Measuring and Assessing Latent Variation in Alliance Design and Objectives:
Supplementary Information (SI) Appendix

Benjamin W. Campbell*  
April 22, 2019

Contents

1 Historical Expectation, Dynamic Roles, & System Features 3
  1.1 Vienna System (1816-1848) 3
  1.2 Bismarckian System (1849-1890) 3
  1.3 Pre-World War I System (1891-1918) 4
  1.4 Interwar System (1919-1945) 5
  1.5 Bipolar System (1946-1991) 6
  1.6 Liberal International System (1992-2002) 7

2 Model Explanation 10
  2.1 Estimation Procedure 10
  2.2 Model BIC 11
  2.3 Model Validation 11

3 Variables & Descriptions 13
  3.1 Alliance Treaty Network 13
  3.2 Edges 13
  3.3 Alternating K-Stars & GW Degree 15
  3.4 Triangles 17
  3.5 Defensive Commitments 17
  3.6 Offensive Commitments 19
  3.7 Secret Provisions 19
  3.8 Degree of Institutionalization 20
  3.9 National Capabilities 20
  3.10 Revisionism 22
  3.11 Regime Type 22

4 Comparison to Other Approaches 24

*BWC: Doctoral Candidate in Political Science, The Ohio State University, e: campbell.1721@osu.edu
1 Historical Expectation, Dynamic Roles, & System Features

The roles that exist within the alliance network and the pressure states experience to adopt them are undeniably dynamic. In other words, the roles adopted likely vary as a function of time-varying system-level attributes that augment which roles are made available to states and states’ incentive structures to adopt certain roles. Consider the Cold War, which provided a host of incentives for states to adopt the role of Balancer and to form alliances as a means of balancing and constraining either the United States or the Soviet Union. Following Soviet dissolution, however, the incentives provided by the system to be a Balancer decreased while simultaneously opening up the possibility that states could become Reformers.

In this section I discuss the six diplomatic periods in which there may be significant change in the role structure: the Vienna System, Bismarckian System, Pre-World War I System, Interwar System, Bipolar System, and Liberal International System.

1.1 Vienna System (1816-1848)

Following the Napoleonic Wars, the victorious great powers of Austria, Great Britain, Prussia, and Russia convened in Vienna to create a new order based upon the principles of sovereignty, concert diplomacy, and peaceful dispute resolution. The norms of this system were relatively strong and reflected a broad agreement that alliances formed during this time should uphold this new order (Taylor 1954; Jervis 1985; Schroeder 1992, 1996). This system is marked with relative peace and is often said to persist, with some hiccups, until it is uprooted by the revolutions of 1848. These revolutions, particularly those in Italy, Germany, France, and Poland radically transformed the international system, sparking off a new system marked by nationalism and the establishment of the alliance system engineered by Otto von Bismarck.

The network of alliance treaties for a typical year during this period, 1819, is presented in Figure 1. During this period it is unlikely that states would adopt the role of Aggregator as international conflict and preparation for it is typically unthinkable according to Vienna norms (Jervis 1985, 71). Jervis (1985) notes that frequent meetings, communication and transparency, third party negotiation, and the denial of unilateral gains among other norms all made conflict much less likely during this period (72-73). The one exception might be if alliances were formed to fight and uphold the Vienna System (Schroeder 1996, 736-740). Additionally, given the system’s tendency to preserve autocrats, particularly the effort of the Prussians, Austrians, and Russians, it does not seem that the Reformers role should be prevalent. Revolution was blamed for the Napoleonic Wars and the chief powers at the time felt that stability would be produced if revolutions were prevented (Jervis 1985, 65). During this period alliances were more commonly used to consolidate relationships, such as the pan-Germanic case (Schroeder 1996, 599-606), or preserve a balance of power (Taylor 1954; Schroeder 1996). As such, roles consistent with the logic of Balancers, Consolidators, and potentially Aggregators should emerge.

1.2 Bismarckian System (1849-1890)

Following the revolutions of 1848, states became increasingly nationalistic. States had realpolitik motivations for foreign policy. Perhaps the greatest example of this shift was the alliance system orchestrated by Otto von Bismarck upon his accession to the Prussian Foreign Ministry in 1862. Bismarck’s system is summarized by a simple axiom: In a system of five great powers, be in the dominant coalition (Taylor 1954; Craig and George 1995). Based upon the direction of the wind,
this would give him the opportunity to be in either coalition, aligned with France and Great Britain or Austria and Russia. This system had a relatively long reach, and is blamed by many to have lead to the two bloc system which created the necessary conditions for World War I (Craig and George, 1995). Nevertheless, it is often said to end upon his succession by von Capri in 1890 (Craig and George, 1995).

This produced an interesting alignment, presented by the treaty network of 1883 in Figure 2, where states would not be expected to adopt roles beyond those outlined in the CAM, like the Aggregator or Balancer role (Craig and George 1995, 35-40). Given these realpolitik motives, these roles are those most incentivized and commonly available. However, with the continuation of pan-Germanism and Bismarck’s push for German unification during this period (Hartshorne, 1950), the presence of Consolidators also seems likely.

1.3 Pre-World War I System (1891-1918)

With von Capri’s replacement of Bismarck, the reins of a relatively unstable alliance system were handed to diplomats that some consider to be professionally incompetent (Taylor, 1954; Kissinger, 1994). This highly competitive system of alliances possessed all the necessary conditions for a general conflict that eventually erupted into World War I (Craig and George, 1995, 40-43). This system clearly came to an end with the termination of World War I, which ushered in a distinct and often perplexing period, the Interwar Period (Carr, 1939; Siverson and McCarty, 1982; Kissinger, 1994; Craig and George, 1995). This bloc structure is made clear through Figure 3.
Figure 2: Alliance Treaty Network, Bismarckian System, 1883. Nodes are sized and colored according to degree centrality.

The effect of this period on role structure should mirror that of the Bismarckian System, a period well explained by the CAM (Craig and George, 1995). While the Bismarckian System might be marked by more Balancers than Aggregators, the Pre-World War I System might be marked by more Aggregators than Balancers as arms buildups were prevalent during this time (Craig and George, 1995).

1.4 Interwar System (1919-1945)

For many the Interwar System, which began in 1919 upon termination of World War I, represents a peculiar system informed by the idealism of Wilson and persistent skepticism (Carr, 1939; Kissinger, 1994). While the system is widely known by lasting failure of the Paris Conference, the Treaty of Versailles, and the League of Nations, it was also marked by the failure of diplomacy (Carr, 1939; Kissinger, 1994). This system came to an end as World War II began and ended.

The Interwar alliance system, presented in Figure 4 has perplexed many (Siverson and McCarty, 1982). In their survey of the literature, Siverson and McCarty (1982) find a great degree of variation in alliance formation patterns during this period, with some arguing that alliances were not random but heterogeneous (Siverson and Duncan, 1976; Job, 1976) and others saying that patterns were random and unstable (Li and Thompson, 1978). This lead Siverson and McCarty (1982) to conclude that the primary reason for alliance formation had little to do with international pressures, but more within-nation pressures that may originate from “national power” and/or a changes in nation’s regime (32). This is consistent with historians’ views of alliances during this
time which hold that states formed alliances to accomplish many different objectives \cite{Craig1995}. During the Interwar Period, the roles available to states may have expanded beyond the simple CAM roles which dominated the lead-up to World War II \cite{Kissinger1994} to include Reformers as states' regimes transformed \cite{Siverson1982} and there was increased pressure for states to liberalize. Additionally, Consolidators may be prevalent as the Paris Conference and Treaty of Versailles sought to establish new institutions to prevent conflict \cite{Carr1939, Kissinger1994, Craig1995}.

1.5 Bipolar System (1946-1991)

As World War II ended and the bipolar post-war order was established, the formation of a new alliance system was inevitable \cite{Holsti1970, Kissinger1994}. The system’s bipolar structure, as made clear by the 1980 alliance network presented in Figure 5, is its dominant ordering characteristic \cite{Holsti1970, Waltz1993, Kissinger1994}. This relatively peaceful system persisted with little change in the order’s grand and organizing principles for nearly 50 years until the Soviet Union’s dissolution \cite{Friedman1989, Mearsheimer1990, Waltz1993, Kissinger1994}. While there was certainly great power competition, it never manifested itself in direct military confrontation \cite{Friedman1989, Mearsheimer1990}.

During the Bipolar System, alliances were largely formed in accordance with the Balancer, Aggregator, and/or Consolidator roles. Reformers seemed unlikely during this time as alliances were often formed with states sympathetic to either Communism or democracy \cite{Waltz1993}. Indeed, the Bipolar System is often held my modern alliance scholars as the golden days of the CAM \cite{Morrow1991, Waltz1993, Morrow2000}. However, it is also clear that alliances during

\[ \text{Figure 3: Alliance Treaty Network, Pre-WW1 System, 1913. Nodes are sized and colored according to degree centrality.} \]
Figure 4: Alliance Treaty Network, Interwar System, 1930. Nodes are sized and colored according to degree centrality.

this time too may have been used to deepen economic or political relationships, particularly between proteges and the United States or the Soviet Union (Waltz 1993; Lebow 1994; Slater 2008).

1.6 Liberal International System (1992-2002)

Upon the dissolution of the Soviet Union, the United States was left as the remaining hegemon. This “unipolar moment” radically transformed international politics, creating a new set of constraints on balancing behavior (Brooks and Wohlforth 2008) while ushering a new order based upon liberal international norms (Mandelbaum 2005; Ikenberry 2012, 2014). This new system of alliances is illustrated in the 2000 treaty network presented in Figure 6.

During this system it seems relatively unlikely that states would form alliances to accomplish traditional capability aggregation-based objectives like balancing or conflict preparation. Instead, states should be much more likely to adopt the role of Consolidator or Reformer. For example, many military alliances such as NATO or the OAS were repurposed to focus upon the propagation of the Liberal International System by bringing new states into the fold and spreading economic and political liberalism (Pevehouse 2005; Cranmer, Desmarais, and Campbell 2018). Take the renewed mission of the OAS following the Cold War, where the focus became promoting a regional security community and order based upon cooperative security and a respect for democracy and trade (Thérien, Mace, and Gagné 2012). As such, many of these conventional alliances may
transition to look more like Reformers than Balancers or Aggregators. Additionally, Consolidators may also be prevalent as many emerging countries, such as those in Africa, Asia, and the Middle
Figure 6: Alliance Treaty Network, Liberal International System, 2000. Nodes are sized and colored according to degree centrality.

East have formed alliance-based institutions for the objective of consolidating relationships and promoting regional cooperation.
2 Model Explanation

Note: This discussion closely follows that of [Campbell (2018)].

The ego-TERGM is the temporal extension of the ego-ERGM, a model derived by [Salter-Townshend and Murphy (2015)] and introduced by [Box-Steppensmeier, Campbell, Christenson et al. (2017)]. The ego-ERGM is a finite mixture model that assumes that each cross-sectional ego-network from some broader network has probabilities associated with belonging to a certain number of specified clusters, \( G \). In other words, for all elements \( g \in G \) there is some probability that a node belongs to a particular role. There is a naive probability for each ego-network that they are assigned to a particular cluster, \( \tau_g \), which sums to one across all \( g \in G \). This set of baseline probabilities creates a vector \( \tau \) (Underlined because it is a vector, bolded characters are matrices). These role assignments are conditioned upon a set of role specific parameters, \( \theta_g \), that are modeled through an ERGM estimated through maximum pseudolikelihood estimation. \( \theta \) refers to the \( G \times H \) matrix of model parameters where \( H \) is equal to the number of terms specified and \( \theta_g \) refers to the \( g^{th} \) row of matrix \( \theta \). Assuming independence across ego-networks, an assumption rarely met (but discussed in greater detail by [Campbell (2018)]), the likelihood is derived:

\[
P(Y \mid \tau, \theta, Y) = \prod_{i=1}^{N} \left[ \sum_{g=1}^{G} \tau_g \exp\{\theta'_g h(Y) - \gamma(\theta_g)\} \right]
\] (1)

The ego-TERGM builds upon this by pooling over all time periods \( t \in T \) and conditioning each iteration on a \( q^{th} \)-order autoregressive process. The ego-TERGM, as derived by [Campbell (2018)] has a likelihood presented in Equation 2. The model attempts to determine the probability of some longitudinal network \( Y_T \) and a \( Z \) matrix of node-based role assignment probabilities conditioned upon the global proportion of cluster assignments \( \tau \), a \( G \times H \) matrix of model parameters \( \theta \) where \( G \) refers to the number of roles to be fit and \( H \) refers to the length of model terms specified, and a \( q^{th} \) order auto-regressive process for the network \( Y_{t-q} \). \( \theta \) is composed of \( G \) rows of role-based parameters \( \theta_g \):

\[
P(Y_T, Z \mid \tau, \theta, Y_{t-q}) = \prod_{i=1}^{N} \prod_{g=1}^{G} \prod_{t=1}^{T} \tau_g \exp\{\theta'_g h(Y_{i,t}, Y_{i,t-1}, ..., Y_{i,t-q}) - \gamma(\theta_g)\}^{Z_{ig}}
\] (2)

From this, the log-likelihood can be derived:

\[
\log[P(Y_T, Z \mid \tau, \theta, Y_{t-q})] = \sum_{i=1}^{N} \sum_{g=1}^{G} \sum_{t=1}^{T} Z_{ig} \log[\tau_g \exp\{\theta'_g h(Y_{i,t}, ..., Y_{i,t-q}) - \gamma(\theta_g)\}]
\]

\[
= \sum_{i=1}^{N} \sum_{g=1}^{G} \sum_{t=1}^{T} Z_{ig}\{\log \tau_g + \theta'_g h(Y_{i,t}, ..., Y_{i,t-q}) - \log \gamma(\theta_g)\}
\] (3)

2.1 Estimation Procedure

The matrix of latent and unobserved cluster assignments, \( Z \), and the estimation of group-level parameters, \( \theta \), are mutually dependent. As such, the ego-TERGM requires a recursive and iterative estimation procedure. As such, an Expectation-Maximization (EM) algorithm is used, similar to
that of Salter-Townshend and Murphy (2015). The algorithm proceeds according to the pseudocode in Algorithm 1. The algorithm is written as introduced by Campbell (2018):

1. Let \( u = 0 \) reflect initial estimates for the initial \( k \)-means clustering initialization, where \( \hat{\pi}^{(0)} \) and \( \hat{\theta}^{(0)} \) represent initial estimates of the mixing proportions and centroids.

2. E-Step: This step is repeated for all nodes and for all groups. Compute the expected role assignments for a particular node in a particular group, \( \hat{Z}_{ig} \), based upon current model parameters where \( P(Y_{it} | \hat{\theta}_g) \) can be understood as the conditional probability of observing the longitudinally observed ego-network \( Y_{IT} \) conditioned on group-level parameters and mixing parameters:

\[
\hat{Z}_{ig}^{(u)} = \frac{\hat{\pi}_g^{(u-1)} P(Y_{IT} | \hat{\theta}_g^{(u-1)})}{\sum_j \hat{\pi}_j^{(u-1)} P(Y_{IT} | \hat{\theta}_j^{(u-1)})}
\] (4)

3. M-Step: Maximize the expected complete data log likelihood to yield new and updated parameter estimates:

\[
(\hat{\pi}^{(u)}, \hat{\theta}^{(u)}) = \arg\max_{\pi, \theta} \sum_{i=1}^{N} \sum_{g=1}^{G} \sum_{t=1}^{T} \hat{Z}_{ig}^{(u)} \log[\pi_g P(Y_{it} | \theta_g)]
\] (5)

4. Check for convergence, if it has converged then stop, otherwise, move to the next iteration and increment \( u \), returning to Step 1. Convergence occurs if the log-likelihood changes by less than \( 10^{-6} \).

2.2 Model BIC

To assess model fit, users should rely upon a combination of theoretical intuition and in-sample fit assessment tools. BIC for the ego-TERGM, often referenced in the manuscript, is calculated as:

\[
BIC = 2\hat{L} - (GH + G - 1)\log(NT)
\] (6)

2.3 Model Validation

Validating the cluster assignments for the ego-TERGM in a formal way is difficult as there are no previously observed labels that can be used for out-of-sample validation. As such, in this piece, and in general, two forms of validation are used. First, the scree plot of model BIC by value of \( G \) is examined to find the best within-model fit. While this routine cannot really be used for the importance of model terms given the collinearity of ERGM terms (Hunter, Goodreau, and Handcock, 2008), it is effective in finding the \( G \) values that produce optimal fit (Salter-Townshend and Murphy, 2015). Second, external validation can occur as long as the analyst has some theoretically motivated intuition about the roles that particular actors should adopt. Another means of validation, stability validation, could be used to assess how stable the findings are to the time-periods included in the network. Such an approach could fit a model to the full longitudinal network and compare the results of that model to a model fit on a subset of the networks.

For the analyses here, selection of \( G \) for each period is informed by BIC. Scree plots for each period are presented in Figure 7. Overall, the \( G \) values chosen are typically those recommended
Algorithm 1 Fitting procedure for ego-TERGM model.

\begin{enumerate}
\item \(N \leftarrow\) ego network list
\item \(G \leftarrow\) expected roles
\item \(H \leftarrow\) length(terms)
\item for \(n \in N\) do
\begin{enumerate}
\item Calculate change statistics for \(n\) and its offset term \(\omega\)
\item Estimate TERGM via bootstrapped MPLE on ego-network \(n\)
\item Save coefficients, change statistics, and offset terms
\end{enumerate}
end for
\item Find \(G\) initial clusters of the \(\text{length}(N) \times H\) matrix TERGM coefficients using \(k\)-means
\item Extract \(\mu\) cluster centroids from \(k\)-means
\item Calculate \(\text{length}(N) \times G\) matrix, \(Z_{u=0}\), of cluster assignment probabilities
\item \(\text{loglikelihood}_{u=0} \leftarrow NaN\)
\item \(\text{steps} \leftarrow\) Maximum number of EM steps
\item \(\theta_{u=0} \leftarrow \mu\)
\item \(\alpha \leftarrow\) convergence value
\item \(\tau_{u=0} \leftarrow\) initial mixing proportions
\item for \(u \in \text{steps}\) do
\begin{enumerate}
\item while \(\text{loglikelihood}_u - \text{loglikelihood}_{u-1} < \alpha\) do
\begin{enumerate}
\item for \(i \in \text{length}(N)\) do
\begin{enumerate}
\item for \(g \in G\) do
\begin{enumerate}
\item Update \(Z_u\) based upon \(\theta_{u-1}\), \(\omega\), and \(\tau_{u-1}\) as per Equation 4
\end{enumerate}
\end{enumerate}
\end{enumerate}
\end{enumerate}
\item Standardize \(Z_u\) for row sum of 1
\item Update \(\tau_u\) as the global mixing proportions
\item Update \(\text{loglikelihood}_u\) using \(\tau_u\), \(Z_u\), \(\omega\), and \(\theta_{u-1}\) as per Equation 5
\item Update \(\theta_u\) using \(\theta_{u-1}\), \(Z_u\), \(\omega\), and \(\tau_u\) as per Equation 5
\end{enumerate}
end while
\end{enumerate}
Figure 7: Scree Plot for $G$ Selection by BIC. Largest and lightest blue circles indicate optimal model fit.

by the manuscript. This is discussed in the manuscript. Additionally, the sensitivity of the results to the inclusion of particular covariates is another form of internal validation. This sensitivity analysis is presented later in this Appendix in Section 5. External validation for my analyses relies upon historical intuition and knowledge about the foreign policies of particular countries during the periods examined. The background influencing interpretation of results and the labels assigned are provided in Section 1. As mentioned, stability validation is difficult as one cannot truly compare the results of one model to the results of a model fit on a subset of networks when nodes drop-in or out of the network.

3 Variables & Descriptions

3.1 Alliance Treaty Network

The network of interest is annualized and composed of all alliance treaties ([Leeds, Ritter, Mitchell et al.] 2002) between all state system members ([cow] 2017) per year. Given that the network is a network of treaties as opposed to commitments, it is treated as an undirected network. These networks are visualized, at varying points, in Figures 1 through 6.

3.2 Edges

The $edges$ term is the ERGM/TERGM equivalent of an intercept term and captures the baseline expectation for tie formation ([Handcock, Hunter, Butts et al.] 2008). In other words, a baseline ERGM or TERGM that only includes this term would be equivalent to an Erdos-Renyi or Bernoulli graph where each dyad has some naive probability of tie formation. When considering this in dyadic perspective, change statistics are used. Change statistics take on a value equal to the
value of the statistic when an edge is present minus the value of the statistic when an edge is not present. In other words, a change statistic shows how a variable’s value changes if an edge is toggled on or off. Change statistics are used for both the ego-TERGM and bootstrapped estimated TERGM (Leifeld, Cranmer, and Desmarais, 2014). The logistic mean function transformation of the estimated ERGM/TERGM coefficient for the edges term will give a baseline probability that two nodes form an edge. When conditioning out any other network effects that may inform tie formation, this is interpreted as a residualized Erdos-Renyi graph wherein it is the baseline probability of tie formation when all other dyad change statistics are held at zero. When considered in the aggregate as a sum of subgraph product, the edges term takes on a value equal to the number of observed edges in the network and speaks to its density. Through taking the logistic mean function of the estimated ERGM/TERGM coefficient, the original density of the network can be returned.

Within the context of the network considered, an edge is defined when any two countries are party to an alliance treaty as defined by the Alliance Treaty Obligations and Provisions (ATOP) dataset for a Gregorian calendar year (Leeds, Ritter, Mitchell et al., 2002). This differs from many modern studies which only view an event as the presence of a particular type of alliance commitment (such as an offensive commitment or a defensive commitment). Figure 8 presents the number of edges in this network over time. The trajectory of edges present during each year increases, which is to be expected as many states join the system as time progresses.
3.3 Alternating K-Stars & GW Degree

*Alternating K-Stars* is an ERGM/TERGM statistic that is often used when conventional k-star statistics may be prone to degeneracy ([Handcock, Hunter, Butts et al., 2008](#)). Introduced by [Snijders, Pattison, Robins et al., 2006](#), this term adds one statistic to the model or change statistic equal to a geometrically weighted \( \lambda \) sequence of k-star statistics. To assist in interpretation, Alternating K-Stars is used when distinguishing between roles in the ego-TERGM while Geometrically Weighted (GW) Degree (where the weight is noted \( \alpha \)) is used when fitting the TERGMs. As GW Degree is often regarded as more easily interpreted, it is included in the TERGMs to assess the role generative process as model term effects are explicitly discussed. These terms are otherwise replaceable and produce mathematically equivalent models ([Snijders, Pattison, Robins et al., 2006](#); [Handcock, Hunter, Butts et al., 2008](#)). A lower geometric weight typically refers to a greater degree of penalization for high values, which may make sense as states are unlikely to ally with all, or even most, states in the system. Both of these terms, overall, capture a tendency towards preferential attachment wherein states accumulate many alliances. When one thinks about a hub-and-spoke network model, the hub would have a fairly high value on these statistics.

Figures 9 and 10 show the values for these statistics as sums of subgraph products for each network-year. As to be expected, consistent with the edges discussion, the value of these statistics increases over time which illustrates perhaps a greater tendency towards high-degree nodes as time progresses. However, again, this may be expected as the number of states in the system increases with time.
Figure 10: GW Degree Sum of Subgraph Products Over Time. Line shows the value of a geometrically weighted degree term ($\alpha = 0.1$) in the alliance treaty network for a given year.
3.4 Triangles

*Triangles* is an ERG statistic that measures the number of triangles, or fully connected triads where $i$, $j$, and $k$ are all connected, present within a network (sum of subgraph products) or in the case examined here, the number of triangles incident on an edge (change statistic) (Handcock, Hunter, Butts et al., 2008). Given that this network is undirected, variations on this statistic are unnecessary. The triangles term is ubiquitous in network science and often used to capture tight clustering or the formation of cliques. Figure 11 shows the number of triangles per network-year. Overall, as to be expected, the number of triangles increases over time as more states enter the system.

3.5 Defensive Commitments

To capture the importance of defensive commitments in differentiating role assignments, an edge covariate for the presence of a defensive commitment is used. Data for this variable is sourced from ATOP (Leeds, Ritter, Mitchell et al., 2002). As a sum of subgraph product, this term reflects the total number of defensive commitments among states for a given network year. As a dyad-wise change statistic, this term captures the number of defensive commitments incident upon an edge. For more detail, the reader is directed to Handcock, Hunter, Butts et al. (2008). Figure 12 shows the relative frequency of defensive commitments in dyad year perspective, indicating that while they are uncommon relative to dyad years without commitments, they are still prevalent in absolute terms. Overall, 42,590 dyad years have a defensive commitment while 651,868 do not.
Figure 12: Dyad Years with Defensive Commitment. Bars show relative frequency for dyad years where there is no defensive commitment (value of zero) and where there is a defensive commitment (value of one).
3.6 Offensive Commitments

The presence of offensive alliance commitments may also be a useful marker for sorting alliances; to fulfill their role-based goals, states may engineer their local alliance networks to include many offensive commitments. Data for this variable is available from Leeds, Ritter, Mitchell et al. (2002). These dynamics are captured using an ERG edge covariate (again, as discussed in Handcock, Hunter, Butts et al. (2008)) for whether a dyad nested within a network-year contains an offensive alliance commitment. Figure 13 illustrates that in both absolute and relative terms, very few dyad years possess an offensive commitment. Only 2,574 dyad years have an offensive commitment while 691,884 do not.

3.7 Secret Provisions

Aggregators (and perhaps Balancers to a lesser degree) may be more likely to construct a local alliance network containing secret provisions than Reformers or Consolidators. To capture this dynamic, an ERGM edge covariate is used to assess the prevalence of secretive alliance provisions in each state’s local ego-network. This variable retains its coding from Leeds, Ritter, Mitchell et al. (2002) in which a dyad is public but has a provision requiring at least one article remain secret. A value of two refers to the highest level of secrecy wherein the entire treaty is private and secret. Figure 13 illustrates that in both absolute and relative terms, very few dyad years possess an offensive commitment. Only 2,574 dyad years have an offensive commitment while 691,884 do not.
Figure 14: Dyad Years with Secret Provisions. Bars show relative frequency for dyad years where there are no secret provisions (value of zero), some secret articles (value of one), or a fully secret treaty (value of two).

3.8 Degree of Institutionalization

Consolidators and Reformers may often include provisions establishing institutions in their alliances as it assists them in helping to accomplish their role-based objectives. Based upon the data from Leeds, Ritter, Mitchell et al. (2002), an additive index of the degree of institutionalization was created by taking the sum of the institutionalization scores (ATOP variable ORG1PURP and ORG2PURP) for an alliance dyad. Any values of one on either ORGPURP variables were recoded to zero as troop placements are not to inform these role based dynamics. Higher values on this edge covariate refer to higher degrees of institutionalization between alliance members. The majority of cases do not include provisions establishing any institutional mechanisms, as illustrated in Figure 15.

3.9 National Capabilities

Some states may, by their role-based objectives, be incentivized to form states that are similar, or different, with respect to national capabilities. To measure this dynamic, Composite Index of National Capability (CINC) scores are used, with data sourced by Singer, Bremer, and Stuckey (1972). To capture the difference or similarity in country year CINC scores within a dyad year the
Figure 15: Dyad Years by Degree of Institutionalization. Bars show relative frequency of dyad years by rate of institutionalization.
absolute difference in scores is calculated and used as a change statistic, as discussed by Handcock, Hunter, Butts et al. (2008). While I do not present the distribution of these differences, the distribution of country-year CINC scores is presented in Figure 16 to show how relatively homogenous or heterogenous these scores are in aggregate.

### 3.10 Revisionism

As discussed in the manuscript, alliances formed with revisionist states may potentially assist in sorting between Aggregators or Balancers. Based upon data from the Correlates of War Project’s MID dataset a binary indicator is used to indicate states that have ever been considered revisionist within a country year (Palmer, d’Orazio, Kenwick et al., 2015). To capture the tendency to form homophilous or heterophilous alliances with respect to revisionism, the same absolute difference method of calculating the CINC difference change statistic is used here (Handcock, Hunter, Butts et al., 2008). Figure 17 shows the relative frequency of country years where an actor is listed as revisionist with respect to at least one dispute.

### 3.11 Regime Type

Last but not least, states may form alliances with states of different or similar regime types in an effort to accomplish their alliance-based objectives. Using data from the POLITY project (Marshall, Jaggers, and Gurr, 2002), a new indicator is created for whether a state is a democracy (country
Figure 17: Frequency of Country Years by Revisionism Status. Bars show relative frequency of dyad years by revisionist (value of one) or status quo (value of zero).
year POLITY 2 value greater than 6) or a non-democracy. While this is a brute measure, such simplifications are common and theoretically defensible (Marshall, Jaggers, and Gurr 2002; Gleditsch and Ward 2006; Cranmer, Desmarais, and Campbell 2018). Such simplifications also assist in model identifiability and estimation. To capture a tendency towards regime-type homophily in alliance ties, a node match term is used as described by Handcock, Hunter, Butts et al. (2008). Figure 18 shows the relative proportion of states in the system that are democratic by network year.

4 Comparison to Other Approaches

Several other approaches could be used to understand the roles that actors adopt within networks. Such commonly used methods include community detection or equivalence-based clustering. In this section, I use the alliance network for the year 2000 to make the case that the results uncovered by the ego-TERGM differ from other approaches in significant ways. The 2000 network was chosen as it reflects a later year that nicely illustrates how differently these approaches behave. To compare the ego-TERGM to a community detection algorithm, I use Walktrap which is a standard algorithm for partitioning a network into a predetermined number of communities. Community detection is about finding subgraphs that are sufficiently distinct from the rest of the graph and that can be partitioned by breaking relatively few edges. Stochastic equivalence, and specifically structural equivalence, is about drawing nodes into clusters according to overlap in their shared alters.
Figure 19: Ego-TERGM Cluster Assignments, Alliance Network 2000. Nodes are colored by common cluster assignment.

Figure 19 presents the ego-TERGM results for the alliance network in 2000 presented in the manuscript. The cluster assignments presented here clearly differ from those presented in Figure 20 for the community detection approach, and Figure 21 for the stochastic equivalence approach. This indicates that the ego-TERGM is capturing different dynamics than community detection or stochastic equivalence.

5 Sensitivity Analysis for Ego-TERGMs

To assess the sensitivity of role assignments to the inclusion of particular covariates, the iterative approach recommended by Box-Steffensmeier, Campbell, Christenson et al. (2017) and Campbell (2018) is used. This approach compares the fully specified model to variations on the model estimated through the iterative exclusion of an included model term. If role assignments are fairly uniform across these models then the model is not sensitive to the inclusion of any one particular covariate. If role assignments do not change, then it would indicate that a variety of structural factors matter in sorting between roles. If role assignments do change, then the assignments uncovered may be particularly sensitive to the inclusion one covariate. In other words, it would indicate that the removed covariate matters in sorting between roles. In this section, this approach is applied to each ego-TERGM estimated on the six periods analyzed.
Figure 20: Walktrap Community Assignments, Alliance Network 2000. Nodes are colored by common community assignment.
Figure 21: Structural Equivalence Cluster Assignments, Alliance Network 2000. Nodes are colored by common cluster assignment.
5.1 Congress System

Figure 22 illustrates results from the sensitivity analysis for the ego-TERGM presented for the Congress Period. The model excluding the “Edges” term was unidentifiable, and as such, the results are not presented. The sensitivity analysis finds that the only covariate that truly seems to sort between notes is the CINC Difference covariate, which captures the absolute difference in CINC scores for a dyad. The exclusion of this covariate indicates that once you fail to account for resources, Aggregators (Role 3) look more like Balancers (Role 2). Overall this analysis illustrates that the model is not particularly sensitive to the inclusion of any one covariate.

5.2 Bismarckian System

In Figure 23, sensitivity analysis results for the ego-TERGM fit on the Bismarckian system are presented. This routine reveals that four covariates appear to matter when sorting between roles. First, the presence of relatively young or old alliances seems to assist in sorting between Consolidators and Balancers. In other words, when you fail to account for alliance age and temporal dependence, Balancers start to look more like Consolidators. When omitting CINC / difference, Balancers start to look more like either Reformers or Aggregators (it is impossible to tell which, however, given the problem of label switching). Given experience from the prior period, it appears most likely to be Aggregators, however. This makes sense as we might expect Aggregators to look more like Balancers when you exclude CINC Difference, a variable which would capture the relative risk of an Aggregator being chain-ganged into conflict or being improperly defended. Third,
Figure 23: Bismarckian Period Sensitivity Analysis. Groupings refer to the largest to smallest cluster sizes. Dashed lines refer to the cluster sizes for the baseline model.

when you exclude alternating k-stars, which captures a tendency towards preferential attachment, Consolidators start to look more like Balancers. This makes sense as you might expect a tendency towards dense graphs and many high-degree nodes for Consolidators, particularly relative to Balancers. Fourth, and finally, when you fail to account for Revisionism Difference, there are fewer Consolidators and more Balancers. This makes sense as we expect Balancers to deeply consider whether they should form alliances with Revisionist (or Non-Revisionist States), particularly relative to Consolidators who may be less inclined to make alliance decisions based upon the state’s “type.”

5.3 Pre-World War 1 System

The model fit for the Pre-World War 1 Period does not appear to be particularly sensitive to the inclusion of any covariates except the Edges term. When failing to account for the density of ego-networks Aggregators look more like Balancers. In other words, Russia’s ego-network during this time looks more like those of Austria, France, Germany, Italy, Romania, and the United Kingdom on all other covariates included in the model save density. It is unclear whether this is consistent with theoretical expectation as both would be expected to have dense subgraphs marked by large rates of triadic closure.
5.4 Interwar System

Given that the model estimated for the interwar period only includes a $G$ of one, excluding covariates cannot lead to changes in role assignments. As such, a figure for this sensitivity analysis is not included.

5.5 Bipolar System

Figure 25 presents results from a sensitivity analysis for the ego-TERGM fit on the Bipolar System. All covariates included are found to be important in sorting between the roles uncovered. When omitting alliance years, the most common role, Balancers, start to look more like the second most common role, Consolidators. This may make sense as many of the Balancing alliances (NATO, for example) that form during this time may be examples of longer alliances quickly forming after World War II while the Consolidators may come later (such as the OAS). The number of edges within a state’s network also appears to matter. Excluding this covariate makes Balancers much more common. Namely, it is the only identified role. This may indicate that the density of a graph goes a long way towards sorting between Consolidators, Reformers, and Aggregators. Third, omitting the Alternating K-Stars term produces a similar effect. This variable seems to be important in sorting between Aggregators and Balancers, which may be intuitive as Aggregators may be more likely to form a large number of dyadic alliances with their alters, but connectivity among their alters may be less common. Consider the case of Bismarck and his flexible alliance network. Fourth, when one fails to account for regime type, Consolidators become less common.
Figure 25: Bipolar Period Sensitivity Analysis. Groupings refer to the largest to smallest cluster sizes. Dashed lines refer to the cluster sizes for the baseline model. Model omitting CINC Difference was unidentifiable.

This makes sense as political factors may influence their alliance decisions. Finally, when removing Revisionism Difference, Balancers become less common. This logically follows as we would expect some Balancers to not form alliances with Revisionist states, and as such, this variable should inform role assignments.

5.6 Liberal International System

Figure 26 shows how sensitive role assignments are to the covariates included in the Liberal International System. When omitting the age of alliance ties, the most common role, Reformers, becomes less common while the most common role, Aggregators, becomes more common. This may make sense given the post-Cold War shift for many alliances. NATO and the OAS were fairly old alliances with a renewed purposes to push for domestic reforms within member states or prospective member states. It is unclear, however, why Reformers would look more like Aggregators when omitting alliance age. Second, excluding the difference in CINC scores leads the model to have a similar effect, albeit at a slightly lower level. When failing to account for CINC scores, both Consolidators and Reformers become less common, and Aggregators become more common. This may illustrate support for the hypothesis that Balancers and Consolidators are more likely to form alliances with states of disparate capabilities. Little change occurs when omitting the Edges or Alternating K-Stars term. However, more dramatic changes occur when estimating the model without the regime homophily variable. In particular, Reformers become less common and Consol-
Groupings refer to the largest to smallest cluster sizes. Dashed lines refer to the cluster sizes for the baseline model.

Given the importance of regime type heterophily for Reformers’ alliances, this makes sense.

6 Goodness of Fit Diagnostics for Pooled TERGMs

To assess the fit of each pooled TERGM used in the model to assess role generative processes, conventional GOF diagnostics were used (Hunter, Goodreau, and Handcock 2008). For this routine, a series of networks were simulated according to the estimated model parameters. For each of these networks walktrap modularity was calculated. The walktrap modularity of the observed networks (black density) was then compared to that of the simulated networks (red density). If the modularity of the simulated networks approximates that of the observed networks then the structure of the networks are simulated and the models are said to be well fitting. For the Figures 27 through 30 present the GOF diagnostics for the Aggregators, Balancers, Consolidators, and Reformers models respectively. Overall, it appears that these TERGMs fit reasonably well as the observed distribution and simulated distribution are similar.

When possible, multiple statistics were used to assess fit. This, however, was often constrained by model fits and what was made possible through the \texttt{btergm} package (Leifeld, Cranmer, and Desmarais 2017).
7 TERGMs Disaggregated By Period

In the manuscript the TERGMs presented assess the generating process of all ego-networks assigned to roles of a common type over all periods where the particular role was identified. In this section, these TERGMs are disaggregated by systemic period. To assess the role generative process of roles uncovered using the ego-TERGM on a period-by-period basis, a conventional bootstrapped TERGM is fit on all networks associated with a role as conventionally recommended (Campbell, 2018). Often, these attempts may be hindered by networks that do not vary during a period.

For additional discussion of the TERGM estimated through bootstrapped pseudolikelihood, I refer the reader to Cranmer and Desmarais (2010), Desmarais and Cranmer (2010), Desmarais and Cranmer (2012), and Leifeld.
Figure 28: Balancers Model GOF Diagnostics, All Systems. Similar networks with similar structures will have similar modularity densities.

When networks are uniform all edges and potential edges have uniform values on certain variables, preventing model estimation. Such a problem occurred when attempting to assess the role generative structure for the Consolidator and Aggregator roles in the Congress period. The inclusion of particular covariates may often lead to this problem as well. As such, the models presented throughout the manuscript and the SI Appendix are the best fitting identifiable models.

Note, occasionally the models presented in Tables 1 through 6 have bootstrapped confidence intervals that contain extreme values. This is a product of separation and is unavoidable as some networks are perfectly dense. Given that there is little substantive reason to exclude them, and certain covariates are important, these networks are not excluded.

Cranmer, and Desmarais (2017).
Given identification problems, only the role generative structure for the Balancers role is considered. Table 4 presents results for a well-fitting TERGM estimated on all networks associated with the Balancers role. The prevalence of institutionalization and triangles within Prussia and Austria’s alliance network during this time makes sense given their ties to the German Confederation. Likewise, the rarity of offensive alliances within these Balancers’ ego-networks makes sense given the norms of the Vienna System, which sought to disincentivize interstate conflict, not prepare for it. However, the a priori expectation that Balancers would form alliances with similar states with respect to revisionist tendencies or capabilities was not met here. Regardless, historical context would indicate that Prussia and Austria formed alliances in an effort to preserve the Vienna
Figure 30: Reformers Model GOF Diagnostics, All Systems. Similar networks with similar structures will have similar modularity densities.

System. This model fits reasonably well as the observed walktrap modularity approximates the estimated walktrap modularity, as illustrated in Figure 31.

7.2 Bismarckian System

To assess the generative structure of the role uncovered in the Bismarckian System, a pooled bootstrapped TERGM is fit on the pooled role assignment for each role. Results from these models are presented in Table 2. Aggregators do not appear to have any clear or consistently discernible structure for their local alliance networks while Consolidators are more likely to have an alliance network constituted by defensive alliances formed between states of disparate capabilities. Balancers typically have local alliance networks that are particularly cliquish (high levels of triangles)
|                              | Balancers                  |
|------------------------------|----------------------------|
| Edges                        | $-1.45^*$                  |
|                              | $[-2.39; -0.30]$           |
| Triangles                    | $0.82^*$                   |
|                              | $[0.69; 0.93]$             |
| Alternating K-Stars (0.5)    | $-1.41^*$                  |
|                              | $[-2.05; -0.88]$           |
| CINC Difference              | 2.72                       |
|                              | $[-1.63; 6.09]$            |
| Revisionism Difference       | 0.31                       |
|                              | $[-0.43; 1.10]$            |
| Defensive Alliances          | 0.25                       |
|                              | $[-0.10; 0.64]$            |
| Offensive Alliances          | $-1.05^*$                  |
|                              | $[-1.71; -0.61]$           |
| Secret Provisions            | 0.03                       |
|                              | $[-0.39; 0.42]$            |
| Degree of Institutionalization| $0.71^*$                   |
|                              | $[0.38; 1.07]$             |
| Alliance Years               | $0.08^*$                   |
|                              | $[0.06; 0.11]$             |

* 0 outside the confidence interval

Table 1: Pooled TERGM Results for Role Generative Structure, Vienna System. Confidence intervals reflect 95% confidence intervals estimated through bootstrapped pseudolikelihood. 500 replications. Models for the Aggregator and Consolidator roles were unidentifiable.
and constituted by a tendency towards preferential attachment. Counterintuitively, Balancers are likely to form defensive or offensive alliances containing secret provisions with states possessing disparate capabilities. More intuitively, however, Reformers are likely to form less cliquish alliances with states of disparate capabilities. Figures 32 through 35 illustrate that these models fit reasonably well.

7.3 Pre-World War 1 System

The generative structures of the roles identified in the Pre-World War 1 System are estimated through period-specific pooled TERGMs with results presented in Table 3. Balancers typically have alliance networks constituted by triangles, preferential attachment, and defensive and offensive
Table 2: Pooled TERGM Results for Role Generative Structure, Bismarckian System. Confidence intervals reflect 95% confidence intervals estimated through bootstrapped pseudolikelihood. 500 replications.

|                          | Aggregators | Consolidators | Balancers | Reformers |
|--------------------------|-------------|---------------|-----------|-----------|
| Edges                    | 1.05        | -2.46         | -8.25*    | 1.53      |
|                          | [-2.61; 4.71] | [-786.87; 797.46] | [-10.70; -7.28] | [-1.75; 62.44] |
| Triangle                 | -0.52       | 2.43*         | -1.29*    |           |
|                          |             | [-1180.26; 0.61] | [2.17; 3.49] | [-185.31; -0.00] |
| GW Degree (0.1)          | -0.38       | 20.29*        |          |           |
|                          |             | [-9.50; 8.74] | [14.65; 33.90] |           |
| CINC Difference          | -524.75     | 39.30*        | 19.22*    | 17.67*    |
|                          |             | [27.55; 45651.45] | [14.14; 25.37] | [9.34; 5735.43] |
| Revisionism Difference   | 0.10        | -0.23         | 0.13      |           |
|                          |             | [-3.05; 3.25] | [-1.00; 0.59] | [-421.75; 3.25] |
| Defensive Alliances      | 3.79*       | 1.12*         |          |           |
|                          |             | [1.38; 1463.64] | [0.71; 1.59] |           |
| Offensive Alliances      |             | 1.78*         |          |           |
|                          |             | [0.43; 3.38] |           |           |
| Secret Provisions        |             | 1.83*         |          |           |
|                          |             | [1.37; 2.57] |           |           |
| Peace Years              | 0.32        | -0.17*        | 0.29*     |           |
|                          |             | [-0.93; 1.58] | [-0.29; -0.10] | [0.21; 46.85] |

* 0 outside the confidence interval

Alliances. As for aggregators, there does not appear to be any consistent structure. Figures 36 and 37 illustrate that these models fit well with respect to many statistics, including dyadwise shared partners, geodesic distance, walktrap modularity, and ROC and PR AUC.

7.4 Interwar System

An ego-TERGM fit on the Interwar System only identifies one role, which appears to be most consistent with the Consolidators role. The generative process for networks during this time is presented in Table 4. During the interwar period, Consolidators’ local alliance networks are marked by many triangles, a tendency towards preferential attachment, and the formation of defensive alliances with states of disparate capabilities. This model fits reasonably well with respect to walktrap modularity, as per Figure 38.

7.5 Bipolar System

To assess the role generative process of Balancers, Consolidators, and Aggregators identified during the Bipolar System, a well-fitting (as per Figures 39 through 41) bootstrapped TERGMs are fit on the pooled networks for each role. Counterintuitively, status quo or revisionist Balancers appear to be more likely to form offensive or defensive alliances with revisionist or status quo states. Balancers are also more likely to form alliances with those of disparate capabilities and to include provisions creating institutions. The TERGM estimated for Consolidators reveals results consistent with theoretical expectation – Consolidators are more likely to form alliances with states of disparate capabilities. They are more likely to form tightly clustered alliance communities and to include
|                               | Balancers          | Aggregators       |
|-------------------------------|--------------------|-------------------|
| Edges                         | $-9.45^*$          | $-0.01$           |
|                               | $[-13.62; -4.51]$  | $[-126.97; 3877868124062776.00]$ |
| Triangles                     | $3.74^*$           |                   |
|                               | $[1.18; 5.71]$     |                   |
| GW Degree (0.1)               | $51.09^*$          | $1.84$            |
|                               | $[23.99; 78.28]$   | $[-661973583064276.75; 2037.62]$ |
| CINC Difference               | 3.87               | 10.19             |
|                               | $[-0.86; 8.11]$    | $[-14667773399201700.00; 13.37]$ |
| Revisionism Difference        | 0.04               | $-0.66$           |
|                               | $[-0.68; 0.82]$    | $[-50.73; 2342851756878530.50]$ |
| Defensive Alliance            | $1.94^*$           | 0.14              |
|                               | $[0.99; 2.97]$     | $[-1168668641189714.25; 199.43]$ |
| Offensive Alliance            | $2.67^*$           |                   |
|                               | $[1.59; 19.13]$    |                   |
| Secret Provisions             | $-0.02$            | $0.66$            |
|                               | $[-0.38; 0.33]$    | $[-1493528118803579.50; 25.83]$ |
| Alliance Years                | $0.03^*$           | $0.02$            |
|                               | $[0.00; 0.11]$     | $[-153166972616500.56; 20.30]$ |

* 0 outside the confidence interval

**Table 3: Pooled TERGM Results for Role Generative Structure, Pre-WW1 System.**
Confidence intervals reflect 95% confidence intervals estimated through bootstrapped pseudolikelihood. 500 replications.

|                               | Consolidators       |
|-------------------------------|---------------------|
| Edges                         | $-1.29^*$           |
|                               | $[-1.55; -1.04]$    |
| Triangles                     | $0.15^*$            |
|                               | $[0.12; 0.20]$      |
| GW Degree (0.5)               | $1.10^*$            |
|                               | $[0.67; 1.52]$      |
| CINC Difference               | $8.45^*$            |
|                               | $[6.64; 10.14]$     |
| Revisionism Difference        | $-0.11$             |
|                               | $[-0.31; 0.09]$     |
| Defensive Alliances           | $2.26^*$            |
|                               | $[1.69; 2.95]$      |
| Secret Provisions             | 0.83                |
|                               | $[-0.13; 1.73]$     |

* 0 outside the confidence interval

**Table 4: Pooled TERGM Results for Role Generative Structure, Interwar System.**
Confidence intervals reflect 95% confidence intervals estimated through bootstrapped pseudolikelihood. 500 replications.
Figure 32: Aggregators Model GOF Diagnostics, Bismarckian System. Similar networks with similar structures will have simulated network statistics (box plots) approximated observed network statistics (line).

provisions establishing institutions. Aggregators show no strongly discernible patterns other than a tendency towards triadic closure.

7.6 Liberal International System

The generative process for each of the roles identified during the Liberal International System is presented in Figure 6. The local alliance networks of Reformers show a tendency towards triadic closure. Reformers also appear to privilege the use of defensive alliances to accomplish their objectives. Aggregators’ local networks show a tendency towards triadic closure and preferential attachment. The alliances formed to aggregate capabilities typically appear to be defensive and
|                                | Balancers  | Consolidators | Aggregators  |
|--------------------------------|------------|---------------|--------------|
| Edges                          | $-5.41^*$  | $-0.41$       | $-2.76^*$    |
|                                | $[-5.83; -5.06]$ | $[-1.46; 1.69]$ | $[-3.83; -1.19]$ |
| Triangles                      | $0.93^*$   | $0.52^*$      | $0.60^*$     |
|                                | $[0.83; 1.05]$ | $[0.39; 0.64]$ | $[0.50; 0.71]$ |
| GW Degree (0.1)                | $7.34^*$   |               |              |
|                                | $[6.67; 8.31]$ |               |              |
| CINC Difference                | $13.08^*$  | $173.29^*$    | $-59.57$     |
|                                | $[11.98; 14.31]$ | $[151.35; 199.84]$ | $[-104.99; 13.00]$ |
| Revisionism Difference         | $0.13^*$   |               |              |
|                                | $[0.03; 0.23]$ |               |              |
| Defensive Alliance             | $1.33^*$   |               | $0.41$       |
|                                | $[1.12; 1.55]$ |               | $[-0.30; 0.85]$ |
| Offensive Alliance             | $1.68^*$   |               |              |
|                                | $[1.42; 2.02]$ |               |              |
| Secret Provisions              | $-0.97$    |               |              |
|                                | $[-2.50; 0.04]$ |               |              |
| Degree of Institutionalization | $0.77^*$   | $7.30^*$      |              |
|                                | $[0.46; 1.01]$ | $[8.58; 10.34]$ |              |
| Alliance Years                 | $0.06^*$   |               |              |
|                                | $[0.05; 0.07]$ |               |              |

* 0 outside the confidence interval

Table 5: Pooled TERGM Results for Role Generative Structure, Bipolar System. Confidence intervals reflect 95% confidence intervals estimated through bootstrapped pseudolikelihood. 500 replications.
Figure 33: Consolidators Model GOF Diagnostics, Bismarckian System. Similar networks with similar structures will have simulated network statistics (box plots) approximated observed network statistics (line).

formed between states of disparate capabilities. Consolidators only show a strong tendency towards triadic closure, a feature consistent with theoretical expectation. GOF diagnostics presented in Figures 42 through 44 are consistent with relatively well fitting models.

8 Notes on Computation

All statistical computing was handled using the R Statistical Computing Environment [Team et al., 2013]. Tables were compiled using the texreg package [Leifeld 2013]. All ego-TERGMs were estimated using v.0.3.0 of the egoTERGM package.
Figure 34: Balancers Model GOF Diagnostics, Bismarckian System. Similar networks with similar structures will have similar modularity densities.
Figure 35: Reformers Model GOF Diagnostics, Bismarckian System. Similar networks with similar structures will have simulated network statistics (box plots) approximated observed network statistics (line).
Figure 36: Balancers Model GOF Diagnostics, Pre-WW1 System. Similar networks with similar structures will have simulated network statistics (box plots) approximated observed network statistics (line).
Figure 37: Aggregators Model GOF Diagnostics, Pre-WW1 System. Similar networks with similar structures will have simulated network statistics (box plots) approximated observed network statistics (line).
Figure 38: Consolidators Model GOF Diagnostics, Interwar System. Similar networks with similar structures will have similar modularity densities.
Figure 39: Balancers Model GOF Diagnostics, Bipolar System. Similar networks with similar structures will have similar modularity densities.
Figure 40: Consolidators Model GOF Diagnostics, Bipolar System. Similar networks with similar structures will have simulated network statistics (box plots) approximated observed network statistics (line).
Figure 41: Aggregators Model GOF Diagnostics, Bipolar System. Similar networks with similar structures will have simulated network statistics (box plots) approximated observed network statistics (line).
|                        | Reformers | Aggregators | Consolidators |
|------------------------|-----------|-------------|---------------|
| Edges                  | −8.09*    | −5.84*      | −2.76*        |
|                        | [−103031198931604.97; −7.00] | [−6.41; −5.28] | [−3.83; −1.19] |
| Triangles              | 0.17*     | 0.52*       | 0.60*         |
|                        | [0.15; 13538949033764.99] | [0.41; 0.60] | [0.50; 0.71] |
| GW Degree (0.1)        | 3.71      | 19.09*      |               |
|                        | [−69.21; 10.85] | [14.63; 23.90] |               |
| CINC Difference        | 36.47     | 15.53*      | −59.57        |
|                        | [−8.23; 57.05] | [13.45; 18.01] | [−104.99; 13.00] |
| Revisionism Difference | −0.37     | −0.06       |               |
|                        | [−1.47; 0.40] | [−0.22; 0.09] |               |
| Defensive Alliances    | 9.68*     | 3.08*       | 0.41          |
|                        | [15.10; 167993289170715.75] | [2.75; 3.47] | [−0.30; 0.85] |
| Degree of Institutionalization | −1.64 | −0.32 |               |
|                        | [−8308151061724.33; 7.03] | [−1.14; 0.37] |               |
| Alliance Years         | 10.32*    | 0.07*       |               |
|                        | [9.00; 14474773169506.38] | [0.04; 0.10] |               |

* 0 outside the confidence interval

Table 6: Pooled TERGM Results for Role Generative Structure, Liberal International System. Confidence intervals reflect 95% confidence intervals estimated through bootstrapped pseudolikelihood. 500 replications.
Figure 42: Reformers Model GOF Diagnostics, Liberal International System. Similar networks with similar structures will have simulated network statistics (box plots) approximated observed network statistics (line).
Figure 43: Aggregators Model GOF Diagnostics, Liberal International System. Similar networks with similar structures will have simulated network statistics (box plots) approximated observed network statistics (line).
Figure 44: Consolidators Model GOF Diagnostics, Liberal International System. Similar networks with similar structures will have similar modularity densities.
References

2017. State System Membership List, v2016. *Tech. rep.*, Correlates of War.

Box-Steinnsmeier, Janet M, Benjamin W Campbell, Dino P. Christenson, and Zachary Navabi. 2017. Role Analysis Using the Ego-ERGM: A Look at Environmental Interest Group Coalition. *Social Networks*.

Brooks, Stephen G, and William C Wohlforth. 2008. *World out of balance: international relations and the challenge of American primacy*. Princeton University Press.

Campbell, Benjamin W. 2018. Detecting Heterogeneity and Inferring Latent Roles in Longitudinal Networks using the Ego-TERGM. *Political Analysis* 26 (3):292–311.

Carr, Edward Hallett. 1939. The twenty years’ crisis, 1919-1939: an introduction to the study of international relations.

Craig, Gordon Alexander, and Alexander L George. 1995. *Force and statecraft: diplomatic problems of our time*. Oxford University Press, USA.

Cranmer, Skyler J, and Bruce A Desmarais. 2010. Inferential network analysis with exponential random graph models. *Political Analysis* 19 (1):66–86.

Cranmer, Skyler J, Bruce A Desmarais, and Benjamin W Campbell. 2018. Indirect Effects and Network Pressures Towards Democracy.

Desmarais, Bruce A, and Skyler J Cranmer. 2010. Consistent confidence intervals for maximum pseudolikelihood estimators. In *Proceedings of the Neural Information Processing Systems 2010 Workshop on Computational Social Science and the Wisdom of Crowds*.

———. 2012. Statistical mechanics of networks: Estimation and uncertainty. *Physica A: Statistical Mechanics and its Applications* 391 (4):1865–1876.

Friedman, Thomas. 1989. US voicing fears that Gorbachev will divide West. *New York Times* 16.

Gleditsch, Kristian Skrede, and Michael D Ward. 2006. Diffusion and the international context of democratization. *International organization* 60 (4):911–933.

Handcock, Mark S, David R Hunter, Carter T Butts, Steven M Goodreau, and Martina Morris. 2008. statnet: Software tools for the representation, visualization, analysis and simulation of network data. *Journal of statistical software* 24 (1):1548.

Hartshorne, Richard. 1950. The Franco—German Boundary of 1871. *World Politics* 2 (2):210–250.

Holsti, Kalevi J. 1970. National role conceptions in the study of foreign policy. *International Studies Quarterly* 14 (3):233–309.

Hunter, David R, Steven M Goodreau, and Mark S Handcock. 2008. Goodness of fit of social network models. *Journal of the American Statistical Association* 103 (481):248–258.

Ikenberry, G John. 2012. *Liberal Leviathan: The origins, crisis, and transformation of the American world order*. Princeton University Press.
Jervis, Robert. 1985. From balance to concert: a study of international security cooperation. *World Politics* 38 (01):58–79.

Job, Brian. 1976. Membership in inter-nation alliances, 1815-1965: an exploration utilizing mathematical probability models. *Mathematical models in international relations. New York: Praeger* 74–109.

Kissinger, Henry. 1994. *Diplomacy*. Simon and Schuster.

Lebow, Richard Ned. 1994. The long peace, the end of the cold war, and the failure of realism. *International Organization* 48 (2):249–277.

Leeds, Brett, Jeffrey Ritter, Sara Mitchell, and Andrew Long. 2002. Alliance treaty obligations and provisions, 1815-1944. *International Interactions* 28 (3):237–260.

Leifeld, Philip. 2013. texreg: Conversion of Statistical Model Output in R to LATEX and HTML Tables. *Journal of Statistical Software* 55 (8):1–24.

Leifeld, Philip, Skyler J Cranmer, and Bruce A Desmarais. 2014. xergm: Extensions of Exponential Random Graph Models. *R package version 0.51*.

———. 2017. Temporal exponential random graph models with btergm: Estimation and bootstrap confidence intervals. *J. Stat. Softw*.

Li, Richard PY, and William R Thompson. 1978. The stochastic process of alliance formation behavior. *American Political Science Review* 72 (4):1288–1303.

Mandelbaum, Michael. 2005. *The case for Goliath: How America acts as the world’s government in the twenty-first century*. Public Affairs.

Marshall, Monty G, Keith Jaggers, and Ted Robert Gurr. 2002. Polity IV project: Dataset users’ manual. *College Park: University of Maryland*.

Mearsheimer, John J. 1990. Back to the future: Instability in Europe after the Cold War. *International security* 15 (1):5–56.

Morrow, James D. 1991. Alliances and asymmetry: An alternative to the capability aggregation model of alliances. *American Journal of Political Science* 904–933.

———. 2000. Alliances: Why write them down? *Annual Review of Political Science* 3 (1):63–83.

Palmer, Glenn, Vito d’Orazio, Michael Kenwick, and Matthew Lane. 2015. The MID4 dataset, 2002–2010: Procedures, coding rules and description. *Conflict Management and Peace Science* 32 (2):222–242.

Pevehouse, Jon C. 2005. *Democracy from above: Regional organizations and democratization*. Cambridge University Press.

Salter-Townshend, Michael, and Brendan Thomas Murphy. 2015. Role analysis in networks using mixtures of exponential random graph models. *Journal of Computational and Graphical Statistics* 24 (2):520–538.
Schroeder, Paul W. 1992. Did the Vienna Settlement Rest on a Balance of Power? *The American Historical Review* 97 (3):683–706.

———. 1996. *The transformation of European politics, 1763-1848*. Oxford University Press.

Singer, J David, Stuart Bremer, and John Stuckey. 1972. Capability distribution, uncertainty, and major power war, 1820-1965. *Peace, war, and numbers* 19:48.

Siverson, Randolph M, and George T Duncan. 1976. Stochastic models of international alliance initiation. *Mathematical models in international relations*. *New York: Praeger* 110–131.

Siverson, Randolph M, and Charles McCarty. 1982. Alliances in the interwar era, 1919-1939: A reexamination. *Western Political Quarterly* 35 (1):24–32.

Slater, David. 2008. *Geopolitics and the post-colonial: rethinking North-South relations*. John Wiley &amp; Sons.

Snijders, Tom AB, Philippa E Pattison, Garry L Robins, and Mark S Handcock. 2006. New specifications for exponential random graph models. *Sociological methodology* 36 (1):99–153.

Taylor, Alan John Percivale. 1954. *The struggle for mastery in Europe, 1848-1918*. Oxford University Press.

Team, R Core, et al. 2013. R: A language and environment for statistical computing.

Thérien, Jean-Philippe, Gordon Mace, and Stefan Gagné. 2012. The Changing Dynamics of Inter-American Security. *Latin American Policy* 3 (2):147–163.

Waltz, Kenneth N. 1993. The emerging structure of international politics. *International security* 18 (2):44–79.