

ENDOGENOUS NETWORKS AND INTERNATIONAL COOPERATION

ABSTRACT. The rise of social network analyses in the social sciences has allowed empirical work to better account for interdependencies among actors and among their actions. However, this work has been, to a large extent, descriptive: it has treated these actions as exogenous and immutable. In many cases these networks describe actions like alliance formation or trade – phenomena that are the outcome variables for programs of social scientific research. In this paper, I attempt to account for both interdependencies and the endogenous nature of networks by incorporating formal theory; helping answer the question of how these networks arise by looking at the incentives of actors to form links with each other. I discuss the appropriate solution concept for a network formation game, and present an algorithm for finding the equilibrium of these networks computationally as well as ways to compare the theoretical networks to observed ones in order to evaluate the fit of the theory. I apply these methods to the study of international cooperation – a subject where both the interdependencies and purposive nature of actors must be accounted for. The theoretical network is able to reproduce a number of important observed characteristics. Still, there are more factors that must be accounted for if we want to understand how the network of international cooperation is formed.

INTRODUCTION

In recent years the field of international relations, and political science more generally, has increasingly employed social network analysis. Such analyses allow scholars to model and account for interdependencies rather than assuming our variables are purely atomistic. What has received less attention is the fact that many of the networks we are interested in are, themselves, the products of strategic processes, and these processes are in fact the sorts of behavior that social scientists study. In this paper, I propose to apply formal and game theoretic logics to the formation of networks, and in particular networks of interstate cooperation.

I proceed by discussing some of the history of network analysis, and its relationship to formal theory. I then discuss the solution concepts used to determine a theoretically stable network, introduce an algorithm for finding this equilibrium, as well as a simple theory of interstate cooperation. Finally, I discuss the theoretical equilibrium that is found when applying these techniques and theories to the fifty richest countries in the world and discuss how it compares to a similar empirical network.

Network theory began with graph theory, which goes back as far as Euler’s work on the bridges of Königsberg in the early 18th Century. In 1959, Erdős & Rényi (1959) introduced random graphs, allowing the first probabilistic interpretation of graphs. From this idea grew the notion, and discipline, of network science. In network analysis, the basic concern is to capture the interdependencies among the elements, or nodes, through the study of the linkages among them. Through the study of the structure of relationships, sociology, political science, and economics have begun to use networks to make important advances. Granovetter (1973) discovered the “strength of weak ties” which are shown to be essential in holding different networks together. White (2002) brought together the idea that markets are embedded in

Date: January 6, 2016.

networks, based upon his early work on block structures (White, 1963). Watts (2003) made popular the network idea of small worlds and scale-free networks. Bearman, Moody, and Stovel (2004) showed how such networks play out in the sexual networks in US high schools . An overview of these contributions is to be found in Ward et al. (2011).

Much of the applications of this work, especially in international relations examines networks composed of strategic actions. Scholars examine phenomena like alliance networks, networks of trade and foreign aid, or networks of conflict and contagion. These are networks where each link is the sort of strategic action that social scientists have devoted great effort to theorizing about, but much of network analysis treats these links as exogeneous. It instead focuses on either describing the networks, or using the characteristics of the networks as a variable in another analysis. I propose that application of formal models and rational choice can help to endogenize these networks, and can help us better understand how they come to be.

Work, which combines the network insights about interconnectivity with game theoretic work that focuses on strategic interaction, is just beginning. A review of this burgeoning effort is provided by Jackson & Zenou (2013). There are works on choice in networks under uncertainty, as typified by work of Matthew Jackson and his collaborators (Jackson & Yariv, 2007; Jackson & Watts, 2002b,a; Galeotti et al., 2010), while another thread of this research follows the idea of looking at equilibria as a function of fixed, known networks (Patacchini & Zenou, 2008). A third thread in this literature focuses on strategic complements and substitutes (Chwe, 1999, 2000; McPherson et al., 2001; Jackson & Lopez-Pintado, 2013; Brambor et al., Winter 2006). I follow this latter thread.

In general the work on formal models of networks has fallen into two categories. In the first category, Jackson (2008) and Goyal & Vega-Redondo (2013), have worked on deriving rules of behavior based on formal set-ups that are not readily apparent, especially due to the complicated non-linearity of effects on networks. In Siegel (2009) as well as in Bramoullé & Kranton (2007), formal work has shown how, when games are conditioned by network structures, the effects of the network on the equilibrium are often complex and rely on non-obvious characteristics of the network: Bramoullé & Kranton find that presence of cooperative equilibria are driven by a network's lowest eigenvalue, while Siegel shows that the link between network ties and voter participation is non-monotonic. In some models, the network conditions what actions players can take, whereas in others it simply conditions the flow of information to the players. Even in the second case, scholars have found that the addition of a network component can have a stark effect on the results, conditioning the level of cooperation in Larson (2011), weakening the relationship between inter-group ties and ethnic peace in Larson (2016), or determining whether a principal is willing to expend resources to monitor her agents in Patty & Penn (2014).

The other category – including Metternich et al. (2013)'s work on explaining conflict in Thailand and Chyzh's work on trade intermediaries – has taken the implications of formal models of networks and tested them empirically (Chyzh, 2013, 2016). In a similar vein, in this issue Maoz & Zoyce (2016) use an agent based model to derive rules for network formation and change, and test the implications of that agent based model.

This project attempts to go further by looking at the formal results in a larger, more complex network based on empirical data. I expand this idea to a much larger network, both

in terms of scope and extent. I examine interactions among the fifty most wealthy countries, looking both at the actual interactions, and those predicted by a simple formal model, I find that this relatively simple model can reproduce many of the characteristics of the observed network, and this can yield important insights about the determinants of international cooperation. I argue that using computational techniques and models of network formation can give us tools to test strategic theories when network interdependencies exist, and can help better account for the strategic nature of many studies of networks in the social sciences.

NETWORKS, UTILITY FUNCTIONS, AND SOLUTION CONCEPTS

It is important for students of international relations to understand networks, and this requires theories of network formation. The need to understand networks has been outlined by a number of scholars, with Dorff & Ward (2012) providing a recent example. If we ignore the interdependencies underlying international interactions, we severely limit our ability to understand and predict these interactions. If we believe that both empirical relations are non-dyadic, that we need to account for networked interdependencies, and we believe in motivating our empirical work with robust theoretical models, then there is a sharp need for us to be able to model how networks form. Thus, in this section I discuss how to determine which networks are stable solutions to network formation games, and then in the next section I go deeper into one such network formation game, and show the value of preliminary results in the formation of an international cooperation network.

Nash Equilibria and networks. We can conceive of network formation as an n -player game, where each of n players decides whether or not to propose a link with the other $(n - 1)$ players. Each player has a utility function that can be based on their endowments, the endowments of their partners, those of their partners' partners, and so on. If the desire to form a link is mutual – if player i wants to form a link with player j *and* player j wants to form a link with player i – then link ij is formed, if either player objects, the link is not formed. Each player's actions will consist of providing a list of players who they wish to link with, and the final outcome will be a network with all mutual links.

For a network to be a Nash Equilibria of a network formation game, no player can unilaterally do better by adding or deleting any number of players from the aforementioned list. If, holding the rest of the network constant, a player i would prefer not to form a link with player j , yet they are linked in the network, it is not an equilibrium because the player would strictly prefer an identical list of partners, but with player j was removed. If player i and player j were not linked in a network, and player j included player i on her list, then if player i preferred to be linked to player i , this network would not be an equilibrium of the network formation game.

The need for links to be mutual hamstrings the usefulness of Nash Equilibria as a solution concept in network formation games. Consider the following extremely simple game: each player earns \$1 for every link formed. Now the obvious solution to this game is for each player to link to each other player, maximizing their profits. However, consider a game where no player includes any other player on their list of potential links. These strategies also comprise a Nash Equilibrium: no player can unilaterally do better by adding any players to the list, because this will not result in the formation of additional links. Similarly, any other non-full network could be a (deeply inefficient) equilibrium of this game. Because Nash Equilibria require only

that a player in isolation cannot improve her lot, these equilibrium will never check whether players could jointly do better by adding links. Thus, it might be fruitful to turn to solution concepts that are robust to groups of players acting jointly.

Pairwise Stability. The inability of a player to unilaterally generate a link in a network formation game pushes us to a solution concept other than Nash Equilibrium: in particular, a solution concept allowing players to act jointly. The simplest of these concepts is Pairwise Stability. As defined by Jackson & Wolinsky (1996), a network (g) is pairwise stable if both of the following are true.

- (1) $\forall ij \in g, u_i(g) \geq u_i(g - ij)$ and $u_j(g) \geq u_j(g - ij)$
- (2) $\forall ij \notin g, \text{ if } u_i(g + ij) > u_i(g) \text{ then } u_j(g + ij) < u_j(g)$

All neighbors in the network prefer that network to one that deletes their link, and if any member of an unconnected dyad would prefer a network where they were connected, their partner must prefer to remain unconnected.

The shortcoming of pairwise stability is that, while it can deal with a pair of players acting in unison, it can only compare a network to another that is one link away in either direction. To see the flaw in this solution concept, consider the following utility function for a five player network formation game – if a player forms one or more links they pay a cost of \$10; at the end of the game each player earns \$1 for each link formed. This game has two pairwise stable networks, the empty network and the full network, but clearly the empty network is a more plausible result than the full one: any player in the full network could easily increase their payoff by changing their list to be an empty one. The ability for players to unilaterally remove any number of links would allow us to weed out this implausible equilibrium. We thus turn to a solution concept which combines the requirements of Nash Equilibrium and Pairwise Stability.

The solution concept: Pairwise Nash Equilibrium. Our solution concept, termed Pairwise Nash Stability, is simply that a network must be both a Nash Equilibrium of the network formation game and a Pairwise Stable Network. A network is considered a Pairwise Nash Equilibrium if:

- (1) No player can improve their payoff by unilaterally deleting any number of links that they are a party to.
- (2) If there is no link between two players, then creating such a link can only improve the utility of one player if it decreases the utility of the other (Jackson, 2008).

The pairwise Nash concept does have a few flaws. One is that a Pairwise Nash Equilibrium will not always be Pareto-efficient: if improving payoffs would involve three actors behaving in concert, further equilibrium refinements are needed.¹ In addition, a pairwise Nash Equilibrium may not always exist: every possible network may be part of an improving cycle. An improving cycle is a set of networks C where, for some network formation game, any $g \in C$ and $g' \in C$ there are a set of rational individual deviations that take us from g to g' : for instance, decisions made by players to improve their utility in a game (either by adding or

¹In particular, strong stability which requires equilibria be robust to any number of actors changing their strategies.

deleting links) take us from \mathbf{G} to \mathbf{G}' , from there to \mathbf{G}'' , and then back to \mathbf{G} . If every possible network is in the improving cycle, there will be no Pairwise Stable Nash Equilibrium.

Thus, choosing a solution concept requires dealing with two competing demands: on the one hand we want to choose a solution concept where equilibria are not suboptimal or unrealistic, and on the other hand we want requirements that are not so stringent that nothing satisfies the solution concept. Pairwise Nash Equilibria will sometimes give second-best outcomes, but this will be driven by many of the collective action problems and fear of individual deviation that we find important in the study of international politics. Further, while Pairwise Nash Equilibrium can feature improving cycles, these are rare, especially when dealing with continuous convex utility functions, and failure to find an equilibrium may be the consequence of chaotic behavior and incentives given to players, rather than a failing of the solution concept. Finally, I find that Pairwise Nash Equilibria rules allow for international cooperation (the Pairwise component), but not while losing sight of the fact that each actor is a rational maximizer (the Nash component).

An example of the solution concepts. Consider the following set of networks and payoffs, from a simple three player system from Jackson (2008; 375). Note here that the particular links in this network do not matter, simply that each player's utility is determined by the total number of links in the network. If there are no links, each player gains 0 utility, if there is one link, the linked players gain 1 utility and the non-linked player gains 0. If there are two links, all players gain -2 utility, and if there are three links, all players gain -1 . This network has three Pairwise Nash Equilibrium outcomes – any pair of players announce links with each other.² Neither linked player here would choose to delete this link as they would lose utility. Further, no player would like to add an additional link as they would then lose more utility. The empty network is a Nash Equilibrium but not a pairwise equilibrium: if no player announces any other link, then deviation will never be profitable, but two players working in concert could both improve by forming a link. The network with two links is the worst of all possible worlds, and is clearly not an equilibrium, and the deletion of a single link would benefit all players, making it neither pairwise stable nor Nash. Finally, the full network is Pairwise Stable, as removing a link would decrease all players' utility and adding one is obviously impossible, but it is not a Nash equilibrium as any player could do better by unilaterally severing two links. This example clarifies the benefits of Pairwise Nash equilibrium over other solution concepts, but we find the equilibrium network by examining each possible network; if the network were a few nodes larger, or the nodes were not identical, this would be a much more onerous task, thus we turn to computational techniques.

Finding an equilibrium algorithmically. It is prohibitively difficult to find an analytic solution to a moderately complex network formation game. What is particularly challenging is that actors' desires to add a link are conditional on all the other links in the network due to concavity and potential diminishing marginal returns. As a result this paper turns to a computational solution via a program that finds a stable equilibrium of the game. The program uses the following algorithm:

²There are more strategy profiles, which involve the linked pair choosing whether or not to announce a link with the third player, but this link will never be realized in equilibrium.

- (1) Generate a network \mathbf{G} . This could be a random network, or a theoretically interesting one (like an empty or full network). Specify a given utility function for each actor.
- (2) Calculate each actor's utility in the network.
- (3) Iterate over every empty link $ij = 0$, in a random order and determine whether $u_i(\mathbf{G} + ij) > u_i(\mathbf{G})$ AND $u_j(\mathbf{G} + ij) > u_j(\mathbf{G})$, if so set $\mathbf{G} = \mathbf{G} + ij$ and return to step 2, else continue.
- (4) Iterate over every actor i in the network, in a random order,³ comparing the utility function for every subset of that actor's links. If utility is maximized by a subset smaller than the full set, replace that actor's links with the subset and return to step 2, if the maximum is found in the network \mathbf{G} , we have found our pairwise stable network.

This equilibrium can, given time, find a Pairwise Stable Nash Equilibrium network. However, there are two remaining issues: multiple equilibria may exist and improving cycles can occur. Multiple equilibria will persist because Pairwise Nash Equilibria will find local maxima, as well as the global maximum to be stable. We want to find all equilibrium networks, especially if we want to compare these theoretical networks to observed ones. I accomplish this by running the algorithm multiple times using different (and random) starting networks—generated by making the links of the starting network the results of a Bernoulli(p) trial with differing values of p . Asymptotically this should give us the set of Pairwise Nash networks for a given utility function. When multiple equilibria are found, the researcher can choose criteria for which to focus on. For example, when studying interstate cooperation (as this paper does) one might choose to focus on a Pareto dominant game given findings that cheap talk enables players to coordinate on the efficient outcome (Blume & Ortmann, 2007).

We also may see a problem with an improving cycle, where the algorithm takes us from \mathbf{G} to \mathbf{G}' , from there to \mathbf{G}'' , and then back to \mathbf{G} .⁴ This is both a technical problem and a substantive one. Substantively a network formation game may have no Pairwise Nash Equilibrium because every possible network is part of an improving cycle. This implies that no set of actions by the players in the game will be stable, and we will expect to see great change and upheaval in the network over time. We would like to know if, for a set of parameters, no stable network exists. However, it is also the case that a given game may have certain networks that are part of an improving cycle, and other networks that are on a path which leads to a stable network. Here, improving cycles become more of a technical problem. If the algorithm reaches a network that is part of the cycle, it will run indefinitely. I account for this by keeping a list of the stages the network has passed through, and if we return to one of these networks, we jump to a new, random network, and begin the algorithm anew. Before running the algorithm, one can also specify a threshold for the maximum tolerable cycle length: this is the number of different networks that are established to be part of a cycle before the algorithm terminates, or moves on to a different utility function. If we are particularly interested in finding out whether a given set of parameters has an equilibrium, this cycle length should be 2^N (the number of possible networks), but if we are just interested in finding an equilibrium network given approximate parameters, examining the entire network space will be prohibitively costly, in both computational power and time.

³These random orders are generated anew each time through this step of the process.

⁴Or possibly forms a cycle with even more steps.

To summarize the process, we look at the utility each actor is getting from a network. We then compare that utility to all of the possible networks with one more link and see if both actors would be better off forming that link, and continue to do this until no more links are profitable. We then investigate whether any actor would like to unilaterally remove any number of links: we do this by looking at an actors neighbors, and seeing if that actor would be better off with any subset of those neighbors. Once we can iterate through all actors with no one choosing to add or delete any link, we have a pairwise stable network. If there are any issues with multiple equilibria, or improving cycles preventing us from finding an existing network, we attempt to avoid them by jumping to a different, random, network. This can require many iterations for a larger network.

MODELING INTERSTATE COOPERATION

The literature on international cooperation has received less attention when compared to the abundance of work on international conflict. Most efforts to understand international cooperation have been focused on an institutional framework such as alliances, the law of the seas, or the environment for example. Some of the literature which seeks to study the phenomena of international integration and security communities looks at cooperation more generally: Lipson (1984) noted the ability of states to cooperate and institutionalize these cooperations in economic, as compared to security dimensions, and Keohane (1998) pointed to the ability of international institutions to facilitate cooperation in the midst of security concerns, while Grieco (1988) posited the risks of relative gains as a barrier to cooperation Snidal (1991) demonstrated how even when states are concerned with relative gains for security reasons, cooperation is rational even if a state's partner benefited more. Exemplar research on this thread would certainly include great works in the tradition of Young (1989) which explained when states would choose to subject themselves to rules and constraints, and how those enabled greater cooperation – Garrett (1992)'s account of state's limiting their sovereignty in search of economic growth in the European Economic Community, Litfin (1994)'s work on international norms and environmental cooperation, and Kaul et al. (1999)'s discussion of institutions needed to protect and continue the provision of global public goods are a few examples – but the work of Karl W. Deutsch, investigating how to foster a cooperative security community in Western Europe, is perhaps seminal in this category (Deutsch et al., 1957). Additional works have looked a cooperation at the most general level, notably (Fearon, 1998)'s argument that we should avoid generating separate theoretic models for different types of cooperation and instead look at them all under a bargaining framework and Ward (1982)'s examination on how cooperation and conflict interrelate over time.

Some recent efforts have undertaken to look at networks as a way of understanding cooperation, but this thread of research is rather thin: Alter & Meunier (2009) looked at networks of international agreements, Warren (2010) at different types of alliance networks, Ward et al. (2013) at trade networks, while Gomez et al. (1999) tried to map research cooperation by looking at citation and co-authorship networks, and Holm et al. (1996) investigated at networks of international business relationships and how they effect interfirm cooperation and profitability. In this very issue Haim (2016) has investigate the effect of alliance networks on patterns of interstate trade. Suffice it to say that the theory of international cooperation is

comparatively undeveloped, and those studies that use networks are generally empirical rather than theoretical.

A model of cooperation. When considering what to include in this model of network formation, I borrow much from the extant literature on international cooperation. In particular, I focus on the roles of similar domestic institutions, geographic proximity, and the links between conflict and cooperation. It is a model of cooperation generally, heeding the argument in Fearon (1998) that ‘understanding problems of international cooperation as having a common strategic structure is more accurate and perhaps more theoretically fruitful.’

Joint democracy and homophily. There has been a large and fruitful research agenda concerning the link between domestic regime types and conflict. Scholars have found, quite consistently that democracies are less likely to engage in conflict with other democracies, and to a lesser degree, that autocracies are less likely to engage in conflict with other autocracies (Maoz & Russett, 1993; Bueno de Mesquita & Smith, 2009). There have been similar, albeit less prominent, findings about the link between regime type and international cooperation. Leeds (1999) finds empirical purchase for the idea that democracies are less likely to cheat and less likely to enter agreements with risk of cheating and so democracies are more likely to cooperate with other democracies, and autocracies with other autocracies, while Bliss & Russett (1998) find support for the existence of greater trade cooperation between democracies. Warren (2016) finds that states with similar domestic politics are likely to ally with each other, and that states allied with democracies are likely to become more democratic. The tendency of states to cooperate with similar regime types maps well onto the concept of homophily in network analysis. In fact, Maoz et al. (2006) found in the international alliance network that a model of homophily on domestic institutions was quite useful in explaining the formation of links. Similar regime types may be more likely to cooperate because of a shared risk and tolerance for cheating, or because domestic pressures reward them for cooperating with ‘good’ regimes more than bad ones, but it seems both plausible and empirically justified that states should gain somewhat more utility when they cooperate with like regimes.

Contiguity. In recent decades, the proliferation of regional cooperative associations – the EU, ASEAN, NAFTA to give a few examples – has highlighted the fact that in international cooperation, proximity matters. Partly, this is because many cooperative activities are cheaper when conducted over short distances: trade requires less costs to ship and it is easier to project military power in support of a neighboring ally than a distant one. It is also the case that proximate states are more likely to have similar interests that require cooperations: balancing against an aggressive mutual neighbor, dealing with refugees from a conflict in the region, creating a larger economy of scale. Thus, when cooperation is with a more proximate stake, there are often higher benefits, and usually lower costs.

Conflict and cooperation. While conflict and cooperation are often seen as opposite actions in the international sphere, some research has shown that their relationship is closer and more complicated than that (Ward, 1982). On the one hand, when cooperating with a foe, states must be concerned with relative gains: there is a risk that a rival will take the benefits of cooperation and if they gained more, use that advantage to triumph in a future conflict (Grieco, 1988). On the other hand, if two states have conflicts, it is a signal that they both have interests in common,

and the risk for future conflict. The first means that they have possible areas for cooperation, and the second increases the benefit of cooperation as a way to potentially forestall violent conflict – consider the recent agreement between the United States (and the rest of the P5+1) and Iran, the threat of conflict made cooperation more likely, and if conflict were impossible, no agreement would have been necessary. The presence of conflict makes cooperation more risky – if it fails the consequences may be more severe – but it also significantly increases the potential gains from cooperation and costs of non-cooperation.

Modeling cooperation. Given these stylized facts, I attempt to generate a network of interstate cooperation. To provide a test of concept, I look at the 50 states with the highest level of GDP, and attempt to generate a network of cooperation that would satisfy the Pairwise Nash Stability solution concept. To do this, however, there is a need to translate the thoughts and literature about cooperation into a formal utility function, and determine the functional forms that maps each of these factors into utility and state action.

The goal for this utility function is not to provide a novel theoretical model of international cooperation, or to provide a perfect prediction of empirical cooperation. Rather, the goal here is a test of the network formation concept, can we generate a model with some interesting theoretical insights about cooperation, that shares some similarities with observed patterns of cooperation. We can then learn both from the things the model gets right, and what it gets wrong. This can show the utility of network formation models, and motivate further exploration, both to develop more ambitious theories and to create models that have more close fits to the empirical world.

The first jumping off point for our utility function is Jackson’s work on homophily networks. There the utility of a network to a player is

$$\alpha \log(s) + \beta \log(s + d) + \gamma \log(d) \tag{1}$$

where s represents the number of links to nodes of the same type and d the number of links to nodes of different types. If $\alpha > \gamma$, a network will exhibit homophily, whereas if $\gamma > \alpha$ it will exhibit heterophily, these will be mediated by the relative sizes of α, γ to β , as larger values of β will mute the effect of type. I first adapt this utility function to the case of interstate cooperation by mapping the ‘same’ and ‘different’ types to the subject of domestic political institutions. In particular, the distinction is between consolidated democracies on the one hand (Polity score larger than six) and non-democracy on the other. The next issue is that in interstate politics, not all states are equal—material resources matter and all else being equal, we would expect states to prefer cooperating with richer states rather than poorer ones. So I replace s and d , with the sum of the capabilities (normalized to between 0 and 1) of similar and different types of states. This maps to the finding in the literature that states are more likely to cooperate with other states if they have similar domestic institutions, and it also incorporates decreasing returns to cooperation, allowing the network to play a role in state’s decision making processes.

I choose to have geographic proximity affect state utility through reducing the costs of war. In particular, this is because many types of cooperative behavior are easier, and require fewer material resources when done over a smaller distance. Thus, we have one cost of conflict for non-contiguous states ($c > 0$) and a second, lower cost gc where $g \in (0, 1)$ for those states that share a common border.

Our final factor brings in the possibility for state interactions that are conflictual. We look at event data for conflict above a certain threshold, and then we add (or subtract) utility to a player for cooperating with those states that the state is in conflict with (the additional utility is based on state capability). One could make the argument either way: a player has more zero-sum concerns about cooperating with potential enemies, but they also have higher benefits to improving the relationship, and more common areas of interest to make cooperation fruitful. So our final utility function is

$$U_i(G) = \alpha \log\left(\sum_{\text{same}} k_j\right) + \beta \log\left(\sum k_j\right) + \gamma \log\left(\sum_{\sim\text{same}} k_j\right) \\ + \kappa \sum k_j \mathbf{I}_{\text{conflict}} - N_{\text{close}}^i c g - N_{\text{far}}^i c$$

where k_j is the capability of state j , $\mathbf{I}_{\text{conflict}}$ is an indicator function that takes on a value of 1 if two states have above a certain threshold of conflict and 0 else, N_{close}^i and N_{far}^i are counts of the number of contiguous and non-contiguous states cooperating with state i .

Data. To test the utility of this model of international cooperation, I take a subset consisting of the world’s 50 richest states. I limit my analysis to these states because increasing the number of states raises the time and computational power required for this analysis factorially. The states are chosen based on their economic prosperity, represented by total GDP in constant 2000 dollars, from the year 2012. This also provides our measure of the benefits of cooperation. For these 50 countries, I take the square root of their total GDP and then normalize it to 1 by dividing by the largest value (that of the United States).

Our measure of political homophily comes from the Polity Project’s Democracy score (Jagers & Gurr, 1995). This score is a 0-10 variable that measures the extent to which executive recruitment is open, there are constraints on the actions of the state’s executive, and there is robust political competition. The utility function includes a dyadic variable for contiguity. To obtain this variable I used CShape’s dataset (Weidmann et al., 2010), which includes values for the shortest distance between the borders of two states. If a state’s borders are less than 150 miles from those of another state, these states are considered contiguous and the cost of a link is adjusted accordingly. I also include a dyadic variable for the number of conflictual events. Here we use the counts of events coded as conflictual according to the CAMEO codebook, in the ICEWS Data in the year 2013: this is a dataset of machine coded stories taken from news articles.

Description of an equilibrium network. For this set of countries, we sought to find a Pairwise Nash Equilibrium with the parameters $\alpha = 1.4$, $\beta = 0.6$, $\gamma = 0.7$, $c = 0.075$, $\kappa = 0.5$, $g = 0.45$ and the threshold for conflictual action set at 200. The values of α , β , and γ lead the network to exhibit homophily, as a state gains twice as a state will gain 1.54 times as much utility only cooperating with similar states than they would if states with the same level of wealth were all different. The size of this effect will change once there is cooperation with both similar and different states, but the relative homophile will remain. The values of g imply that the benefits from cooperation with a neighbor need be less than half as high for it to be worth the cost. The value of κ implies that states will choose to cooperate with most of the richer states with whom they also have conflict, though the high-threshold for conflictual action means that this will only inhere when the relative intensity of this conflict is quite large.

The choice for the cost of cooperation was done with an eye towards the overall size of the network: if the cost were too low states would all cooperate with all other states, too high and the network would be too sparse for real analysis, this value allowed for a medium sized network. The results change, but not radically when one makes small changes to the different parameters, and this sensitivity is shown in Figure 4.

The equilibrium network is depicted in Figure 1.⁵ The 50 states had a total of 158 cooperative edges, with the most connected states being Germany and the United States (both had a degree of 11). The United States is part of a cluster with the Germany, Japan and China, with each linked to the other. The other notable clusters in the theoretic network were generally geographic: there was a cluster in the Middle-East with Iran, the UAE, Saudi Arabia, and Turkey, and interestingly, a Pacific cluster centered on Indonesia, Malaysia, the Philippines and Singapore.

[Figure 1 about here.]

TESTING THE THEORETICAL NETWORKS

For our empirical analysis, I use the records of cooperative events since the year 2001 as coded by the ICEWS dataset (Gerner, 2009). ICEWS scans vast quantities of textual information, and with a bag-of-words approach, uses a developed ontology to glean information about who does what to whom for all countries in the world from 2001 to the present. This can be further parsed to determine whether a given story indicated that two states were materially cooperating. We used this source to provide counts of the cooperative events, which we then reduced to a binary measure of cooperation by comparing them to a threshold. To facilitate comparison, we used an empirically derived threshold of 118 – allowing the empirical and theoretical networks to have nearly same density of links (158 for the theoretical network, 159 for the empirical one). This represents the 159 dyads with the highest volume of cooperation in news articles in the year 2013, as determined by the aforementioned approach. The results are depicted in Figure 2.

I have a few reasons for using this event data, as opposed to other common measures of cooperation such as trade or alliance data. If we want to measure a strategic model of cooperative network formation, we need the cooperation being measured to be cooperative acts by the executive decision maker of a country. We can ensure that we only use the ICEWS stories counting acts by a countries government or leader, whereas using counts of economic or technological cooperation will include many actions taken by individual citizens or researchers. We also would like the cooperative events to correspond to the factors we believe predict cooperation, which pushes against using alliance networks or NGO co-membership because these are often the products of past cooperative environments, inertia, and bureaucratic incentives. Finally, it is important to have a measure of cooperation with gradations so that we do not risk

⁵The total list of predicted edges is too long to fully present, but is available upon request. A few of the edges are: (Turkey with Australia, Iran, Indonesia, Sweden, Greece, Switzerland, Argentina, Russia), (Italy with United Kingdom, India, France, United States, Austria, China, Switzerland, Japan, Germany), (France with United States, Netherlands, United Kingdom, Switzerland, China, Belgium, Japan, Germany, Spain), (Israel with Saudi Arabia, Iran, South Africa, Egypt, Denmark, Singapore, Thailand), (Iran with Pakistan, Saudi Arabia, United Arab Emirates, United States, Russia), (Germany, with United States, Poland, United Kingdom, Switzerland, Sweden, China, Belgium, Japan), (United States with Mexico, India, United Kingdom, China, Japan, Russia), (Switzerland with Austria) .

having either total cooperation or no cooperation in our empirical record; by using event data we are able to match the level of cooperation observed empirically to the level in the theoretical network and are thus able to focus comparison on patterns of cooperation, rather than having it be swamped by differences in scale.

[Figure 2 about here.]

Looking at the empirical network a few things become evident: first, while both networks have the United States as the most connected node, the empirical data is much more centralized. In particular, the United States has 37 links, and it is followed by China with 27, Russia with 20, Japan with 18, and Iran with 17. This highlights a few problems with the link between the theoretical and empirical model: they are capturing somewhat different things. By looking at the countries with the most cooperative events, we are looking at the cases with the highest intensity of cooperation, whereas the theoretical model looks not for the best overall cooperation, but the cooperation best for each individual state, thus the noting of state cooperation in the Pacific and Latin America for example. The other issue is a possible reason to either switch to a different empirical dataset, or significantly increase κ : the newsworthiness of cooperation is in large part a function of the risk of conflict, so cooperation between the US and Russia or Israel-Iran is much more likely to be picked up than cooperation between France, Germany and Italy. Therefore, while only about one third of the links are in common between the two networks, another third of the links in the theoretical network are ranging from plausible to insightful predictions of cooperation.

Measures of network comparison. One useful test of the models descriptive ability would be looking at the characteristics of a theoretic network and how it compares to the observed ones. Authors working on Exponential Random Graph Models (hence ERGM) have proposed a form that looks at four descriptive statistics of graphs to determine how similar two graphs are. These four statistics are *geodesic distance distribution*, the proportion of dyads in the matrix that have a shortest distance of k links; the *edgewise shared partner distribution*, the distribution of $EP_0, EP_1 \dots EP_{n-1}$ where EP_k is the number of links between two nodes that have exactly k neighbors in common, divided by the total number of links in the graph; the *degree distribution*, which is the distribution of the number of nodes with exactly k relationships; and the *triad census distribution*, the proportion of triads with exactly 0, 1, 2 and 3 edges between them. The statistics for our model of cooperation are described in Figure 3.

[Figure 3 about here.]

These network statistics drive home the commonalities of the two networks as well as where they diverge. The edgewise shared partner distribution and the triadic census vary little between the empirical and theoretical networks. However, while they share the same number of links, the theoretical network is much less centralized than the empirical one. This is shown most prominently in the degree distribution, while the observed network has a few very central states, and a large number of more peripheral ones, while in the theoretical network most states have a moderate number of links. The same seems to hold in the geodesic distance, where a greater number of dyads have no possible path between them, but those states with a path are more likely to have a distance of one or two, compared to the greater connectivity, but longer mean distance of the theoretical network. One possible issue here is that the use of concave functions in the utility terms are poorly-suited to the data generating process – diminishing

returns may not hold for cooperation, as state A cooperating with state B may not reduce the benefit of cooperating with state C. The shared partner distribution and triadic distribution, show remarkable similarities between the two networks.

Single value measures of network fit. While the four measures of network similarity in the above section ably characterize different facets of network structure, there is also value in providing single values that characterize how similar two networks are. In table I I present two such measures. I follow Koutra et al. (2011) in looking at the ‘eigenvalue method’, a way of comparing two graphs by looking at the distance between the eigenvalues of the Laplacians. More formally, it generates Laplacians, by first creating a diagonal matrix, where the diagonals are each node’s degree, and it subtracts the adjacency matrix from the degree matrix. Then the method finds the sum of squared distances between these eigenvalues, for the eigenvalues containing 90% of the networks energy. The other single value metric of network comparison I look at follows Papadimitriou et al. (2009), and looks at the difference in node ranking between two networks. This method computes Spearman’s ρ for the ranking of each node. It does so by computing:

$$\rho = 1 - \frac{2 \sum_{i=1}^n d_i^2}{n(n^2)/3} \quad (2)$$

where d_i is the difference in rank between nodes and n is the number of nodes.

[Table 1 about here.]

The single value measures of network similarity show, that while the theoretical network captures some valuable facets of interstate cooperation, it still diverges in important ways from the empirical network. The networks are strikingly dissimilar according to the ‘eigenvalue method’ where small values indicate greater similarity. In particular, here, the greater centrality and degree of a few states leads to an extremely large dissimilarity in the eigenvalues. On a more positive note, the networks are markedly similar in their ranking of the nodes. By this measure, -1 would indicate reverse correlation with the rankings, 1 would indicate a perfect fit, and 0 would be no relation between the rankings, and so a *Spearman’s* ρ greater than 0.5 indicates serious commonality between the rankings.

Sensitivity to alternate parameter values. One important question that might be asked is whether the results shown in this paper are knife-edged, whether we are presenting and discussing a network equilibrium relying on very precise parameter values. To investigate this possibility, I reran our algorithm and found equilibrium networks holding all else equal and changing one of five main parameters ($c, g, \alpha, \beta, \gamma$) increasing or decreasing its value by a factor of 0.05 . We can then compare these new equilibrium networks to our previous theoretical equilibrium. These alternative networks range in size from 129 when (γ is decreased to $.665$) to 168 (when c is increased to $.07125$).⁶ The network statistics are compared in Figure 4. As you can see from the figure, the degree distribution is notably sensitive to the particular parameter values chosen, but the other measures seem comparatively stable. This provides some reassurance that the inferences drawn about the similarities between the triadic distribution and shared partner distribution in our empirical and theoretical networks are stable to small

⁶An example of the odd non-monotonicities, as pointed out by Siegel, in network equilibria that make computational methods desirable.

changes in parameter values, and that if I have made small misspecifications of the different parameters it will not lead to inference that is far off-base.

[Figure 4 about here.]

CONCLUSION

Even with models of strategic interactions as a prism, it is clear that countries cooperate with one another in ways that are not as well understood as we would like. Our work shows that while there is leverage in the use of games on networks, empirical patterns still have many secrets to be unwrapped. The empirical data look a lot like a scale-free network with the US as a hub, and do not look like the utility of interaction with the hub, either through first or higher order effects, is yet adequately captured. This particularly shines through in the theoretical network's serious dissimilarity to the empirical network in the eigenvalues. Clearly, there is some advantage to looking at the network dependencies, but there is also need to incorporate some inertial, initial conditions. Further, it may be the case that the theoretical network captures relationships that the empirical work does not. A valuable next step, would be to scan the parameter space and determine which sets of parameters generate an equilibrium network most similar to the observed empirical network. In particular, this would involve individually varying the parameters by small amounts while comparing the individual equilibria to the theoretical network using a loss function combining the different measures of network similarity (or using one of the two single number measures of similarity). This maximization process would also involve mutation and "hill-jumping" – randomly changing some parameters to try to avoid returning particularly unimpressive local minima. This would not solve all the problems with the model, as omitted variables and incorrect functional forms would not be corrected, but it would allow the model to move from simply providing a proof of concept, and yielding some insights about cooperation, to explicitly testing given theoretical models of network behavior and cooperation.

Data replication: The dataset, codebook, and code for the generation of theoretical networks and empirical analysis in this article can be found at <http://www.prio.org/jpr/datasets>.

REFERENCES

- Alter, Karen J & Sophie Meunier (2009) The politics of international regime complexity. *Perspectives on politics* 7(01): 13–24.
- Bearman, Peter S; James Moody & Katherine Stovel (2004) Chains of affection: The structure of adolescent romantic and sexual networks. *American Journal of Sociology* 110(1): 44–91.
- Bliss, Harry & Bruce Russett (1998) Democratic trading partners: The liberal connection, 1962-1989. *Journal of Politics* 60(4): 1126–1147.
- Blume, Andeas & Andreas Ortmann (2007) The effects of costless pre-play communication: Experimental evidence from games with pareto-ranked equilibria. *Journal of Economic Theory* 132: 274–290.
- Brambor, Thomas; William R Clark & Matt Golder (Winter 2006) Understanding interaction models: improving empirical analyses. *Political Analysis* 14(1): 63–82.
- Bramoullé, Yann & Rachel Kranton (2007) Public goods in networks. *Journal of Economic Theory* 135: 478–494.

- Bueno de Mesquita, Bruce & Alastair M Smith (2009) Political survival and endogenous institutional change. *Comparative Political Studies* 42(2): 167–197.
- Chwe, Michael S.-Y (1999) Structure and strategy in collective action. *American Journal of Sociology* 105(1): 128–156.
- Chwe, Michael S.-Y (2000) Communication and coordination in social networks. *Review of Economic Studies* 67: 1–16.
- Chyzh, Olga (2013) Tell me who your friends are: An endogenous model of international network formation and effect on domestic political outcomes. Dissertation, from personal correspondence.
- Chyzh, Olga (2016) Dangerous liaisons: An endogenous model of international trade and human rights. *Journal of Peace Research* 53: To Appear.
- Deborah Gerner, Omur Yilmaz, Philip S (2009) Conflict and mediation event observations (cameo): An event data framework for a post cold war world. In *International Conflict Mediation*, Jacob Bercovitch and Scott Sigmund Gartner (eds.).
- Deutsch, Karl W; Sidney A Burrell; Robert A Kann; Jr. M Lee; Martin Lichterman; Raymond E Lindgren; Francis L Loewenheim & Richard W. V Wagenen (1957) *Political Community and the North Atlantic Area*. Princeton, N.J.: Princeton University Press.
- Dorff, Cassy & Michael D Ward (2012) Networks, dyads, and the social relations model. Accessed at mdwardlab.com/sites/default/files/NetworksNotDyads.pdf.
- Erdős, Paul & Alfréd Rényi (1959) On random graphs. *I, Publicationes Mathematicae* 6: 290–297.
- Fearon, James D (1998) Bargaining, enforcement, and international cooperation. *International Organization* 52(02): 269–305.
- Galeotti, Andrea; Sanjeev Goyal; Matthew O Jackson; Fernando Vega-Redondo & Leeat Yariv (2010) Network games. *The Review of Economic Studies* 77(1): 218–244.
- Garrett, Geoffrey (1992) International cooperation and institutional choice: The European Community’s internal market. *International Organization* 46(02): 533–560.
- Gomez, Isabel; Maria T Fernandez & Jesus Sebastian (1999) Analysis of the structure of international scientific cooperation networks through bibliometric indicators. *Scientometrics* 44(3): 441–457.
- Goyal, Sanjeev & Fernando Vega-Redondo (2013) Network formation and social coordination. Available at SSRN: <http://ssrn.com/abstract=369460> or <http://dx.doi.org/10.2139/ssrn.369460>.
- Granovetter, Mark S (1973) The strength of weak ties. *American Journal of Sociology* 78(6): 1360–1380.
- Grieco, Joseph M (1988) Anarchy and the limits of cooperation: A realist critique of the newest liberal institutionalism. *International Organization* 42(3): 485–507.
- Haim, Dotan (2016) Alliance networks and trade: The effect of indirect political alliances on bilateral trade flows. *Journal of Peace Research* 53: To Appear.
- Holm, Desiree B; Kent Eriksson & Jan Johanson (1996) Business networks and cooperation in international business relationships. *Journal of International Business Studies* 27(3): 1033–1053.
- Jackson, Matthew O (2008) *Social and Economic Networks*. Princeton, N.J.: Princeton University Press.

- Jackson, Matthew O & Dunia Lopez-Pintado (2013) Diffusion and contagion in networks with heterogeneous agents and homophily. *Network Science* 1(1): 49–67.
- Jackson, Matthew O & Alison Watts (2002a) The evolution of social and economic networks. *Journal of Economic Theory* 106(2): 265–295.
- Jackson, Matthew O & Alison Watts (2002b) On the formation of interaction networks in social coordination games. *Games and Economic Behavior* 41(2): 265–291.
- Jackson, Matthew O & Asher Wolinsky (1996) A strategic model of social and economic networks. *Journal of Economic Theory* 71(1): 44–74.
- Jackson, Matthew O & Leeat Yariv (2007) Diffusion of behavior and equilibrium properties in network games. *American Economic Review* 97(2): 44–74.
- Jackson, Matthew O & Yves Zenou (2013) Games on networks. Forthcoming in Handbook of Game Theory, Vol 4.
- Jagers, Keith & Ted R Gurr (1995) Transitions to democracy: Tracking democracy’s ‘third wave’ with the polity III data. *Journal of Peace Research* 32(4): 469–482.
- Kaul, Inge; Isabelle Grunberg & Marc Stern (1999) *Global public goods: International cooperation in the 21st century*. Oxford University Press.
- Keohane, Robert O (1998) International institutions: Can interdependence work? *Foreign Policy* 110: 82–194.
- Koutra, Danai; Ankur Parikh; Aaditya Ramdas & Jing Xiang (2011) Algorithms for graph similarity and subgraph matching. Presented at the Ecological Inference Conference, June 17-18, 2002, Harvard University, Center for Basic Research in the Social Sciences.
- Larson, Jennifer (2011) A failure to communicate: The role of networks in inter- and intra-group cooperation.
URL: <https://people.cs.umass.edu/wallach/workshops/nips2011css/papers/Larson.pdf>
- Larson, Jennifer (2016) Interethnic conflict and the potential dangers of cross- group ties. *Journal of Peace Research* 53: To Appear.
- Leeds, Brett A (1999) Domestic political institutions, credible commitments, and international cooperation. *American Journal of Political Science* 43(4): pp. 979–1002.
- Lipson, Charles (1984) International cooperation in economic and security affairs. *World Politics* 37(01): 1–23.
- Litfin, Karen (1994) *Ozone Discourses: Science and Politics in Global Environmental Cooperation*. Columbia University Press.
- Maoz, Zeev; Ranan D Kuperman; Lesley Terris & Ilan Talmud (2006) Structural Equivalence and International Conflict: A Social Networks Analysis. *Journal of Conflict Resolution* 50(5): 664–689.
- Maoz, Zeev & Bruce M Russett (1993) Normative and structural causes of democratic peace, 1946-1986. *American Political Science Review* 87(3): 624–38.
- Maoz, Zeev & Kyle Zoyce (2016) The effects of shocks on international networks: Changes in the attributes of states and the structure of international alliance networks. *Journal of Peace Research* 53(3): To Appear.
- McPherson, Miller; Lynn Smith-Lovin & James M Cook (2001) Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27: 415–444.

- Metternich, Nils; Cassy Dorff; Max Gallop; Simon Weschle & Michael D Ward (2013) Antigovernment networks in civil conflicts: How network structures affect conflictual behavior. *American Journal of Political Science* 57(4): 892–911.
- Papadimitriou, Panagiotis; Ali Dasdan & Hector G Molina (2009) Web graph similarity for anomaly detection. Presented at the Ecological Inference Conference, June 17-18, 2002, Harvard University, Center for Basic Research in the Social Sciences.
URL: http://ilpubs.stanford.edu:8090/836/2/webgraph_similarity.pdf
- Patacchini, Eleonora & Yves Zenou (2008) The strength of weak ties in crime. *European Economic Review* 52(2): 209–236.
- Patty, John W & Elizabeth M Penn (2014) Sequential decision-making information aggregation in small networks. *Political Science Research and Methods* 2(2): 243–271.
- Siegel, David A (2009) Social networks and collective action. *American Journal of Political Science* 53(1): 122–138.
- Snidal, Duncan (1991) Relative gains and the pattern of international cooperation. *The American Political Science Review* 85(3): 701–726.
- Ward, Michael D (1982) A preliminary model of conflict and cooperation in foreign policy behavior. In: Gerhardt Bruckmann (ed.) *Environmental Aspects of Global Modeling*. Laxenburg, Austria: IIASA, 249–274.
- Ward, Michael D; John Ahlquist & Arturas Rozenas (2013) Gravity’s rainbow: A dynamic latent space model for the world trade network. *Network Sciences* 1(1).
- Ward, Michael D; Katherine Stovel & Audrey Sacks (2011) Network analysis and political science. *Annual Review of Political Science* 14: 245–264.
- Warren, Camber (2010) The geometry of security: Modeling interstate alliances as evolving networks. *Journal of Peace Research* 47(6).
- Warren, Camber (2016) Modeling the co-evolution of international and domestic institutions: Alliances, democracy, and the complex path to peace. *Journal of Peace Research* 53: To Appear.
- Watts, Duncan J (2003) *Six Degrees: The Science of a Connected Age*. New York, N.Y.: Norton.
- Weidmann, Nils B; Doreen Kuse & Kristian S Gleditsch (2010) The geography of the international system: The cshapes dataset. *International Interactions* 36(1): 86–106.
- White, Harrison (1963) *An Anatomy of Kinship: Mathematical Models for the Structure of Cumulated Roles*. Englewoods Cliffs, NJ: Prentice Hall.
- White, Harrison (2002) *Markets from Networks: Socioeconomic Models of Production*. Princeton, NJ: Princeton University Press.
- Young, Oran R (1989) *International cooperation: Building regimes for natural resources and the environment*. Cornell University Press.

LIST OF FIGURES

- 1 Equilibrium network for top 50 states by GDP. Nodes are colored to correspond to geographic region as detailed on the map. 19
- 2 Observed cooperation among top 50 states. Links are formed based on news stories of cooperative events as coded for the ICEWS project by JABARI. The threshold for cooperation was set at 118 acts to facilitate a similar density in theoretical and empirical networks. Nodes are colored to correspond to geographic region. 20
- 3 Comparison of theoretical and observed network on four measures of network similarity. The observed network's statistics are drawn in black, the theoretical ones in blue. 21
- 4 Comparison of theoretical network and ten alternatives on four measures of network similarity. The theoretical network is drawn in blue, the alternatives in gray. The alternatives are equilibrium with all parameters equal except one of $c, g, \alpha, \beta, \gamma$ which has been increased or decreased by 5% its original value. 22

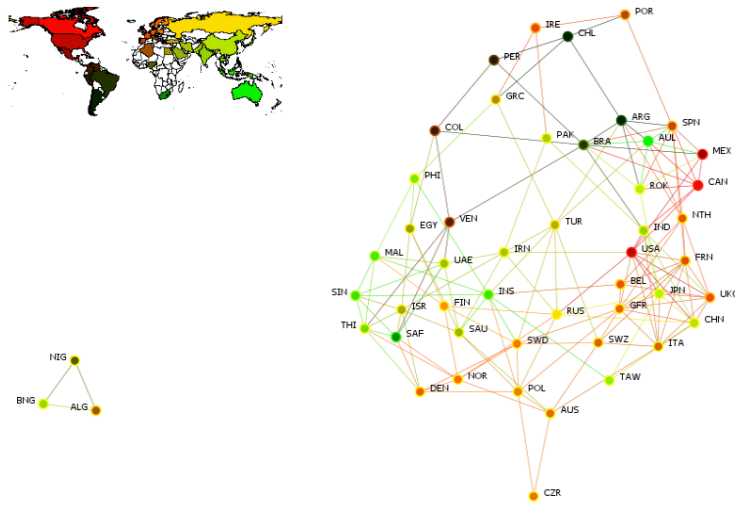


Figure 1. Equilibrium network for top 50 states by GDP. Nodes are colored to correspond to geographic region as detailed on the map.

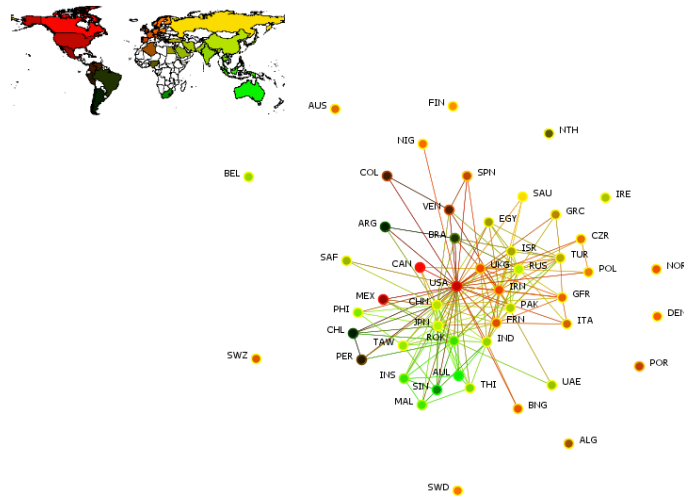


Figure 2. Observed cooperation among top 50 states. Links are formed based on news stories of cooperative events as coded for the ICEWS project by JABARI. The threshold for cooperation was set at 118 acts to facilitate a similar density in theoretical and empirical networks. Nodes are colored to correspond to geographic region.

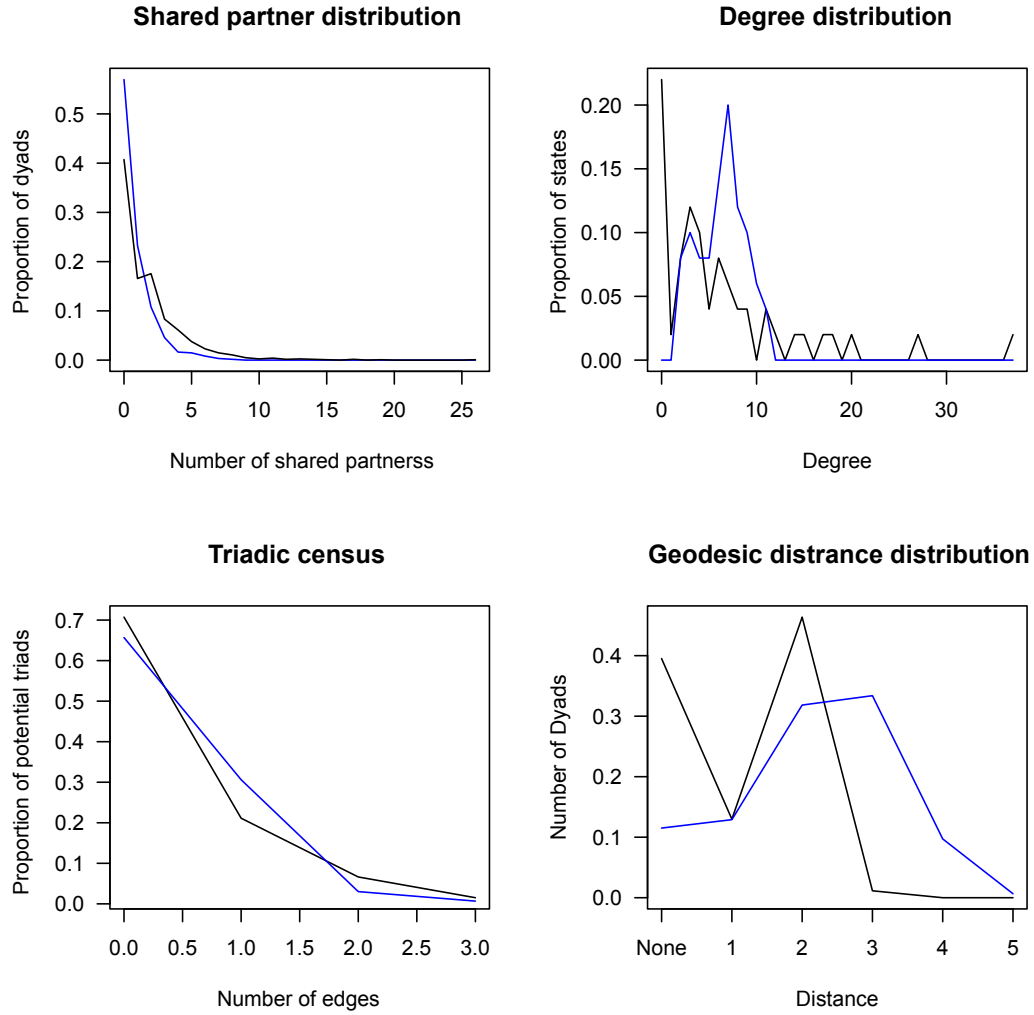


Figure 3. Comparison of theoretical and observed network on four measures of network similarity. The observed network's statistics are drawn in black, the theoretical ones in blue.

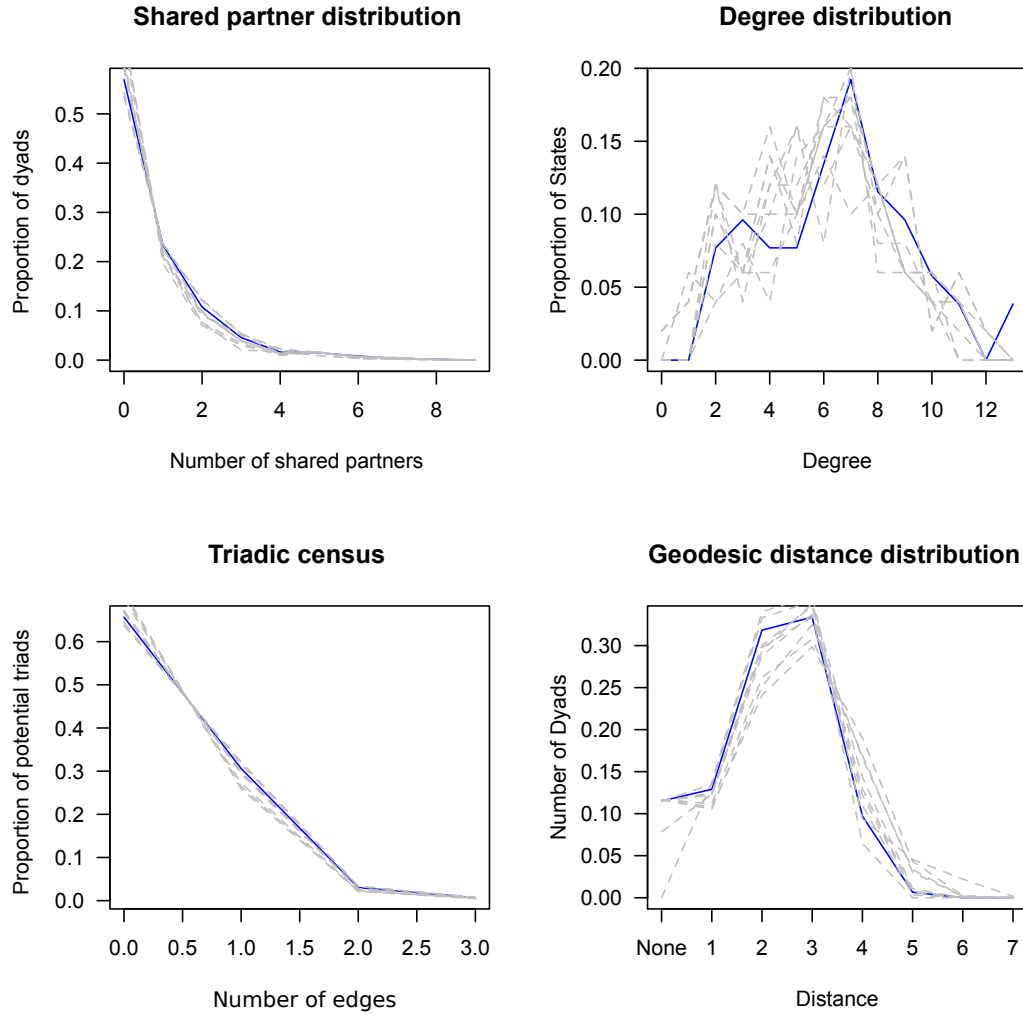


Figure 4. Comparison of theoretical network and ten alternatives on four measures of network similarity. The theoretical network is drawn in blue, the alternatives in gray. The alternatives are equilibrium with all parameters equal except one of $c, g, \alpha, \beta, \gamma$ which has been increased or decreased by 5% its original value.

LIST OF TABLES

- I Measures of Network Similarity. The Eigenvalue Method compares two networks on the squared distance between their Laplacian's Eigenvalues, where identical networks would score 0 and dissimilar networks would score highly. Spearman's ρ is used to compare the ranks of nodes in a network based on their out-degree. Identical networks would have a score of 1, perfectly reversed rankings would have -1 , and unrelated networks would in expectation score 0.

24

Method	Score
Eigenvalue Method	1104.87
Spearman's ρ	0.53

Table I. Measures of Network Similarity. The Eigenvalue Method compares two networks on the squared distance between their Laplacian's Eigenvalues, where identical networks would score 0 and dissimilar networks would score highly. Spearman's ρ is used to compare the ranks of nodes in a network based on their out-degree. Identical networks would have a score of 1, perfectly reversed rankings would have -1 , and unrelated networks would in expectation score 0.