as their sources of expert scientific information about climate change, which actors or organizations that they collaborate with on a regular basis regarding the issue of climate change, and which policy actors they identified as being opposed to their organization's stance on climate change politics. It is the second network question—regarding sources of expert scientific information—that makes up the foundation of this research. A copy of the survey, as well as the entire public US dataset and a codebook, can be found at: http://www.drfisher.umd.edu/ under the 'Present Projects' tab.

The authors also collected data on organizational and structural characteristics (such as the age of the organization, the organization's number of employees, and so forth) from a systematic review of the websites and published materials of each policy actor. Due to their unique nature among other types of actor, many of these variables are inappropriate for the Congressional offices in our sample, and are therefore missing.

Missing Data

We made a number of decisions about handling missing data. First, non-respondents were removed from the network, reducing our size from the 100 actors in the roster to the 64 respondents. Only two of the 36 actors removed were mentioned more than 10 times by the 64 respondents. While this approach is clearly not an ideal method, it is standard in network analysis *(62)* and appropriate missing data methods are still being developed *(56)*. Additionally, we had cases where respondents left questions blank. In order to use the ERGM attribute methods in Statnet, we could not leave any entries missing. Instead, we replaced the NA's with the average value of the attribute. This action would have the least impact on average for the sender and heterophily terms.

For the echo chambers term, this choice is more complex. An echo chamber occurs when a transitive triad exists among three actors that all responded with the exact same rating for the given variable. By constructing a new attribute category, we've added a new category that can increase the count of echo chambers. In the empirical data, there were a few transitive triads among the actors in this missing category, but these amount to 2% of the 563 total transitive triads in the network. Since this process was applied to all three of the existing attribute categories (Anthropogenic, Cap and Trade, and Binding), it cannot explain our results regarding why echo chambers are only found in one of the three. Additionally, we ran a separate model where each respondent with an 'NA' response for Binding was given a separate category. The coefficient for Binding transitive triads did not change in either size or significance from the model presented in the paper.

ERG Models

With Exponential Random Graph (ERG) models, we are able to test for the presence of these terms in empirical data over and above competing mechanisms *(63)*. The ERG model is expressed as:

$$P(Y = y|X) = \frac{\exp(\theta^T g(y,X))}{\kappa(\theta)} \quad [1]$$

This model simulates the probability of observing a given network (*y*) using a reference distribution of alternate networks ($\kappa(\theta)$), a set of network statistics ($g(y,X)$), and a set of coefficients for these statistics (θ^T). The models were fit using the statnet package for R *(64)*.

We include control variables that represent other important network processes. The first of these is the general tendency to seek information from specific types of policy actor. Since the directionality of our ties points towards the recipient of the information (the direction of information flow), this measure is the likelihood of a given actor type to be chosen as an information source compared to a reference group (Fig 3e of the manuscript). Respondents in our study were classified into five different types of policy actors: business organizations (the omitted reference category in the ERG model), congressional offices, administrative offices, non-governmental organizations, and professional and university scientists. Similarly, we expect that other organizational characteristics (lobbying budget, number of full time employees) affect each policy actor's likelihood of being cited as a source. We include these terms for the log of these values as well. Since these attributes are measured as continuous variables, their coefficients are interpreted as the relative likelihood of a tie as the sender's attribute level increases (Fig 3f of the manuscript).

Beyond these control variables, there are other structural characteristics that must be controlled. In our model, we condition on the in-degree of each policy actor. In-degree measures the number of information sources mentioned by each respondent. Rather than trying to simulate this distribution, conditioning on the empirical numbers ensures that every simulated network has the same distribution of responses. This process ensures that any significance in the model is not due to variation in the memory of the respondents or the time they allotted for the survey. Conditioning on this distribution also greatly aided in model convergence. The final mechanism included that is not directly related to the echo chamber is centralization, or the popularity of an actor. We include terms for 2- and 3-star network structures (Fig 3b and 3c of the manuscript), which together represent the tendency for actors that are already listed as sources to gain additional ties (with diminishing returns if the 3-star coefficient is negative). This process is interpreted as a tendency towards popularity or centralization in the network *(63)*. We fit a variety of models using the terms specified, but only present the best fitting model (determined by the Bayesian Information Criterion (BIC) value – a model selection criterion).

Methods and Model Fit

The ERG models were fit using the Statnet software package for the R programming language *(64)*. This method models the empirical data via Markov-Chain Monte-Carlo simulation to produce estimates of the sufficient statistics supplied to the model. We first attempted to fit models without conditioning on the indegree distribution, but no such model converged. By simulating from these, we saw that the model was unable to capture anything resembling the skewed indegree distribution shown in Figure S2. Conditioning on this distribution permitted the models to converge.

Appropriate model convergence is essential for verifying the interpretability of the model. The model adequacy check examines how well the simulation mixes over the sample space and whether these simulated networks produce normal distributions centered at the empirical values for each statistic in the model. For each of the 17 model terms, comparing the empirical values to the distribution of each value produced by 10,000 sample networks drawn from the MCMC simulations, resulting in no p-values less than 0.1. See Figures S3-S8 for visual images of the results. The value of the given statistic in each of the 10,000 simulations appears on the left-hand side. The distribution on the right has been centered with 0 indicating the empirical value. Since none of the empirical values for our sufficient statistics are significantly different from the distributions generated via the MCMC (a different test than that showing significance in network generation) our model has converged *(64; see also for examples of poorly converged chains).*

The goodness of fit measures simulate from the model and compare empirical network statistics not included in the model to the results from the simulation. This simulation gives an indication of how well the model reproduces features of the empirical network beyond the specific statistics it was intended to capture. In Figures S9-S11, statistics not included in the model (out-degree distribution, edgewise shared partners, and minimum geodesic distance) are simulated from the model (the box plots) and compared to the empirical values (the thicker, darker line). If the empirical values fall within the simulated values (as most of ours do), it demonstrates that the model is able to recover correctly network properties outside those included in the model. These results give confidence in the fit of the overall model to the empirical data.

As we note in the paper, we present the model that best fit the data (by lowest BIC value). We also present corresponding model adequacy checks and goodness of fit. However, this model was only one of many models we ran. Twelve models—including the model we provide in the paper—are presented in Table S1. These models add complexity to our understanding of the relationship between homophily and echo chambers. We see in both the Binding and Cap and Trade terms that homophily decreases when the transitive triad term is included for the model, however this is not true in the Anthropogenic case. Finally, as the table demonstrates, whenever included, the Binding transitive triad term was always significantly positive, even in the presence of other transitive triad terms. In other words, our paper may present the best-fitting model, but the results are robust across a range of models.

REFERENCES

52. The Center for Public Integrity (2009) *The climate change lobby*. (Database prepared by Center: http://www.publicintegrity.org/investigations/climate_change/)

53. UNFCCC, Conference Of The Parties, Fifteenth Session, Copenhagen List of Participants, Part 1 and Part 2 (7–18 December 2009, available at: http://unfccc.int/documentation/documents/advanced_search/items/3594.php)

54. Marsden, P. in *The Sage Handbook of Social Network Analysis*. Scott, J., Carrington, P.J., eds. (Sage Publications, 2011).

55. Kossinets G. Effects of missing data in social networks. *Social Networks* **28,** 1-31 (2006).

56. Shalizi CR, Rinaldo A. Consistency under sampling of exponential random graph models. *The Annals of Statistics* **41**, 508-535 (2013).
57. Leifeld, P., Schneider, V. Information exchange in policy networks. *Am. J. of Pol. Sci.* **53,** 731-744 (2012).
58. Heaney, M.T. Brokering health policy: Coalitions, parties, and interest group influence. *J of Health Polit., Pol., Law* **31,** 887-944 (2006).
59. Lewis, J.M. Being around and knowing the players: Networks of influence in health policy. *Soc. Sci. & Med.* **62,** 2125-2136 (2006).
60. Fisher, Dana R. *National Governance and the Global Climate Change Regime.* Rowman & Littlefield Publishers, Inc. (2004).
61. Fisher, Dana R. "Bringing the Material Back In: Understanding the United States Position on Climate Change." *Sociological Forum,* Volume **21, 3,** 467-494 (2006).
62. Goodreau, S.M., Kitts, J.A., Morris, M. Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography* **46**, 103-125 (2009).
63. Lusher, D., Koskinen, J., Robins, G. *Exponential Random Graph Models for Social Networks* (Cambridge University Press, 2012).
64. Goodreau, S., Handcock, M.S., Hunter, D.R., Butts, C.T., Morris, M. A Statnet Tutorial. *J. of Stat. Soft.* **24,** 1-27 (2008).

Figure S1: Distribution of Responses to Attitudinal Questions

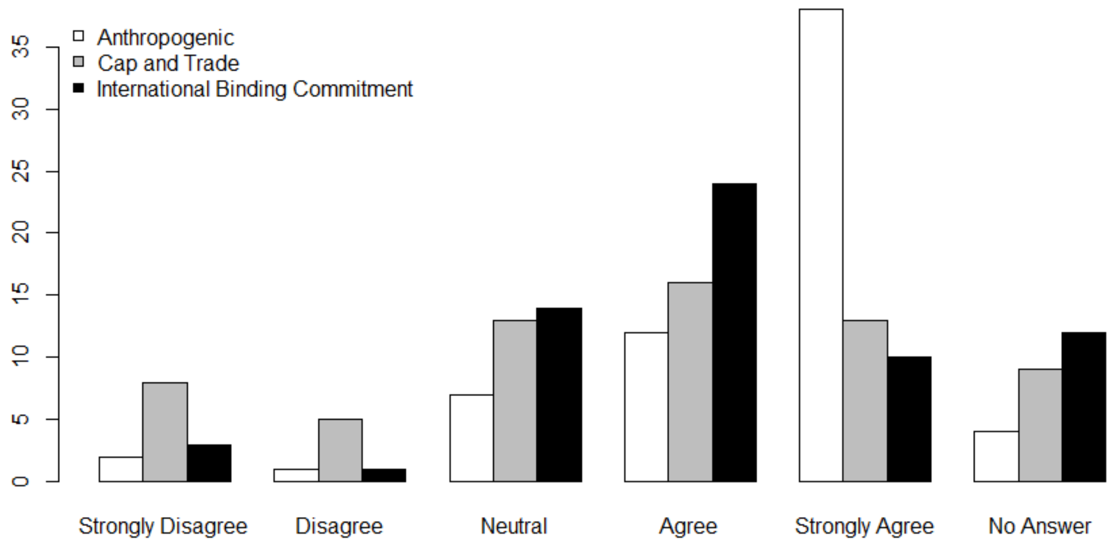
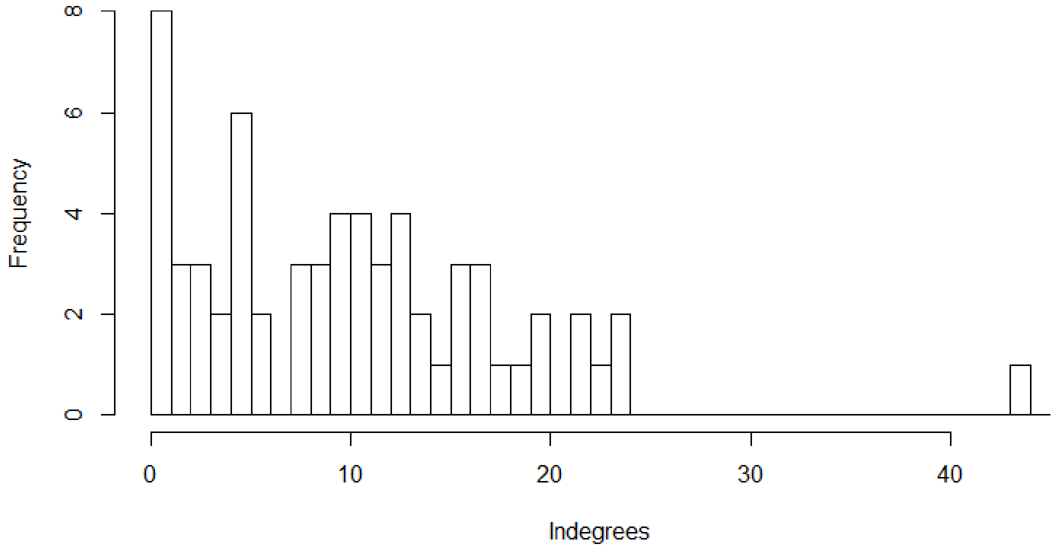
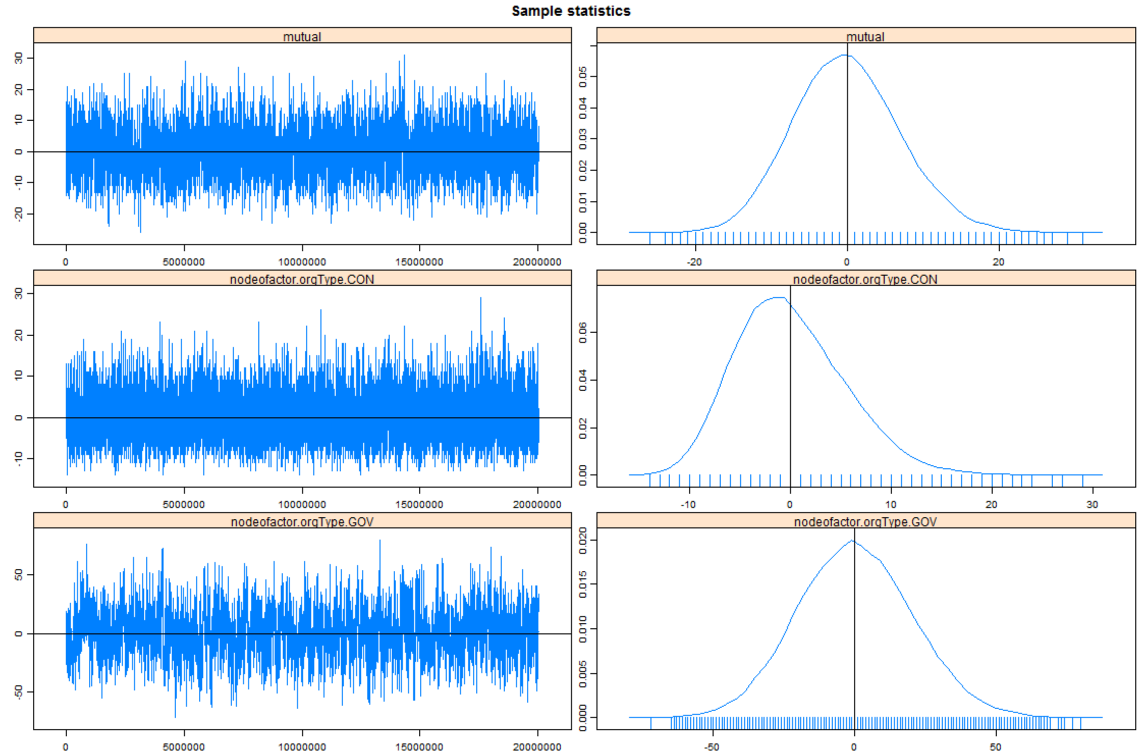
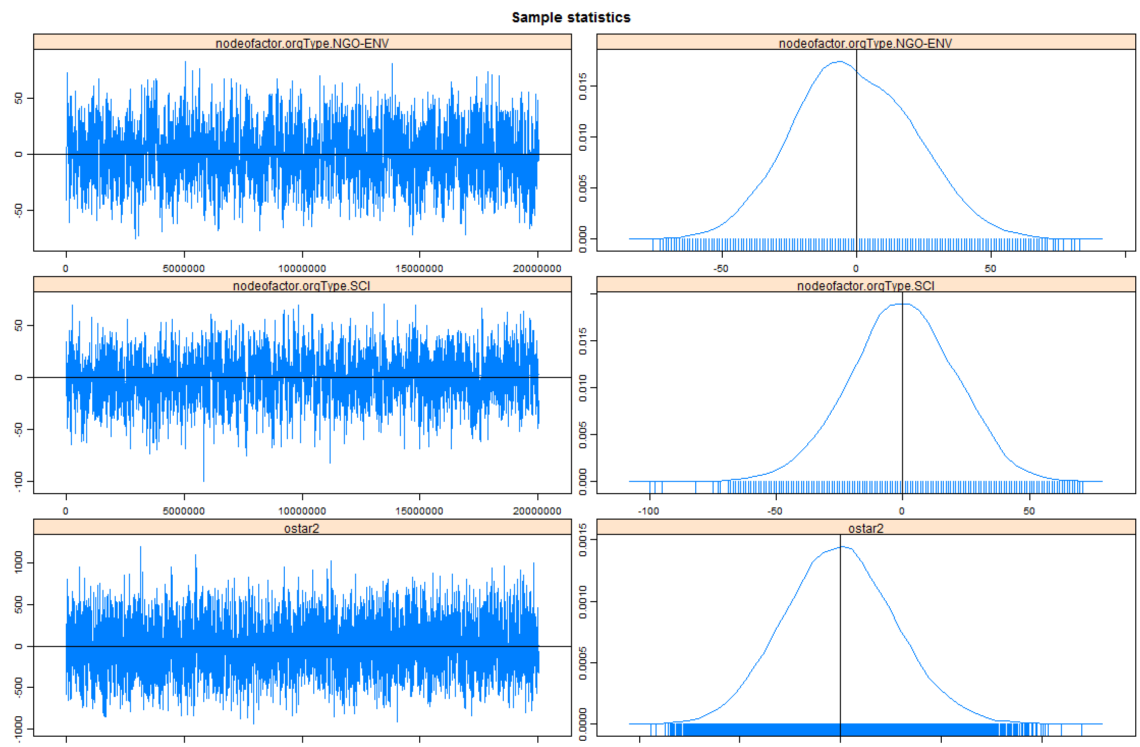


Figure S2: Indegree Distribution is the number of political actors each respondent mentioned as an information source.

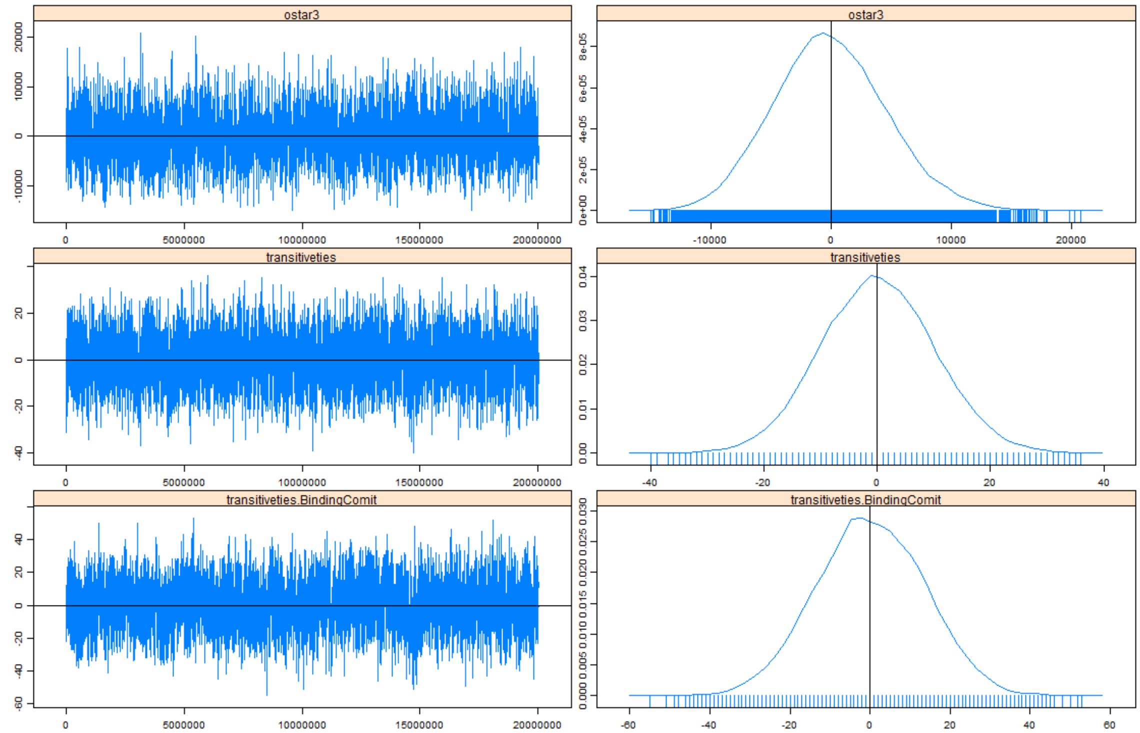


Figures S3-S8: Diagnostic Plots of the Markov Chain Monte Carlo Simulation

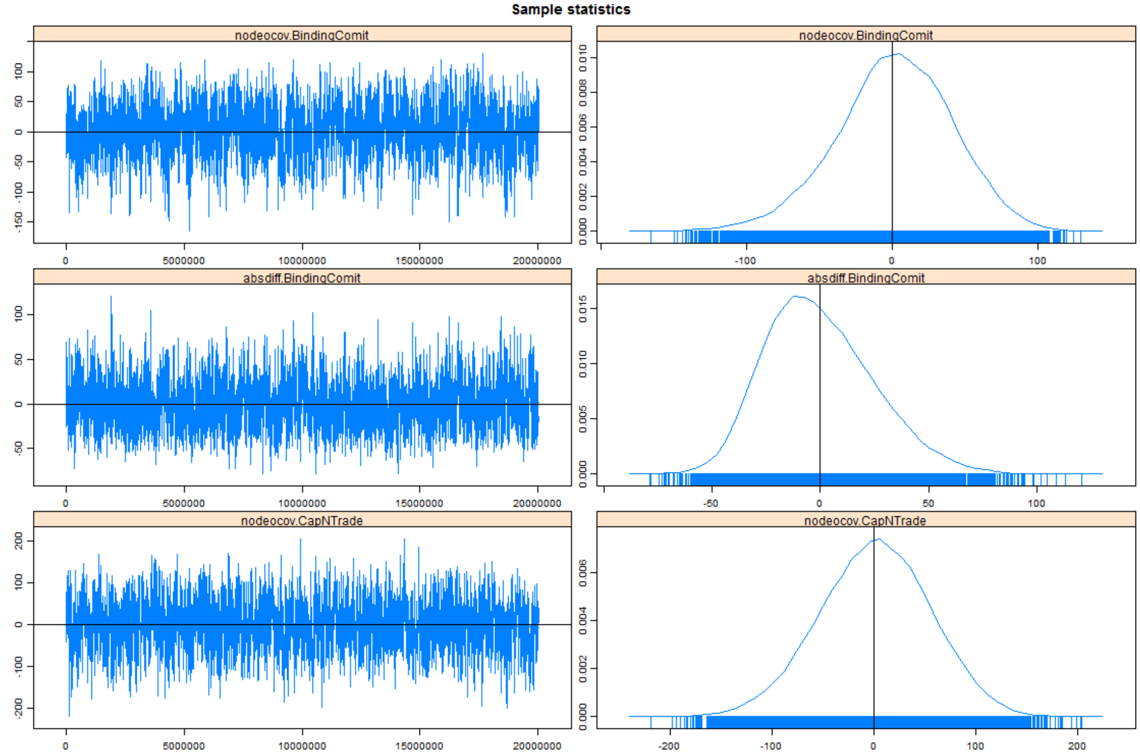




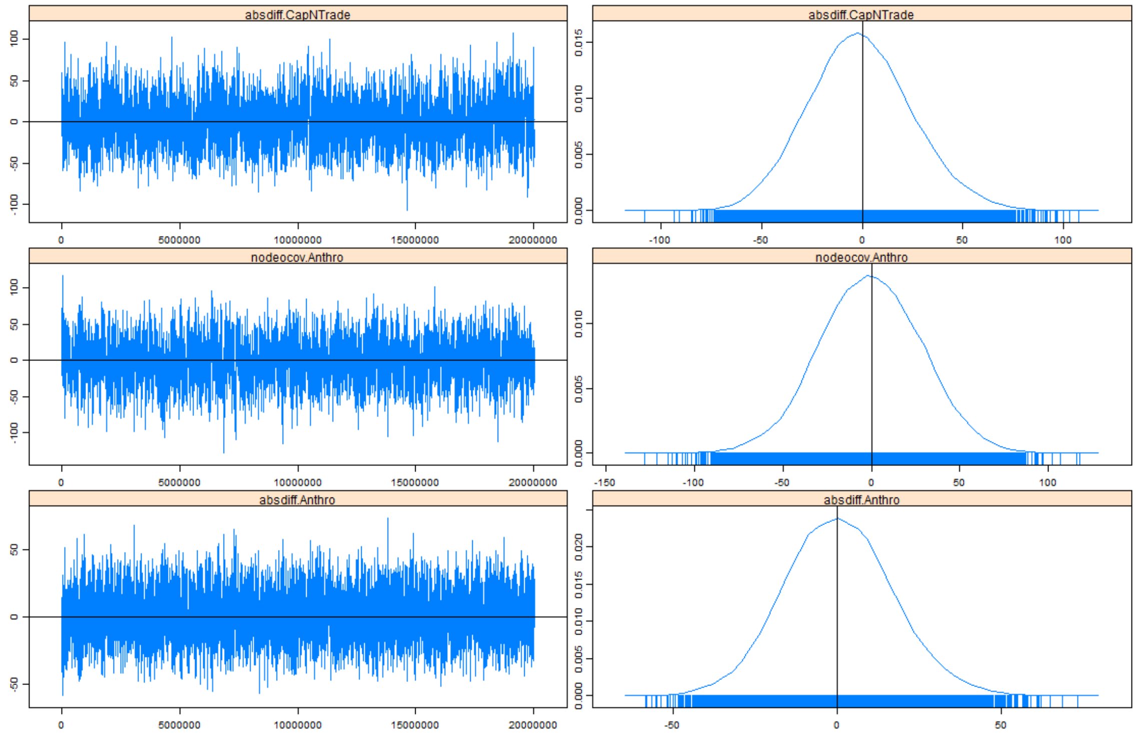
Sample statistics



Sample statistics



Sample statistics



Sample statistics

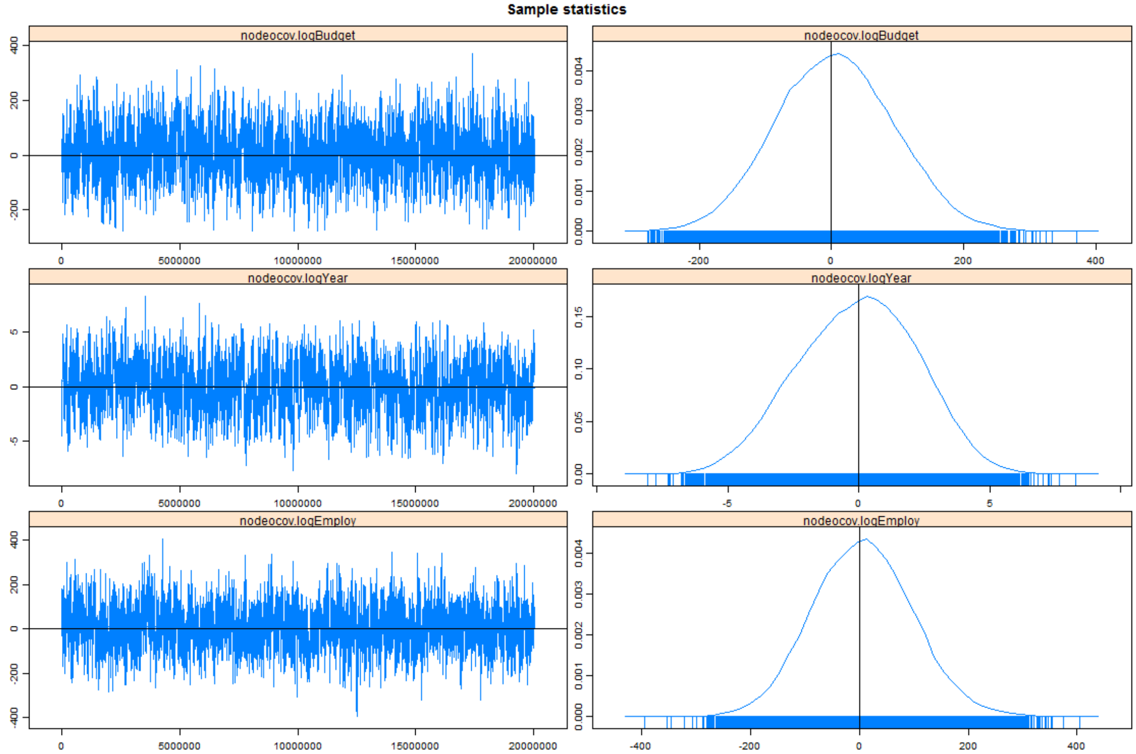


Figure S9: Simulated fit of the Out-Degree distribution

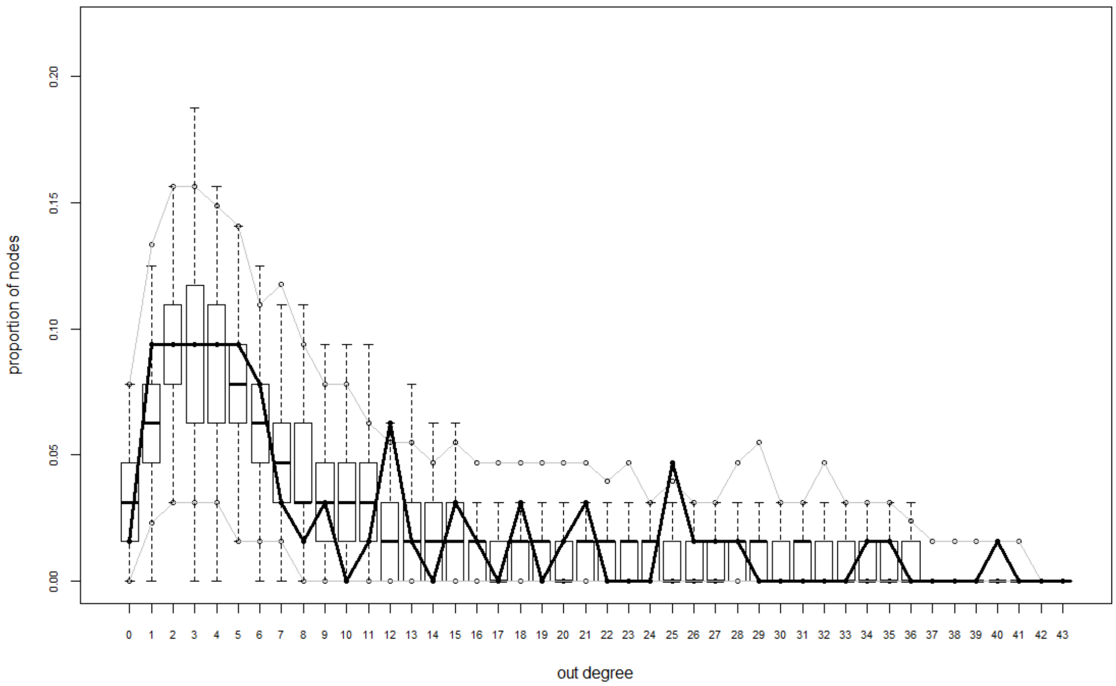


Figure S10: Goodness of Fit of Edge-wise Shared Partners

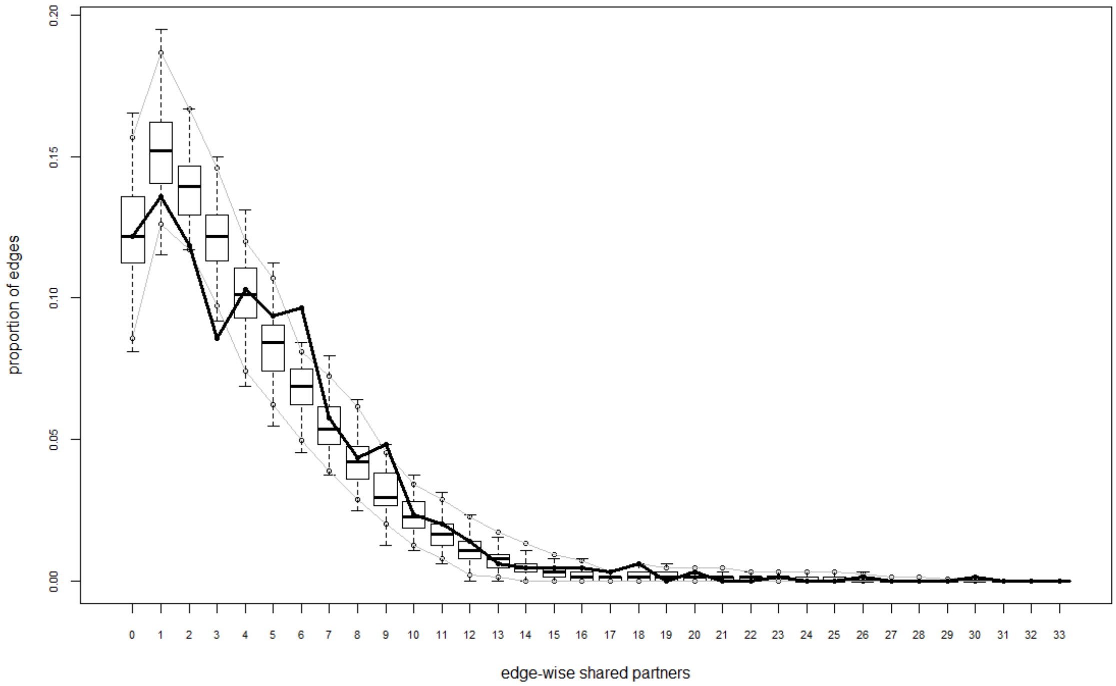


Figure S11: Goodness of Fit of Geodesic Distance

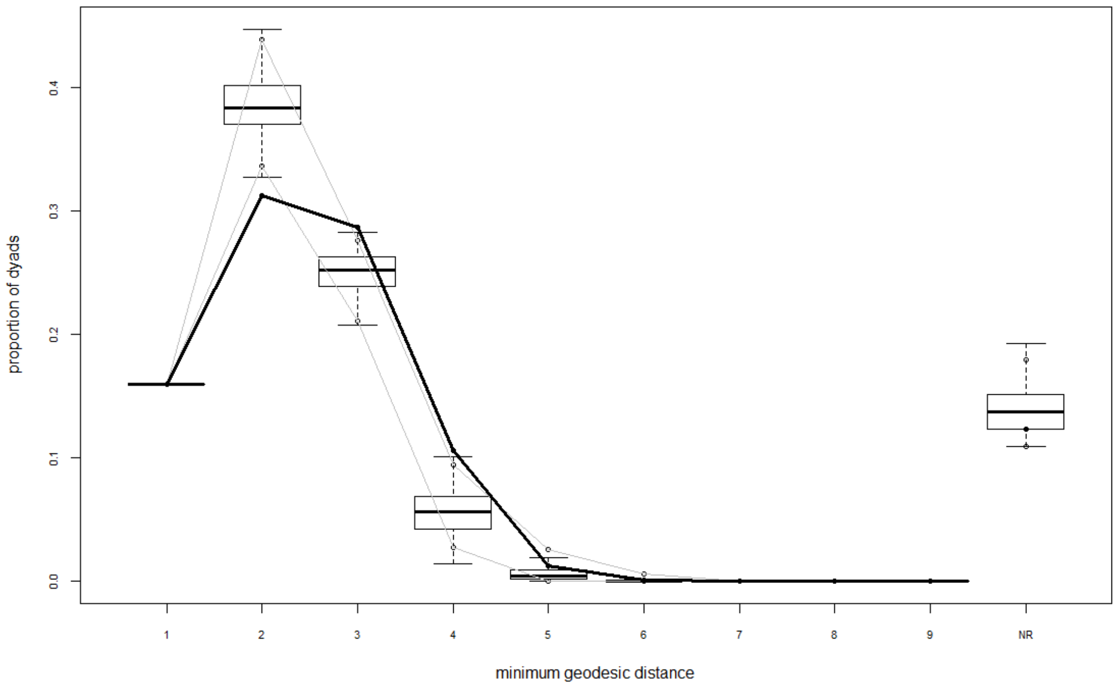


Table S1: All Models

	1†	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Structural Characteristics												
Binding: Transitive Triads	.27(.09)**	.28(.1)**			.26(.1)**	.25(.1)*	.27(.1)**	.3(.1)**		.25(.1)*		.28(.1)**
Binding: Heterophily	-.05(.06)	-.04(.06)	-.15(.06)**	-.16(.06)**	.06(.06)	-.07(.06)	-.06(.06)	-0.06(.07)	-0.19(.07)**	-0.12(.07)	-.17(.07)*	-.09(.07)
Binding: Sender	-.05(.03)	-.04(.03)	-.05(.04)	-.06(.04).	.04(.03)	-.05(.04)	-.04(.04)	-0.06(.04)	-0.07(.04)	-0.08(.04)*	-.07(.04)	-.06(.04)
Anthropogenic: Transitive Triads						-.02(.1)	-.01(.1)				.03(.09)	-.06(.09)
Anthropogenic: Heterophily	-.38(.09)***	-.39(.09)***	-.38(.09)***		-.38(.09)***	-.38(.1)***	-.39(.1)***	-.4(.09)***	-.39(.09)***		-.38(.1)***	-.42(.1)***
Anthropogenic: Sender	-.2(.06)***	-.2(.06)***	.019(.06)***		-.2(.06)***	-.19(.06)***	-.19(.06)**	-.23(.06)***	-0.22(.06)***		-.23(.06)***	-.22(.06)***
Cap and Trade: Transitive Triads			.21(.11).	.23(.11)*	.18(.11)	.17(.11)	.18(.11)		.25(.11)*	.2(.11).		.2(.11).
Cap and Trade: Heterophily	-.08(.04)	-.08(.04).	-.03(.04)	-.04(.05)	-.05(.05)	-.05(.05)	-.05(.05)	-.08(.05)	-0.01(.05)	-0.05(.05)	-0.07(.05)	-.03(.06)
Cap and Trade: Sender	.04(.03)	.04(.03)	.03(.03)	.03(.03)	.03(.03)	.03(.03)	.03(.02)	.06(.03).	0.04(.03)	.03(.03)	.05(.03)	.05(.03)
Reciprocity		-.13(.19)			-.16(.18)		-.16(.19)					
Transitive Triads	.03(.11)	.08(.13)	.08(.11)	.08(.1)	.06(.13)	.02(.12)	.07(.13)	.07(.12)	.1(.12)	.05(.12)	.13(.12)	.07(.12)
Out 2-Star	.17(.02)***	.17(.02)***	.17(.02)***	.17(.02)***	.17(.02)***	.17(.02)***	.17(.02)***	.09(.00)***	.09(.004)***	.09(.00)***	.09(.004)***	.09(.004)***
Out 3-star	-0.004(.001)***	-.004(0)***	-0.004(.001)***	-.004(0)***	-.004(0)***		-.004(.001)***					
Node Characteristics												
Actor Type (Ref. Category: Business)												
Science	.35(.09)***	.35(.09)***	.34(.1)***	.33(.08)***	.34(.09)***	.34(.09)***	.34(.09)***	.52(.11)***	.52(.12)***	.5(.11)***	.53(.12)***	.52(.11)***
Environmental NGO	.11(.07)	.11(.07)	.13(.07).	.12(.07)	.11(.07)	.1(.07)	.1(.07)	.26(.1)**	.29(.1)**	.26(.1)**	.31(.10)**	.26(.1)**
Government	.19(.07)**	.19(.08)*	.18(.8)*	.18(.07)*	.2(.07)**	.2(.08)**	.2(.08)*	.38(.1)***	0.36(.1)***	0.39(.1)***	.36(.10)***	.38(.1)***
Congress	-.22(.2)	-.21(.19)	.22(.2)	-.3(.2)	-.22(.19)	-.22(.2)	-.22(.2)	-.40(.23).	-0.39(.24).	-0.5(.23)*	-.4(.24).	-.38(.23)
Log of Budget	-.01(.01)	-.01(.01)	0(.01)	.004(.01)	-.004(.01)	-.01(.01)	-.01(.01)	.01(.02)	-0(.02)	.01(.02)	-.001(.02)	-.01(.02)
Log of FT Employmees	.01(.01)	.01(.01)	0(.01)	0(.01)	.01(.01)	.01(.01)	.01(.01)	0.01(.01)	0.01(0.01)	.01(.01)	.003(.01)	.02(.01)
BIC	-490.2	-489.4	-485.9	-484.8	-482.9	-482.8	-475.9	-475.9	-472.3	-467.8	-467.3	-463

† We present Model 1 in the article.
* Significant at p< 0.05
** Signifncant at p<0.01
*** Significant at p<0.001