International Arms Transfers and Conventional Trade: A Multilayer Network Approach

Daniel Seussler[*] Dennis Klein[†]

Submission Date: 23 May 2022

A Statistical Consulting Report.
Topic provided by Prof. Dr. Göran Kauermann (Department of Statistics, LMU),
in Collaboration with Prof. Dr. Paul W. Thurner (Institute for Political Science, GSI-LMU);
supervised by Cornelius Fritz and Dr. Marius Mehrl.

Replication files are available on Github: https://github.com/dennis-klein/arms-trade-networks.

[*]Corresponding Email Address: daniel.seussler@campus.lmu.de
[†]Corresponding Email Address: dennis.klein@campus.lmu.de

Abstract

Researchers repeatedly claim that alliances lead to an intensification of trade and service flows. However, formal alliances tend to be the exception in international relations. In contrast, arms trade networks can be conceptualized as relationships that are likewise based to a large extent on trust. The question then arises whether such networks induce trade and service flows to the same extent, which is the respective sequential dynamics, and to what extent such relationships are conflict mitigating and stabilizing. Based on data from the Stockholm International Peace Research Institute (SIPRI) and the Centre d'Études Prospectives et d'Informations Internationales (CEPII) we study the interlinkages of conventional trade and flows of Major Conventional Weapons (MCWs). The data shows both temporal-spatial dependencies as well as network structures, which necessitate the use of inferential network analysis. To model the co-evolution and cross-network interdependence of the two networks, we employ both cross-sectional and longitudinal models for multilayer networks. In summary, we find nuanced improvements of model fit with explicit modelling of inter-network interdependencies *vis-à-vis* a independent modelling approach.

Contents

1 Introduction

Contemporary literature on international weapons transfers after World War II is centered around the political considerations and security implications of such trade (Akerman and Seim, 2014; Comola, 2012; Thurner et al., 2019; Pamp et al., 2021; Fritz et al., 2021; Willardson and Johnson, 2022). Economic interests on the other hand, although ubiquitous in public perception of weapon trade,[1] have not been regarded as decisive. Yet, economic incentives are acknowledged in theoretical considerations, as costly Research & Development and production costs can be distributed over a higher number of units (Smith, 2009; Levine et al., 1994). Thus, disentangling economic incentives from political drivers of the arms trade is an ongoing line of research.

Besides, it has been hypothesised that political alliances lead to the intensification of conventional trade, yet, formal alliances are usually the exception (Haim, 2016). Arms trade relationships, on the other side, can be considered as relationships requiring a high amount of trust and cooperation, begging the question of whether conventional trade between states can be informative about arms transfers or vice versa. In this report, we apply methods from statistical network analysis to both the network of weapons and conventional trade relationships between countries in the international system. We use panel data on trades of Major Conventional Weapons (MCWs), such as aircraft, artillery, and tanks, and conventional bilateral trade flows in the period 1995–2018 between 114 countries, to study the interdependence of both domains of trade. Following Thurner et al. (2019), we conceptualise a network between a set of countries (called nodes), denominating a trade tie (edge) if the value traded exceeds a certain threshold.

From a statistical point of view, network data - also referred to as dyadic or relational data - pose a unique estimation problem. Generally, regression approaches require the assumption of conditional independence of observations, that, if wrongly assumed, risks faulty inference. We, therefore, follow the contemporary literature and employ methods for the statistical analysis of networks i.e., methods accounting for and measuring network dependencies. Although it would be possible to designate one network as an independent variable and use the second one as an exogenous predictor, both networks may be, and likely are, endogenous. Hence, the aim of this report is to review two approaches in statistical network analysis that incorporate multiple networks (i.e., layers) in regression models of network data, so-called multilayer network approaches and assess the application of this research design to active research topics in the international arms trade.

[1] See for example https://www.economist.com/asia/2022/02/12/south-korea-wants-to-become-one-of-the-worlds-biggest-arms-exporters

We employ two distinct statistical models, a cross-sectional Exponential Random Graph Model for the year 2003 and a dynamic network model, the Stochastic-actor Oriented Model with their respective multilayer extensions. Both approaches model the presence (or absence) of a tie, we therefore binarise the conventional trade network following a notion of import and export dependency. Both networks show a consistent presence of network inter-dependencies. In general, however, the inclusion of a multilayer component to the model describing both networks yield little predictive power *vis-à-vis* a model excluding the component. Therefore, we conclude, the inclusion of exogenous covariates for political and economic proxies explain the inter-linkages of weapon transfers and conventional trade similarly well.

The rest of the report is structured as follows. We first discuss the definition of networks in this report and in particular the novel application of networks encompassing several domains of interaction. In section 2, we present the available data sources for weapon transfers and conventional trade and produce descriptive statistics. Theory and application for the two multilevel approaches in the study of international arms trade are discussed in section 3 and section 4, section 5 concludes.

1.1 Network Analysis in Political Science and Economics

Network Analysis has been successfully applied to the study of a variety of political phenomena, amongst others, alliance ties (Cranmer et al., 2012; Camber Warren, 2010), conflict dynamics (Dorff et al., 2020; Olivella et al., 2022), and arms trade (Kinne, 2016; Thurner et al., 2019). Data in international relations studies is often available at the dyad-level, denominating the presence/absence, category, or continuous quantity of a relationship between two entities, such as states, conflict parties, or individuals. Network data is sometimes referred to as relational data and requires specific statistical approaches. This is because the observations are embedded in the network structure and dependencies arising from the relational nature of the social system at hand, such as popularity and clustering effects, have to be accounted for. Statistically speaking, the conditional independence assumption of observations fundamental for general regression approaches does not hold and may lead to a biased inference (see e.g. Hoff and Ward, 2004; Cranmer and Desmarais, 2016).

The line of research on statistical network analysis and its applications in the study of weapons trade, at the Geschwister-Scholl-Institute for Political Science and the Institute of Statistics, University of Munich, started with the article *Network Interdependencies and the Evolution of the International Arms Trade* (Thurner et al., 2019). Similar to the approach presented in this report, the authors estimate a Temporal Exponential Random Graph Model (TERGM) for international weapons transfers and analyse the associated political and economic drivers. Lebacher et al. (2021b) present an extension to the TERGM framework to account for time-varying network- and actor-specific effects. (Pamp et al., 2021) adapt a two-stage Heckman sample selection approach to the network framework. Before, arms trade was conceptualised as a binary trade relationship, i.e. *if* trade between two countries was realised (intensive margin). The extension, conditional on the trade relationship, allowed for an analysis of the intensity of trade (extensive margin). While data collection efforts of Major Conventional Weapons transfers are deemed to be fairly reliable with respect to their coverage of actual trade, in the case of Small Arms and Light Weapons, this does not hold. Lebacher et al. (2021a) apply a censored regression approach accounting for this reporting uncertainty. Such methods constitute extensions to spatial econometric modelling, Lebacher et al. (2020) is a further example of this empirical approach. Recently, in a relational event model framework, Fritz et al. (2021) estimate economic and political factors associated to the rate of international deliveries of combat aircraft trades.

In this case study, we observe yearly networks of the amount traded in weapons and conventional goods for a selection of countries. That the network of arms trade exhibits network interdependence is commonly acknowledged (see above). In particular, compared to a network drawn at random, we expect fewer out-ties, more in-ties, and lower reciprocity; as most countries import, not export weapons, and if they export, may not generally import (Thurner et al., 2019). The number of in- or out-ties can be interpreted as popularity effects. On a hyper-dyadic level, we expect countries that maintain a weapons trade relationship to share common partners, as this is concomitant to security ties and cliques (Thurner et al., 2019; Beardsley et al., 2020). The latter network effect is from the class of edge-wise shared partner (ESP) statistics, three of them of relevance to this study are illustrated in figure 1. Definition and interpretation of network structures vary across modelling frameworks, we discuss our selection along with each model below.

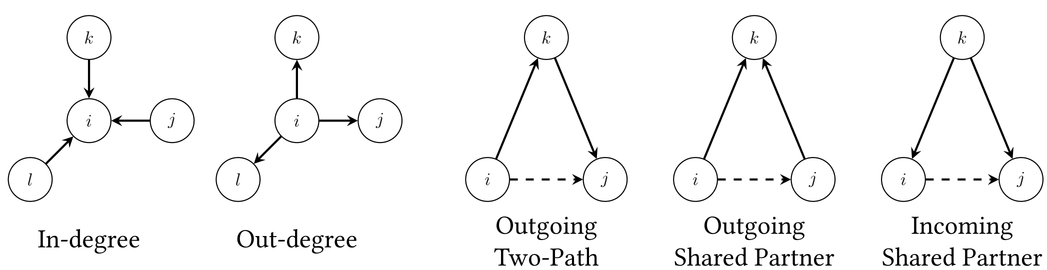


Figure 1: Exemplary endogenous network structures.

International (conventional) trade is primarily studied by means of gravity models, which do not account for network dependence. Ward et al. (2013) argue this to be empirically inadequate and present a latent space approach for bilateral trade data, showing out-performance in predictive metrics. On theoretical grounds, Chaney (2014) argues for a 'triadic search' present in the establishment of trade ties. If country i has established trade with k, it can use its presence to establish new commercial relationships with j, a partner of k. This would correspond to the edge-wise shared partner statistic Outgoing Two-Path, figure 1. More in line with our empirical approach, Herman (2022) estimates Temporal Exponential Graph Models (TERGMs) presenting evidence of the influence of network structures on trade relationships. From a purely descriptive perspective, Río-Chanona et al. (2017) analyse the dynamics of input-output networks based on the World Input-Output Database (Timmer et al., 2015) and conclude that correlational measures may uncover interesting network dynamics. More recently Smith and Sarabi (2022a,b) provide descriptive studies of international trade networks and intra firm trade. Read collectively this corroborates the importance of empirical approaches that account for the dependencies present in political and economic networks and underlines the use of statistical network analysis for inferential or descriptive studies of such.

As discussed above, several approaches from statistical network analysis have been developed for network data. Yet, these approaches are usually limited to a single type of relationship. This requirement may become too restrictive, in particular, if different networks could be influencing each other and may be relevant for the observed outcome. Next, we introduce the idea of multilayer networks.

1.2 Multilayer Networks: Actions across different Domains

Kivelä et al. (2014) state:

> In most natural and engineered systems, a set of entities interact with each other in complicated patterns that can encompass multiple types of relationships, change in time and include other types of complications. Such systems include multiple subsystems and layers of connectivity, and it is important to take such 'multilayer' features into account to try to improve our understanding of complex systems.

It is precisely the idea of multiple subsystems (henceforth domains), which demand the application of tailored empirical approaches. Actions (or behaviours) across different domains may explain (or mask) effects and incentives of interest to the researcher. Reduction to a single layer of observation may be a crude approximation and can lead to wrong conclusions, as Atkisson et al. (2020) and Górski et al. (2017) argue from an anthropological and sociological perspective, and De Domenico et al. (2015) generally.

Numerous concepts of 'multilayer' networks exist, see for example Kivelä et al. (2014) or Boccaletti et al. (2014) for a complete taxonomy. For political and economic networks two classes are of particular interest, multi-node and multiplex networks. In a multi-node network, different categories of actors (nodes) are associated with different layers and ties within or between layers are categorically different. For example, conflict networks with state and non-state actors may be defined with two layers, one for each actor type. Conflicts on the 'state' layer are therefore interstate disputes and non-state conflict are coded in the 'non-state' layer . Then, between-layer ties, i.e. conflicts between state and non-state actors, are of civil conflict nature.[2] On the contrary, multiplex networks presume the same set of actors across layers, each layer encompassing a different domain of interaction. In this formalisation, between-layer ties connect the same actor across layers and carry no substantive interpretation.

We argue that the composition of the above presented research question may benefit in similar fashion from the conception as multi-layered network. In particular, we are interested in the interdependence of two networks – MCWs trade and conventional trade – *without* fixing one layer as exogenous. That is, we consider the emergence of both networks simultaneously. Previously, studies centred around descriptive statistics, such as centrality measures, and exploratory correlational analyses (Alves et al., 2019; Río-Chanona et al., 2017; Bonaccorsi et al., 2019). We discuss popular approaches that provide statistical inference with respect to multilayer structures. Such statistical approaches for multilayer networks have received substantial research interest in the past years and new methods are added continuously (Kivelä et al., 2014). First, we present an extension to the exponential random graph model (ERGM, Chen, 2021b,a), and second, an approach capable of modeling time-dependency in repeated measurements of networks, the stochastic actor-oriented model (SAOM, Snijders et al., 2010, 2013). While both are popular approaches to conduct inference with network data, both require different sets of structural assumptions, rendering a direct comparison difficult. Most critical, by construction the SAOM adopts a strict actor-agency point of view, which is not available in the ERGM framework. We discuss the implications for inference and the results in each section.

[2]See Chen (2021b) for an application to the global conflict system in the post-Cold War period. The author refers to multi-node networks as 'node-coloured' networks.

2 Research Setup

2.1 Data Sources

2.1.1 Trade of Major Conventional Weapons

The Stockholm International Peace Research Institute (SIPRI) collects data on bilateral trade of major conventional weapons from publicly available information (Stockholm International Peace Research Institute, 2021a,b). The arms transfers database includes international exchanges of aircraft, air defence systems, artillery, missiles, and satellites, amongst others. Data on country-to-country transfers, available since 1950, are provided in a common unit, the trend-indicator value (TIV). It ensures comparability over time and encapsulates the real production value, specifically correcting for possible political-incentivised granted discounts. In contrast to the literature focused on security factors, we aggregate transfers based on the *order* date. The time between order and delivery has been increasing in recent decades, thereby introducing substantial lag in the decision to trade, which may or may not be economically motivated, and the decision to deliver, which should be predominantly regulated by security concerns.

Table 1: Major Conventional Weapons Transfers (1995-2018)

Armament Category	Count	TIV (Deal Unit)
Aircraft	4019	50408
Missiles	3573	1299
Armoured vehicles	2552	1838
Sensors	1971	10940
Engines	1881	2551
Ships	787	53513
Artillery	766	571
Air defence systems	409	7711
Naval weapons	335	2110
Other	264	524
Satellites	9	450

For the period 1995–2018, and the 114 countries included in the analyses, table 1 provides an overview aggregated by the armament category. Major conventional weapons, the class of weapons discussed in this report, consists of aircraft, missiles, ships, amongst others, and represent the primarily the heavy machinery of a military. In particular, Small Arms and Light Weapons (SALW) such as handguns, rifles and ammunition or anti-tank missiles, are not within the scope of this class of weapons. Generally, transfers of such military capability is accompanied by greater public scrutiny and political consideration, reducing possible biased reporting compared to SALW or dual-use weapon systems.

2.1.2 International Trade

Several databases for conventional bilateral trade flows exist, but complete panel data for a longer period of time is surprisingly hard to find. We settled for the *CEPII BACI: International Trade Database at the Product-Level* (Gaulier and Zignago, 2010) but review some other options below. *CEPII BACI* is, as some of the alternatives, based on *UN COMTRADE* and provides data for over 200 countries and 5000 products over the period 1994–2019. While *UN COMTRADE* provides detailed data, "accounting for more than 95% of the world trade" (Gaulier and Zignago, 2010, p. 9), it only provides the values as reported by the importer and exporter. Thus, for many bilateral combinations, two observations are provided. The primary benefit of those extensions is the reporting of a single harmonised value that provides a valid comparison and the imputation of missing values. Gaulier and Zignago state that this results in "more than 200 countries in BACI whereas in COMTRADE there are 130 on average in the period 1994-2007" (p.11). The respective methodology to obtain harmonised statistics is given in the accompanying publication.

Alternatives are readily available. First, the International Monetary Fund's *Direction of Trade Statistics*. It provides annual panel data starting in 1947 on Imports and Exports for "all IMF member states, some non-member countries, the world, and major areas".[3] Imports are provided as "cost, insurance, and freight" (CIF), exports as "free on board" (FOB). Where data is not available, it may have been imputed with time-series estimates. However, the coding of the zeros appears to be problematic. Gleditsch (2002) states "On closer inspection, many of the trade flows of exactly zero in the DOT [IMF's Direction of Trade] data seem problematic. To maintain a rectangular data structure, many missing observations appear to have been substituted with zeros. These structural zeros are probably better treated as missing observations rather than true zeros." Second, the *International Trade and Production Database for Estimation (ITPD-E)* (Borchert et al., 2021) from the U.S. International Trade Commission. It covers 243 countries and provides a disaggregation of trade flows on the industry level covering agriculture, mining, energy, manufacturing, and services. No data imputation is performed providing 'raw' estimates. The main limitation is the limited temporal coverage, providing only data for the years 2000–2017. The database *NBER - United Nations World Trade Flows: 1962-2000* (Feenstra et al., 2005) provides a third option. Unfortunately, it as well is limited in temporal scope and because we hypothesise changes in the relative importance of economic and political considerations in the early 2000s, it is not suitable for this project. The same applies to the Expanded GDP and Trade Dataset

[3] See https://data.imf.org/?sk=9D6028D4-F14A-464C-A2F2-59B2CD424B85.

Table 2: Summary statistics of exogenous covariates (1995–2017)

Name	N	Mean	SD	Min	Max
MCW Trade Tie (TIV, '000s)	391690	0.00	0.05	0.00	10.26
Conventional Trade Tie (2010 USD '000 000s)	391690	602.25	5353.08	0.00	406338.63
GDP per capita (log)	3013	24.69	1.98	20.38	30.49
Military Expenditure (log)	3013	20.55	2.26	14.41	27.37
Difference in Polity (abs)	391690	6.70	5.52	0.00	20.00
Alliance	391690	0.23	0.42	0.00	1.00
Distance in km (log)	17030	1.76	0.79	-4.88	2.99

(Gleditsch, 2002). If sectoral trade between countries is of main interest, the World Input-Output Database (Timmer et al., 2015) covers 43 countries for the period 2000–2014. For this case study the selection of countries is too limited.

Lastly, *UN COMTRADE* United Nations (2022) could be employed directly.[4] If this is a sensible strategy depends on the research question at hand and if inharmonious exporter/importer data can be integrated into the analysis.

2.1.3 Other Covariates

We include common exogenous covariates for gravity trade models. Gross Domestic Product (GDP) in aggregate and per capita terms is taken from The World Bank (2021a,b) and converted to constant 2010 USD. We impute values with linear imputation if at least half of consider period is observed. Yearly data on military expenditure is collected by Stockholm International Peace Research Institute (2019). We use data about alliances from Leeds et al. (2002). The Polity V Index (Marshall and Gurr, 2020) scales the political system of any country between -10 (Autocracy) to 10 (Democracy). We use the time-series corrected index and impute linearly if we have at least three values. To account for geographic distances of countries we include the capital-to-capital distance in km.[5] In table 2 we present summary statistics for the exogenous covariates for both approaches discussed below.

2.2 Setup & Computational Tools

Due to varying data availability across different data sets, a selection of countries is made. We only include countries that are present in the international system for the complete period of observation (1995-2018) and where data on conventional trade is available. This leaves 114 countries,

[4] UN COMTRADE can be freely accessed at https://comtrade.un.org/. However, to download data without restrictions on parameters and amount access to a Premium data API is necessary.

[5] See http://ksgleditsch.com/data-5.html.

a complete list can be found in the appendix A.1.1. As explained in the next section, current approaches in statistical network analysis are usually defined with binary outcomes as a dependent variable. Therefore, unfortunately, both the weapons transfer network and the trade network have to be binarised. We code an edge between country i and j in year t if SIPRI records an arms transfer relationship for these two countries in the respective year, i.e. we apply a threshold of 0. While higher thresholds may be used, as e.g. in Thurner et al. (2019), the network already is of rather low density. The density of a network is defined as the ratio of present ties to possible ties, which, if in the extremes, could introduce issues in the estimation process.

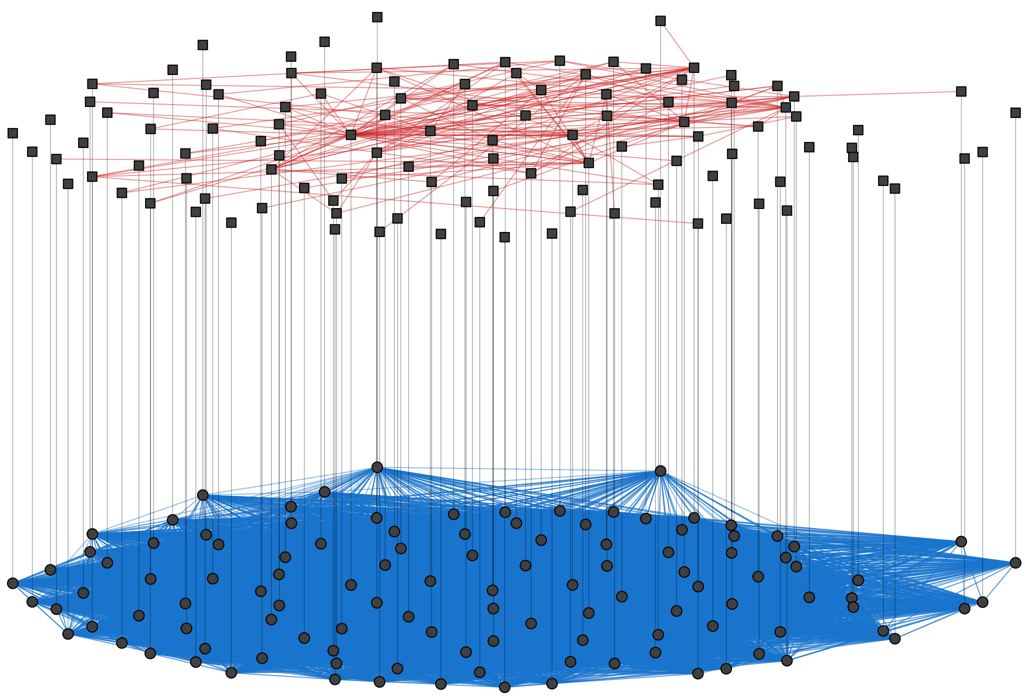


Figure 2: Multilevel network of arms and conventional trade in the year 2003.
Two-layer multiplex network of arms trade (red, upper layer) and conventional trade (blue, lower layer) in the year 2003. Vertical lines indicate the same actor across domains.

We encounter the opposite problem with the network of conventional trade. This network has a rather high density with little temporal variability. To thin out the conventional trade network but still maintain interpretable models, we keep edges that are at least 1 percent of trade (to GDP ratio) value for either the importer, exporter, at least one of both, or both, and discuss this selection in each model below. This is a major limitation of this approach. In particular, the coefficients might be sensitive to the chosen threshold. This is exacerbated by the observation, that the relationship of trade volume and trade partners follow a power law distribution. Usually few countries make up the greater part of a country's trade value, hence small changes in the threshold may have strong effects on network density. Ideally, one would conduct robustness checks concerning this threshold, but this is out of the scope of this report, and we discuss the implications in section 5.

Figure 2 illustrates the multiplex nature of the two networks for the year 2003, the year chosen for the static analysis below (cf. ERGM section 3). It already indicates the above-mentioned issue of high network density in the conventional trade network. Effectively every country trades with every other country, although of course to different extents. Figure 3 plots the network density over the considered time period. It corroborates the need for a threshold to make this data workable with the proposed methods of statistical inference.

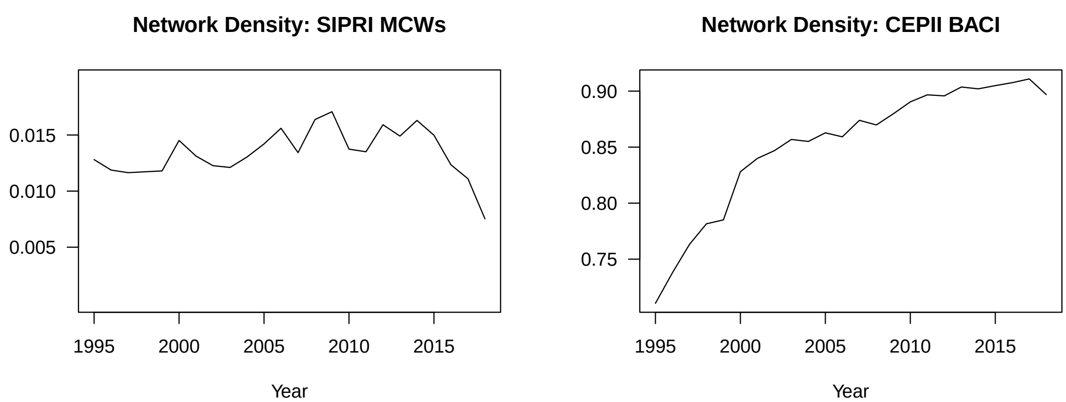


Figure 3: Network density of arms trade and conventional trade over the period 1995–2018.
Network density is defined as the ratio of all present ties in the network to all possible ties. We observe a balanced panel of 114 countries over the period 1995–2018. See appendix A.1.1 for a complete list of included countries.

The analyses are implemented in the R language (R Core Team, 2021). Network statistics are computed with the R-package `sna` (Handcock et al., 2019). For the empirical analyses below we use the `rsiena` package (Snijders et al., 2021) and the `ergm` package (Handcock et al., 2021) from the `statnet` suite of R packages for network analysis (Handcock et al., 2019). Additional multilayer effects are provided in the `multilayer.ergm` package (Chen, 2021a). Replication files for all analyses conducted in this report are available on in a Github repository.[6]

[6] See https://github.com/dennis-klein/arms-trade-networks.

3 Exponential Random Graph Models

The Exponential Random Graph Model (ERGM) is a popular approach in statistical network analysis and has been applied successfully in the study of different political studies (for a general introduction see for example Cranmer et al., 2021). Rather than choosing the dyadic value as an outcome, which may be endogenous to network dependencies, the model assumes the complete network as an observation from a data generating process. This specification circumvents any local dependence issues. Note that there is no straightforward actor-oriented interpretation as in the stochastic actor-oriented model, which we will discuss in section 4.

3.1 Theory

Let $y \in \mathbb{R}^{n \times n}$ be the matrix describing a network of n actors, where y_{ij} takes the value 0 or 1 depending on whether an (directed) tie between actors i and j exists. The set of all possible networks is denoted by $\mathcal{Y}$. Naturally, the matrix corresponding to an undirected network is symmetric. For any given observed network y we write the probability of the realised network as

$$P\left(Y = y \mid \theta\right) = \kappa(\theta)^{-1} \exp\{\theta^\top \mathbf{s}(y)\},\ y \in \mathcal{Y} \tag{1}$$

where $\kappa(\theta)$ is the normalisation constant which sums over $\mathcal{Y}$, i.e.

$$\kappa(\theta) := \sum_{\tilde{y} \in \mathcal{Y}} \exp\{\theta^\top \mathbf{s}(\tilde{y})\}. \tag{2}$$

Thus, the probability of the network y is modelled as a log-linear function of endogenous network statistics and exogenous covariates, which are captured in the vector $\mathbf{s}(y)$. Except for very small networks, this normalisation constant is usually intractable. Therefore, statistical inference generally relies on approximate Monte Carlo approaches for maximum likelihood or estimators based on a pseudo-likelihood. We briefly review both approaches and discuss our selection in section 3.1.2.

The endogenous network statistics can be selected by either postulating a specific network dependence one would like to account for or by substantive theory one would like to test in this framework. Yet, some structures are considered essential and their inclusion in network models is obvious. That is, we control for network density by the inclusion of a statistic capturing the number of ties in the network. Reciprocity, defined as the number of reciprocated ties in the network, accounts for the tendency to reciprocate ties in the network. The local propensity to send out or receive ties is captured in in- or out-degree distributions, which can be defined in different ways. We use geometrically weighted degree counts as presented in Snijders et al. (2006):

$$\mathbf{s}_{\text{in-degree}}(y \mid \alpha) := \sum_{k=0}^{n-1} e^{-\alpha k} \mathbf{ideg}_k(y) \tag{3}$$

with $\mathrm{ideg}_k(y)$ the number of nodes with in-degree k. This specification requires a decay parameter $\alpha > 0$ which controls the rate of decrease in weights. The out-degree statistic is defined analogously. Triadic network structures, constellations involving three actors, can also be defined differently, and are included to account for clustering effects. We follow the specification in Hunter (2007) and choose geometrically weighted counts of the Outgoing Two-Path statistic:

$$\mathbf{s}_{\text{GWESP OTP}}(y \mid \alpha) := e^{-\alpha} \sum_{i=1}^{n-2} \left(1 - (1 - e^{-\alpha})^i\right) p_i. \tag{4}$$

The decay-value α is defined as above and controls the weighting of the distribution of counts from Outgoing Two-path, the p_i, in the network. Instead of the Outgoing Two-path statistic, other constellations could be used, such as Outgoing Shared Partners. These statistics may be of interest in theoretical considerations. We restrict our introduction to these statistics of interest because we are restricted by the availability of terms in the `multilayer.ergm` package, which we discuss below. The extension for other (user-defined) network statistics of interest is possible using the `ergm.userterms` package (Handcock et al., 2019) but requires an in-depth understanding of the `ergm` package implementation and C. Coefficients of a fitted ERGM can be interpreted locally (on dyad-level) analogously to a logistic regression as the effect on the log-odds of a tie.

$$\begin{aligned} \mathrm{Odds}_\theta \; (y_{ij} = 1 \mid y_{ij}^c) &= \frac{\exp\{\theta^\top \mathbf{s}(y^+)\}/\kappa(\theta)}{\exp\{\theta^\top \mathbf{s}(y^-)\}/\kappa(\theta)} \qquad (5) \\ &= \exp\{\theta^\top \Delta_{ij} \, \mathbf{s}(y)\} \qquad (6) \end{aligned}$$

where

$$\Delta_{ij} \, \mathbf{s}(y) = \mathbf{s}(y^+) - \mathbf{s}(y^-). \tag{7}$$

In the case of the endogenous network statistics we need to additionally consider the value change of the statistic. To interpret the odds of tie $i \to j$ let y^+ be the network with $i \to j$ toggled *on* and y^- the network with $i \to j$ toggled *off* and all else equal. The difference in the statistics is the change statistic Δ_{ij}, which is the multiplier of the estimated effect on the log-odds. As most network statistics are essentially counts of occurrences of network structures, we can interpret a positive coefficient as the network having a propensity for the formation of the corresponding structure. This does also hold for geometrically weighted statistics. In effect, this weight downweighs higher numbers of for example edgewise shared partners in the distribution. We therefore interpret positive coefficients of the GWESP statistic as a tendency for triadic closure within the network (Cranmer et al., 2021).

While the ERGM framework provides the foundation to study a single network as the outcome, it cannot encode multiple types of actors[7] or relationships in the analysis. As discussed in section 1.2, if the probability of tie creation depends on the different types of actors, or the actions of actors in a different domain, it could bias our understanding of tie creation in both networks under study.

3.1.1 *Multilayer Network Extension*

The multilayer approach presented allows for a flexible implementation of multiplex and multi-relational networks in the ERG class of models. In its specification we follow Chen (2021b) where the corresponding ERGM statistics are implemented in the accompanying `multilayer.ergm` package (Chen, 2021a). To encode a multiplex network with different relations in the ERGM framework, we decompose the previous outcome network Y in a block matrix. This keeps compatibility for estimation within the `ergm` package and the included convenience functions. However, it should be noted that different approaches and packages for multilayer structures in ERGMs exist, e.g. the `mlergm` package (Stewart and Schweinberger, 2021) or the `ergm.multi` package (Krivitsky et al., 2020; Krivitsky, 2022). A comparison to those methods is beyond the scope of this report.

Now, the block matrices define inter- and intra-layer ties between actors. A network with k layers is encoded as

$$y := \begin{bmatrix} y^{1,1} & y^{1,2} & \cdots & y^{1,k} \\ y^{2,1} & y^{2,2} & \cdots & y^{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ y^{k,1} & y^{k,2} & \cdots & y^{k,k} \end{bmatrix}. \tag{8}$$

The matrix $y^{l,l}$ encodes the intra-layer ties in the l-layer; the off-diagonal matrices $y^{l,m}$ the ties between actors in the layer l and m. Note that it is not required that each layer consists of the same composition of actors. In particular, actor types, such as state and non-state actors in conflict studies, could be introduced by layer membership, lending inter- and intra-layer ties context-specific interpretations. Naturally, in the case of multiplexity, the off-diagonal matrices are the identity matrices of size n, i.e. the inter-layer only indicates which actors are equal across layers.

[7]Although it is possible to categorise types of actors as nodal covariate, it is not straightforward to restrict ties between types of actors in the classical ERG model.

Additionally, an actor-specific covariate *layer.mem* codes layer membership: $1, 2, \ldots, k$. This also preserves many network statistics within the `ergm` package, because statistics concerning degree distributions or reciprocity can be included conditional on layer membership. Unfortunately this does also restrict the decay parameter for any given statistic to be the same on different layers. The `multilayer.ergm` package adds additional statistics, in particular concerning the multilayer structure of the network and the construction of layer-specific covariates. With multiplex networks, the between-layer ties are fixed, as these define the same actor across layers and are not supposed to change by definition. To account for this, the sampling distribution, both in the estimation and goodness-of-fit procedure, has to be constrained. The `ergm` package provides an option for this.

In the case of directed multiplex networks, the cross-layer dependence terms are of particular interest. These are shown in figure 4 following the naming scheme of Chen (2021b). For each configuration, each row is a layer and between-layer ties are indicated by a grey line which connects the same actor across layers. The directed ties are indicated by an arrow. Taking into account the presence of directed ties on opposite sides of two layers, a total of 5 combinations exist. This statistic is always defined for two designated layers and counts for each combination the number of occurrences.

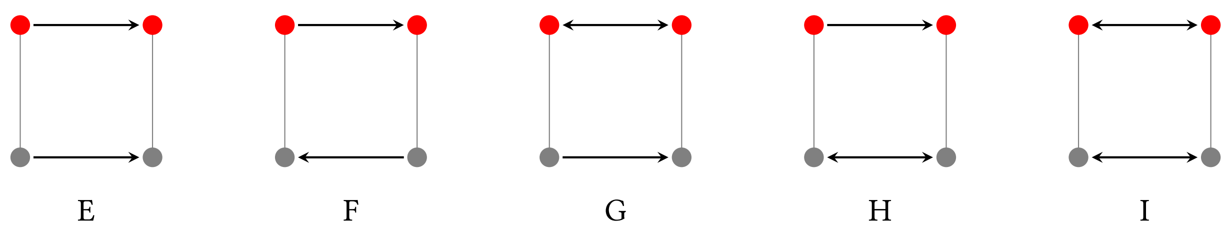


Figure 4: Cross-layer interdependence in a two-layer network by Chen (2021b).
Following the taxonomy of Chen (2021b): Red and gray indicate layer membership, the gray vertical line indicates the same actor across network layers.

A full account of the additional statistics for directed and undirected networks can be found in the corresponding vignette of the `multilayer.ergm` package. Besides a selection of the cross-layer network statistics, we include the geometric weighted edge-wise shared partner statistic (GWESP) – layer-specific – which is provided in the package. This statistic counts only Outgoing Two-Path partners (see figure 1). Unfortunately, this restricts the network features of interest that can be studied with this approach, but it is interpretable as a general account of the presence of clustering in each layer (Cranmer et al., 2021).

3.1.2 Estimation and Goodness of Fit

Inference on the parameters of an ERG model can be conducted in different ways and is not dependent on the multilayer extension discussed above. While it is theoretically possible to state the likelihood in equation 1, the normalising constant is intractable for all but the simplest examples. A comprehensive survey of inference methods for ERGMs can be found in Hunter et al. (2012). We use Stochastic Approximation (Snijders, 2002) to obtain first estimates and to tune the estimation procedure. In a second step we employ these estimates as initialization values for the Monte Carlo Maximum Likelihood algorithm (MCMLE, Geyer and Thompson, 1992; Hunter and Handcock, 2006). The default approach is informed by MPLE estimators as initial values, this variant, however, does not converge in timely fashion. Additionally, this leads to an overall improved result, with respect to criteria applied to the estimation procedure such as the quality of the sampling.

We did also explore different estimation approaches and selected our combined approach considering desirable properties of the resulting estimator and computational feasibility. First, the MPLE estimator does not come with the high computational cost of Monte Carlo based methods, in general, however, the quality of the estimator in the presence of strong network dependencies is not known. Contrastive Divergence – which currently is in an experimental implementation in the `ergm` package – failed to deliver estimates different to the initial MPLE estimates in this particular application. The MPLE estimates are often used as starting values for simulation based methods, this is the case for Stochastic Approximation too.

Ideally, to select appropriate nuisance parameters α for the geometrically weighted degree and ESP statistics one would estimate curved exponential random graph models, where α is estimated as a free parameter (Hunter, 2007; Hunter and Handcock, 2006). Unfortunately, the current version of the `multilayer.ergm` package does not support this approach for the layer-specific geometrically weighted statistics and we select fixed parameters as discussed below.

Because the stochastic approximation and MCMLE algorithm are both Monte Carlo based methods, visual diagnostics of trace plots are checked for stationarity and good mixing. Additionally, we conduct a model fit assessment. A flexible approach is to compare observed network statistics to the corresponding statistics of simulated networks provided the fitted model (Hunter et al., 2008). For example, the distribution of the in-degree statistics of simulated networks should contain the in-degree distribution of the observed network. If it does not or the sampled statistic deviate in evident form, we reject the model because of lack-of-fit. This ensures that the model captures the topology of the network under consideration, checked statistics are usually not included in the model, i.e. for example the complete degree distribution rather than the geometrically weighted one.

As an alternative, we can validate the estimated models in an out-of-sample comparison in terms of their predictive capacity. Given the fitted model at time point t and assuming stationarity of the underlying process, we predict the network at time point $t+1$, either by taking a single simulation or the average of many. Because we are predicting a binary outcome, we can use the Hamming distance in the former case and the Brier Score in the latter. Receiver Operator Characteristic (ROC) curves are also commonly used, but since the classes (1/0) are unbalanced Precision-Recall (PR) curves may be more informative. As it is not guaranteed that optimisation for the area under curve (AUC) of the ROC optimises the area under curve of the PR curve (Davis and Goadrich, 2006), we only show Precision-Recall results.

3.1.3 Dynamic Networks

In its initial form, the ERG model is applied to a static network, usually a snapshot of an underlying, possibly evolving, dynamic process. Applying this model to (pooled) panel data may mask temporal heterogeneity or dependency in the parameters under study. In particular, we expect changes in the relative significance of economic and security considerations in the late 90s and early 2000s (Thurner et al., 2019), which is not testable in this specification. We explored options to implement a temporal exponential random graph model (TERGM, Hanneke et al., 2010; Leifeld et al., 2018), along the lines of the analyses in the aforementioned Thurner et al. (2019) but extending the model with a multilayer extension. In a TERGM, time-dependency is introduced by time-dependent summary statistics and to achieve implementation, a Markov assumption of first-order is stated. Let $y^{(t)}$ be the network at time point $t = 1, \ldots, T$, then $y^{(t)}$ is independent of $y^{(1)}, \ldots, y^{(t-2)}$, conditional on $y^{(t-1)}$. This yields the following model for an observed network at a time t:

$$P\left(Y = y^{(t)} \mid \theta, y^{(t-1)}\right) = \kappa(\theta)^{-1} \exp\left\{\theta^\top \mathbf{s}\left(y^{(t)}, y^{(t-1)}\right)\right\}, \quad t = 2, \ldots, T. \tag{9}$$

The factor $\kappa(\theta)$ is the normalization constant as before, and the vector of summary statistics is extended to include statistics based on the previous network. One computational implementation of temporal ERG models is the `btergm` package (Leifeld et al., 2021). Due to the increased complexity statistical inference is usually achieved with maximum pseudo-likelihood estimators (MPLE) and a bootstrap-based correction of confidence intervals (Desmarais and Cranmer, 2012, 2010).

In this case study, the implementation of the `btergm` package proved to be prohibitive. The multilayer implementation as by Chen (2021b) requires specific constraints on between-layer ties, as these are merely indicative of the same actor across layers. I.e., if these are not fixed, these are included in the (pseudo-) likelihood and the estimation procedure is therefore poised to yield biased inference. Setting constraints as required was not possible and we did not proceed with this approach further. On a separate note, as the `btergm` package was not designed for such constraints, a post-hoc correction of the confidence intervals and goodness-of-fit assessments would have required considerable recoding of the included functions. Although a direct comparison to the stochastic actor-oriented model is tempting, benchmarking a simulation study is not straightforward and is subject to considerable debate (see for example the discussions in Lerner et al., 2013; Leifeld and Cranmer, 2019; Block et al., 2019, 2022)

3.2 Application

We study major conventional weapons transfers and conventional trade flow as a two-layer multiplex network with the ERG model. As a case study for this report we select the year 2003 and expect this year to be of interest because of the Iraq war.

3.2.1 Model Specification

Let $k := 2$, then the matrix $y^{1,1}_{2003}$ encodes transfers of MCWs and $y^{2,2}_{2003}$ conventional trade relations with 0 or 1 in the year $t = 2003$. Thresholds are selected as described in section 2. We compare two definitions for trade binarisation: import and export dependency. That is, we code a trade tie as present (1) if the trade value surpasses 1% of the total trade of the importer (or exporter), else as absent (0). By definition of a multiplex network, we set the off-diagonal matrices $y^{1,2}_t$ and $y^{2,1}_t$ as identity matrices, linking the same actor across networks. Following Thurner et al. (2019) we introduce all covariates with a temporal lag of $t = 1$.

Given the focus on the multilayer approach, the inclusion of cross-layer statistics are of particular interest, which are included to account for the occurrence of patterns spanning multiple layers (shown in figure 4). We select the statistics 'E', 'F' and 'H'. The excluded statistics have reciprocal arms trade relationships and as such are seldom, we do not expect much additional information of this structure. To appropriately capture additional network dependence present in the data, we include for each layer the above discussed network statistics, in particular, network density (edges), reciprocity, geometrically weighted in- and out-degree statistics with 'defaultish' decay parameter $\log(2) \approx 0.69$. To account for the tendency to form triadic relationships in the network, we include the Outgoing Two-path statistic (geometrically weighted) with the same decay. This decay value is the default value in the `rsiena` package and can be tuned for improved model fit: initial estimations during tuning with Stochastic Approximation show multi-modal distribu-

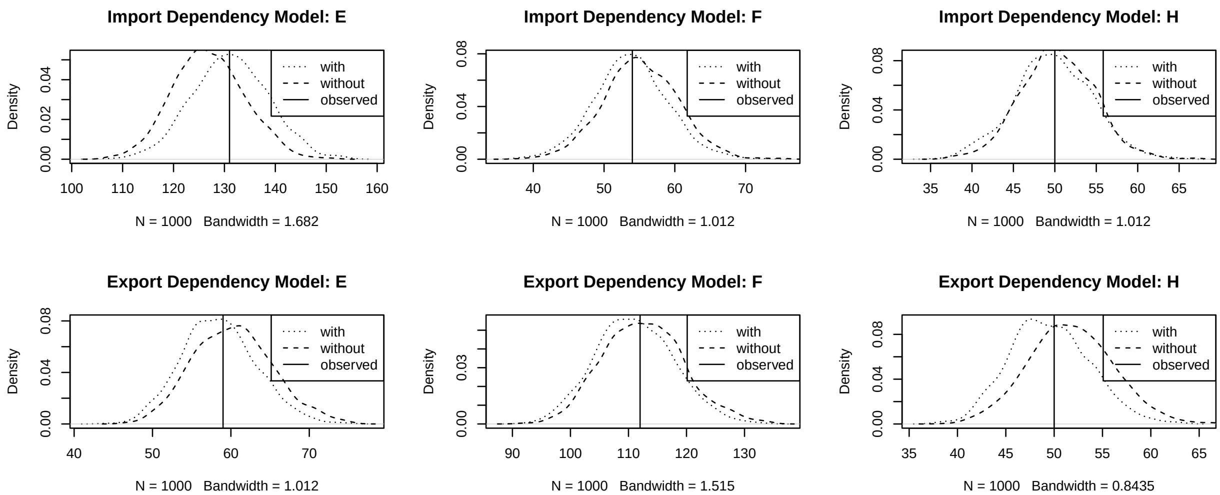


Figure 5: Goodness-of-fit assessment of the cross-layer network statistics.
Model specification with or without cross-layer network statistics. See figure 4 for the specification of the cross-layer terms. Results based on 10,000 simulations.

tions for some sampled statistics in different model fits. We therefore eased the down-weighting of the degree counts and set those decay parameters to 1.5, which resulted in satisfactory trace plots.

In addition, we include exogenous covariates of interest that previous literature showed to be associated to and account for political and economic drivers of arms transfers and conventional trade. That is, we include total GDP in 2010 US$ and Military Expenditure to account for market sizes of importer and exporter, as such it enters the model for each layer twice and log-transformed. Distance, constructed by the capital-to-capital distance in km is log-transformed and accounts for trade-gravity considerations. We further account for institutional differences with the absolute difference of polity values and an indicator if the two countries share an alliance in the given year. Finally we include a Path Dependency Indicator which assumes the value 1 if both countries traded in the three years preceding the year of the analysis.

3.2.2 Validation

Before we present the results we briefly discuss the validation of our model specification. In the first step, we estimate the presented models with the Stochastic Approximation algorithm, a Monte Carlo based method for approximate maximum likelihood inference. Generally, the trace plots seem to indicate a stationary distribution and sufficiently well-mixed sampled statistics. Initially, the trace plots showed a high amount of auto-correlation and occasionally non-centred statistics. This is likely because of the large network size, we, therefore, increased the thinning

96-fold and the sample size four-fold compared to the default values. This did improve the trace plots (not shown), the algorithm returns a successful estimation. In a second step, we used the estimated values as the initial values for the MCMLE algorithm. Again, the traceplots indicate a stationary distribution and sufficiently well-mixed sampled statistics. The trace plots for sample statistics for the import dependency model with cross-layer terms can be found in the appendix, section A.2.2. We do not include the trace plots for the other models. However, the trace plots are qualitatively the same and can be found in the accompanying code repository. The effective sample size, i.e. a statistic providing a heuristic for the sample size corrected for correlation of the drawn samples, is sufficiently high.

In total we consider four models, given by the choice of import or export dependency as threshold-approach in the conventional trade layer, and the inclusion or exclusion of cross-layer effect terms. As we impose constraints on the between layer ties, if we exclude cross-layer dependence terms we effectively estimate two ERGMs with each layer as a separate outcome network without consideration of possible endogeneity. We are therefore particularly interested if the inclusion of those cross-layer terms - and thus the multilayer approach - provide additional improvements to the model fit, its predictive performance or hypothesis testing of substantive theory. From here on out we refer to the model including cross-layer terms as "with" and else "without".

We first document the improvement in model fit for the cross-layer terms in figure 5. As described in section 3.1.2, we simulate 1000 networks from each fitted model and compare the distribution of those two network statistics from the simulated to the observed statistic (vertical line). The changes are nuanced across the different terms in the four models.

Next, we consider general statistics that characterise the topology of any given network. For this we restrict our attention to the import dependency model. The first line in 6 shows for each layer the number of edges in both the model with and without cross-layer effects. The boxplot characterise the distribution of the computed statistics from sampled networks and the red dots mark the observed statistic. Here, there is excellent agreement of both. The four rows characterise each the geodesic distances of one of the two layers for both import dependency models. The right column indicates the distribution for the value Inf which is a general indicator for the extend of unconnected components of the network. Again, we see the observed statistics are well within the distributions and as such indicate adequate goodness-of-fit.

Figure 7 repeats the same procedure for the indegree and outdegree statistics. The second row, the indegree statistics for the conventional trade layer, exhibit irregular behaviour. We explain this with the intricacies of the threshold approach employed, as due to the distribution of trade

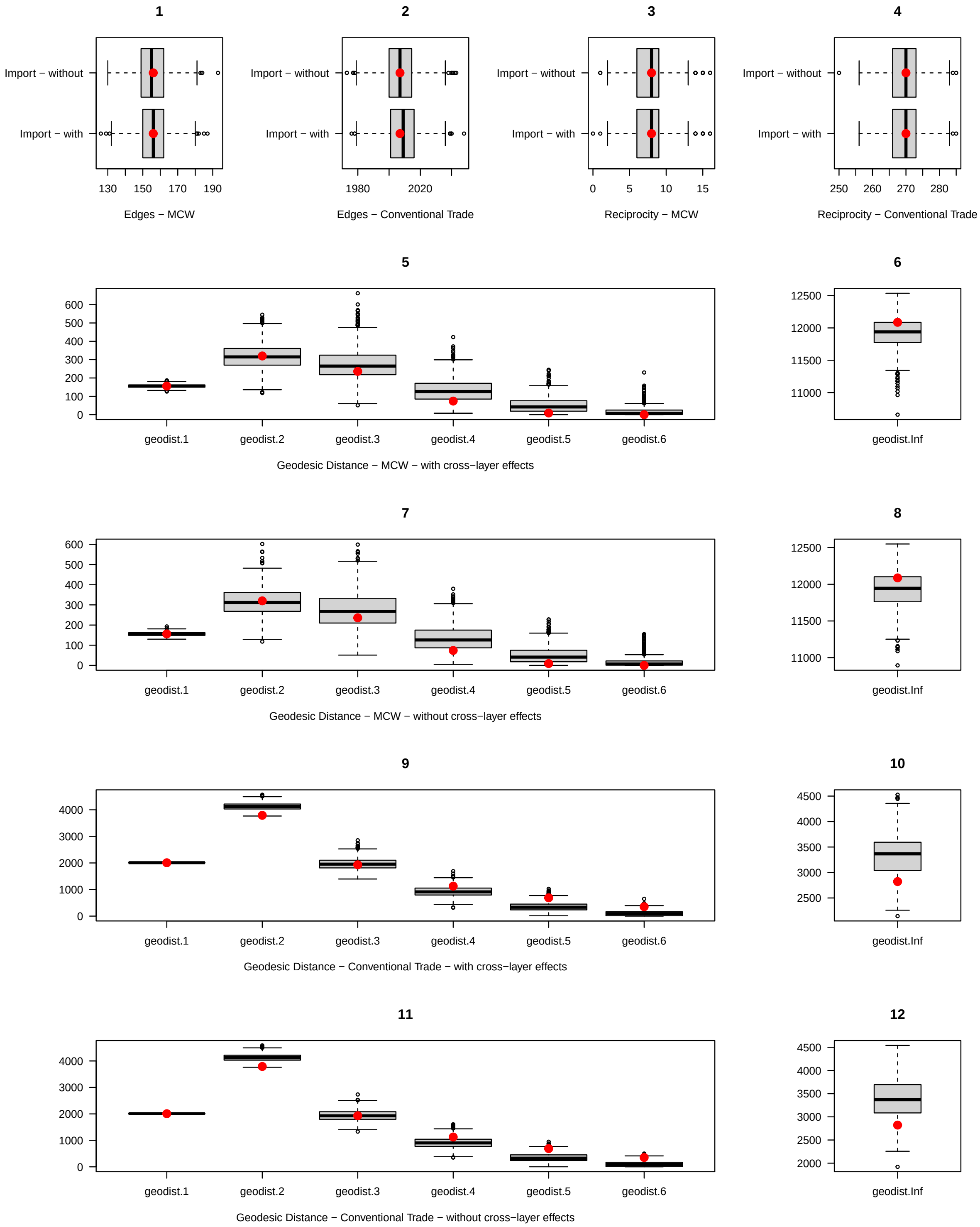


Figure 6: Goodness-of-fit assessment for the ERGM: density, reciprocity and geodesic distance.

Boxplots indicate sampled distributions, the red dot marks the observed statistic. Plots 1-4 show edges (network density) and reciprocity for each layer separately. Plots 5-12 show the geodesic distance in the network, with the the value Inf shown separately on the right.

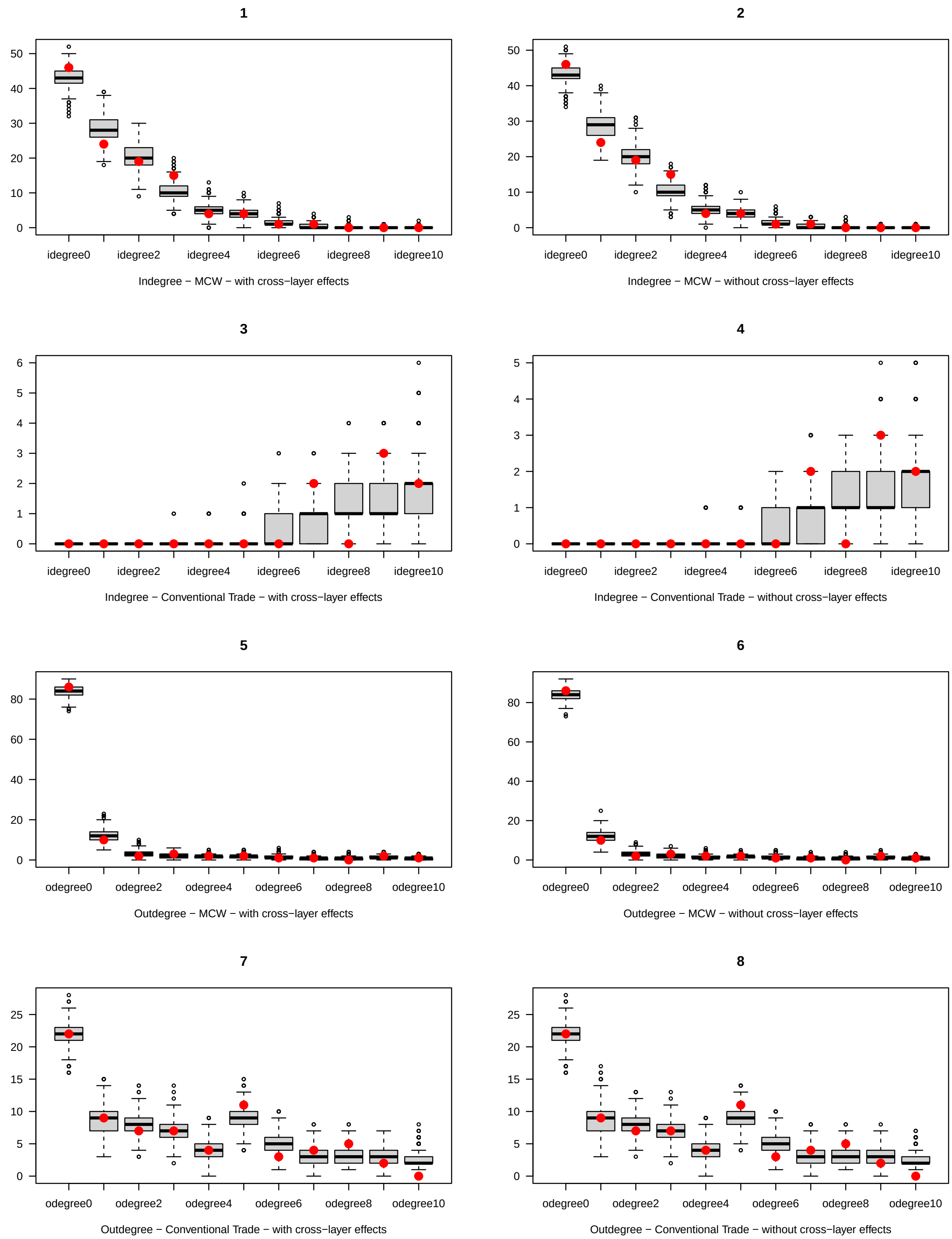


Figure 7: Goodness-of-fit assessment for the ERGM: indegree and outdegree. Boxplots indicate sampled distributions, the red dot marks the observed statistic.

value over trade partners the resulting indegree distribution is artificially 'capped'. Finally. figure 8 shows the same qualities for the edge wise shared partner and dyad wise shared partner counts. We conclude our model fits well. In general, the models with and without cross-layer statistics perform fairly similar, with no particular apparent changes in model fit visible. This is corroborates the discussion of figure 5.

Finally, we validate our model in out-of-sample fashion. For this, we use the fitted model for the year 2003 and predict the year 2004 using the exogenous covariates of this time step. Table 3 compares the out-of-sample performance of the model specifications excluding and including cross-layer effects, for both outcome models – import and export dependency – with respect to the Precision Recall Area under Curve (PR AUC) and the Brier Score. The performance of the model including cross-layer network effects is virtually identical to the model excluding it.

	Model	Metric	without	with
1	Import Dependency	PR AUC	0.8239	0.8242
2	Import Dependency	Brier Score	0.0195	0.0194
3	Export Dependency	PR AUC	0.8018	0.8012
4	Export Dependency	Brier Score	0.0192	0.0192

Table 3: Out-of-sample validation for the MERGM
Comparison of out-of-sample validation with and without cross-layer effects.

3.2.3 Results

The regression results are shown in table 4. The table is split in three parts: layer-specific covariates for arms trade and conventional trade as well as cross-layer network effects central to this case study. We present two models, one with import dependency as threshold approach in conventional trade (left) and one with export dependency (right). We control for network density and degree distribution. The first observation is a tendency for reciprocity in the arms trade, the same coefficient for conventional trade is closer to 0 and not statistically significant. The GWESP Outgoing Two Path statistic is small and not statistically significant for arms trade. For conventional trade the estimated effect is large, positive and statistically significant. Recall that we threshold the trade ties and as such only keep trade ties with highest dependency for the importer or exporter. The estimated effect indicates a strong tendency for clustering which is in line with a strong trade integration of common trade partners.

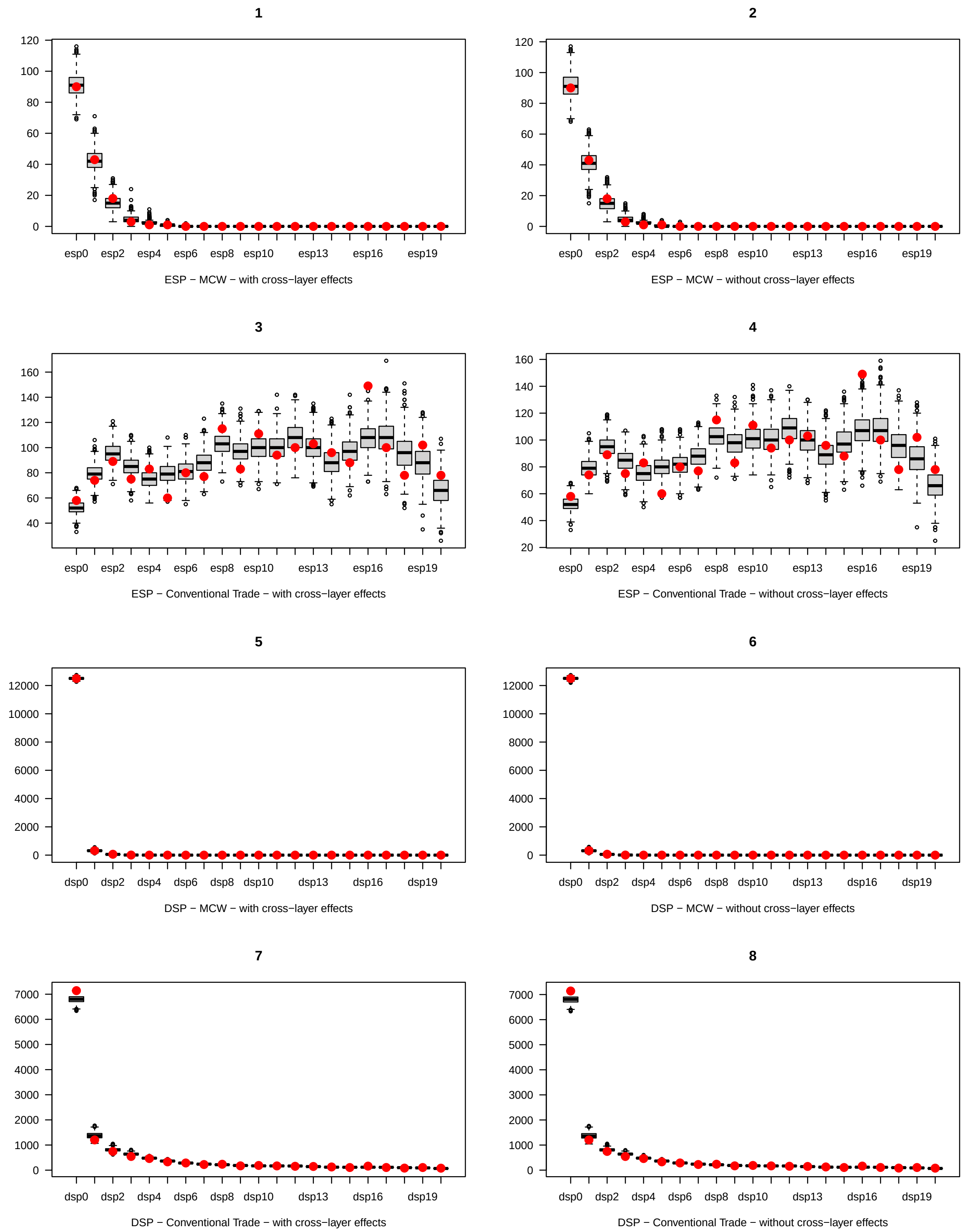


Figure 8: Goodness-of-fit assessment for the ERGM: ESP and DSP counts. Boxplots indicate sampled distributions, the red dot marks the observed statistic.

	Import Dep.	Export Dep.
Layer 1: Arms Trade		
Edges	$-6.40\ (5.53)$	$-8.63\ (5.34)$
Reciprocity	$1.32\ (0.64)^{*}$	$1.42\ (0.63)^{*}$
Gw Indegree (d = 1.5)	$-2.27\ (1.15)^{*}$	$-2.37\ (1.07)^{*}$
Gw Outdegree (d = 1.5)	$-3.17\ (0.89)^{***}$	$-3.36\ (0.88)^{***}$
GWESP Outgoing Two-path (d = 0.69)	$0.07\ (0.16)$	$0.06\ (0.15)$
Distance (log)	$-0.14\ (0.22)$	$-0.23\ (0.21)$
GDP in (log)	$0.00\ (0.31)$	$0.19\ (0.28)$
GDP out (log)	$-0.30\ (0.25)$	$-0.33\ (0.25)$
Alliance	$0.09\ (0.47)$	$0.09\ (0.46)$
Polity Diff. (abs)	$-0.03\ (0.03)$	$-0.02\ (0.03)$
Military Expenditure in (log)	$0.21\ (0.29)$	$0.14\ (0.29)$
Military Expenditure out (log)	$0.33\ (0.21)$	$0.34\ (0.22)$
Path Dependency	$1.67\ (0.36)^{***}$	$1.76\ (0.35)^{***}$
Layer 2: Conventional Trade		
Edges	$-16.03\ (3.26)^{***}$	$-13.92\ (2.44)^{***}$
Reciprocity	$0.14\ (0.34)$	$0.45\ (0.27)$
Gw Indegree (d = 1.5)	$-5.19\ (3.80)$	$0.43\ (0.83)$
Gw Outdegree (d = 1.5)	$0.20\ (0.82)$	$-1.71\ (1.65)$
GWESP Outgoing Two-path (d = 0.69)	$0.69\ (0.15)^{***}$	$0.58\ (0.13)^{***}$
Distance (log)	$-0.58\ (0.15)^{***}$	$-0.51\ (0.13)^{***}$
GDP in (log)	$0.56\ (0.18)^{**}$	$0.12\ (0.12)$
GDP out (log)	$0.06\ (0.16)$	$0.40\ (0.14)^{**}$
Alliance	$0.87\ (0.29)^{**}$	$0.09\ (0.26)$
Polity Diff. (abs)	$0.04\ (0.02)^{*}$	$-0.02\ (0.02)$
Military Expenditure in (log)	$-0.02\ (0.14)$	$-0.18\ (0.10)$
Military Expenditure out (log)	$-0.19\ (0.13)$	$0.02\ (0.11)$
Path Dependency	$2.13\ (0.08)^{***}$	$1.92\ (0.07)^{***}$
Cross Layer Network Effects		
E	$0.76\ (0.51)$	$-0.07\ (0.61)$
F	$-0.11\ (0.72)$	$-0.00\ (0.34)$
H	$-0.21\ (0.70)$	$-0.26\ (0.61)$
AIC	1458.68	1760.66
BIC	1675.12	1977.10
Log Likelihood	-700.34	-851.33

Estimates based on Monte Carlo MLE. Standard Errors in parenthesis.$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

Table 4: MERGM results for two-layer network of weapons and import (left) or export (right) trade dependency in the year 2003.

The value for the capital-to-capital distance (log) is in line with theoretical expectations derived from trade gravity models, it is negative, large in magnitude and statistically significant implying lower odds for trade (dependency) at increasing distance. For the arms trade layer the value is still negative – although indistinguishable from zero – suggesting less influence of economic transaction costs dependent from distance. In both layers the indicator covariate for previous trade relationships is positive and large in magnitude indicating high system inertia and strong predictive power of previous trade relationships on future trade.

Concerning arms trade, the other exogenous covariates are all statistically insignificant. The signs of the Alliance indicator, absolute difference in Polity value and military expenditure are directional as conjectured but estimated with a high uncertainty and as such interpretation is tentative. For example, the estimation for GDP out (sender-specific) is not compatible with substantive theory, as a higher GDP is considered to be associated with a higher industrial advancement and usually (arms) industries adept for export.

Regrettably, the estimates for the cross-layer network effects are small in magnitude and with wide confidence intervals, in line with the implications of the validation exercise before (i.e. figure 5). It suggests the addition of those multilayer effects add little to the understanding of the domain interdependence beyond the inclusion of the regular network structures and exogenous covariates. This observation is corroborated by the fact that the Bayesian Information Criterion is slightly lower for the models excluding cross-layer network effects, indicating that the increased model complexity of the multilayer structure cannot achieve a superior model fit.

We briefly review the results for the model without cross-layer network effects, appendix table 7. In general, the estimates are in very close agreement, indicating only nuanced information added by the inter-linkages of the two domains - or, put differently - no strong misfit if the domain endogeneity is not considered.

3.2.4 Discussion

In this section, we discussed the Exponential Random Graph Model and its extension for multilayered networks. We applied the model to a two-layered network encompassing ties constructed from transfers of Major Conventional Weapons on one layer and trade flows – characterised by their implied dependency – on the second layer. The multilayer approach did provide modest improvements compared to a modelling approach without regard for the inter dependency of domains. This may be due to several limitations of the presented approach, we will discuss those in further detail in section 5.

Finally, the networks under consideration constitute a cross-sectional analysis of a evolving process over time. This may prove too restrictive if the theory is precisely the change of influences of different factors over time. Consequently, we discuss a version of the stochastic actor-oriented model next, which provides a modelling approach for a network process observed at discrete time points.

4 Stochastic Actor-Oriented Model

4.1 Theory

4.1.1 Basic Description

The Stochastic Actor-Oriented Model (SAOM) for network change (Snijders, 1996) is a model for the dynamical change in networks over time. For this, we consider network data collected at multiple points in time, called waves. Based on Snijders et al. (2010, 2013), we provide a brief conceptual description of the model, before we discuss the formal framework. The model is implemented in the R package `RSiena` (Snijders et al., 2021). This framework enjoys large support in the sociological literature, but has also been frequently applied to research questions in political science, see for example Kinne (2016); Kinne and Bunte (2020).

In the SAOM framework, the position of a single actor in a network is assessed and the decision to create, maintain or end a relationship with other actors at a given point in time is modeled. The decision is based on the current state of the network. This state most importantly contains the current states of the relationships between the other actors and can also include additional exogenous relationships and characteristics of the involved actors. The second core assumption of the SAOM is the Markov property of the network evolution process. In this case, it can be interpreted as actors making their decision only based on the current situation they face in the network and not on information from earlier periods.

The central decision about which ties to build is controlled by the 'objective function'. For this function, higher values correspond to a preference for certain ties, low values result in the opposite. On the most basic level, the estimated model then assigns an effect of structures or covariates as a contribution to this function. From an intuitive perspective, this strongly resembles the idea of a utility function from economics. The actor chooses a new outgoing tie to reach a network state that has a higher objective function value than before. By doing this, the actor expresses a preference for structures or covariate effects that have a positive contribution to this function. For the sake of a complete introduction to the model, we note here that there exist three main types of the objective function in the RSiena framework that change the interpretation of the model significantly – evaluation, creation and endowment.

We restrict our approach to the 'evaluation' setting, both in the theory and in our empirical specification. We note, however, that there may be use cases for other objective functions in different research designs. An extension of the model allows us to jointly model the co-evolution of multiple networks as separate dependent variables. The main adaptions are that, first, an actor gets a separate evaluation function for each network, and second, the effects can now in theory depend on the state of both networks at the time of decision. In our two-network case, it leads to the following formulation of the evaluation function. This makes it clear that the actor decision is made separately for each network.

The objective function determines the decision of one single actor. However, the SAOM models the decision of all actors between two panel waves. If it would allow all actors to take the decision at once, the resulting network after an actor decision would potentially differ from the previous along many ties. Complex structural changes could arise between waves without any of the actors intended to build them. Therefore, the model assumes that decisions are made sequentially with only one potential tie change. The time at which a decision is taken and also the respective actor is chosen at random.[8]

In this report we pursue multilevel modelling approach, we, therefore, use an extension of the model that allows us to describe the evolution of the two networks simultaneously. That is, how they affect each other while also considering the effect of the covariates. In the SAOM framework this is often referred to as 'network co-evolution' and therefore we use the term below. The interpretation changes to actors that face more complicated decisions considering the effects of the decisions on both networks.

Having introduced the model at an intuitive level, the next section provides the mathematical setup. Moreover, it allows us to precisely understand how the effects work and how to interpret them in our quantitative results.

4.1.2 Mathematical Formulations

Like above, we follow closely the theoretical setup of Snijders et al. (2010) and the RSiena manual (Snijders et al., 2021, Chapter 5). Let $y_t \in \{0,1\}^{n\times n}$ be our network with n actors representing the existence of directed ties between them and y_t the realisations with matrix entries $y_{ij,t}$. In contrast to the ERGM model, however, the model includes multiple waves of network data which corresponds to the time index t. Note that therefore the notation differs slightly from the definitions in section 3. We refer to data on covariates in the networks as $x_{i,t}$ for actors and $x_{ij,t}$ for information on the edges between them. Time indices are dropped when they are clear from the context.

To model the actor decision, we take a given state of a network y_0 and consider all potential

[8] The time until the next decision between to observation point is exponentially distributed with a parameter estimated within the model. For more information on the concept of the related 'rate function' see (Snijders et al., 2021, section 5.1)

networks $\mathcal{Y}$ that the decision can induce – those networks that have at most one change in the outgoing ties of the actor or no tie change at all. As mentioned before, we define actor i's objective function $u_i(y_0, y)$ for any potential new state y in $\mathcal{Y}$ given the current state y_0. The probability that actor i chooses y as the new state is then given by

$$\frac{\exp\left(u_i\left(y_0, y\right)\right)}{\sum_{y' \in \mathcal{Y}} \exp\left(u_i\left(y_0, y'\right)\right)} \tag{10}$$

Within the SAOM framework, this objective function can potentially be a complicated expression depending on the types of network effects and on whether we use effects on evaluation, creation or endowment. For a network y, we call the evaluation function $f_i(y)$ and define the objective function as the change between the two sequential network evaluations:

$$u_i\left(y_0, y\right) = f_i\left(y\right) - f_i\left(y_0\right) \tag{11}$$

Because of the exponential term in the formula for the probability above, we can equivalently define the objective as only the evaluation of the new state to make things easier. We arrive at:

$$u_i\left(y_0, y\right) = f_i\left(y\right) \tag{12}$$

In our simplified version, we now define the evaluation function – the core part of the model – as a weighted sum of what we call 'effects':

$$f_i(y) = \sum_k \beta_k s_{ik}(y) \tag{13}$$

An effect $s_{ik}(y)$ could may depend on the entire network but often only reflects the local surroundings of the actor in the network. The direction and magnitude of the effect is then determined by the β_k coefficients which are to be estimated. However, the interpretation depends also on the functional form and scale of the effect (e.g., is it a binary indicator or an intensity).

As noted above, the definition of the objective function as the difference between the old and new evaluation is equivalent to our case. In combination with the linear form of the evaluation function, it shows that the *change statistic*

$$\Delta_{ki} s(y) = s_{ki}(y) - s_{ki}(y_0) \tag{14}$$

i.e., the change in the effect intensity from the current state to the next, is the primary determinant for the contribution to the objective function and, therefore, the probabilities for the new states.

An extension of the model allows us to jointly model the co-evolution of multiple networks $y^(1), y^(2)$ as separate dependent variables. The main adaptions are that, first, an actor gets a separate evaluation function for each network, and second, the effects can now in theory depend on the state in both networks at decision time. In our two-network case, it leads to the following formulation of the evaluation function. This makes it clear that the actor decision is made separately for each network, for example:

$$f_i^{y^{(1)}}\left(y^{(1)}, y^{(2)}\right) = \sum_k \beta_k^{y^{(1)}} s_{ik}^{y^{(1)}}\left(y^{(1)}, y^{(2)}\right) \tag{15}$$

4.1.3 Network Effects

This section we discuss the effect function, different sub-types and explain the effects we use in our model. From equation 13 above, we have the effect functions $s_i(y)$ on a given network. In our model, we distinguish between covariate and structural effects.[9] We further split up the latter in what we call within and between (network) effects which will be discussed in more detail below.

Covariate effects

Covariate effects reflect the influence of covariates of the focal actor which faces a decision (ego) or the potential partner (alter) on tie evaluation. Accordingly, they depend on node-level covariates that can be constant or varying over time. The corresponding time indices are dropped here because the covariates value always refers to the current network state in y.

$$\begin{aligned} \text{Ego:} \quad & s_{i,ego}(y) = \sum_{j \in N_i} x_i y_{ij} \\ \text{Alter:} \quad & s_{i,alt}(y) = \sum_{j \in N_i} x_j y_{ij} \end{aligned} \tag{16}$$

The ego effect can be interpreted as a tendency to build up ties if the own value of the covariate is high and the alter effect as a tendency to build up ties with partners that have high values. The depicted effects in equation 16 refer to node-level covariates. However, the idea naturally extends to dyad-level covariates by replacing x_i or x_j with x_{ij}.

[9] We restrict the analysis to the effects we use in our main model specifications.

Structural: within effects

Structural effects can capture dyadic, triadic, or even more complex network structures. We call the effects of structures in a network on the evaluation in the *same* network 'within effects'. In contrast, we call them 'between effect' on the evaluation in a network if they depend on structures in the *other* network. Since there is a large number of potential effects to be included, we follow recommendations in the RSiena manual (Snijders et al., 2021) and the general literature on these models.

Equation 17 depicts the structural between effects we use in our models. Similar to the covariate effects they can be interpreted as counting some structure in the local surrounding of the network and therefore as a tendency to build up this particular structure. We group them in three categories:

- **Basic effects**
  The Outdegree effect counts the number of ties and therefore expresses a tendency to build up ties in general. It can be viewed as an intercept (Snijders et al., 2021). Reciprocity counts the number ties that are reciprocal for the actor.

- **Local triadic structures**
  Figure 9 depicts the three effects of this category that we include in our model. Since there can be many directed triadic relations between three nodes and some of them might capture similar structures, it is hard to argue for an optimal set of included effects.

- **Dynamics in indegrees and outdegrees**
  The Indegree Popularity caputures a tendency to build up ties with nodes that have many indegrees. Indegree Activity describes a preference building up ties when the focal actor has many indegrees, and Outdegree Activity a preference for potential partners with many outdegrees. The square roots in the effect formulas are not necessary but tend to result in better fitting models (Snijders et al., 2021).

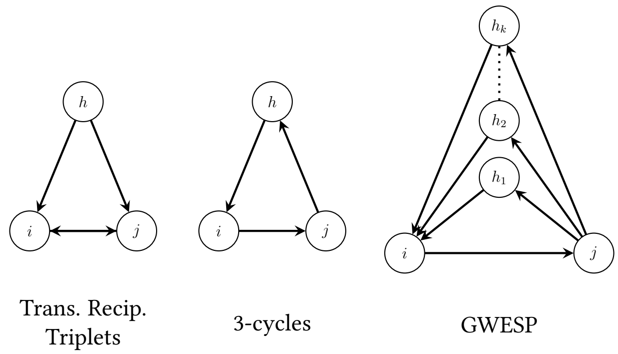


Figure 9: Illustrations for the three local triadic structural effects.
Taken from Snijders et al. (2021) with minor adaptations.

$$\left.\begin{array}{ll} \text{Outdegree:} & s_i(y) = \sum_j y_{ij} \\ \text{Reciprocity:} & s_i(y) = \sum_j y_{ij} y_{ji} \end{array}\right\} \text{basic effects}$$

$$\left.\begin{array}{ll} \text{Trans. Recip. Triplets:} & s_i(y) = \sum_{j,h} y_{ij} y_{ji} y_{ih} y_{hj} \\ \text{3-cycles:} & s_i(y) = \sum_{j,h} y_{ij} y_{jh} y_{hi} \\ \text{GWESP:}^{10} & s_{i,\alpha}(y) = \sum_{k=1}^{n-2} e^{\alpha} \left(1 - (1 - e^{-\alpha})^k\right) EP_{ik} \end{array}\right\} \text{local triadic structures} \tag{17}$$

$$\left.\begin{array}{ll} \text{Indegree Popularity:} & s_i(y) = \sum_j y_{ij} \sqrt{\sum_h y_{hj}} \\ \text{Indegree Activity:} & s_i(y) = \sum_j y_{ij} \sqrt{\sum_j y_{ji}} \\ \text{Outdegree Activity:} & s_i(y) = \sum_j y_{ij} \sqrt{\sum_h y_{jh}} \end{array}\right\} \text{dynamics in degrees}$$

Structural: between effects

The between effects we include and which are of particular interest for us are defined in equations 18. The Entrainment effect reflects a positive influence in the same direction of ties between networks, in equation 18 formulated as from network $y^{(}2)$ on network $y^{(}1)$. Reciprocity is the reciprocal equivalent e.g., ties on one network lead to reverse ties in the other. We use only very basic effects here to clearly focus on our main question: Is there an effect from arms transfer on

[10] The parameter α controls the weighting of the GWESP effect. EP_{ik} ("edgewise shared partners") counts the number of partners j of i such that there are exactly k different two-paths via other nodes from i to j. We leave the value at the default of $\log(2)$.

trade or vice versa?

$$\begin{aligned} \text{Entrainment } y^{(2)} \rightarrow y^{(1)}: \quad & s_i^{(1)}(y) = \sum_j y_{ij}^{(1)} y_{ij}^{(2)} \\ \text{Reciprocity } y^{(2)} \rightarrow y^{(1)}: \quad & s_i^{(1)}(y) = \sum_j y_{ij}^{(1)} y_{ji}^{(2)} \end{aligned} \tag{18}$$

4.1.4 Parameter Interpretation

The interpretation of the parameters for the effects of a fitted model relates to the above described 'decision process' of the focal actors. As mentioned, the actor can change at most one of his ties (build or drop) or do nothing. The evaluation function values of the resulting n networks determine a probability distribution on them. The higher the evaluation (or 'utility'), the higher the probability. Let $\hat{y}, \tilde{y}$ be two possible networks, then the log-odds can be expressed as

$$\text{log-odds}_i(\hat{y}, \tilde{y}) = \beta_1 \Delta s_{i,1}(\hat{y}, \tilde{y}) + \beta_2 \Delta s_{i,2}(\hat{y}, \tilde{y}) + \ldots + \beta_k \Delta s_{i,k}(\hat{y}, \tilde{y}) \tag{19}$$

where $\Delta s_{i,j}$ are the change statistics. In particular, the interpretation is network specific.

4.2 Application

4.2.1 Specification

After laying out the theoretical background for the Stochastic Actor-oriented Models, we fit several models to our data. To obtain binary data, we again apply a thresholds to thin out the dense conventional trade network. For our SAOM specification, we used the ratio of imports to GDP and a threshold of 0.4%.[11] The import per GDP ratio influences the interpretation of our models. Specifically, building up ties now corresponds to increasing imports from another country. Compared to other statistics that can be used for the threshold like for example total trade flow (import and exports), the trade to GDP ratio better fits the actor perspective of the SAOM. A country might have more autonomy in importing goods from a trade partner than exporting to it.

We use the R package `RSiena` to fit the following two specifications: In our main model we include all the covariate, within and between effects from section 4.1.3. Here, we still estimate both networks in one model. In addition, we fit a 'restricted' model which only differs from the main model in the absence of the between effects. However, since there are no interaction

[11]We choose this threshold according to a desired density in the trade network and the goodness of fit for the resulting SAOM model. The applied value of 0.4% results in a density of about 10% in the trade network.

between the models, this is equivalent to estimating two separate models (Snijders et al., 2021). Exogenous covariates are included as in the ERGM specifciation and we include the complete observation period.

We fit these two models to investigate two questions: First, do the coefficients of the SAOM change when properly modelling the potential endogeneity between imports and arms trade compared to two independent models? And second, what are the basic mechanisms between ties in the two networks (Entrainment and Reciprocity)?

4.2.2 Results

Table 6 shows our results for the restricted and the multilayer SAOM. We see little to no change in the coefficients for the covariate and within effects between the two model types. This suggest that the coefficients are not affected by large biases resulting from the inappropriate modelling of the two networks independently from each other.

In general, the direction of the coefficients is as expected from substantive theory. We see negative estimates for capital distance, larger in magnitude for import (dependency). Polity Difference in absolute terms is negative too for arms-trade, but statistically not significantly different from 0 for the trade network. In contrast to the ERGM model, the Alliance indicator is statistically significant at the .01 level, however – against expectation, negative for arms trade.

We see no evidence for reciprocity in the arms network, as expected, and evidence for reciprocity in the import dependence network. Jointly with the statistical significance of the GWESP Term, this indicates a strong trade integration of trade dependent nations as posited in trade theory.

Concerning the between effects, we find highly significant evidence for an reciprocal effect from arms trade to import dependency. In particular, for our data and modelling approach, arms trades from country A to B tend to be followed by increased import dependency of country B from country A. The other effects between the networks are not statistically significant different from no effect.

4.2.3 Goodness of Fit

We evaluate the fit of our model by comparing certain network statistics in our data versus their counterpart in simulated networks from our model, as done in section 3.2.2 for the ERGM. Specifically, we assess the outdegree and indegree distributions as well as the Davis and Leinhardt triad

Estimated Coeffcients for the Stochastic Actor-Oriented Models

Main model and restricted model without between networks effects

Effect	Restricted Model[1]		With Between Effects[2]	
	Arms	Trade	Arms	Trade
Covariates				
Capital Distance	-0.36 (0.02)***	-0.41 (0.02)***	-0.36 (0.02)***	-0.40 (0.02)***
Abs. Polity Difference	-0.53 (0.06)***	-0.08 (0.05).	-0.52 (0.06)***	-0.06 (0.05)
Alliance	-0.18 (0.04)***	0.55 (0.03)***	-0.18 (0.04)***	0.53 (0.03)***
GDP Alter (log)	-0.03 (0.02)	0.05 (0.03).	-0.01 (0.02)	0.05 (0.03).
GDP Ego (log)	0.28 (0.04)***	-0.31 (0.03)***	0.27 (0.04)***	-0.29 (0.04)***
Military Exp. Alter (log)	0.13 (0.02)***	0.08 (0.02)***	0.13 (0.02)***	0.07 (0.02)***
Military Exp. Ego (log)	0.47 (0.07)***	0.03 (0.02)*	0.46 (0.07)***	0.02 (0.02)
Within				
Outdegree	-4.84 (0.15)***	-3.43 (0.08)***	-4.86 (0.16)***	-3.41 (0.08)***
Reciprocity	0.26 (0.13).	1.01 (0.06)***	0.27 (0.12)*	1.04 (0.06)***
Transit. Recip. Triplets	-0.15 (0.22)	0.03 (0.01)*	-0.16 (0.21)	0.02 (0.01)*
3-cycles	-0.04 (0.09)	-0.13 (0.01)***	-0.03 (0.09)	-0.12 (0.01)***
GWESP	0.76 (0.16)***	1.25 (0.05)***	0.74 (0.16)***	1.29 (0.05)***
Indegree Popularity (sqrt)	0.51 (0.04)***	0.10 (0.01)***	0.51 (0.04)***	0.09 (0.01)***
Indegree Activity (sqrt)	-0.49 (0.18)**	0.15 (0.03)***	-0.45 (0.16)**	0.13 (0.03)***
Outdegree Activity (sqrt)	0.02 (0.03)	-0.06 (0.02)***	0.03 (0.03)	-0.06 (0.02)***
Between				
Conventional Trade			-0.08 (0.05)	
Recip. Conventional Trade			0.03 (0.05)	
Arms Trade				0.02 (0.19)
Recip. Arms Trade				1.08 (0.17)***

[1]Overall max. conv. ratio: 0.31; all t-conv. ratios below 0.05.

[2]Overall max. conv. ratio: 0.18; all t-conv. ratios below 0.04.

Table 6: Results of the multilayer SAOM 1995–2017.

The columns 'Arms' and 'Trade' refer to the network on that the effects is observed i.e., on the corresponding evaluation function of the network. The convergence ratios of the models indicate an appropriate convergence of the parameters. We briefly discuss them in section 4.2.3.

census classification.[12] Detailed plots can be found in the appendix A.3, figure 12. In general, our multilayer model seems to fit the data reasonably well, with the exception that the SAOM shows issues in capturing the low density (outdegrees) of the arms trade network.

From a technical perspective, we also consider the convergence of our coefficients. The so-called t-convergence ratio is calculated for each effect based on their empirical effect function values compared to values in simulated networks. We refer to the RSiena Manual (Snijders et al., 2021, Section 6.3.2) and note that the t-convergence and maximum convergence ratios indicate adequate convergence in the coefficients in our model.

Due to the long time span of our longitudinal network data, we also test for time heterogeneities of the coefficients with the `RSiena` package. For the detailed procedure see for example Lospinoso et al. (2011) and again the RSiena Manual (Snijders et al., 2021). We find strong evidence for time heterogeneity for most of the effects in our model (not shown). This does not invalidate our results over the whole period but motivates models that account for and identify possible time trends in the coefficients. In the next section, we provide a variation of our model extended to capture possible time-dependent dynamics.

4.2.4 Sliding Windows Model

To investigate time heterogeneities in the coefficients, we follow a similar approach to Kinne (2016); Thurner et al. (2019). We fit our main specification for a collection of sliding windows over a span of four years 'sliding' over the observed time period. This allows us to estimate different coefficient values over time. The results can be seen in figures 10 (covariate effects) and 11 (structural within and between effects). We plot the coefficient values against the starting period of their four-year windows. In general, we see that most effects are subject to time dynamics. Especially the structural effects show very different values across the time periods. The identified reciprocity between arms trade and imports seems to increase in recent years. Additionally, we might interpret the results for the Entrainment effect from imports to arms trade also slowly becoming stronger (negative) over the last years although due to large uncertainty of the estimates the results should be interpreted with caution.

4.2.5 Discussion

In this section, we discussed and presented an application for of the Stochastic Actor-oriented Model to model the interlinkages of the international arms trade and conventional trade flows.

[12] This concept classifies directed triads into 16 categories and therefore allows for a comparison between simulation and actual data by counting the occurrences.

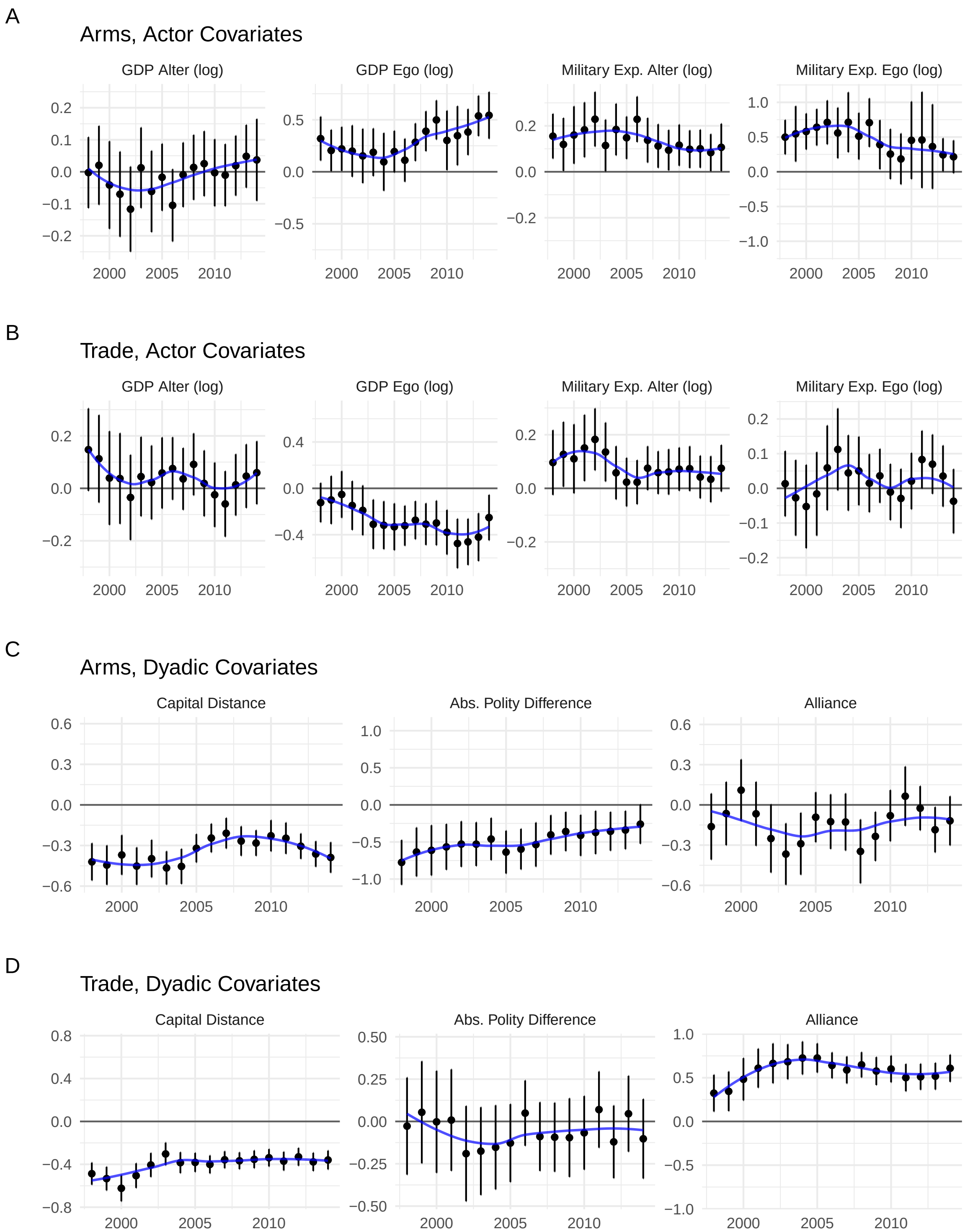


Figure 10: SAOM results sliding windows specification: covariates

Exogenous covariate effects for the SAOM sliding sindows specification. Estimated effect sizes are plotted against the starting year of the respective 4-year window together with 95% confidence intervals. Panel A, B: Estimated coefficients for actor characteristics on the tie evaluation on the arms trade and conventional trade network. “Ego” refers to the focal actor and “Alter” to the potential partner for tie evaluation; Panel C, D: Coefficients for the effect of exogenous characteristics of the dyadic relationship between two actors on one-sided tie evaluation.

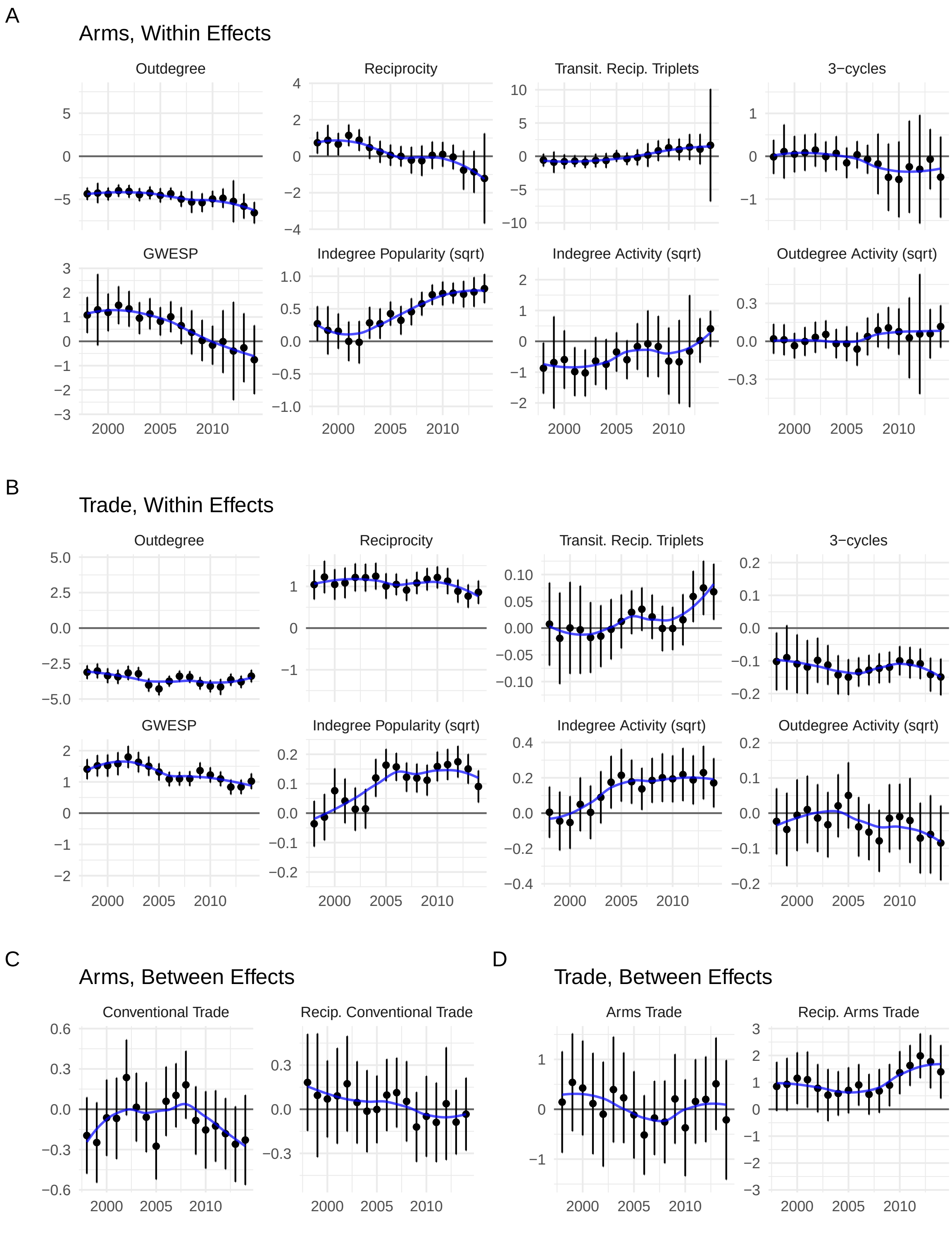


Figure 11: SAOM results sliding windows specification: structural effects
Structural network effects for the SAOM sliding windows specification. Panel A, B: Structural network effects on tie evaluation in the same network ("within"). Panel C, D: Structural effects from on network on the tie evaluation in the other ("between"). See also the description of figure 10.

This approach, in contrast to the cross-sectional ERGM approach presented in section 3, models the network (co-) evolution and allows for inference over a repeated series of network measurements.

In general, the results are in agreement with the results of the ERGM. Explicitly adding 'between' network effects, like the cross-network effects in the ERGM, does not yield a considerable improvement in model fit or uncovers biased estimates *vis-à-vis* a model without explicit inclusion. We discuss the implications for future research next.

5 Conclusion

In this report, we showcase two recent approaches of statistical network analysis for multilayer networks to study the international trade of arms and the inter-linkages with conventional trade. Using data collected from the Stockholm International Peace Research Institute and international trade data from CEPII, we construct two time series of networks, one for trade in Major Conventional Weapons and one for conventional trade flows.

In the first approach, we employ a recent addition to the Exponential Random Graph Model which adds a multilayer structure to the network as outcome. This approach is cross-sectional, we, therefore, select the year 2003 as our observation of analysis. Throughout the model validation both model specifications show comparable estimates and goodness-of-fit, indicating only nuanced improvements - if at all - throughout the statistics of interest. Put differently, the inter-linkage of of weapons transfers and trade dependency - import and export - is well captured in the model specifications without cross-layer effect terms.

The longitudinal SAOM approach corroborates this observation, and structural and covariate network effects do not change much when jointly modelling both domains. In contrast to the ERGM however, we find tentative evidence for a reciprocal effect of arms trade on import (dependency).

In summary, the multilayer structure added modest improvements to the model fit and the predictive capacity. This may be due to several limitations of the presented approach. First, binarisation, ultimately, provides a crude measure for dependency, may carry little information, and be too dependent on thresholds chosen in this report. Second, if the multilayer structure carries no information *beyond* the exogenous covariates included, the chosen approach may yield no further information for the validation or rejection of the substantive theory. It is difficult to determine the comparative weight of each of those limitations.

Still, the conceptual leap of multiple domains with inter-linkages may be proved to be useful in analyses of political and economic networks if the outcome network and layer provide a more informative measure of the present relationship. In the case presented here, one could think of substituting the conventional trade relationship with measures of bilateral lending. Re-framing the research question set up at the beginning of this report, specific sectoral trade may be more informative of the economic and political inter-linkages in the arms trade. In particular, in light of the Russian invasion of Ukraine, trade ties in the agricultural and energy sector and the implied dependency thereof would fit aptly in a multilayer framework.

Finally, while this report adopted an approach directed towards conducting inference, much can be learned from purely descriptive studies in the multilayer framework. Such studies are important, as *ex ante* not much is known about the inter-dependencies in the multilayer networks under consideration, and therefore, postulating any correlational structure may risk false specification. Descriptive approaches may uncover correlations of interest and provide a starting point for the development of substantive theory, upon which more capable statistical models may be motivated.

List of Figures

List of Tables

References

Akerman, A. and A. L. Seim (2014, August). The global arms trade network 1950–2007. *Journal of Comparative Economics 42*(3), 535–551.

Alves, L. G. A., G. Mangioni, I. Cingolani, F. A. Rodrigues, P. Panzarasa, and Y. Moreno (2019, February). The nested structural organization of the worldwide trade multi-layer network. *Scientific Reports 9*(1), 2866.

Atkisson, C., P. J. Górski, M. O. Jackson, J. A. Hołyst, and R. M. D'Souza (2020, May). Why understanding multiplex social network structuring processes will help us better understand the evolution of human behavior. *Evolutionary Anthropology: Issues, News, and Reviews 29*(3), 102–107.

Beardsley, K., H. Liu, P. J. Mucha, D. A. Siegel, and J. F. Tellez (2020, April). Hierarchy and the Provision of Order in International Politics. *The Journal of Politics 82*(2), 731–746.

Block, P., J. Hollway, C. Stadtfeld, J. Koskinen, and T. Snijders (2022, March). Circular specifications and "predicting" with information from the future: Errors in the empirical SAOM–TERGM comparison of Leifeld & Cranmer. *Network Science*, 1–12.

Block, P., C. Stadtfeld, and T. A. B. Snijders (2019, February). Forms of Dependence: Comparing SAOMs and ERGMs From Basic Principles. *Sociological Methods & Research 48*(1), 202–239.

Boccaletti, S., G. Bianconi, R. Criado, C. I. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin (2014, November). The structure and dynamics of multilayer networks. *Physics Reports 544*(1), 1–122.

Bonaccorsi, G., M. Riccaboni, G. Fagiolo, and G. Santoni (2019, December). Country centrality in the international multiplex network. *Applied Network Science 4*(1), 1–42.

Borchert, I., M. Larch, S. Shikher, and Y. V. Yotov (2021, August). The International Trade and Production Database for Estimation (ITPD-E). *International Economics 166*, 140–166.

Camber Warren, T. (2010, November). The geometry of security: Modeling interstate alliances as evolving networks. *Journal of Peace Research 47*(6), 697–709.

Chaney, T. (2014, November). The Network Structure of International Trade. *American Economic Review 104*(11), 3600–3634.

Chen, T. H. Y. (2021a). multilayer.ergm: Exponential Random Graph Models for Multilayer Networks.

Chen, T. H. Y. (2021b, April). Statistical inference for multilayer networks in political science. *Political Science Research and Methods 9*(2), 380–397.

Comola, M. (2012, February). Democracies, Politics, and Arms Supply. *Review of International Economics 20*(1), 150–163.

Cranmer, S. J. and B. A. Desmarais (2016, June). A Critique of Dyadic Design. *International Studies Quarterly 60*(2), 355–362.

Cranmer, S. J., B. A. Desmarais, and J. H. Kirkland (2012, July). Toward a Network Theory of Alliance Formation. *International Interactions 38*(3), 295–324.

Cranmer, S. J., B. A. Desmarais, and J. W. Morgan (2021). *Inferential network analysis.* Analytical methods for social research. Cambridge New York, NY Port Melbourne New Delhi Singapore: Cambridge University Press.

Davis, J. and M. Goadrich (2006, June). The relationship between Precision-Recall and ROC curves. In *Proceedings of the 23rd international conference on Machine learning*, ICML '06, New York, NY, USA, pp. 233–240. Association for Computing Machinery.

De Domenico, M., A. Solé-Ribalta, E. Omodei, S. Gómez, and A. Arenas (2015, April). Ranking in interconnected multilayer networks reveals versatile nodes. *Nature Communications 6*(1), 6868.

Desmarais, B. and S. Cranmer (2012, February). Statistical mechanics of networks: Estimation and uncertainty. *Physica A: Statistical Mechanics and its Applications 391*(4), 1865–1876.

Desmarais, B. A. and S. J. Cranmer (2010, December). Consistent confidence intervals for maximum pseudolikelihood estimators. In *Proceedings of the Neural Information Processing Systems 2010*, pp. 1865–76.

Dorff, C., M. Gallop, and S. Minhas (2020, April). Networks of Violence: Predicting Conflict in Nigeria. *The Journal of Politics 82*(2), 476–493.

Feenstra, R., R. Lipsey, H. Deng, A. Ma, and H. Mo (2005, January). World Trade Flows: 1962-2000. Technical Report w11040, National Bureau of Economic Research, Cambridge, MA.

Fritz, C., P. W. Thurner, and G. Kauermann (2021, September). Separable and semiparametric network-based counting processes applied to the international combat aircraft trades. *Network Science 9*(3), 291–311.

Gaulier, G. and S. Zignago (2010, October). BACI: International Trade Database at the Product-Level. The 1994-2007 Version. Working Papers 2010-23, CEPII.

Geyer, C. J. and E. A. Thompson (1992). Constrained Monte Carlo Maximum Likelihood for Dependent Data. *Journal of the Royal Statistical Society. Series B (Methodological) 54*(3), 657–699.

Gleditsch, K. S. (2002). Expanded Trade and GDP Data. *The Journal of Conflict Resolution 46*(5), 712–724.

Górski, P. J., K. Kułakowski, P. Gawroński, and J. A. Hołyst (2017, November). Destructive influence of interlayer coupling on Heider balance in bilayer networks. *Scientific Reports 7*(1), 16047.

Haim, D. A. (2016, May). Alliance networks and trade: The effect of indirect political alliances on bilateral trade flows. *Journal of Peace Research 53*(3), 472–490.

Handcock, M. S., D. R. Hunter, C. T. Butts, S. M. Goodreau, P. N. Krivitsky, S. Bender-deMoll, and M. Morris (2019, June). statnet: Software Tools for the Statistical Analysis of Network Data.

Handcock, M. S., D. R. Hunter, C. T. Butts, S. M. Goodreau, P. N. Krivitsky, and M. Morris (2019, May). ergm.userterms: User-specified Terms for the statnet Suite of Packages.

Handcock, M. S., D. R. Hunter, C. T. Butts, S. M. Goodreau, P. N. Krivitsky, M. Morris, L. Wang, K. Li, S. Bender-deMoll, C. Klumb, M. Bojanowski, B. Bolker, and C. Schmid (2021, July). ergm: Fit, Simulate and Diagnose Exponential-Family Models for networks.

Hanneke, S., W. Fu, and E. P. Xing (2010, January). Discrete temporal models of social networks. *Electronic Journal of Statistics 4*(none), 585–605.

Herman, P. R. (2022, February). Modeling complex network patterns in international trade. *Review of World Economics 158*(1), 127–179.

Hoff, P. D. and M. D. Ward (2004). Modeling Dependencies in International Relations Networks. *Political Analysis 12*(2), 160–175.

Hunter, D. R. (2007, March). Curved Exponential Family Models for Social Networks. *Social Networks 29*(2), 216–230.

Hunter, D. R., S. M. Goodreau, and M. S. Handcock (2008, March). Goodness of Fit of Social Network Models. *Journal of the American Statistical Association 103*(481), 248–258.

Hunter, D. R. and M. S. Handcock (2006, September). Inference in Curved Exponential Family Models for Networks. *Journal of Computational and Graphical Statistics 15*(3), 565–583.

Hunter, D. R., P. N. Krivitsky, and M. Schweinberger (2012, October). Computational Statistical Methods for Social Network Models. *Journal of Computational and Graphical Statistics 21*(4), 856–882.

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

Kinne, B. J. and J. B. Bunte (2020, July). Guns or Money? Defense Co-operation and Bilateral Lending as Coevolving Networks. *British Journal of Political Science 50*(3), 1067–1088.

Kivelä, M., A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter (2014, September). Multilayer networks. *Journal of Complex Networks 2*(3), 203–271.

Krivitsky, P. N. (2022, February). ergm.multi: Fit, Simulate and Diagnose Exponential-Family Models for Multiple or Multilayer Networks.

Krivitsky, P. N., L. M. Koehly, and C. S. Marcum (2020, September). Exponential-Family Random Graph Models for Multi-Layer Networks. *Psychometrika 85*(3), 630–659.

Lebacher, M., P. W. Thurner, and G. Kauermann (2020, April). Exploring dependence structures in the international arms trade network: A network autocorrelation approach. *Statistical Modelling 20*(2), 195–218.

Lebacher, M., P. W. Thurner, and G. Kauermann (2021a, August). Censored regression for modelling small arms trade volumes and its 'Forensic' use for exploring unreported trades. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 70*(4), 909–933.

Lebacher, M., P. W. Thurner, and G. Kauermann (2021b, January). A dynamic separable network model with actor heterogeneity: An application to global weapons transfers. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 184*(1), 201–226.

Leeds, B., J. Ritter, S. Mitchell, and A. Long (2002, July). Alliance Treaty Obligations and Provisions, 1815-1944. *International Interactions 28*(3), 237–260.

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

Leifeld, P., S. J. Cranmer, and B. A. Desmarais (2018, February). Temporal Exponential Random Graph Models with btergm: Estimation and Bootstrap Confidence Intervals. *Journal of Statistical Software 83*, 1–36.

Leifeld, P., S. J. Cranmer, and B. A. Desmarais (2021, June). btergm: Temporal Exponential Random Graph Models by Bootstrapped pseudolikelihood.

Lerner, J., N. Indlekofer, B. Nick, and U. Brandes (2013, December). Conditional independence in dynamic networks. *Journal of Mathematical Psychology 57*(6), 275–283.

Levine, P., S. Sen, and R. Smith (1994, January). A model of the international arms market. *Defence and Peace Economics 5*(1), 1–18.

Lospinoso, J. A., M. Schweinberger, T. A. B. Snijders, and R. M. Ripley (2011, July). Assessing and accounting for time heterogeneity in stochastic actor oriented models. *Advances in Data Analysis and Classification 5*(2), 147–176.

Marshall, M. G. and T. R. Gurr (2020). POLITY5: Political Regime Characteristics and Transitions, 1800-2018 Dataset Users' Manual.

Olivella, S., T. Pratt, and K. Imai (2022, February). Dynamic Stochastic Blockmodel Regression for Network Data: Application to International Militarized Conflicts. *Journal of the American Statistical Association*, 1–14.

Pamp, O., M. Lebacher, P. W. Thurner, and E. Ziegler (2021, September). Explaining destinations and volumes of international arms transfers: A novel network Heckman selection model. *European Journal of Political Economy 69*, 102033.

R Core Team (2021). R: A Language and Environment for Statistical Computing.

Río-Chanona, R. M. d., J. Grujić, and H. J. Jensen (2017, January). Trends of the World Input and Output Network of Global Trade. *PLOS ONE 12*(1), e0170817.

Smith, M. and Y. Sarabi (2022a, July). How does the behaviour of the core differ from the periphery? – An international trade network analysis. *Social Networks 70*, 1–15.

Smith, M. and Y. Sarabi (2022b, February). Mapping intra firm trade in the automotive sector: a network approach. *arXiv:2202.00409 [econ, q-fin, stat]*. arXiv: 2202.00409.

Smith, R. (2009). *Military economics: the interaction of power and money*. Basingstoke: Palgrave Macmillan.

Snijders, T. A. (1996). Stochastic actor-oriented models for network change. *Journal of mathematical sociology 21*(1-2), 149–172.

Snijders, T. A. B. (2002). Markov Chain Monte Carlo Estimation of Exponential Random Graph Models. *Journal of Social Structure 3*.

Snijders, T. A. B., A. Lomi, and V. J. Torló (2013, May). A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social Networks 35*(2), 265–276.

Snijders, T. A. B., P. E. Pattison, G. L. Robins, and M. S. Handcock (2006). New Specifications for Exponential Random Graph Models. *Sociological Methodology 36*, 99–153.

Snijders, T. A. B., R. Ripley, C. Steglich, J. Koskinen, N. Niezink, V. Amati, C. Stadtfeld, J. Hollway (IHEID), P. Block, R. Krause, C. Greenan, J. Lospinoso, M. Schweinberger, M. Huisman, K. Boitmanis, F. Schoenenberger, M. Ortmann, M. Hoffman, R. Hellpap, and A. Uzaheta (2021, November). RSiena: Siena - Simulation Investigation for Empirical Network Analysis.

Snijders, T. A. B., G. G. van de Bunt, and C. E. G. Steglich (2010, January). Introduction to stochastic actor-based models for network dynamics. *Social Networks 32*(1), 44–60.

Stewart, J. and M. Schweinberger (2021, August). mlergm: Multilevel Exponential-Family Random Graph Models.

Stockholm International Peace Research Institute (2019). Military Expenditure Database. (Accessed 2020-09-03).

Stockholm International Peace Research Institute (2021a). Arms Transfers Database. (Accessed 2020-03-09).

Stockholm International Peace Research Institute (2021b, September). *SIPRI Yearbook 2021: Armaments, Disarmament and International Security*. Oxford, New York: Oxford University Press.

The World Bank (2021a). GDP (constant 2015 US$) | Data.

The World Bank (2021b). GDP per capita (constant 2015 US$) | Data.

Thurner, P. W., C. S. Schmid, S. J. Cranmer, and G. Kauermann (2019, August). Network Interdependencies and the Evolution of the International Arms Trade. *Journal of Conflict Resolution 63*(7), 1736–1764.

Timmer, M. P., E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries (2015, August). An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production: User Guide to World Input-Output Database. *Review of International Economics 23*(3), 575–605.

United Nations (2022). UN COMTRADE.

Ward, M. D., J. S. Ahlquist, and A. Rozenas (2013, April). Gravity's Rainbow: A dynamic latent space model for the world trade network. *Network Science 1*(1), 95–118.

Willardson, S. L. and R. A. Johnson (2022, March). Arms transfers and international relations theory: Situating military aircraft sales in the broader IR context. *Conflict Management and Peace Science 39*(2), 191–213.

A Appendix

A.1 Additional Data Resources

A.1.1 List of Countries

Albania
Algeria
Argentina
Australia
Austria
Bahrain
Bangladesh
Belgium
Benin
Bolivia
Brazil
Bulgaria
Burkina Faso
Burundi
Cambodia
Cameroon
Canada
Central African Republic
Chad
Chile
China
Colombia
DR Congo
Congo
Cote dIvoire
Cuba
Cyprus
Denmark
Dominican Republic
Ecuador
Egypt
El Salvador
Ethiopia
Fiji
Finland
France
Gabon
Gambia
Germany
Ghana
Greece
Guatemala
Guinea
Guyana
Honduras
Hungary
India
Indonesia
Iran
Ireland
Israel
Italy
Jamaica
Jordan
Kenya
South Korea
Kuwait
Laos
Liberia
Libya
Madagascar
Malawi
Malaysia
Mali
Mauritania
Mauritius
Mexico
Mongolia
Morocco
Mozambique
Myanmar
Nepal
Netherlands
New Zealand
Nicaragua
Niger
Nigeria
Norway
Oman
Pakistan
Panama
Papua New Guinea
Paraguay
Peru
Philippines
Poland
Portugal
Romania
Russia
Rwanda
Saudi Arabia
Senegal
Sierra Leone
Singapore
South Africa
Spain
Sri Lanka
Sudan
Sweden
Switzerland
Tanzania
Thailand
Togo
Trinidad and Tobago
Tunisia
Turkey
Uganda
United Arab Emirates
United Kingdom
United States
Uruguay
Viet Nam
Zambia
Zimbabwe

A.2 ERGM: Additional Analyses

A.2.1 Supplementary Results for the Analysis of the Year 2003

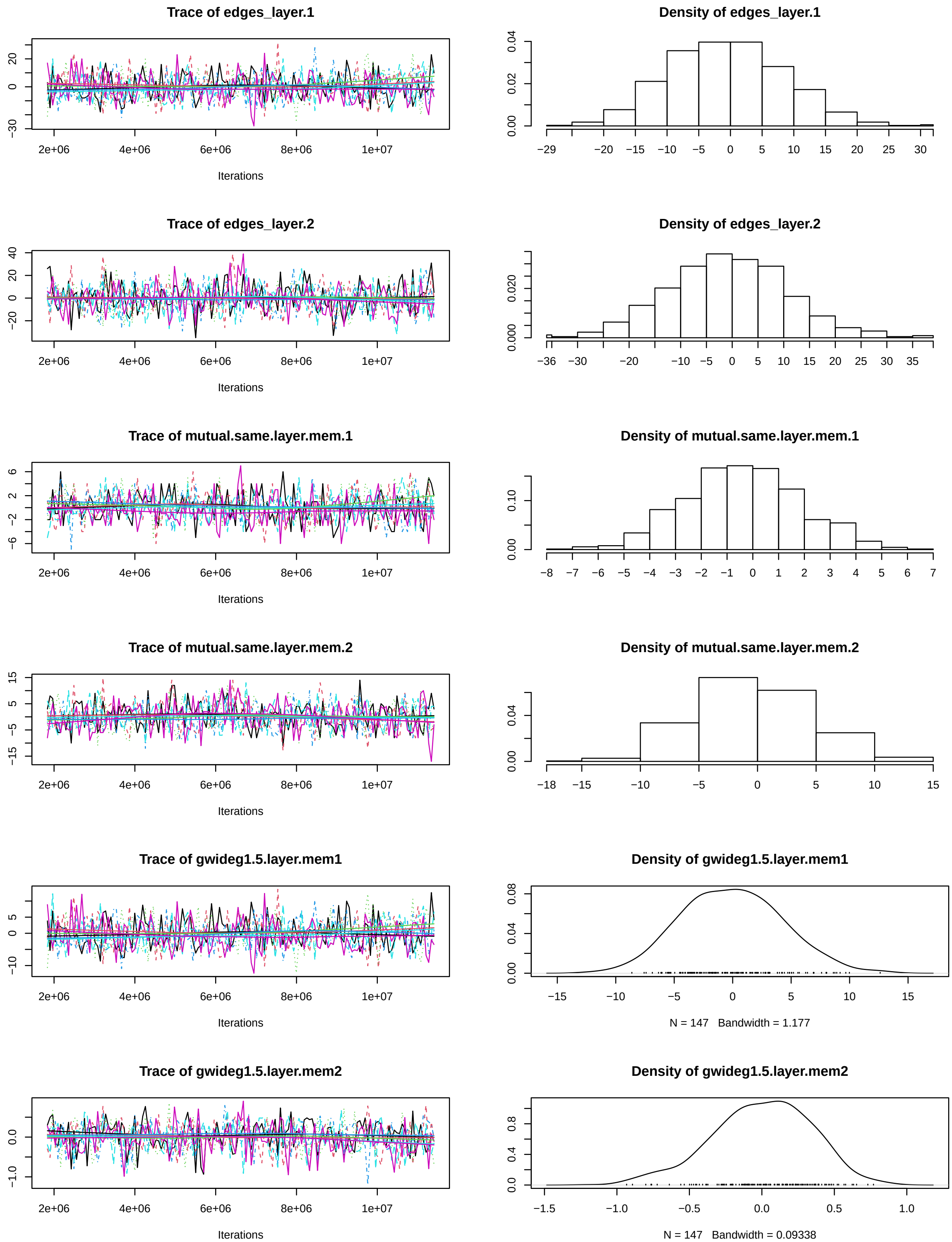
In this section we provide additional analyses for the multilayer ERGM analysis in section 3.2.3.

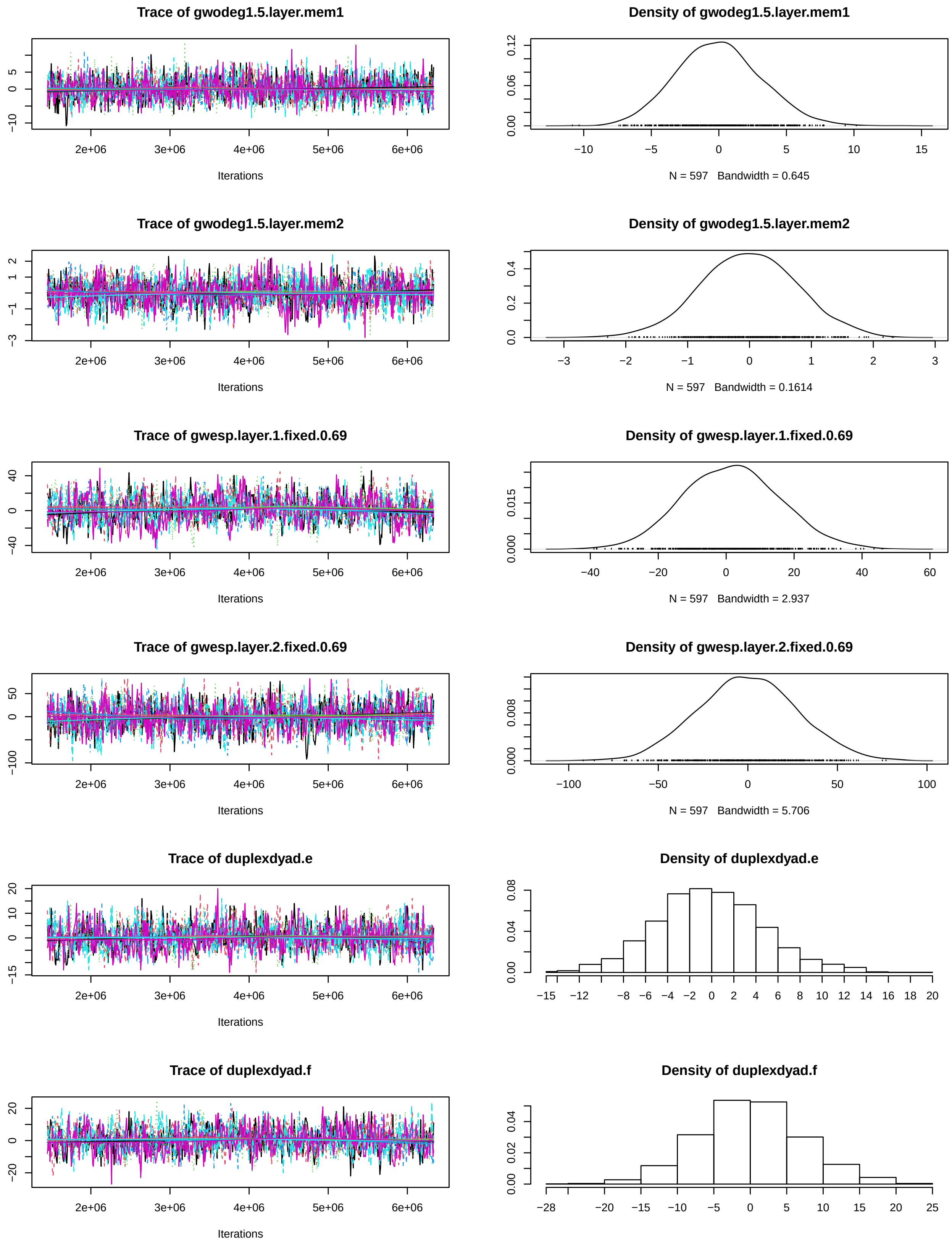
	Import Dep.	Export Dep.
Layer 1: Arms Trade		
Edges	$-7.93\ (5.10)$	$-7.61\ (5.06)$
Reciprocity	$1.34\ (0.62)^{*}$	$1.32\ (0.60)^{*}$
Gw Indegree (d = 1.5)	$-2.43\ (1.05)^{*}$	$-2.45\ (1.05)^{*}$
Gw Outdegree (d = 1.5)	$-3.37\ (0.89)^{***}$	$-3.37\ (0.85)^{***}$
GWESP Outgoing Two-path (d = 0.69)	$0.06\ (0.15)$	$0.07\ (0.15)$
Distance (log)	$-0.17\ (0.20)$	$-0.18\ (0.20)$
GDP in (log)	$0.19\ (0.27)$	$0.18\ (0.28)$
GDP out (log)	$-0.37\ (0.25)$	$-0.37\ (0.24)$
Alliance	$0.03\ (0.45)$	$0.05\ (0.46)$
Polity Diff. (abs)	$-0.02\ (0.03)$	$-0.02\ (0.03)$
Military Expenditure in (log)	$0.13\ (0.29)$	$0.13\ (0.29)$
Military Expenditure out (log)	$0.36\ (0.22)$	$0.35\ (0.22)$
Path Dependency	$1.74\ (0.34)^{***}$	$1.73\ (0.36)^{***}$
Layer 2: Conventional Trade		
Edges	$-16.13\ (3.21)^{***}$	$-13.82\ (2.45)^{***}$
Reciprocity	$0.13\ (0.33)$	$0.43\ (0.27)$
Gw Indegree (d = 1.5)	$-5.27\ (3.79)$	$0.44\ (0.82)$
Gw Outdegree (d = 1.5)	$0.21\ (0.85)$	$-1.74\ (1.61)$
GWESP Outgoing Two-path (d = 0.69)	$0.68\ (0.14)^{***}$	$0.58\ (0.13)^{***}$
Distance (log)	$-0.60\ (0.15)^{***}$	$-0.51\ (0.13)^{***}$
GDP in (log)	$0.57\ (0.18)^{**}$	$0.11\ (0.12)$
GDP out (log)	$0.05\ (0.16)$	$0.41\ (0.14)^{**}$
Alliance	$0.84\ (0.29)^{**}$	$0.10\ (0.25)$
Polity Diff. (abs)	$0.04\ (0.02)^{*}$	$-0.02\ (0.02)$
Military Expenditure in (log)	$-0.03\ (0.14)$	$-0.18\ (0.10)$
Military Expenditure out (log)	$-0.18\ (0.13)$	$0.01\ (0.11)$
Path Dependency	$2.13\ (0.08)^{***}$	$1.93\ (0.07)^{***}$
AIC	1456.13	1754.85
BIC	1650.19	1948.90
Log Likelihood	-702.07	-851.43

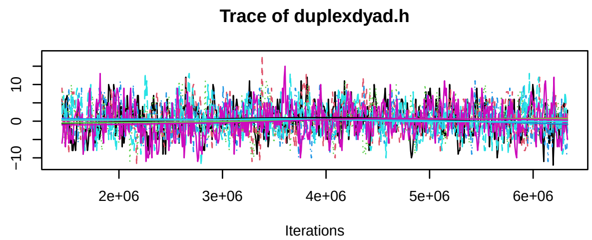
Estimates based on Monte Carlo MLE. Standard Errors in parenthesis.$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$.

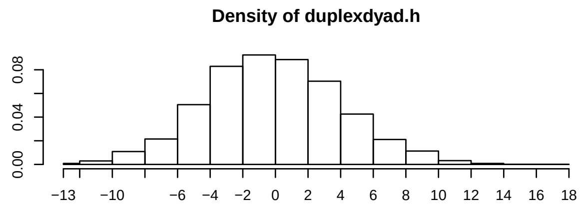
Table 7: MERGM results for two-layer network of weapons and import (left) or export (right) trade dependency in the year 2003 - estimated without cross-layer effects.

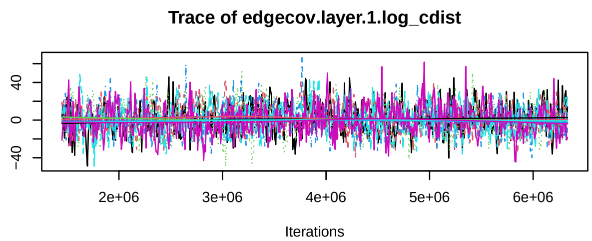
A.2.2 Markov Chain Monte Carlo Diagnostics for the Import Dependency Model 2003

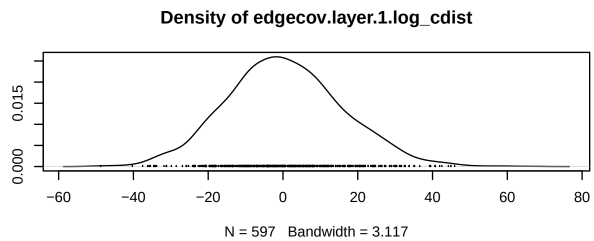


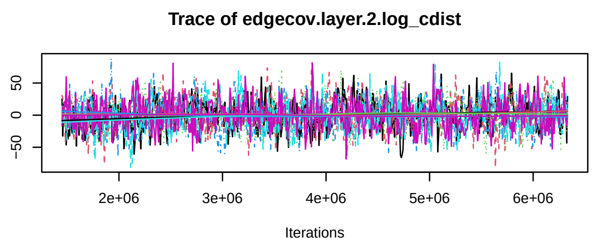

Trace of gwodeg1.5.layer.mem1
Iterations
Density of gwodeg1.5.layer.mem1
N = 597 Bandwidth = 0.645
Trace of gwodeg1.5.layer.mem2
Iterations
Density of gwodeg1.5.layer.mem2
N = 597 Bandwidth = 0.1614
Trace of gwesp.layer.1.fixed.0.69
Iterations
Density of gwesp.layer.1.fixed.0.69
N = 597 Bandwidth = 2.937
Trace of gwesp.layer.2.fixed.0.69
Iterations
Density of gwesp.layer.2.fixed.0.69
N = 597 Bandwidth = 5.706
Trace of duplexdyad.e
Iterations
Density of duplexdyad.e
Trace of duplexdyad.f
Iterations
Density of duplexdyad.f

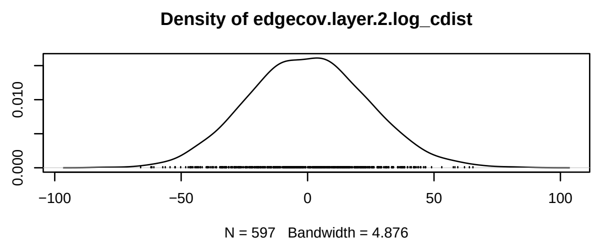

Trace of duplexdyad.h
Iterations

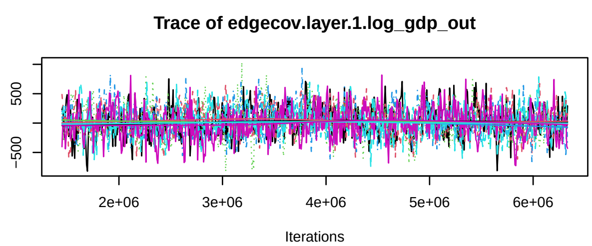

Density of duplexdyad.h

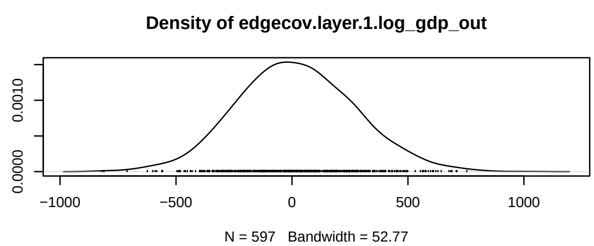

Trace of edgecov.layer.1.log_cdist
Iterations

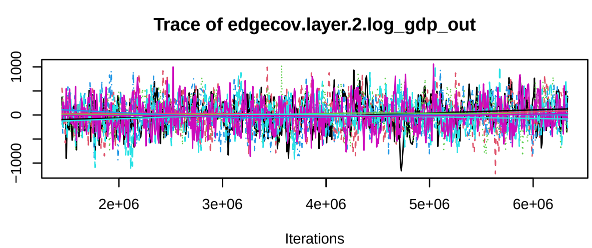

Density of edgecov.layer.1.log_cdist
N = 597 Bandwidth = 3.117

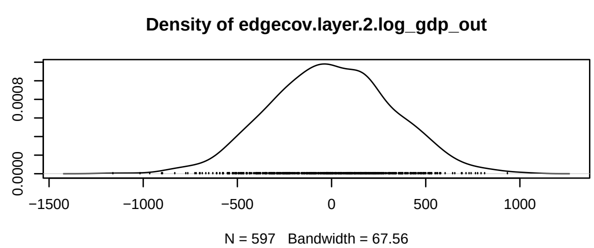

Trace of edgecov.layer.2.log_cdist
Iterations

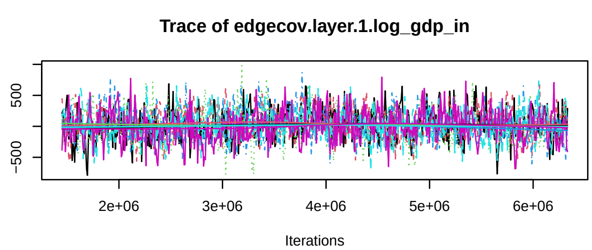

Density of edgecov.layer.2.log_cdist
N = 597 Bandwidth = 4.876

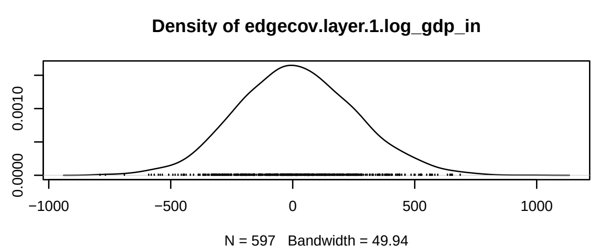

Trace of edgecov.layer.1.log_gdp_out
Iterations

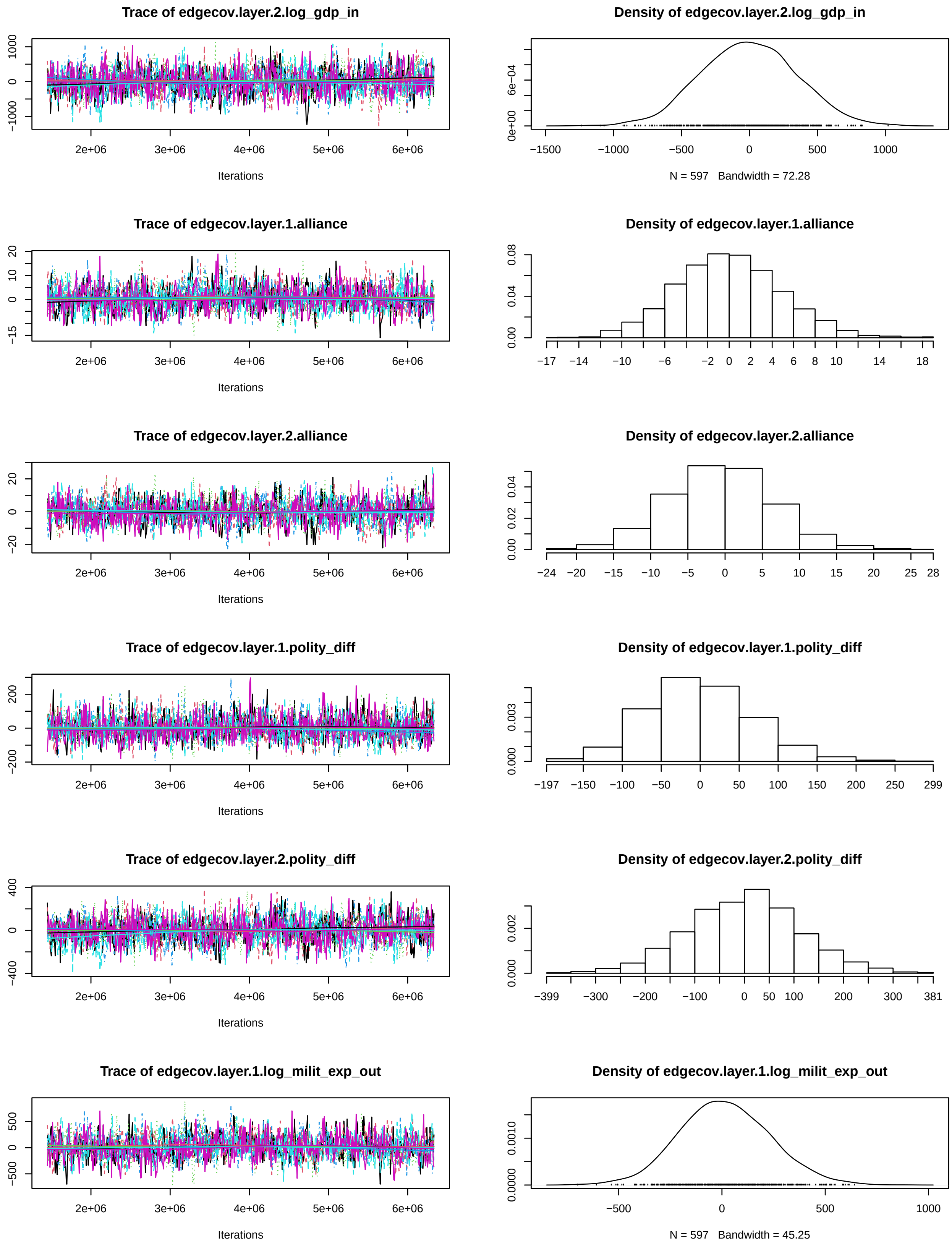

Density of edgecov.layer.1.log_gdp_out
N = 597 Bandwidth = 52.77

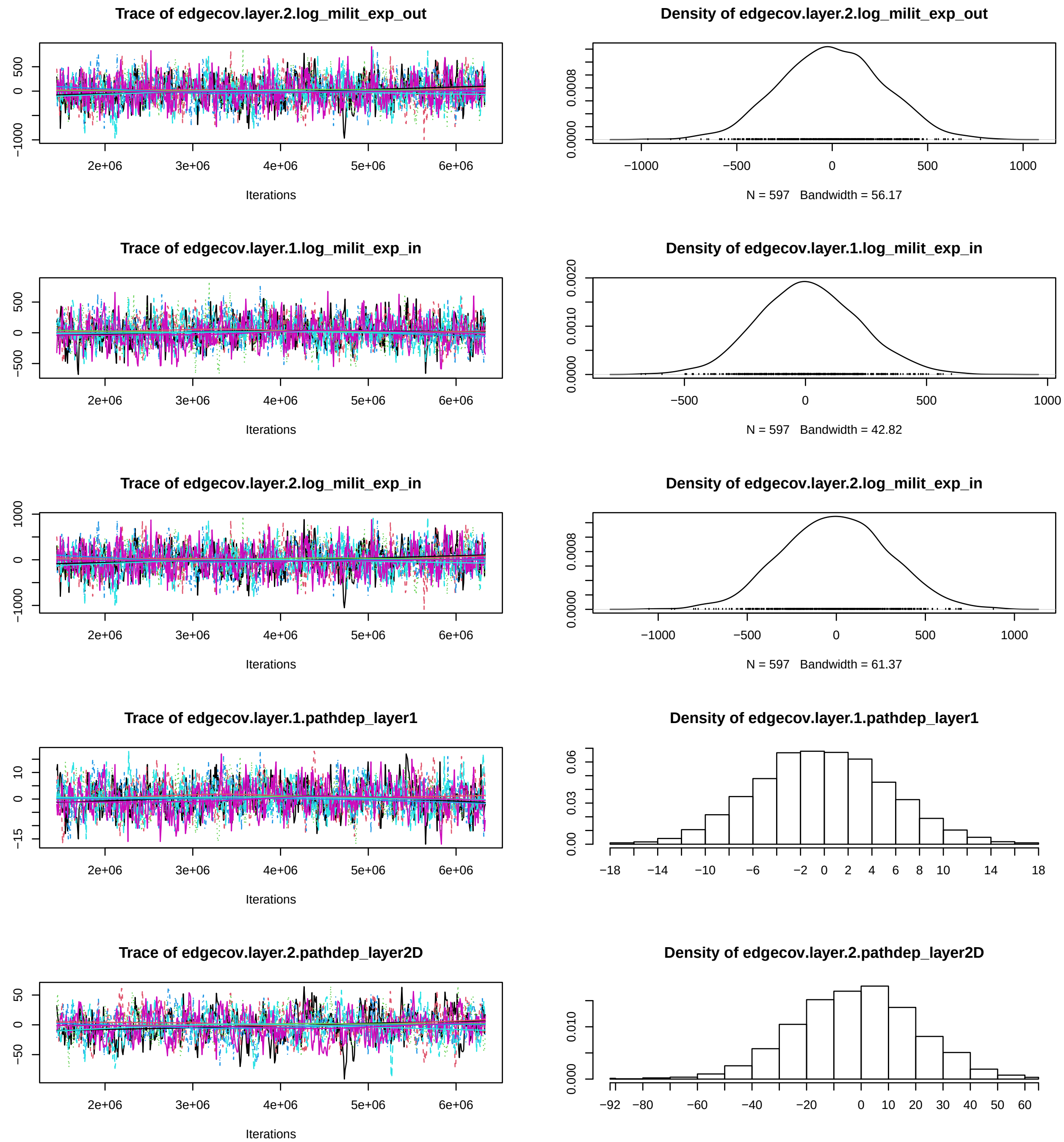

Trace of edgecov.layer.2.log_gdp_out
Iterations


Density of edgecov.layer.2.log_gdp_out
N = 597 Bandwidth = 67.56


Trace of edgecov.layer.1.log_gdp_in
Iterations


Density of edgecov.layer.1.log_gdp_in
N = 597 Bandwidth = 49.94


Trace of edgecov.layer.2.log_gdp_in
Iterations
Density of edgecov.layer.2.log_gdp_in
N = 597 Bandwidth = 72.28
Trace of edgecov.layer.1.alliance
Iterations
Density of edgecov.layer.1.alliance
Trace of edgecov.layer.2.alliance
Iterations
Density of edgecov.layer.2.alliance
Trace of edgecov.layer.1.polity_diff
Iterations
Density of edgecov.layer.1.polity_diff
Trace of edgecov.layer.2.polity_diff
Iterations
Density of edgecov.layer.2.polity_diff
Trace of edgecov.layer.1.log_milit_exp_out
Iterations
Density of edgecov.layer.1.log_milit_exp_out
N = 597 Bandwidth = 45.25


Trace of edgecov.layer.2.log_milit_exp_out
Iterations
Density of edgecov.layer.2.log_milit_exp_out
N = 597 Bandwidth = 56.17
Trace of edgecov.layer.1.log_milit_exp_in
Iterations
Density of edgecov.layer.1.log_milit_exp_in
N = 597 Bandwidth = 42.82
Trace of edgecov.layer.2.log_milit_exp_in
Iterations
Density of edgecov.layer.2.log_milit_exp_in
N = 597 Bandwidth = 61.37
Trace of edgecov.layer.1.pathdep_layer1
Iterations
Density of edgecov.layer.1.pathdep_layer1
Trace of edgecov.layer.2.pathdep_layer2D
Iterations
Density of edgecov.layer.2.pathdep_layer2D

A.3 SAOM: Goodness of fit

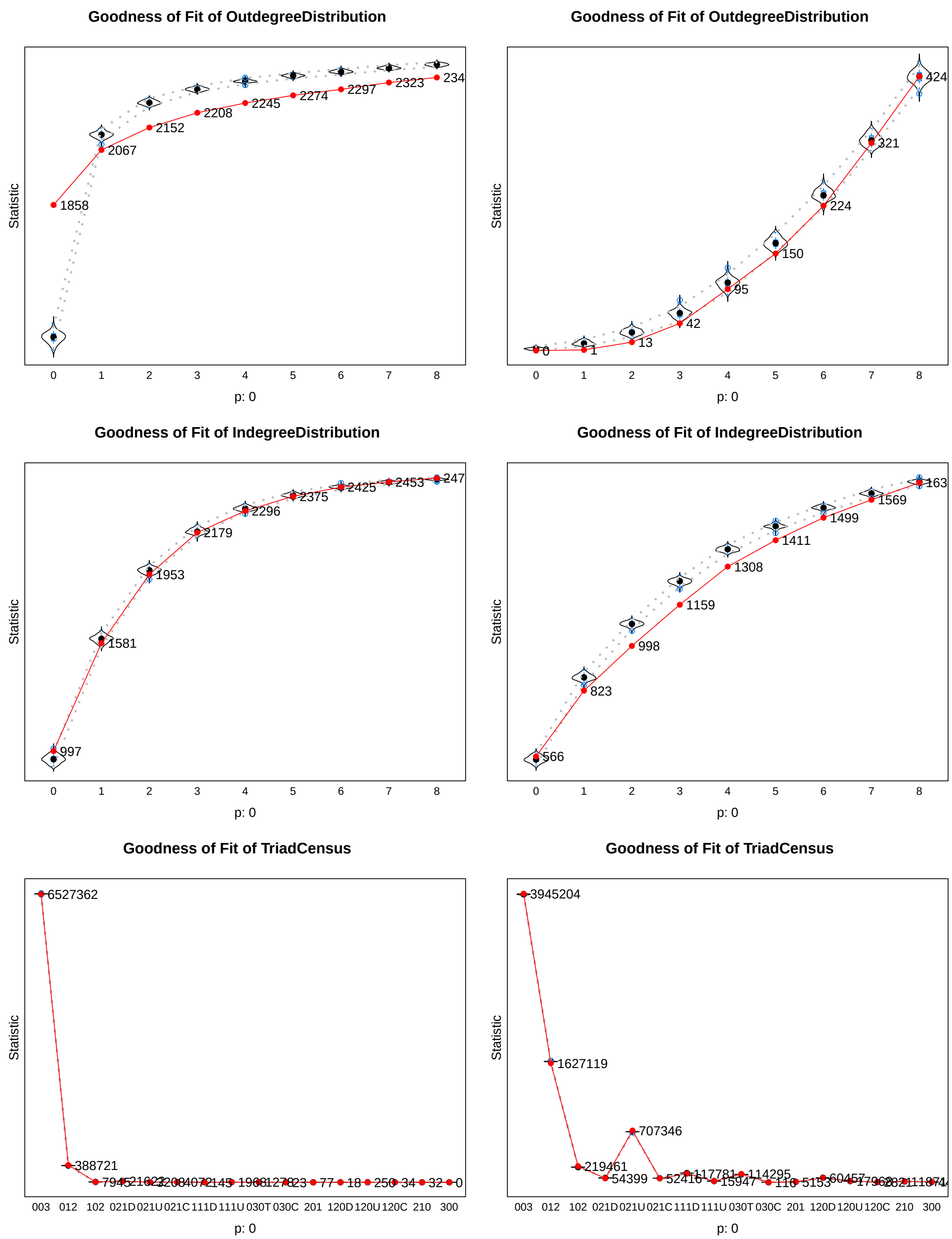


Figure 12: Goodness of Fit for different network statistics – multilayer SAOM.
We simulate 100 networks and plot the empirical distribution functions for selected network statistics. Observed statistics in red, violin plots refer to the distribution from the simulations. TriadCensus: Davis and Leinhardt triad census classifcation. First column: arms trade network, second column: import dependency network.