Appendix for "Free Riding, Network Effects, and Burden Sharing in Defense Cooperation Networks"

Contents

A1 Computational Model

Consider a finite set of nodes or agents, $N = \{1, 2, \ldots, n\}$, where connections or ties between agents are defined by the $n \times n$ network matrix, $g \in \{0, 1\}$. A nonzero entry, $g_{ij} = 1$, indicates that a network tie exists between row agent i and column agent j. The matrix is symmetric, such that $g_{ij} = g_{ji}$. Because agents cannot form ties with themselves, $g_{ii} = 0 \; \forall \; i \in N$. Substantively, the g matrix represents the DCA network at a given moment in time. Further consider an $n \times 1$ matrix, $r \in \{1, \ldots, M\}$, which defines a node-level or individual behavior, scaled across M ordinal categories. The r matrix represents country-level defense effort in a given year, where larger r_i values indicate that the focal agent i expends more effort on defense (e.g., in terms of spending as a percentage of GDP).

The agent-based model (ABM) of network-behavior coevolution assumes that agents are myopically utility maximizing (Snijders and Steglich 2015: 233). That is, agents select network ties and behaviors in a way that maximizes their immediate subjective utility. They do so by optimizing two distinct objective functions, corresponding to the g network and r behavior. First consider the g network. When given an opportunity to make a change to g, an agent i may create a new tie, terminate an existing tie, or make no change at all. Let g^+ be the network that exists after i has been given an opportunity to change its ties. If i chooses not to change its ties, $g^+ = g$. Otherwise, g^+ differs by one tie, such that, for a given j partner, $g^+_{ij} = 1 - g_{ij}$. Assume that i

chooses the network g^+ that maximizes the function $f_i^{\text{net}}(g, g^+, r) + \epsilon_i^{\text{net}}(g, g^+, r)$, where ϵ_i^{net} is a random disturbance term that represents unexplained change. If we further assume that ϵ_i^{net} follows a type I extreme value or Gumbel distribution (Maddala 1983), which is a common assumption in random utility models (Snijders 2001: 363), then the choice probabilities for i can be expressed as

$$P_i(g^+) = \frac{\exp(f_i^{\text{net}}(g^+, r))}{\sum_{k=1}^{n} \exp(f_i^{\text{net}}(g^{+k}, r))}, \tag{A1}$$

where the sum in the denominator refers to all possible g^{+k} states of the network, or the options available to i for toggling its network ties (Snijders and Steglich 2015: 233). This probability is the same as that used in multinomial logit models (Maddala 1983), and it allows agents to maximize their utility by choosing whichever tie offers the greatest payoff.

Because DCAs are nondirected, agent i cannot unilaterally impose network ties but must have its proposed ties accepted by their respective targets. Assume that i's preferred g^+ network involves a new ij tie, and let g^- represent the network *without* the ij tie. The target node j must then choose between g^+ and g^-. The probability that j will accept the tie proposed by i is given by

$$P_j(g^+) = \frac{\exp(f_j^{\text{net}}(g^+, r))}{\exp(f_j^{\text{net}}(g^-, r)) + \exp(f_j^{\text{net}}(g^+, r))}, \tag{A2}$$

where $\exp(f_j^{\text{net}}(g^+, r))$ is the agent-oriented network objective function applied to j rather than i. Thus, the more that g^+ increases node j's utility relative to g^-, the more likely j is to accept the offer. This method of modeling symmetric ties is known as "unilateral initiative and reciprocal confirmation" (Snijders and Pickup 2016), and it is a reasonable approximation of real-world treatymaking processes (Kinne 2013). The same choice probabilities have been used to model preferential trade agreements (Manger et al. 2012), military alliances (Warren 2010, 2016), and bilateral agreements across multiple issue areas (Kinne 2013).

The ABM models behavior choice probabilities in a similar way. When given an opportunity, i selects an r_i behavior so as to maximize the function $f_i^{\text{beh}}(g, r, r^+) + \epsilon_i^{\text{beh}}(g, r, r^+)$, where r^+ is the $n \times 1$ matrix of behaviors that results after i has decided whether to change its behavior, and ϵ_i^{beh} is again a random disturbance. The probability of i changing its behavior is thus

$$P_i(r^+) = \frac{\exp(f_i^{\text{beh}}(g, r^+))}{\sum_{k=1}^{M} \exp(f_i^{\text{beh}}(g, r^{+k}))}, \tag{A3}$$

where the sum in the denominator refers to all possible next behavior states (Steglich et al. 2010: 350). By assuming an ordinal behavior variable, the ABM can model both network relations and individual behavior from a common statistical framework—i.e., a continuous-time Markov chain with a discrete outcome space (Niezink et al. 2019: 296).

The ABM implements these functions as linear combinations of effects,

$$f_i^{\text{net}}(g, g^+, r) = \sum_{h}^{L^g} \beta_h^{\text{net}} s_h^{\text{net}}(i, g, g^+, r) \tag{A4}$$

$$f_i^{\text{beh}}(g, r, r^+) = \sum_h^{L^r} \beta_h^{\text{beh}} s_h^{\text{beh}}(i, g, r, r^+), \quad \text{(A5)}$$

where the statistics s_h are specified by the user on the basis of theory, and L^g and L^r indicate the number of unique terms in each function. These statistics may include endogenous effects of the r behavior (e.g., autocorrelation); features of the g network, such as closed triangles; or exogenous monadic and/or dyadic covariates. The β_h parameters are weights that determine the extent to which agents attempt to achieve a network-behavioral state that yields large values for the corresponding s_h statistics. For example, if the β_1^{net} parameter for the statistic s_1^{net} is positive, agents choose ties in a way that increases the calculated value of s_1^{net} in the simulated networks. If β_1^{net} is negative, agents work to decrease the value of that statistic. Table A1, reproduced from the main paper, summarizes the ABM terms, including formal definitions of each statistic.

Table A1: Summary of ABM Terms

Variable	Parameter	Name	Definition	Description
Defense spending equation				
r_i	α^{beh}	*Constant*	r_i	Baseline defense spending behavior, or cost of defense effort
d_i	γ	*DCA Degree*	$r_i \sum_j^n g_{ij}$	Effect of bilateral DCAs on i's defense effort
q_i	ψ	*DCA Dense Triads*	$r_i \sum_{j,k}^n g_{ij} g_{ik} g_{jk}$	Effect of DCA triangles on i's defense effort
z_i	η	*DCA Triads Effort*	$q_i r_i \sum_j^n g_{ij} r_j$	Effect of DCA triangles conditional on partners' defense effort
c_i	π	Monadic covariate	$r_i c_i$	Exogenous influences at the country level
DCA network equation				
$g_{i\bullet}$	α^{net}	*Density*	$\sum_j^n g_{ij}$	Baseline tendency to form ties, or cost of DCAs
a_{ij}	τ	*Total Degree*	$\sum_j^n g_{ij}(g_{j\bullet} + g_{i\bullet})$	Selection of partners based on total # of DCAs signed
b_{ij}	δ	*Transitive Triads*	$\sum_{j<k}^n g_{ij} g_{ik} g_{jk}$	Selection of partners based on closure of triangles
r_j	ζ	*Defense Spending$_j$*	$\sum_j^n g_{ij} r_j$	Selection of high-spending partners
c_j	ϕ	Monadic covariate	$\sum_j^n g_{ij} c_j$	Exogenous country-level influences on partner selection
w_{ij}	ξ	Dyadic covariate	$\sum_j^n g_{ij} w_{ij}$	Exogenous country-pair influences on partner selection

The opportunity for agent i to change either a network tie or its behavior is determined by two separate rate functions drawn from the exponential distribution with parameters $\lambda_i(\rho_g, g)$ and $\lambda_i(\rho_r, r)$. Any such opportunity for an agent to make a change is known as a "micro step" (Steglich et al. 2010: 348). The network and behavior co-evolve in continuous time as a result of a series of micro steps, where agents sequentially change their ties or behavior. Once an agent takes a micro step, the network and associated behaviors update accordingly for the next agent, such that network-behavior coevolution follows a Markov process (Steglich et al. 2010).

The ρ_g and ρ_r parameters play a crucial role in identifying equilibrium network and behavior properties (Stadtfeld et al. 2020). These are user-specified "basic parameters" that determine the average number of opportunities an agent has to make changes to its ties or behavior in a single iteration of the ABM (Snijders and Steglich 2015: 233). Higher values of ρ offer agents more opportunities for change, which in turn facilitates the emergence of statistical equilibria—

i.e., equilibrium distributions where mean behavior remains stable even as the model continues to iterate (De Marchi and Page 2014). For all results shown in the main paper, we use ρ_g and ρ_r values of at least 100. Opportunities to make changes are uniformly distributed across agents, and are equally divided between networks and behavior. Section A1.2 below shows that these values are sufficient to generate stable equilibria.

We calibrate the ABM using observed empirical data on the DCA network and defense effort for the year 2000.[1] The initial g matrix is thus the observed DCA network, where $g_{ij} = 1$ indicates a DCA in force, and the initial r matrix is country-level defense spending as a percentage of GDP, discretized at 1% increments, plus a residual category for spending above the 10% level, $[0, 0.01, \ldots, 0.1, 1]$, which yields an 11-point scale of defense effort ($M = 11$). discretization is necessary to implement the behavior choice probabilities expressed in Eq. A2. Conversion of continuous data to discrete data is common in applications where continuous data would impose a substantial computational burden, such as machine learning (Catlett 1991; Chmielewski and Grzymala-Busse 1996). In this case, a continuous behavior variable implies a virtually infinite number of choices in the probability function, which is computationally impossible. Section A3 below discusses discretization further and shows that, when extended to the empirical analysis, discretization has no substantive impact on the results.

The full ABM consists of (1) the g network and r behavior, with year 2000 values designated as the $t = 0$ initial state of a stochastic process; (2) the rate functions $\lambda(\rho, g)$ and $\lambda(\rho, r)$; and (3) the choice probabilities $P(g)$ and $P(r)$. Together, these components of the model define a continuous-time Markov chain over a discrete outcome space, where all possible combinations of network ties and behaviors constitute the state space (Steglich et al. 2010: 355).

As a Markov chain, the model can be fully described by the steps in a single iteration of the algorithm, as follows:

- Begin with time$= t_0$, and set $r = r_0$ and $g = g_0$ at observed year 2000 values
  1. For all $N = \{1, \ldots, n\}$ agents, sample waiting times from $\lambda(\rho, g)$ and $\lambda(\rho, r)$
  2. Determine from waiting times whether the next micro step will be a network or behavior change, and find the $i \in \{1, \ldots, n\}$ node with the shortest waiting time
     a. If i has an opportunity to change its network ties:
        a1. Sample prospective j targets in proportion to $f_i^{\text{net}}(g^+, r)$
        a2. If i chooses to add or remove a tie, and j accepts, then $g_{ij}^+ = 1 - g_{ij}$; otherwise, $g^+ = g$
        a3. Update the network matrix, $g = g^+$
     b. If i has an opportunity to change its behavior:
        b1. Determine the value of r_i that maximizes $f_i^{\text{beh}}(g, r^+)$
        b2. If i chooses to increase or decrease its behavior, then $r_i^+ = r_i \pm 1$; otherwise $r^+ = r$
        b3. Update the behavior matrix, $r = r^+$
  3. Increment time $=$ time $+ t$
  4. Return to step 1 and repeat until time $= t_{end}$
- Return final network and behavior matrices

[1] The choice of year is inconsequential to the results presented here.

A1.1 Parameter values

As described in the main paper and summarized above in Table A1, we specify the full behavior objective function as,

$$f_i^{\text{beh}}(g, r, r^+) = \alpha^{\text{beh}} r_i + \pi c_i + \gamma d_i + \psi q_i + \eta z_i, \tag{A6}$$

and the network objective function as,

$$f_i^{\text{net}}(g, g^+, r) = \alpha^{\text{net}} g_{i\bullet} + \phi c_i + \xi w_{ij} + \zeta r_j + \tau a_{ij} + \delta b_{ij}. \tag{A7}$$

Table A2 summarizes the parameter values used in the simulations, corresponding to Figures 5, 6, and 7 in the main paper. In setting parameter values, we draw on existing empirical research on defense spending and DCAs (e.g., Kinne 2018; Whitten and Williams 2011). In the behavior equation, r_i is agent i's current level of defense spending, and the α^{beh} parameter thus reflects governments' baseline tendency toward spending on defense. We set α^{beh} at a negative value to reflect the nonzero costs of defense products (Sandler and Hartley 2001: 873). The quantity c_i is an $n \times 1$ random variable with an exponential distribution, which represents exogenous, country-specific demands for defense spending, such as variations in national income and/or exposure to security threats (Sandler 1993). We set the corresponding π parameter at a constant positive value to reflect exogenous upward pressures on defense spending.

Table A2: ABM Terms and Parameter Profiles

Statistic	Parameter	Model 1	Model 2	Model 3	Model 4*	Model 5
Behavior (defense spending)						
r_i	α^{beh}	-0.5	-0.5	-0.5	-0.5	-0.5
d_i	γ	$[-0.05, \ldots, 0]$	$[0, \ldots, 0.05]$	0.025	0.025	0.025
q_i	ψ	0	0	$[-0.005, \ldots, 0.005]$	$[-0.005, \ldots, 0.005]$	$[-0.005, \ldots, 0]$
z_i	η	0	0	0	0	$[-0.0001, \ldots, 0]$
c_i	π	0.75	0.75	0.75	0.75	0.75
Network (DCAs)						
$g_{i\bullet}$	α^{net}	0	0	-4	-6	-4
a_{ij}	τ	0	0	0.1	0.35	0.1
b_{ij}	δ	0	0	0.5	-0.75	0.5
r_j	ζ	0	0	0.025	0.025	0.025
c_j	ϕ	0	0	0.1	0.1	0.1
w_{ij}	ξ	0	0	0.1	0.1	0.1
ABM results in...		Fig. 5(a)	Fig. 5(b)	Fig. 6(a)	Fig. 6(b)	Fig. 7

* The a_{ij} quantity is defined only as target (j) nodal degree in this model.

The crucial quantity d_i is calculated as agent i's nodal degree in the DCA network and thus corresponds to spillin—or anticipated contributions from one's partners—in the public-goods model (Conybeare et al. 1994). The associated parameter, γ, moderates the effect of DCA partnerships on i's own defense spending. As discussed in the main paper, we vary γ at incremental negative values, $\gamma \in \{-0.05, \ldots, 0\}$, in ABM Model 1, which yields standard public-goods expectations (Figure 5(a) of the main paper). In Model 2, we instead set γ at incremental positive values $\gamma \in \{0, \ldots, 0.05\}$, which accounts for detection and punishment (Figure 5(b)).

The other crucial quantity in the behavior equation is q_i, which is a count of the number of transitive triads in agent i's local network. In both ABM Model 3 and Model 4—results illustrated in Figures 6(a) and 6(b), respectively, of the main paper—we set the corresponding ψ parameter at incremental values $\psi \in \{-0.005, \ldots, 0.005\}$, which reflect a wide range of potential influence for dense local networks on country-level defense effort. The z_i statistic is a conditional variant of transitive triads, such that the influence of i's dense local network depends on the defense spending of i's j DCA partners. The parameter η determines i's triangle-based responsiveness to the spending of those partners. Considering ψ and η together, as in Figure 7 of the main paper, allows us to derive distinct hypotheses for the proposed efficiency and free-riding mechanisms, respectively.

Because the DCA and behavior equations are functions of one another, the calculated values of the q_i and z_i statistics depend on the structure of the network. And that structure is determined by the network objective function specified in Eq. A7. The $\alpha^{\text{net}} g_{i\bullet}$ term reflects agents' baseline tendency toward forming network ties. We set α^{net} at a negative value to reflect the costs of tie creation, such as negotiating and signing bilateral DCAs (Kinne 2018, 2020; Smaldino et al. 2018). Both c_j and w_{ij} are exponentially distributed random variables—one monadic, the other dyadic. We set their corresponding parameters—ϕ and ξ, respectively—at positive values to reflect exogenous upward pressures on DCA formation. Note that these exogenous terms can easily be expanded into multiple monadic and/or dyadic covariates, with differing positive and negative parameters for each, without substantively altering the results of the ABM. (The empirical analysis effectively expands these terms via inclusion of multiple monadic and dyadic covariates.)

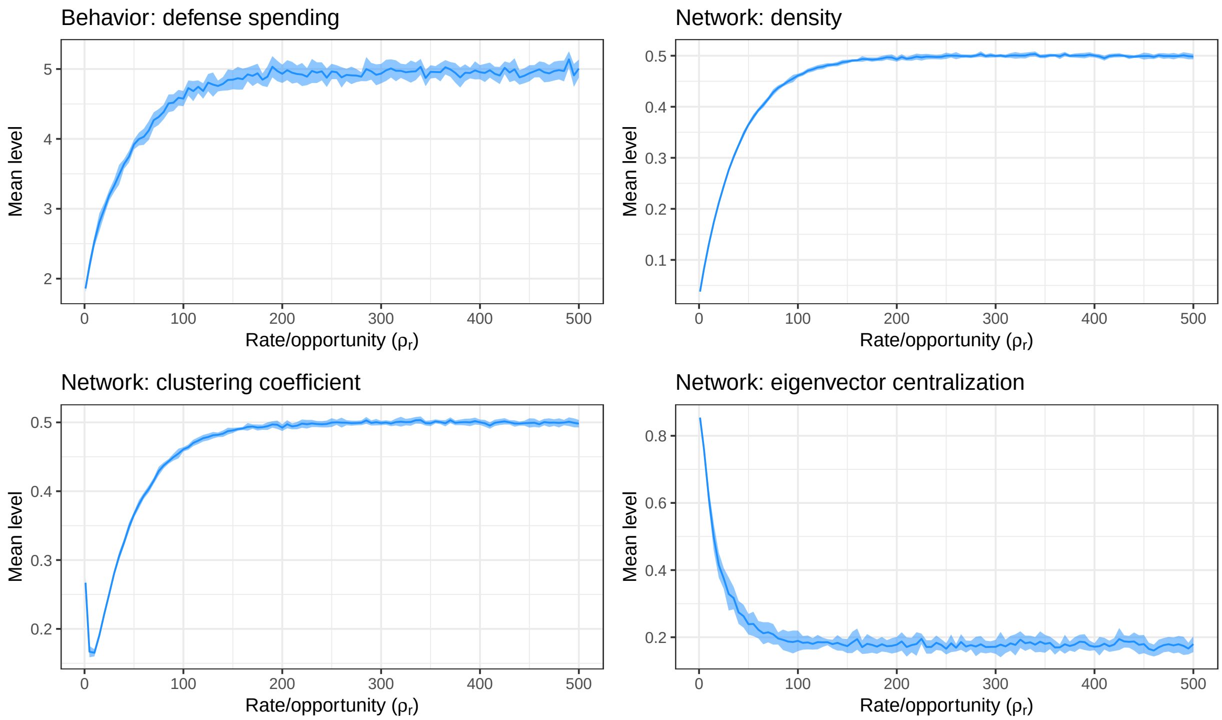
The ζr_j term reflects the tendency of agents to select partners that spend on defense at high levels. Inclusion of this term further accounts for endogenous selection-influence dynamics. That is, the model allows not only for the possibility that an accumulation of DCAs influences defense spending, but also that states select partners on the basis of their spending levels. We set the ζ parameter at a positive value to reflect the attraction of states, ceteris paribus, for high-spending nodes.

Finally, the terms τa_{ij} and δb_{ij} are endogenous network selection effects, corresponding to mutual attraction between high-degree nodes (Maoz 2012; Newman 2002) and tendencies toward transitive closure (Holland and Leinhardt 1971). In ABM Model 3 (Figure 6(a) of the main paper), which generates the main hypothesis regarding the influence of transitive triads on defense spending, we set $\tau = 0.1$ and $\delta = 0.5$. These values reflect the empirical reality that states in the DCA network prefer to form ties both to high-degree nodes and to partners of partners (Kinne 2018). In ABM Model 4 (Figure 6(b)), we specify an alternative network formation process, dominated by preferential attachment and aversion to transitive closure (Barabási and Albert 1999). In that case, we increase τ to 0.35 and decrease δ to -0.75. Note that Model 4 also defines the a_{ij} degree quantity solely in terms of the j target node's degree (i.e., as opposed to mutual ij degree), which better reflects the logic of preferential attachment (Barabási and Albert 1999).

A1.2 Equilibrium analysis

The ABM generates statistical equilibria, where macro-level patterns of network tie formation and mean behavior remain stable even as the model continues to iterate and, at the micro level, individual nodes continue to adjust their ties and behavior (De Marchi and Page 2014: 10–11). The key mechanism in achieving stable equilibria is the rate function, which determines the "rate of change" in network-behavior coevolution, or the number of opportunities that agents have to change their ties and/or behavior. In the ABM, this function is set by the parameters ρ_g for the network and ρ_r for behavior. We simulated the ABM at incremental values of the ρ parameters ranging from 1 to 500, and we found that network-behavior coevolution typically reaches an equilibrium state with rates of 100–150. We thus set ρ_g and ρ_r at no less than 100 in all ABM simulations, depending on computational feasibility (i.e., higher ρ values increase computational intensity).

Figure A1: ABM Equilibrium Analysis



Note: Results based on simulating network-behavior coevolution 10 times each at successively increasing values of ρ_r and ρ_g, and calculating behavior and network statistics across all nodes and all simulations. Dark blue lines are mean levels. Light blue polygons are 99% confidence intervals. Top-left panel shows mean level of defense spending on an $M = 11$ ordinal scale. Top-right panel shows network density, defined as the ratio of existing network ties to possible ties. Bottom-left panel shows clustering, defined as the ratio of existing triangles to possible triangles. Bottom-right panel shows normalized degree-based centralization.

Figure A1 illustrates equilibria with regard to one behavior statistic and three network statistics. The results show not only that the ABM converges relatively quickly on stable macro-level states, but that these states correspond to plausible values of the specified network properties and are not an artifact of model degeneracy (Schweinberger 2011).

A2 Stochastic Actor-Oriented Model

The empirical model, which is a stochastic actor-oriented model (SAOM) of network-behavior coevolution, combines the simulation architecture of the ABM with a method-of-moments estimator (Snijders 2001: 372). In short, the SAOM empirically validates the ABM. See Snijders (2001, 2005) for a rigorous technical treatment of the model; Snijders et al. (2010) for a minimally technical overview; Steglich et al. (2010) for the network-behavior specification of the SAOM; and Snijders and Steglich (2015) for the treatment of SAOMs as ABMs. The below summary of the model draws primarily on Snijders (2001, 2005).

As discussed in the main paper, the empirical versions of the objective functions can be written as

$$f_i^{\mathbf{g}}(\mathbf{g}, \mathbf{r}) = \sum_h^{L^g} \beta_h^{\mathbf{g}} s_{ih}^{\mathbf{g}}(\mathbf{g}, \mathbf{r}), \tag{A8}$$

and

$$f_i^{\mathbf{r}}(\mathbf{g}, \mathbf{r}) = \sum_h^{L^r} \beta_h^{\mathbf{r}} s_{ih}^{\mathbf{r}}(\mathbf{g}, \mathbf{r}), \tag{A9}$$

where $\mathbf{g}$ and $\mathbf{r}$ refer to stacked $n \times n$ and $n \times 1$ matrices, respectively, of empirical DCA and defense spending data, observed over T time periods. (L^g and L^r again indicate the number of unique terms in each function.) While the ABM simulates network-behavior coevolution under various user-specified parameter values in order to generate statistical equilibria, the SAOM instead uses empirically constrained simulations to derive expected values for the method-of-moments estimator. Specifically, the SAOM simulates network-behavior coevolution while stochastically sampling from a parameter space, with the goal of locating those $\hat{\beta}$ parameter estimates that minimize the difference between simulated and observed networks/behaviors (Steglich et al. 2010: 355–358).

The model first uses empirical data to calculate target values for each of the s_{ih} statistics in the above objective functions. These target values are calculated by summing the relevant statistics over all nodes and time periods in the data. For example, target values for the g network are defined as

$$s_h^{\text{obs}} = \sum_{t=1}^{T-1} \sum_{i=1}^{n} s_{ih}(g^{\text{obs}}(t+1)),\ (h = 1, \ldots, L^g), \tag{A10}$$

with the full set of observed L^g target values collected in the s^{obs} vector.

The method-of-moments estimator works by fitting the target values of these statistics to their expected values. Expected values are not known ex ante and must be generated through simulations. As a preliminary, define the distance between any two arbitrary networks, x and y, as the sum of their pairwise differences, $\|x - y\| = \sum_{ij} |x_{ij} - y_{ij}|$. Then define

$$c_t = \|g^{\text{obs}}(t+1) - g^{\text{obs}}(t)\|, \tag{A11}$$

which is the distance between two sequential observations of the empirical g network. For each $t = 1, \ldots, T-1$, take the initial observed network, $g^{\text{obs}}(t)$, as the starting point. Then, for a given parameter vector $\beta = (B_1, \ldots, \beta_{L^g})$ and a rate function $\lambda(\mathbf{g})$, simulate network-behavior coevolution, as described in the above presentation of the ABM, until the first time point, denoted R_t, where $\|g_t^{\text{sim}}(R_t) - g^{\text{obs}}(t)\| = c_t$. Or, in words, run the simulation until the distance between the observed network at t and the simulated network is equal to the distance between the observed network at t and the observed network at $t+1$. Then calculate the following statistics:

$$S_h = \sum_{t=1}^{T-1} \sum_{i=1}^{n} s_{ih}(g_t^{\text{sim}}(R_t)),\ (h = 1, \ldots, L^g). \tag{A12}$$

This procedure yields a random variable, $S = S_1, \ldots, S_{L^g}$. The goal of this estimation procedure is to find the $\hat{\beta}$ vector that satisfies the moment equation,

$$\varepsilon_{\hat{\beta}} S = s^{\text{obs}}, \tag{A13}$$

where the identified parameter estimates yield simulated networks identical to the observed networks. Thus, while the SAOM leverages the same simulation architecture as the ABM, the ABM only uses empirical data to calibrate the model, as a starting point for the coevolutionary process, with parameters fully determined by the user. By contrast, the SAOM simulations are constrained throughout by observed longitudinal networks and behaviors, and the parameters are estimated through method of moments.

The SAOM implements a stochastic Robbins-Monro algorithm to search the parameter space for $\hat{\beta}$ values that produce the best fit. Steglich et al. (2010: 357) document each step of the estimation algorithm. Standard errors are obtained by holding $\beta = \hat{\beta}$ and continuing to simulate the network in order to generate a matrix of simulated covariances for the statistics of the objective functions, and taking the square root of the diagonal elements. Null hypotheses can then be tested with a standard t-statistic, $t_h = \frac{\hat{\beta}_h}{\text{s.e.}(\hat{\beta}_h)}$.

A2.1 Goodness of fit

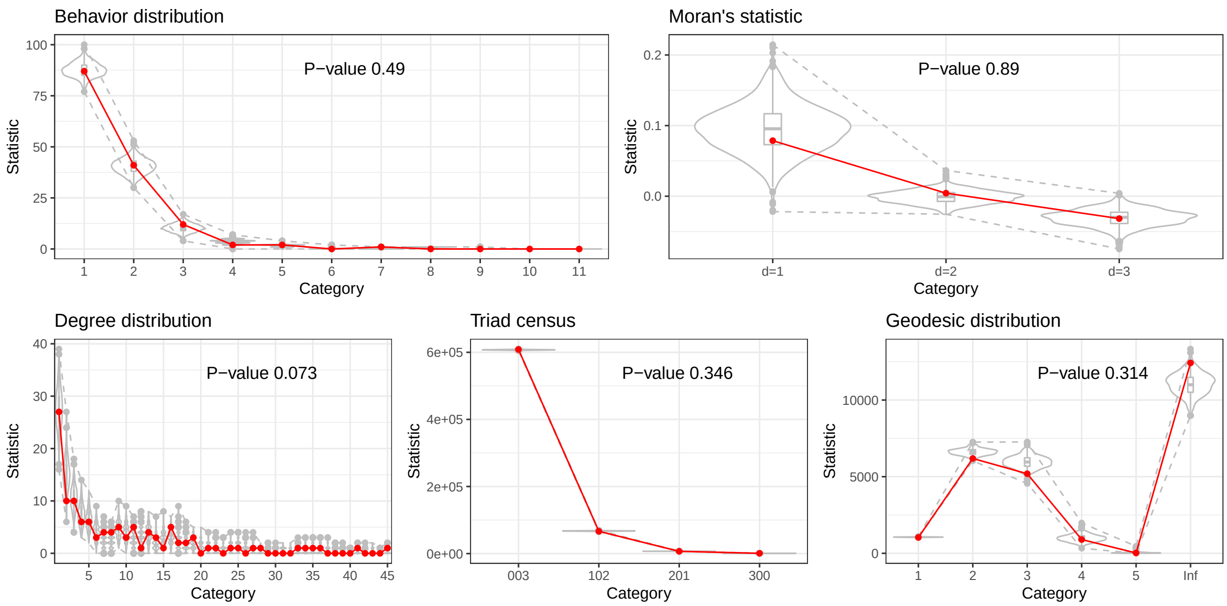
Like other inferential network models (Hunter et al. 2008b; Minhas et al. 2019), SAOMs pose unique challenges in assessing goodness of fit (GoF). Traditional regression statistics, such as R^2, cannot be calculated. The most commonly used approach, especially for models that rely on simulations, is to assess fit by comparing relevant statistical features of the simulated networks with those same features in the observed networks (Hunter et al. 2008a). Ideally, the statistics used for this purpose should not be directly included in the modeling equations. In this way, goodness of fit is a matter of determining how well the model generates simulated networks that share emergent properties and topological features with the real-world networks, even though those features are not explicitly modeled in specified equations (Lospinoso and Snijders 2019). Put differently, the model should generate well-fitting networks at the macro level via specifying network formation processes at the local level.

Figure A2 illustrates goodness of fit for our main empirical model (i.e., SAOM Model 3 in the main paper) across a range of statistics. Note that, in this context, p-values test the null hypothesis that the simulated networks do not differ significantly from the observed networks. Thus, large p-values—specifically, p-values larger than 0.05—indicate that the specified model is a good fit to the observed data. P-values in the range of 0.01–0.05 suggest potential fit problems, while p-values less than 0.01 necessitate reconsideration of the model specification.

The top-left panel of Figure A2 illustrates fit with regard to the behavior distribution, i.e., defense spending. The close match between observed and simulated values, as well as the high p-value, indicate an excellent fit. The top-right panel assesses fit with regard to network-behavior autocorrelation. This diagnostic is especially helpful in SAOMs that include both network and behavior equations (De La Haye et al. 2011; Kretschmer et al. 2018). We calculated Moran's statistic for dependence of behavior r on network g for neighborhoods of order $d \in \{1, 2, 3\}$. The fit is excellent. The dependence between DCA networks and defense spending in the real-world data closely mirrors that in the SAOM simulations. Given our primary focus on defense spending, the results of these first two diagnostics are reassuring.

The remaining panels show goodness of fit with respect to the DCA network, which is largely a

Figure A2: Goodness of Fit



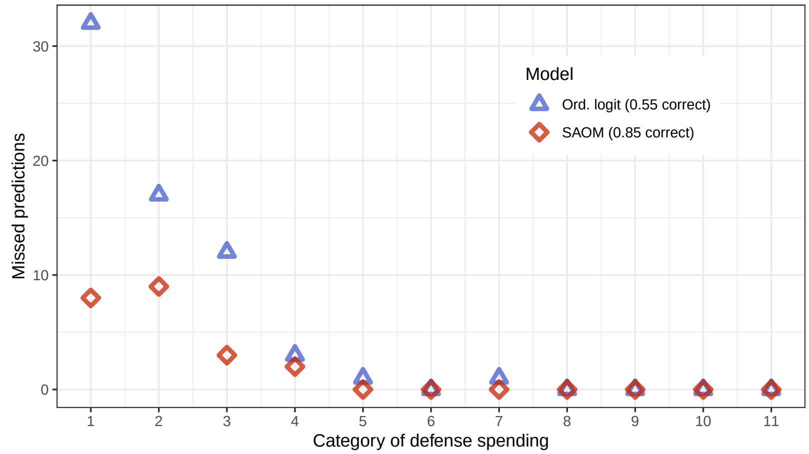
Note: Red dots and lines indicate observed values of the corresponding statistics. Gray violin plots show values in simulated networks. All results derived from Model 3 of the main paper.

matter of the specification of the network objective function. The fits are again uniformly strong, especially for the triad census and geodesic distance. The fit for the degree distribution is the weakest—though the p-value is nonetheless higher than the 0.05 threshold. Given that the other fits are extremely strong, especially with regard to behavior and network-behavior autocorrelation, the fit for degree distribution is not concerning.

We also assessed SAOM fit using out-of-sample prediction (Kinne 2013; Leifeld and Cranmer 2019). Though this approach has recently attracted controversy when applied to SAOMs (Block et al. 2018), we include it here due to the prevalence of out-of-sample validation in analysis of political phenomena (Ward 2016). We used the same method used by Kinne and Bunte (2020). We removed the final year of data from the analysis (2010), re-estimated Model 3 on the 1990–2009 data (training set), and generated out-of-sample predictions for 2010 (validation set). To compare the SAOM results to standard regression approaches, we estimated two additional models: an ordered logit model of defense effort, and a binary logit model of DCA membership with an AR1 autoregressive term.

We first compare the SAOM to the ordered logit model by comparing incorrect predictions for each category of defense spending, as illustrated in Figure A3. The SAOM predicts the true defense spending category for approximately 85% of countries. The ordered logit model, by contrast, correctly classifies only about 55% of countries. Although the logit model contains the same terms as the behavior equation of the SAOM—including terms for DCA degree and local DCA triangles—it performs very poorly. Incorporating the coevolutionary network-behavior dynamics of DCA membership and defense spending sharply improves fit. The SAOM is not merely an academic exercise but instead allows us to predict defense spending across countries with great accuracy.

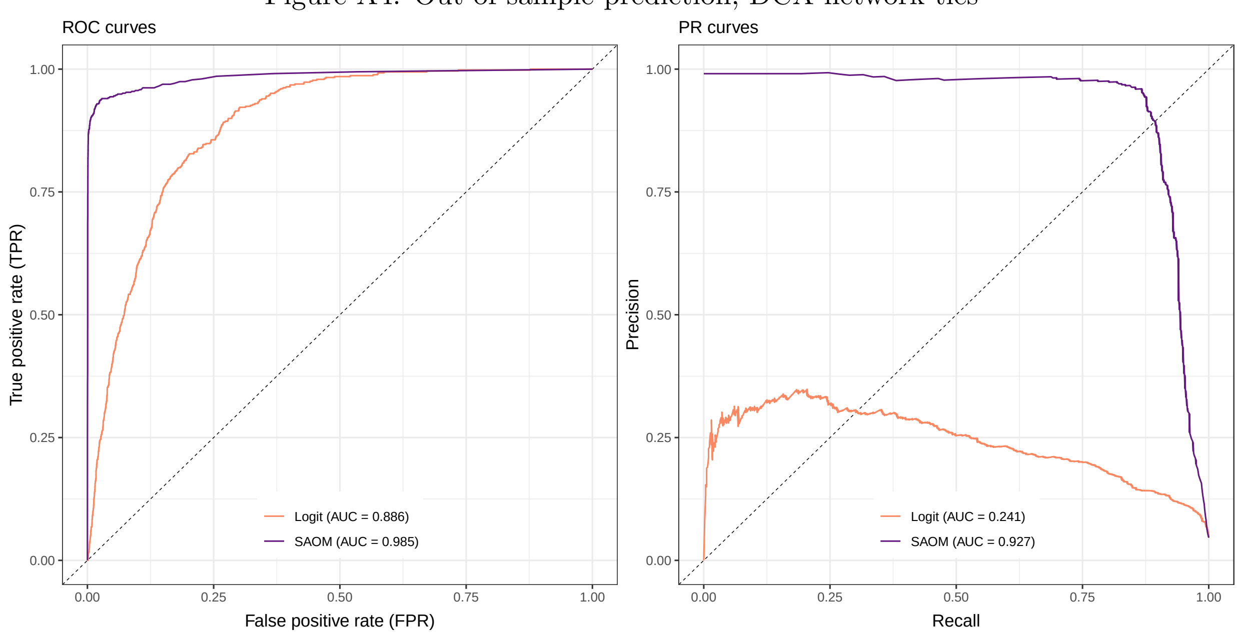
Figure A3: Out-of-sample prediction, defense spending



Note: Out-of-sample prediction of defense spending categories ($M = 11$) for the year 2010, with prior years as the training set. SAOM RMSE=0.397. Logit RMSE=0.828.

Though we are less concerned with out-of-sample prediction of DCA membership—and the predictive benefits of the network approach have been shown elsewhere (Kinne 2013; Minhas et al. 2019; Ward et al. 2007)—we also plot receiver-operating characteristic (ROC) and precision-recall (PR) curves for the out-of-sample predictions of network ties. As Figure A4 illustrates, the SAOM greatly improves on binary logit.

Figure A4: Out-of-sample prediction, DCA network ties



A3 Robustness Checks and Sensitivity Analysis

We estimated a battery of additional models to assess the robustness of our main results. Although DCAD covers the 1980–2010 period, the DCA network is sparse in the 1980s, and largely dominated by the United States. This sparseness poses a challenge for the SAOM. Further, the outsize role of the US suggests a structural break in the network-formation process between the 1980s and 1990s. We thus estimated our main models on data for the period 1990–2010. As a robustness check, we estimated a model on the 1981–2010 period.[2] Model A1 of Table A3 lists the estimates.

We also considered alternative versions of the DCA variable. The main analysis uses the "dcaGeneralV1" variable from DCAD, which includes only general DCAs coded with high confidence (Kinne 2020). We estimated a model using "dcaAnyV1," which includes both general and sector DCAs, again coded with high confidence. And we estimated a model using the most generous DCA variable, "dcaAnyV2," which includes all DCAs, whether general or sector, regardless of coding confidence. The results for these two estimations are in Models A2 and A3 of Table A3.

Model A4 returns to the specification used in the main paper, but with an additional term, *Ally Spillin*, to capture alliance-based defense contributions (Smith 1995: 72). This term is defined as the sum of the defense spending, relative to GDP, of the focal node's alliance partners.

Kinne and Bunte (2020) show that DCAs coevolve with bilateral lending. Incorporating bilateral loans as an additional network layer would entail estimating a SAOM with three separate equations—DCAs, loans, and defense spending—and is thus beyond the scope of this paper, but we nonetheless estimated a model that controls for bilateral loans (Model A5 in Table A3). We included a binary dyadic term, *Bilateral Loan*, in the network equation, and a log-transformed count of country-year bilateral loans in the behavior equation.[3]

Model A6 uses a log transformation of the defense spending variable. Because some countries spend zero dollars on defense, we must add a small value to this variable before taking the log. We add a constant equal to the smallest observed nonzero level of defense spending. We then discretize at integer values of the log-transformed metric.

To account for the potential influence of foreign-policy ideology on defense effort, Models A7 and A8 include UNGA ideal point estimates in the defense spending equation (Bailey et al. 2017). The concern here is that the apparent influence of network structure may be epiphenomenal to ideological clustering.[4] That is, if states enter into DCAs with partners that are ideologically similar to themselves, then dense local networks may be dominated by ideologically aligned states. In such a case, states may reduce spending not because of policy convergence, but because commitments among like-minded states are more credible and trustworthy, thus strengthening perceptions of security (cf. Bearce and Bondanella 2007; Deutsch et al. 1957; Macon et al. 2012; Pauls and Cranmer 2017).

We account for this possibility in two ways. First, we include monadic (i.e., country level) UNGA

[2] For the specific version of the DCA variable we use here, the network is static from 1980 to 1981, and so 1980 must be removed from the analysis.

[3] We thank an anonymous reviewer for suggesting this model.

[4] We thank the IO editors for suggesting this possibility.

ideal points as a covariate, *UNGA Ideal Point*, in the spending equation. Second, we derive a measure of the ideal point distance between i and its triangle partners. The resulting variable, *UNGA Sim. Triangles*, essentially weights i's DCA triangles by i's ideological similarity to its triangle partners. To derive this variable, we inverted the UNGA ideal-point difference metric such that larger values indicate less distance in foreign-policy ideologies. Thus, larger values of *UNGA Sim. Triangles* indicate greater ideological similarity in dense local networks.

The estimate for *UNGA Ideal Point* is positive and significant (A7), indicating that countries aligned more closely with the US-led liberal order are more likely to increase their defense effort. The estimate for *DCA-UNGA Sim. Triangles* is not statistically significant (A8). In both cases, the estimates for *DCA Degree* and *DCA Dense Triads* remain significant in the expected directions.

These results obtain, we believe, because foreign-policy ideology is only loosely related to defense policy coordination. For example, Bailey et al. (2017: 449) caution that UN votes address "issues of global importance" and, further, that "measures based on UN votes are only useful if we believe, theoretically, that a state's position on global issues matters for the outcome under consideration." Defense policy is narrowly focused on security issues of immediate interest. And while macro-level strategic policy—i.e., which threats to prioritize and how to address them—may overlap with ideology, defense policy also includes highly technical concerns like interoperability, logistics & supply, information-sharing protocols, procurement and acquisition standards, and so on. These fine-grained aspects of defense policy facilitate concrete cooperative activities in a way that broad similarities in foreign-policy preferences likely do not.

Table A3: Robustness Checks, DCAs and Defense Expenditures

	Model A1	Model A2	Model A3	Model A4	Model A5	Model A6	Model A7	Model A8
DCA Equation								
Main effects								
Transitive Triads	0.200***	0.203***	0.150***	0.199***	0.199***	0.211***	0.199***	0.196***
	(0.049)	(0.041)	(0.024)	(0.046)	(0.047)	(0.048)	(0.048)	(0.047)
Total Degree	0.050***	0.039***	0.016***	0.051***	0.051***	0.049***	0.051***	0.050***
	(0.005)	(0.004)	(0.003)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)
Defense Spending$_j$	−0.020	0.031	0.028	−0.018	−0.015	0.224*	−0.017	−0.017
	(0.049)	(0.043)	(0.031)	(0.052)	(0.052)	(0.092)	(0.050)	(0.050)
Control variables								
Distance	−0.829***	−0.857***	−0.839***	−0.838***	−0.835***	−0.841***	−0.836***	−0.834***
	(0.062)	(0.060)	(0.046)	(0.064)	(0.066)	(0.063)	(0.066)	(0.064)
Alliance (non-NATO)	0.985***	0.821***	0.581***	0.958***	0.954***	0.910***	0.959***	0.952***
	(0.122)	(0.111)	(0.086)	(0.127)	(0.125)	(0.124)	(0.124)	(0.127)
NATO	−2.945***	−3.184***	−2.430***	−3.249***	−3.244***	−3.369***	−3.252***	−3.210***
	(0.288)	(0.295)	(0.208)	(0.314)	(0.318)	(0.339)	(0.312)	(0.309)
UNGA Ideal Point Diff.	−0.555***	−0.593***	−0.371***	−0.542***	−0.545***	−0.621***	−0.542***	−0.538***
	(0.061)	(0.060)	(0.044)	(0.061)	(0.069)	(0.073)	(0.069)	(0.065)
Trade	0.086**	0.086***	0.021	0.072**	0.070*	0.072**	0.071**	0.072*
	(0.027)	(0.025)	(0.018)	(0.028)	(0.029)	(0.026)	(0.027)	(0.028)
Bilateral Loan					0.064			
					(0.130)			
Democracy$_j$	1.527***	1.338***	0.912***	1.451***	1.454***	1.523***	1.454***	1.441***
	(0.223)	(0.177)	(0.113)	(0.215)	(0.218)	(0.224)	(0.231)	(0.210)
Capabilities$_j$	0.828***	0.757***	0.772***	0.793***	0.788***	0.768***	0.792***	0.789***
	(0.061)	(0.057)	(0.045)	(0.060)	(0.066)	(0.061)	(0.065)	(0.064)
Density	−5.669***	−5.006***	−3.804***	−5.373***	−5.364***	−5.354***	−5.363***	−5.354***
	(0.195)	(0.168)	(0.094)	(0.196)	(0.205)	(0.205)	(0.208)	(0.197)
Defense Spending Equation								
Main effects								
DCA Degree	0.069***	0.075***	0.046***	0.085***	0.080***	0.059**	0.070**	0.080***
	(0.019)	(0.020)	(0.014)	(0.022)	(0.023)	(0.022)	(0.023)	(0.023)
DCA Dense Triads.	−0.019**	−0.016**	−0.007*	−0.022**	−0.020**	−0.013*	−0.018*	−0.024**
	(0.006)	(0.006)	(0.003)	(0.007)	(0.007)	(0.006)	(0.007)	(0.009)
Control variables								
Democracy	−0.366***	−0.490***	−0.448***	−0.481***	−0.499***	−0.206	−0.839***	−0.495***
	(0.109)	(0.134)	(0.132)	(0.130)	(0.129)	(0.113)	(0.151)	(0.131)
GDP Growth	−1.372*	−2.126*	−2.197**	−2.106*	−2.072*	−1.063	−1.656*	−2.128*
	(0.614)	(0.829)	(0.836)	(0.831)	(0.839)	(0.898)	(0.830)	(0.870)
Allies (non-NATO)	−0.008**	−0.009**	−0.009**	−0.009**	−0.009**	−0.003	−0.012***	−0.009**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
NATO Member	0.884***	0.854***	0.879***	0.846***	0.796***	0.416*	0.558***	0.835***
	(0.126)	(0.161)	(0.149)	(0.155)	(0.172)	(0.175)	(0.167)	(0.159)
Military Regime	0.059	0.252	0.242	0.244	0.249	0.459*	0.353	0.240
	(0.137)	(0.186)	(0.194)	(0.187)	(0.184)	(0.229)	(0.190)	(0.188)
MIDs	0.070	0.153	0.137	0.163	0.142	0.173	0.101	0.163
	(0.074)	(0.084)	(0.087)	(0.090)	(0.087)	(0.105)	(0.091)	(0.088)
Spatial Lag	8.763***	7.364**	7.398**	5.315	7.375**	17.611***	8.140**	7.331**
	(2.219)	(2.702)	(2.765)	(3.490)	(2.800)	(4.861)	(2.880)	(2.758)
Ally Spillin				3.749				
				(3.695)				
Bilateral Loans					0.004			
					(0.004)			
UNGA Ideal Point							0.415***	
							(0.085)	
UNGA Sim. Triangle								0.004
								(0.008)
Constant	−1.037***	−1.155***	−1.152***	−1.150***	−1.144***	−0.198**	−1.159***	−1.113***
	(0.056)	(0.074)	(0.076)	(0.076)	(0.074)	(0.071)	(0.080)	(0.101)
Constant^2	0.007	−0.033	−0.035	−0.036	−0.035	−0.486***	−0.046	−0.033
	(0.014)	(0.024)	(0.025)	(0.024)	(0.024)	(0.048)	(0.027)	(0.023)
Iterations	2,971	2,640	2,640	2,661	2,677	2,640	2,661	2,661

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

As discussed above, both the ABM and the empirical model require an ordinal behavior dependent variable. Modeling the coevolution of a discrete network and a continuous individual behavior poses formidable methodological challenges; to our knowledge, there are not yet established empirical models capable of accomplishing this task.[5] This methodological limitation sets up a clear trade-off between a traditional regression approach and the SAOM. While a linear regression model would of course allow for a continuous spending variable, regression models assume identical, independently distributed (i.i.d.) observations and therefore cannot account for essential aspects of the data generating process, including but not limited to (1) coevolution between network ties and individual behavior; (2) endogenous influences within the DCA network itself, such as transitivity and degree effects; and (3) endogenous influences on individual behavior, particularly network-dependent responses to the behavior of other nodes. These phenomena comprise the core of our theory. Further, the SAOM analysis shows that they play a central role in real-world burden sharing. In short, not only would a standard regression model prevent us from testing the hypotheses, but it would also suffer from severe omitted variable bias.

Scholars routinely face data constraints when attempting to model dynamic, highly complex social processes, as the available methodologies often utilize systems of interdependent equations and computationally intensive simulation algorithms. Discretization of continuous data is required for many applications in neural networks (Kim and Han 2000), data mining (Liu et al. 2002), and machine learning (Catlett 1991; Chmielewski and Grzymala-Busse 1996; Dougherty et al. 1995). Discretization is also common in high-profile political science research (e.g., Beck et al. 2002; Enamorado et al. 2019; Gelman and Park 2009; Hainmueller et al. 2019; Knox and Lucas 2021; Marquardt and Pemstein 2018; Weschle 2018). The growing body of research that uses SAOMs to model the coevolution of social networks and individual behavior likewise relies on conversion of continuous to discrete data (e.g., Berardo 2013; Chyzh 2016; Kinne 2016; Manger and Pickup 2016; Wang and Yang 2019; Warren 2016).

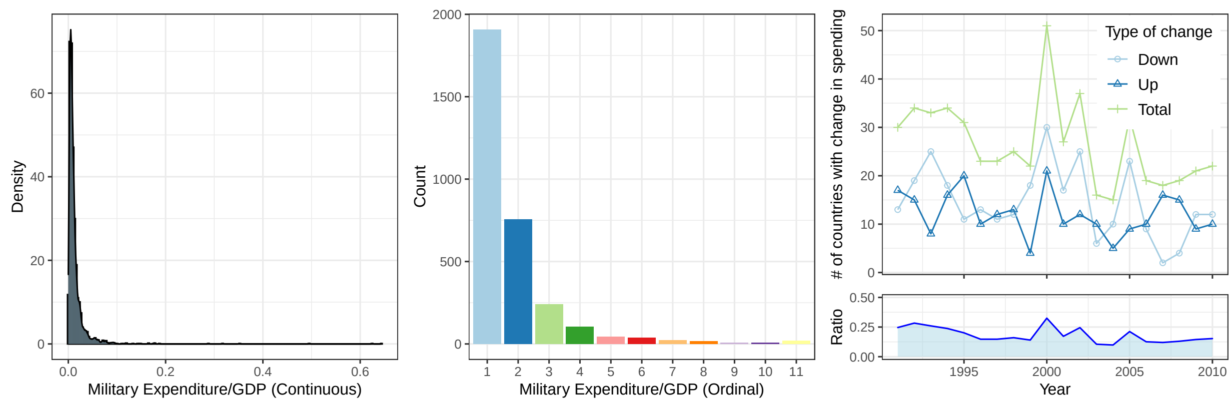
The two most potentially serious problems with discretization are information loss and arbitrary cutpoints (Niezink et al. 2019). To assess the sensitivity of our results to these two issues, we implement a wide variety of discretization techniques, which vary both the cutpoints imposed and the amount of information lost. We also estimate a simple linear regression model of (continuous) defense spending and show that the same divergent DCA effects discussed in the main paper appear in this simpler model.

We first note that our preferred discretization, which bins defense spending into ordinal categories of defense effort based on 1% spending increments, plus a residual category for spending over 10%, has a number of strengths. Using absolute values as cut points retains the underlying distribution of the data, as recommended by Ripley et al. (2021), while also limiting the number of categories to a manageable level. As Figure A5 shows, the skewed distribution of the discretized data mirrors the skewed continuous distribution, which reflects the empirical reality that most countries spend relatively little on defense. Further, binning the small number of extreme values (i.e., for spending over 10%) in a single category minimizes outlier influence. Finally, as illustrated in the right-hand panel of Figure A5, the discretized data exhibit a significant amount of over-time variation. From one year to the next, anywhere from 10% to 30% of countries change their defense spending behavior, with decreases and increases occurring with similar frequency. And because the data

[5] Scholars have recently begun considering precisely this problem (e.g., Niezink and Snijders 2017; Niezink et al. 2019). However, these approaches are exploratory, not widely used, computationally intensive, and frequently suffer from problems like nonconvergence of estimation algorithms.

are binned, these changes likely represent substantive shifts in defense policy rather than trivial changes due to budgetary politics, economic conditions, or other unrelated influences. Overall, the discretized data offer strong face validity as a measure of defense effort.

Figure A5: Ordinal versus Continuous Data



We first considered the sensitivity of the main results to increases and decreases in the number of ordinal categories. Model A9 of Table A4 shows estimates from a model with only five categories of defense effort ($M = 5$), which reduces information even further than our preferred $M = 11$ discretization. Models A10 and A11, by contrast, increase the number of categories to 15 and 25, respectively, which retains more information. In Model A12, we implement a "semi-continuous" specification that uses expenditure data at three decimal places of precision to define over 100 categories of defense effort. The high precision of this metric reduces information loss to trivial levels. This model is extremely computationally intensive and is thus not a practical replacement for the models specified in the main paper, but the results nonetheless support the conclusion that the estimated relationship between network structure and defense spending is not an artifact of information loss. Indeed, for all of these models, the signs of the estimates are unchanged, and the estimates remain significant at conventional levels.

We next considered discretization applied to total defense spending, calculated in constant 2010 US dollars, rather than spending as a percentage of GDP. This approach defines defense effort in terms of a wholly different metric and thus provides a novel set of sensitivity checks on cutpoints and information loss. Model A13 discretizes total defense spending into $M = 11$ categories, and Model A14 discretizes the log transformation of total defense spending at integer values. Interestingly, these specifications affect the estimates for many of the covariates, rendering them weaker or even insignificant. Yet, the estimated relationships between DCAs and defense spending remain stable.

Model A15 implements discretization at decile values, which provides yet another perspective on cutpoint sensitivity and information loss. We disfavor decile discretization for two reasons. First, the true observed distribution of the continuous defense spending variable is right skewed—or, in the case of logged values, normally distributed. Binning observations by decile instead imposes a uniform distribution. Second, individual countries may shift deciles even if they make no change to their defense effort. For example, if other countries substantially increase spending, then a focal country that leaves its spending unchanged may move to a lower decile, implying a change in spending. Given that we explicitly model decreases and increases in defense spending, this feature of the decile data may produce estimates that are misleading, substantively ambiguous, or simply biased. Nonetheless, the estimates from this model clearly show the same basic pattern found elsewhere. (Results for log-transforming the data before binning into deciles are identical to those shown in Model A15, as the log transformation does not affect decile binning.)

Table A4: Sensitivity Analysis of discretization Methods

	Model A9	Model A10	Model A11	Model A12	Model A13	Model A14	Model A15
DCA Equation							
Main effects							
Transitive Triads	0.197***	0.193***	0.200***	0.202***	0.202***	0.240***	0.385***
	(0.044)	(0.049)	(0.049)	(0.048)	(0.047)	(0.053)	(0.106)
Total Degree	0.051***	0.051***	0.050***	0.051***	0.051***	0.047***	0.048***
	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
Defense Spending$_j$	−0.245	−0.028	0.023	0.001	0.019	0.658***	0.963***
	(0.173)	(0.054)	(0.023)	(0.005)	(0.059)	(0.163)	(0.226)
Control variables							
Distance	−0.841***	−0.832***	−0.831***	−0.838***	−0.839***	−0.897***	−1.082***
	(0.067)	(0.069)	(0.066)	(0.067)	(0.063)	(0.072)	(0.123)
Alliance (non-NATO)	0.987***	0.955***	0.925***	0.941***	0.950***	0.915***	1.475***
	(0.129)	(0.118)	(0.124)	(0.125)	(0.125)	(0.126)	(0.199)
NATO	−3.237***	−3.186***	−3.251***	−3.269***	−3.277***	−3.624***	−4.676***
	(0.325)	(0.311)	(0.316)	(0.321)	(0.313)	(0.364)	(0.685)
UNGA Ideal Point Diff.	−0.518***	−0.530***	−0.570***	−0.554***	−0.556***	−0.706***	−0.674***
	(0.064)	(0.068)	(0.064)	(0.071)	(0.066)	(0.079)	(0.085)
Trade	0.068*	0.072*	0.074**	0.073**	0.070*	0.042	−0.042
	(0.027)	(0.028)	(0.027)	(0.027)	(0.027)	(0.028)	(0.038)
Democracy$_j$	1.425***	1.415***	1.463***	1.464***	1.454***	1.467***	1.713***
	(0.213)	(0.212)	(0.221)	(0.223)	(0.209)	(0.216)	(0.317)
Capabilities$_j$	0.805***	0.788***	0.773***	0.786***	0.769***	0.507***	0.252**
	(0.062)	(0.064)	(0.062)	(0.064)	(0.088)	(0.093)	(0.090)
Density	−5.403***	−5.360***	−5.330***	−5.363***	−5.362***	−5.470***	−7.152***
	(0.198)	(0.201)	(0.193)	(0.203)	(0.202)	(0.209)	(0.777)
Defense Spending Equation							
Main effects							
DCA Degree	0.175**	0.089***	0.025**	0.007*	0.070**	0.134***	0.059**
	(0.061)	(0.022)	(0.009)	(0.003)	(0.023)	(0.037)	(0.020)
DCA Triads	−0.063*	−0.023**	−0.006*	−0.002*	−0.017**	−0.033**	−0.014**
	(0.025)	(0.007)	(0.003)	(0.001)	(0.006)	(0.011)	(0.005)
Control variables							
Democracy	−1.239**	−0.517***	−0.083	−0.016	0.008	0.019	0.015
	(0.426)	(0.132)	(0.048)	(0.013)	(0.104)	(0.142)	(0.085)
GDP Growth	−4.174	−2.260**	−0.399	−0.278***	1.213	1.899	1.545**
	(2.183)	(0.824)	(0.321)	(0.083)	(0.717)	(1.049)	(0.573)
Allies (non-NATO)	−0.012	−0.010**	−0.001	−0.000	−0.002	−0.003	0.001
	(0.009)	(0.003)	(0.001)	(0.000)	(0.003)	(0.004)	(0.002)
NATO Member	0.910	0.883***	0.176*	0.043*	0.223	0.433	0.194
	(0.528)	(0.166)	(0.071)	(0.022)	(0.165)	(0.246)	(0.146)
Military Regime	0.502	0.250	0.210*	0.051	0.166	0.225	0.150
	(0.385)	(0.178)	(0.089)	(0.029)	(0.190)	(0.285)	(0.145)
MIDs	0.395	0.124	0.059	0.003	0.044	0.265	0.108
	(0.206)	(0.090)	(0.041)	(0.012)	(0.095)	(0.139)	(0.080)
Spatial Lag	20.514**	5.876*	2.447	0.294	0.969	5.017	−1.695
	(7.549)	(2.737)	(1.526)	(0.365)	(2.987)	(5.093)	(2.619)
Constant	−3.245***	−1.147***	−0.183***	−0.058***	−0.215***	−0.177	−0.041
	(0.245)	(0.074)	(0.027)	(0.009)	(0.063)	(0.091)	(0.051)
Constant^2	0.193	−0.040	−0.039***	−0.002***	−0.049**	−0.322***	−0.001
	(0.117)	(0.024)	(0.005)	(0.000)	(0.015)	(0.052)	(0.009)
Iterations	2,640	2,640	2,640	3,416	2,640	2,640	2,640

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

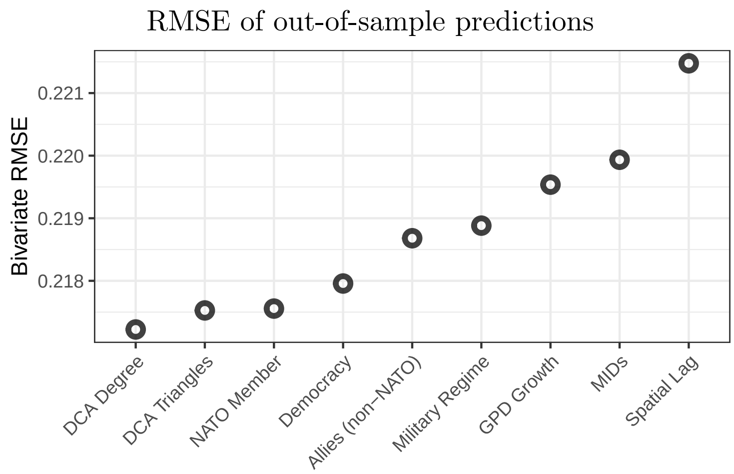
As a final sensitivity check, we estimated a standard linear regression model of defense spending with a (log transformed) continuous dependent variable, and with the DCA network terms specified as exogenous covariates. As with regression models generally, this model assumes i.i.d. random variables. Thus, given the myriad known dependencies in the data, as well as endogenous terms on the right-hand-side of the equation, this model is misspecified. Nonetheless, the goal of this analysis is to assess whether the network terms improve model fit in a regression model in the same way they do in the SAOM.

Figure A6: Linear regression model of defense spending

Regression model

Variable	Estimate (s.e.)
DCA Degree	0.007 (0.003)**
DCA Dense Triads	-0.002 (0.001)*
Democracy	-0.024 (0.014)
GDP Growth	-0.080 (0.088)
Allies (non-NATO)	-0.000 (0.000)
NATO Member	0.048 (0.022)*
Military Regime	0.060 (0.027)*
MIDs	0.019 (0.013)
Spatial Lag	1.305 (0.516)*
Lagged Spending	0.887 (0.008)***
Intercept	0.314 (0.028)***
N	2,944

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$



The left panel of Figure A6 displays the estimated regression coefficients and standard errors. Despite bias in parameter estimates, the signs of the estimates reveal the same basic pattern as in the SAOM. Bilateral DCAs increase defense spending while DCA triangles reduce spending. To assess improvement in fit, we again turn to out-of-sample prediction. We first specify a "baseline" model that includes only a one-period lag of the dependent variable on the right-hand-side of the regression equation. We then individually add model terms and assess each variable's unique contribution to model fit. As the right-hand panel of Figure A6 illustrates, the two network terms, *DCA Degree* and *DCA Dense Triads*, produce the greatest improvement in model fit (i.e., in terms of the root-mean-square error (RMSE) of the out-of-sample predictions). Overall, even with a misspecified model, we find evidence of a divergent influence for DCAs on defense effort, and we find that the network terms substantially improve model fit.

A4 Table of Main Results

Table A5 contains the full, unscaled estimates from SAOM Models 1, 2, 3, and 4 in the main paper. The table also includes convergence diagnostics for each parameter estimate, denoted Φ in the accompanying columns. As a general guideline, values of Φ below 0.1 indicate excellent convergence of the estimation algorithm (Ripley et al. 2021). All diagnostics fall well below that threshold.

Table A5: Stochastic Actor-Oriented Model of DCAs and Defense Expenditures

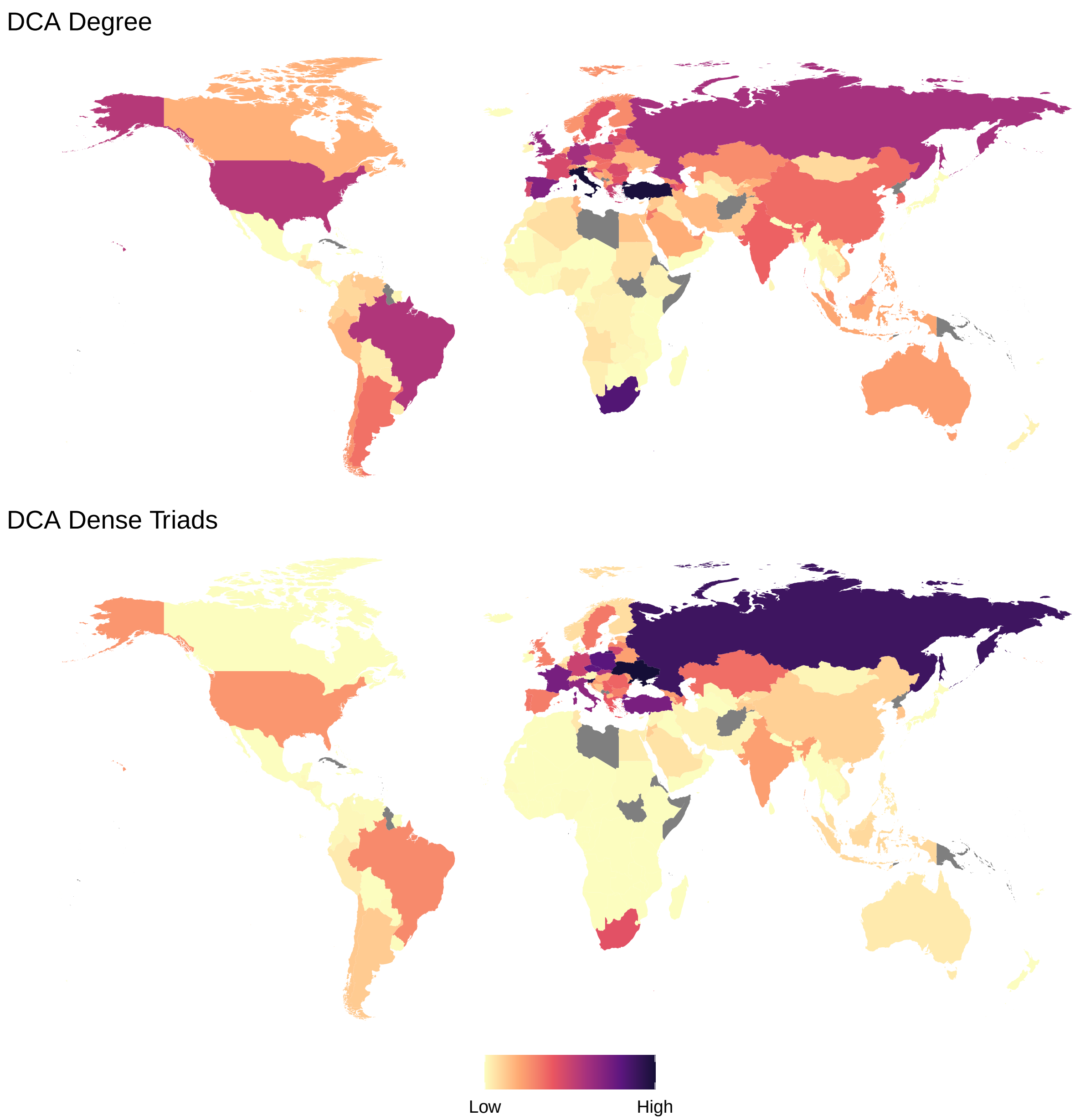
	Model 1	Φ	Model 2	Φ	Model 3	Φ	Model 4	Φ
DCA Equation								
Main effects								
Transitive Triads			0.198***	−0.04	0.199***	0.009	0.198***	0.005
			(0.047)		(0.049)		(0.047)	
Total Degree			0.051***	−0.021	0.051***	0.016	0.051***	0.013
			(0.005)		(0.005)		(0.005)	
Defense Spending$_j$	0.015	0.011	−0.020	0.001	−0.020	0.009	−0.020	−0.001
	(0.054)		(0.053)		(0.052)		(0.050)	
Control variables								
Distance	−0.944***	−0.003	−0.834***	0.033	−0.835***	0.012	−0.835***	0.003
	(0.067)		(0.066)		(0.065)		(0.067)	
Alliance (non-NATO)	1.398***	−0.003	0.956***	−0.009	0.957***	−0.006	0.956***	−0.001
	(0.141)		(0.123)		(0.123)		(0.126)	
NATO	−2.503***	−0.01	−3.237***	−0.03	−3.246***	−0.012	−3.239***	0.018
	(0.268)		(0.321)		(0.312)		(0.318)	
UNGA Ideal Point Diff.	−0.507***	0.009	−0.539***	0.013	−0.539***	0.017	−0.539***	−0.018
	(0.067)		(0.066)		(0.065)		(0.066)	
Trade	0.045	0.013	0.072**	−0.023	0.073**	0.012	0.073*	0.001
	(0.030)		(0.027)		(0.028)		(0.028)	
Democracy$_j$	2.238***	−0.021	1.438***	0.02	1.442***	0.027	1.441***	−0.002
	(0.239)		(0.216)		(0.216)		(0.219)	
Capabilities$_j$	0.946***	0.003	0.789***	−0.018	0.790***	0.007	0.790***	0
	(0.062)		(0.063)		(0.062)		(0.064)	
Density	−3.984***	0	−5.367***	0.009	−5.371***	0.01	−5.364***	0.018
	(0.196)		(0.197)		(0.193)		(0.205)	
Defense Spending Equation								
Main effects								
DCA Degree	0.016	0.004	0.016	0.012	0.085***	0.025	0.085***	0.022
	(0.009)		(0.008)		(0.022)		(0.022)	
DCA Dense Triads					−0.022**	0.015	−0.021**	0.022
					(0.007)		(0.007)	
DCA Triads Effort							0.001	−0.008
							(0.000)	
Control variables								
Democracy	−0.445***	−0.005	−0.445***	0.004	−0.498***	0.02	−0.494***	0.038
	(0.131)		(0.128)		(0.132)		(0.131)	
GDP Growth	−2.076*	−0.026	−2.072*	−0.003	−2.092*	−0.017	−2.162*	0
	(0.819)		(0.808)		(0.834)		(0.842)	
Allies (non-NATO)	−0.007*	−0.021	−0.007*	0.041	−0.009**	−0.005	−0.010**	0.027
	(0.003)		(0.003)		(0.003)		(0.003)	
NATO Member	0.900***	−0.038	0.903***	0.02	0.852***	0.019	0.899***	0.022
	(0.156)		(0.157)		(0.164)		(0.162)	
Military Regime	0.224	−0.002	0.219	0.003	0.235	−0.01	0.241	−0.012
	(0.191)		(0.184)		(0.190)		(0.191)	
MIDs	0.225**	0.018	0.226**	−0.009	0.159	0.016	0.161	0
	(0.083)		(0.086)		(0.088)		(0.086)	
Spatial Lag	6.791*	−0.004	6.776*	−0.007	7.330**	0.034	7.461**	0.019
	(2.736)		(2.817)		(2.803)		(2.860)	
Constant	−1.049***	−0.004	−1.047***	0.016	−1.148***	0.061	−1.153***	0.003
	(0.071)		(0.069)		(0.075)		(0.076)	
Constant^2	−0.030	−0.012	−0.030	−0.013	−0.034	0.016	−0.036	0.006
	(0.024)		(0.023)		(0.024)		(0.024)	
Iterations	9,666		9,858		9,862		9,992	

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

A5 Variable importance by country

Figure A7 shows the relative importance of DCA network variables for individual countries. For some states, such as the less integrated countries of sub-Saharan Africa, *DCA Degree* exercises substantial influence while *DCA Dense Triads* plays a minor role. By contrast, for many countries in regions characterized by dense webs of overlapping defense partnerships, such as Central and Eastern Europe, transitive triads are more influential. Nonetheless, for most countries both factors are important determinants of defense spending. Overall, the importance of the DCA variables indicates that the complex, multilevel relationship between DCAs and defense expenditures is not merely a statistical curiosity. Rather, this relationship is a key driver of states' defense efforts.

Figure A7: Importance of DCA Network Variables by Country, 2010



Note: Variable importance as calculated by the Indlekofer-Brandes method. Estimates based on Model 3.

Appendix References

Bailey, M. A., Strezhnev, A., & Voeten, E. (2017). Estimating dynamic state preferences from united nations voting data. *Journal of Conflict Resolution*, *61*(2), 430–456.

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509–512.

Bearce, D. H., & Bondanella, S. (2007). Intergovernmental organizations, socialization, and member-state interest convergence. *International Organization*, *61*(4), 703–733.

Beck, P. A., Dalton, R. J., Greene, S., & Huckfeldt, R. (2002). The social calculus of voting: Interpersonal, media, and organizational influences on presidential choices. *American Political Science Review*, *96*(1), 57–73.

Berardo, R. (2013). The coevolution of perceptions of procedural fairness and link formation in self-organizing policy networks. *Journal of Politics*, *75*(3), 686–700.

Block, P., Koskinen, J., Hollway, J., Steglich, C., & Stadtfeld, C. (2018). Change we can believe in: Comparing longitudinal network models on consistency, interpretability and predictive power. *Social Networks*, *52*, 180–191.

Catlett, J. (1991). On changing continuous attributes into ordered discrete attributes. *European working session on learning*, 164–178.

Chmielewski, M. R., & Grzymala-Busse, J. W. (1996). Global discretization of continuous attributes as preprocessing for machine learning. *International Journal of Approximate Reasoning*, *15*(4), 319–331.

Chyzh, O. (2016). Dangerous liaisons: An endogenous model of international trade and human rights. *Journal of Peace Research*, *53*(3), 409–423.

Conybeare, J. A., Murdoch, J. C., & Sandler, T. (1994). Alternative Collective-Goods Models of Military Alliances: Theory and Empirics. *Economic Inquiry*, *32*(4), 525–542.

De La Haye, K., Robins, G., Mohr, P., & Wilson, C. (2011). Homophily and contagion as explanations for weight similarities among adolescent friends. *Journal of Adolescent Health*, *49*(4), 421–427.

De Marchi, S., & Page, S. E. (2014). Agent-based models. *Annual Review of Political Science*, *17*, 1–20.

Deutsch, K. W., Burrell, S. A., Kann, R. A., Lee Jr., M., Lichterman, M., Lindgren, R. E., Loewenheim, F. L., & Van Wagenen, R. W. (1957). *Political Community and the {N}orth {A}tlantic Area: International Organization in the Light of Historical Experience*. Princeton University Press.

Dougherty, J., Kohavi, R., & Sahami, M. (1995). Supervised and unsupervised discretization of continuous features. *Machine learning proceedings 1995* (194–202). Elsevier.

Enamorado, T., Fifield, B., & Imai, K. (2019). Using a probabilistic model to assist merging of large-scale administrative records. *American Political Science Review*, *113*(2), 353–371.

Gelman, A., & Park, D. K. (2009). Splitting a Predictor at the Upper Quarter or Third and the Lower Quarter or Third. *The American Statistician*, *63*(1), 1–8.

Hainmueller, J., Mummolo, J., & Xu, Y. (2019). How much should we trust estimates from multiplicative interaction models? simple tools to improve empirical practice. *Political Analysis*, *27*(2), 163–192.

Holland, P. W., & Leinhardt, S. (1971). Transitivity in structural models of small groups. *Comparative Group Studies*, *2*(2), 107–124.

Hunter, D. R., Goodreau, S. M., & Handcock, M. S. (2008a). Goodness of fit of social network models. *Journal of the American Statistical Association*, *103*(481), 248–258.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008b). ERGM: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, *24*(3), 1–29.

Kim, K.-j., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert systems with Applications*, *19*(2), 125–132.

Kinne, B. J. (2013). Network dynamics and the evolution of international cooperation. *American Political Science Review*, *107*(4), 766–785.

Kinne, B. J. (2016). Agreeing to arm: Bilateral weapons agreements and the global arms trade. *Journal of Peace Research*, *53*(3), 359–377.

Kinne, B. J. (2018). Defense cooperation agreements and the emergence of a global security network. *International Organization*, *72*(4), 799–837.

Kinne, B. J. (2020). The defense cooperation dataset (dcad). *Journal of Conflict Resolution*, *64*(4), 729–755.

Kinne, B. J., & Bunte, J. B. (2020). Guns or money? defense co-operation and bilateral lending as coevolving networks. *British Journal of Political Science*, *50*(3), 1067–1088.

Knox, D., & Lucas, C. (2021). A dynamic model of speech for the social sciences. *American Political Science Review*, *115*(2), 649–666.

Kretschmer, D., Leszczensky, L., & Pink, S. (2018). Selection and influence processes in academic achievement: More pronounced for girls? *Social Networks*, *52*, 251–260.

Leifeld, P., & Cranmer, S. J. (2019). A theoretical and empirical comparison of the temporal exponential random graph model and the stochastic actor-oriented model. *Network science*, *7*(1), 20–51.

Liu, H., Hussain, F., Tan, C. L., & Dash, M. (2002). Discretization: An enabling technique. *Data Mining and Knowledge Discovery*, *6*(4), 393–423.

Lospinoso, J., & Snijders, T. A. (2019). Goodness of fit for stochastic actor-oriented models. *Methodological Innovations*, *12*(3), 1–18.

Macon, K. T., Mucha, P. J., & Porter, M. A. (2012). Community structure in the United Nations General Assembly. *Physica A: Statistical Mechanics and its Applications*, *391*(1-2), 343–361.

Maddala, G. S. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press.

Manger, M. S., & Pickup, M. A. (2016). The Coevolution of Trade Agreement Networks and Democracy. *Journal of Conflict Resolution*, *60*(1), 164–191.

Manger, M. S., Pickup, M. A., & Snijders, T. A. (2012). A Hierarchy of Preferences: A Longitudinal Network Analysis Approach to PTA Formation. *Journal of Conflict Resolution*, *56*(5), 853–878.

Maoz, Z. (2012). Preferential Attachment, Homophily, and the Structure of International Networks, 1816–2003. *Conflict Management and Peace Science*, *29*(3), 341–369.

Marquardt, K. L., & Pemstein, D. (2018). Irt models for expert-coded panel data. *Political Analysis*, *26*(4), 431–456.

Minhas, S., Hoff, P. D., & Ward, M. D. (2019). Inferential Approaches for Network Analysis: AMEN for Latent Factor Models. *Political Analysis*, *27*(2), 208–222.

Newman, M. E. (2002). Assortative mixing in networks. *Physical Review Letters*, *89*(20), 208701.

Niezink, N. M., & Snijders, T. A. (2017). Co-Evolution of Social Networks and Continuous Actor Attributes. *The Annals of Applied Statistics*, *11*(4), 1948–1973.

Niezink, N. M., Snijders, T. A., & van Duijn, M. A. (2019). *No Longer Discrete: Modeling the Dynamics of Social Networks and Continuous Behavior* (Vol. 49).

Pauls, S. D., & Cranmer, S. J. (2017). Affinity communities in United Nations voting: Implications for democracy, cooperation, and conflict. *Physica A: Statistical Mechanics and its Applications*, *484*, 428–439.

Ripley, R. M., Snijders, T. A., Boda, Z., Vörös, A., & Preciado, P. (2021). Manual for RSiena University of Oxford, Department of Statistics, Nuffield College.

Sandler, T. (1993). The economic theory of alliances: A survey. *The Journal of Conflict Resolution*, *37*(3), 446–483.

Sandler, T., & Hartley, K. (2001). Economics of Alliances: The Lessons for Collective Action. *Journal of Economic Literature*, *39*(3), 869–896.

Schweinberger, M. (2011). Instability, sensitivity, and degeneracy of discrete exponential families. *Journal of the American Statistical Association*, *106*(496), 1361–1370.

Smaldino, P. E., D'Souza, R. M., & Maoz, Z. (2018). Resilience by structural entrenchment: Dynamics of single-layer and multiplex networks following sudden changes to tie costs. *Network Science*, *6*(2), 157–175.

Smith, R. (1995). The demand for military expenditure. *Handbook of Defense Economics*, *1*, 69–87.

Snijders, T. A. (2001). The Statistical Evaluation of Social Network Dynamics. *Sociological Methodology*, *31*, 361–395.

Snijders, T. A. (2005). Models for Longitudinal Network Data. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (215–247). Cambridge University Press.

Snijders, T. A., & Pickup, M. A. (2016). Stochastic Actor-Oriented Models for Network Dynamics. In J. N. Victor, M. Lubell, & A. H. Montgomery (Eds.), *Oxford handbook of political networks* (221–247). Oxford University Press.

Snijders, T. A., & Steglich, C. E. (2015). Representing Micro–Macro Linkages by Actor-Based Dynamic Network Models. *Sociological Methods and Research*, *44*(2), 222–271.

Snijders, T. A., van de Bunt, G. G., & Steglich, C. E. (2010). Introduction to Stochastic Actor-Based Models for Network Dynamics. *Social Networks*, *32*(1), 44–60.

Stadtfeld, C., Takács, K., & Vörös, A. (2020). The Emergence and Stability of Groups in Social Networks. *Social Networks*, *60*(November 2019), 129–145.

Steglich, C. E., Snijders, T. A., & Pearson, M. (2010). Dynamic Networks And Behavior: Separating Selection From Influence. *Sociological Methodology*, *40*(1), 329–393.

Wang, Y., & Yang, C. (2019). Elections and the development of civil war: An empirical study of Sri Lanka. *Democracy and Security*, *15*(2), 111–136.

Ward, M. D. (2016). Can We Predict Politics? Toward What End? *Journal of Global Security Studies*, *1*(1), 80–91.

Ward, M. D., Siverson, R. M., & Cao, X. (2007). Disputes, Democracies, and Dependencies: A Reexamination of the Kantian Peace. *American Journal of Political Science*, *51*(3), 583–601.

Warren, T. C. (2010). The Geometry of Security: Modeling Interstate Alliances as Evolving Networks. *Journal of Peace Research*, *47*(6), 697–709.

Warren, T. C. (2016). Modeling the coevolution of international and domestic institutions: Alliances, democracy, and the complex path to peace. *Journal of Peace Research*, *53*(3), 424–441.

Weschle, S. (2018). Quantifying political relationships. *American Political Science Review*, *112*(4), 1090–1095.

Whitten, G. D., & Williams, L. K. (2011). Buttery guns and welfare hawks: The politics of defense spending in advanced industrial democracies. *American Journal of Political Science*, *55*(1), 117–134.