Modeling international trade disputes as a statistical network

Adam Scherling

December 15, 2017

Abstract

Trade disputes amongst members of the World Trade Organization are analyzed as a directed network. The European Union and the United States are found to be central players, followed distantly by countries in Latin America and Asia; most countries are not involved in any disputes. Amongst those involved in distputes, statistical models find consistent, significant effects of GDP, GDP per capita, and the volume of imports and exports as a percentage of GDP, for explaining network structure.

Introduction

The World Trade Organization (WTO) is an international organization that governs international trade. Its 164 member countries[1] are all signatories of a set of agreements known as the Uruguary Round agreements which were first signed in 1994 and which came into effect in January 1995. The Uruguay Round agreements outline principles of trade, commitments to low barriers to trade (e.g. low tariffs), and procedures for resolving disputes.[5]

From the time at which the Uruguay Round agreements came into force in January 1995 through the end of 2015, a total of 501 disputes were filed with the WTO.[3] In each dispute, a member country (or group of member countries) accuses another member country (or group of member countries) of violating WTO rules. I describe these disputes as a network. In particular, each member country can be represented as a node, and each dispute can be represented as a directed edge from one node to another, or by multiple directed edges.[2]

[1] Technically, not all members of the WTO are countries; the European Union is also a member of the WTO, as are all its member states.

[2] In cases where a dispute involves more than two countries, a directed edge is created from each complainant to each respondent.

This representation allows me to gain numerous statistical insights about the historical pattern of disputes. For example, I can summarize the network using statistical summaries such as the density of ties, or the distributions of outgoing ties and incoming ties. I can also examine which outside factors can be used to explain the likelihood of a tie, that is, the likelihood that one particular country files a dispute against another particular country.

To this end I include a dataset of country-level economic indicators collected by the World Bank and provided by Kaggle[1], such as GDP, GDP per capita, and the volume of exports and imports. I also include one dyad-level covariate indicating whether each pair of countries has common membership in a regional trade agreement (e.g. NAFTA or MERCOSUR). This data was collected from the WTO website.[4]

Using these variables, I characterize the network and seek to identify which factors are relevant in explaining the presense of ties (trade disputes) between countries in the WTO. The rest of this paper is as follows: first I characterize the network using visual and statistical summaries. I then fit statistical models to the network to explain which covariates determine the likelihood of a complaint. The discussion considers the implications of these findings. Further details about the data are contained in the Appendix.

Analysis and Results

Network Summaries

The best summaries of data are often visual; networks are no exception. Figure 1 shows the network of trade disputes amongst WTO member countries from 1995 to 2015.[3] Due to the large number of countries, it is impractical to label each node. I do, however, color each node by region.[4] Notably, although there were 501 disputes during this period, there are only 203 unique edges; the rest are redundant in the sense that a previous dispute had already been filed by the same complainant against the same respondent.

[3]The network excludes Liechtenstein, Myanmar, and Taiwan, for which the World Bank's economic indicators were unavailable.

[4]My region categorizations come from the World Bank.[2]

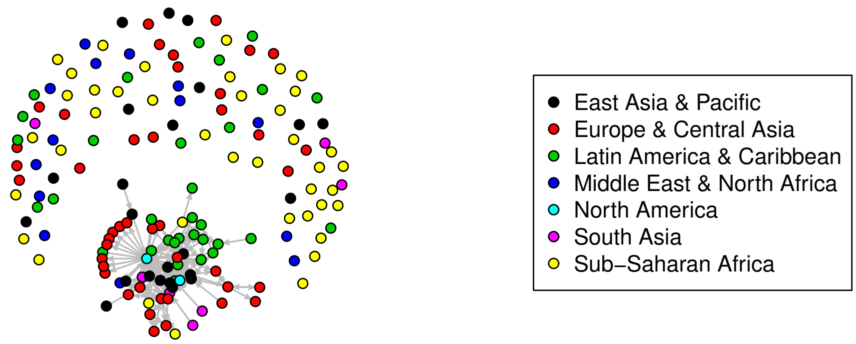


Figure 1: WTO dispute network

A few clear patterns emerge from this plot. First, the majority of nodes are isolates; most countries are not involved in a dispute. Second, some regions are clearly more likely to be involved in disputes. In particular, countries in Sub-Saharan Africa and the Middle East & North Africa are very rarely involved in trade disputes. On the other hand, both North American countries (the United States and Canada)[5] and at least 50% of the WTO member countries in each other region are involved in trade disputes.

Figure 2 below shows the network excluding all isolated points. A few more patterns become clear. First, the non-isolated countries form a single connected component. Additionally, there a small but noticeable amount of clustering by region; the Latin Amerian countries in particular form a fairly coherent cluster. It can also be seen that the United States (the light blue node in the center) plays a fairly central role as the hub of a wheel with many spokes, especially countries in Europe. It is a lopsided wheel, to be sure; the network is much more complex than a simple hub-and-spoke structure.

Due to the large number of isolates, the network that includes all WTO members is very sparse; it has a density of 0.007. If the isolates are all removed the density becomes 0.05. Whether or not to include the isolates is an interesting question. Any model that excludes them is inherently incorrect in the sense that it assumes that the countries which were not in any trade disputes had zero probability of ever being in a dispute and can be completely disregarded. On the other hand, including them in the model is likely to cause issues of model degeneracy and poor fit.

[5]The World Bank classification places the United States, Canada, and Bermuda in North America. Mexico is classified as part of Latin America. Bermuda is not a member of the WTO.

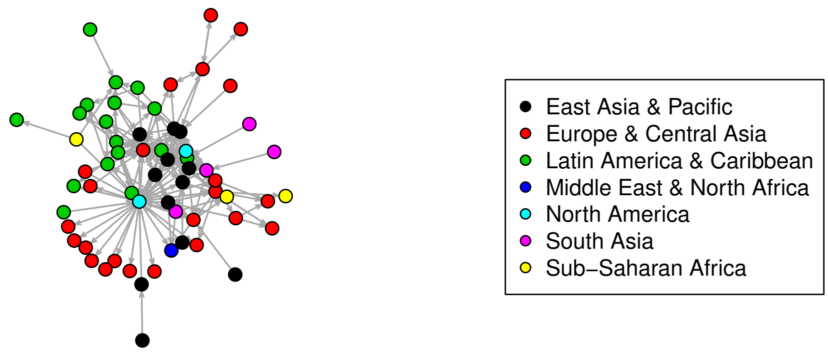


Figure 2: WTO dispute network (isolates removed)

If we were particularly interested in explaining which countries were involved in any disputes (i.e. which countries are not isolates), then a binary classification model would be more appropriate than any network methods. However worthwhile such an effort may be, I consider it to be outside the scope of the current analysis. Therefore I proceed using the abridged network in which all isolates are removed, with the understanding that all of my results will be conditional on having restricted my sample to countries which were involved in disputes.

Since a country may file multiple complaints against another country, it is sensible to keep track of the number of disputes represented by each edge. I call these counts the edge weights. The distribution of edge weights is skewed heavily to the right, as shown in Table 1 below. A little over half of the edges contain only one dispute; on the other hand, there are a handful of edges with a rather large number of disputes. The mean and standard deviation of the edge weights are 2.58 and 3.49, respectively. Perhaps unsurprisingly, the edges with the most repeats are those that include the largest economies. The European Union filed 33 complaints against the United States; in turn the US filed 19 complaints against the EU.[6]

Table 1: Distribution of edge weights

Weight	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**15**	**17**	**19**	**33**
Count	119	30	16	9	5	8	3	3	3	2	1	1	1	1	1

Another way of characterizing the network is by the indegree and outdegree distributions. Whereas the edge weights shown above are calculated amongst all edges (which represent pairs of countries), the

[6]The next largest counts include 17 disputes filed by the US against China, and 15 by Canada against the US.

indegree and outdegree distributions give the number of edges originating from, and directed towards, each individual country. Recall that the number of edges involving a country is different from the number of disputes involving that country. I present the unweighted degree distributions, representing the number of edges, in Figure 3 below. The distribution of disputes (the weighted degree distribution) looks similar, but with more extreme values amongst the countries with large numbers of ties.

Figure 3: Unweighted degree distributions

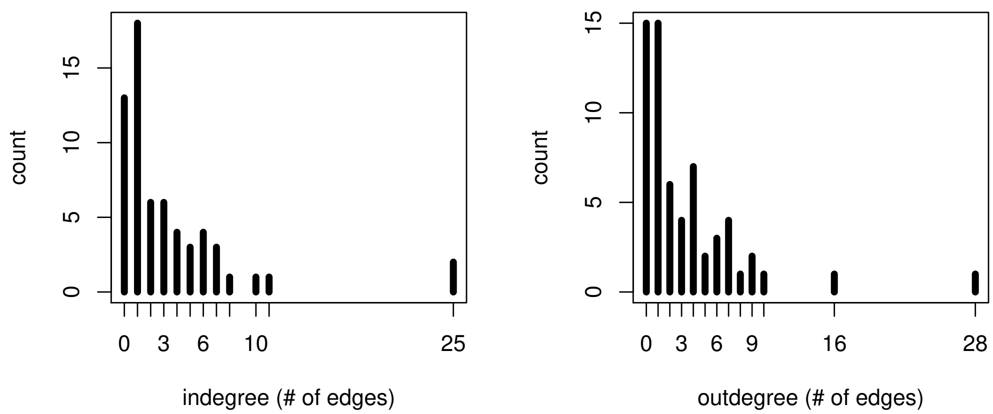


Figure 3 shows that the distributions of indegree and outdegree are quite similar, and that both are skewed to the right. The extreme values, once again, belong to the US and EU, which filed disputes against 28 and 16 distinct countries, respectively. Additionally both the US and EU were the respondent in disputes from 25 distinct countries.

A related question is how correlated the indegree and outdegree distributions are. If they are highly correlated, that would suggest that countries tend to differ from each other simply in the number of disputes that they are involved in, and that they do not differ so much in their tendency to originate disputes or to be the target of disputes. The correlations for the unweighted and weighted distributions are shown in Table 2 below.

Table 2: Degree distribution correlations

	Unweighted degree distributions	Weighted degree distributions
Including US and EU	0.82	0.96
Excluding US and EU	0.54	0.74

For the network as a whole, the indegree and outdegree are highly correlated. This is particularly true if the edge weights are taken into account. However, if the largest economies (the US and EU) are removed, then the correlation drops noticeably. Thus a large portion of the correlation is explained by the fact that these two nodes both have very high indegree and outdegree. This suggests once again that these two WTO members play a central role in the network, as their inclusion or exclusion has a dramatic effect on the properties of the network.

Another measure of centrality which can be considered is the eigenvector centrality. The eigenvector centrality is calculated as the first eigenvector of the sociomatrix[7] representing the network. Eigenvector centrality has the useful property of placing a high value on nodes which are closely linked to other nodes that also have high values. As such it is a good measure of how central a node is in a given network. I calculate the eigenvector centrality of both the unweighted and weighted versions of the WTO dispute network. The countries with the top ten weighted and unweighted eigenvector centralities are shown in Table 3. For reference I also include the degree (indegree + outdegree) of each country.

Table 3: Eigenvalue centralities

Unweighted			Weighted		
country	eigen centrality	degree	country	eigen centrality	degree
european union	0.41	41	european union	0.61	187
united states	0.40	53	united states	0.50	244
japan	0.28	14	canada	0.33	51
canada	0.26	14	brazil	0.24	43
brazil	0.24	17	india	0.20	42
new zealand	0.22	7	korea, republic of	0.17	32
mexico	0.21	14	mexico	0.17	37
chile	0.20	12	china	0.16	47
indonesia	0.19	13	argentina	0.15	42
argentina	0.18	16	japan	0.15	36

Table 3 produces a few interesting results. First, the European Union is shown to be more central than the United States in the WTO dispute network, despite having a significantly smaller degree. Although the United States is tied to more nodes, these nodes are of lesser importance to the network than the nodes which are tied to the EU. To some extent this can be seen in Figure 2; many of the nodes linked to the US are not tied to any other countries.[8] On the other hand, the EU (the red dot just above the US) is in a thicker mesh of ties. It is also of interest that Latin American countries seem to have a prominent role in the network. It is less surprising to see large Asian economies in the

[7]A sociomatrix is a matrix representation of a network in which entry i, j is 1 if an edge exists from node i to node j, and 0 otherwise. For a valued network, the sociomatrix may instead include numeric values representing the strength of each edge.

[8]Rather interestingly, many of these countries linked to the US are individual member states of the European Union.

list of central nodes, nor is it surprising that African countries do not seem to be very central to the network.

Network Models

Having gained a basic understanding of the structure of the network, it is appropriate to approach questions of a greater scientific interest. In particular, I attempt to discern which factors are significant in determining which pairs of countries are tied together in the WTO dispute network. In particular, I consider each country's GDP, GDP growth, GDP per capita, GDP per capita growth, exports and imports (both dollar-valued and as a percentage of GDP), and membership in common regional trade agreements.[9] All of these variables can be reasonably expected to have an influence on the occurence of trade disputes.

Disputes are unlikely to occur if the stakes are low; a country will presumably have a vested interest in its trade with another country if it is going to accuse that country of trade violations. It is thus reasonable to expect that countries with large GDP will both attract and file more disputes; they represent a large market for exports from other countries, and they are likely to export significant amounts of goods to these countries as well. Countries with a large GDP per capita are, regardless of their size, more likely to be involved in international affairs, and their governments will have more highly developed institutions, which may be critical for pursuing international trade disputes. Countries that experience high growth in GDP or GDP per capita may also be important players on the international stage, and will have high stakes in leveraging their exports (or possibly restricting imports) to sustain high levels of growth. The volume of exports and imports might speak more directly as to which countries are involved in international trade, with high volumes likely corresponding to high numbers of disputes as well. Similarly, countries that are members of a common regional trade agreement are likely to have high volumes of trade between them and may thus incur more disputes; or, perhaps, their regional cooperation will reduce trade violations, and fewer disputes will arise.

To test these theories, I fit a few variants of an ERGM (exponential random graph model). The model defines a probability distribution over the set of all possible graphs (networks) with the same number of nodes as the WTO dispute network. For the unweighted network, there are $2^{(62\times61)}$ possible graphs, since there are 62 WTO members included in the network; each possible edge can take on two

[9] See the Appendix for more information about how these variables are calculated.

values, either 0 or 1. The probability of any one of these graphs y is modeled as

$$P(Y = y) = \frac{\exp\{\sum_{k=1}^{K} \theta_k g_k(y)\}}{c(\theta)}$$

where $g(y) = \{g_1(y), ..., g_k(y)\}$ is a set of K graph statistics, $\theta = \{\theta_1, ..., \theta_K\}$ is the set of corresponding parameters, and $c(\theta)$ is a normalizing constant such that the total probability of all possible graphs is equal to 1. The set of graph statistics $g(y)$ may include any statistic derived from the graph, including statistics based on node-level and dyad-level covariates such as my economic indicators from the World Bank. These statistics would then affect the probability of the graph as determined by the parameter θ. Estimation of the model is done by Markov Chain Monte Carlo (MCMC).

For interpreting the parameter θ, it is perhaps helpful to note that the log odds of network y_1 compared to network y_0 are as follows:

$$\log \frac{P(Y = y_1)}{P(y = y_0)} = \sum_{k=1}^{K} \theta_k \left[g_k(y_1) - g_k(y_0)\right]$$

In other words, θ_k is the improvement in the log odds of network y_1 compared to y_0 that is due to statistic g_k if $g_k(y_1) - g_k(y_0) = 1$. If y_1 and y_0 differ only in one tie (suppose y_1 has one tie that is not present in y_0) then the values of θ and g can be used to show how the existence of this tie changes the likelihood of the network. If the log odds are positive, then adding the tie increases the likelihood; if the log odds are negative, then it decreases the likelihood.[10]

With this in mind, I fit an ERGM for the WTO trade network. In addition to the trade-related variables discussed above, I include terms for the network density, the number of mutual ties, and the number of edges within regions. The model coefficients are shown in Table 4. Diagnostics for the MCMC fit are shown in Figure 4. Most of the graph statistics stay centered at 0, and the variances appear to be constant. These two facts would imply that for these statistics, the Markov Chain has reached its stationary distribution, and the MLE estimates should be reliable. The RTA (regional trade agreement) statistics, however, appear to be quite unstable. Furthermore, the "MCMC %" column in Table 4 shows that the MCMC estimation is actually highly unstable for the edges, mutual, and nodematch("region") terms in addition to the RTA terms.[11]

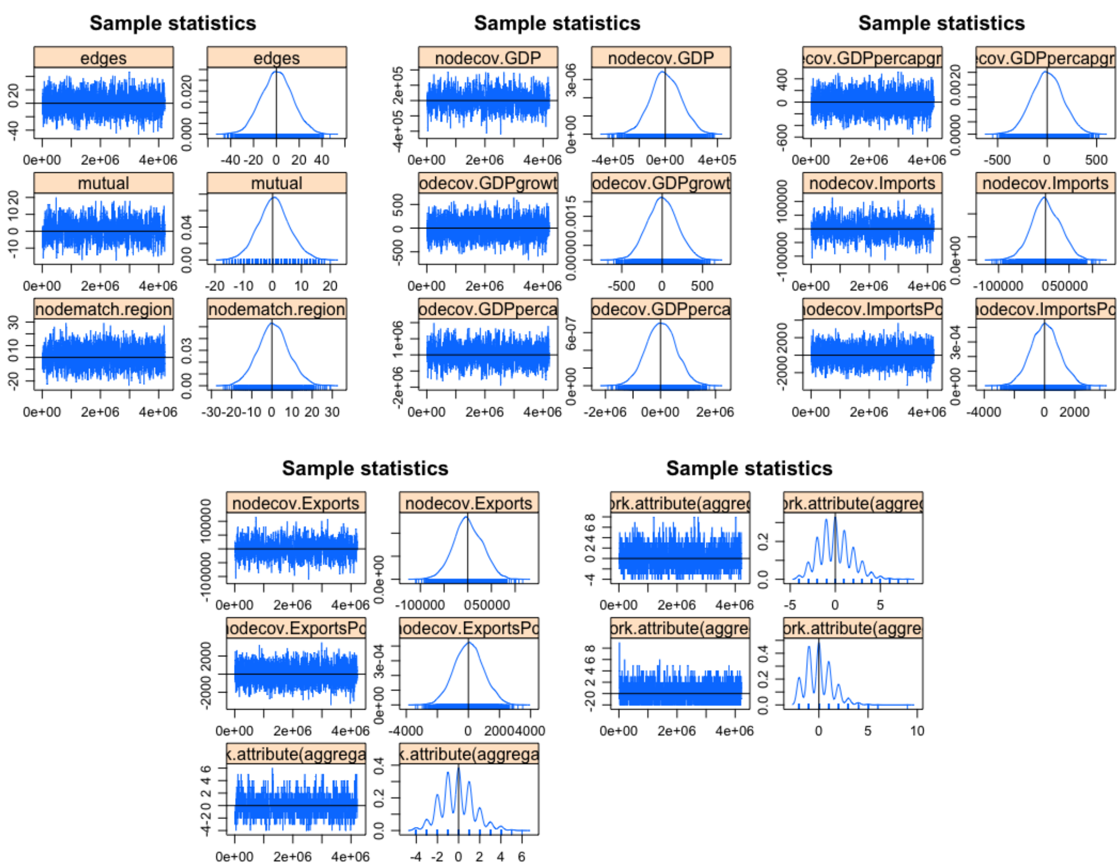
[10]For reference, note that a log odds of 0.69 implies an odds ratio of 2, meaning that network y_1 is twice as likely as network y_0. Therefore if adding a tie to a network leads to log odds of 0.69, this means that inclusion of the tie doubles the probability of observing the network.

[11]The "MCMC %" column denotes the percent of variation in the estimates attributable to Markov Chain instability; levels above 10% should cause concern.

Table 4: ERGM model fit

	Estimate	Std. Error	MCMC %	p-value
edges	-1.973e+00	1.311e-02	97	< 1e-04
mutual	1.344e+00	7.585e-03	99	< 1e-04
nodematch.region	8.833e-01	4.453e-03	97	< 1e-04
nodecov.GDP	2.576e-04	8.191e-05	0	0.001671
nodecov.GDPgrowth	1.585e-01	8.279e-03	3	< 1e-04
nodecov.GDPpercap	-1.723e-05	4.436e-06	0	0.000105
nodecov.GDPpercapgrowth	-2.165e-01	7.135e-03	4	< 1e-04
nodecov.Imports	4.918e-04	1.380e-03	0	0.721569
nodecov.ImportsPct	-3.546e-02	8.538e-03	0	< 1e-04
nodecov.Exports	-6.492e-04	1.132e-03	0	0.566449
nodecov.ExportsPct	1.850e-02	8.215e-03	0	0.024365
dyadcov.rta.mutual	5.102e-01	3.023e-02	100	< 1e-04
dyadcov.rta.utri	2.665e-01	1.263e-02	100	< 1e-04
dyadcov.rta.ltri	-4.714e-01	1.510e-02	100	< 1e-04

Figure 4: MCMC diagnostics



Due to this instability, I fit a second model, shown in Table 5, that excludes these variables. I also exclude the terms for dollar-valued imports and exports, which appear to be insignificant. To account for the loss of all these terms, I add in a term for the geometric-weighted edgewise shared partner distribution. The edgewise shared partner distribution in the WTO dispute network is the distribution of counts of directed edgewise shared partners; for example, country A files a dispute against country B and country A also files a dispute against country C, who in turn files a dispute against country B, then country C is an edgewise shared partner for the directed edge from A to B. In the ERGM, this distribution is then weighted geometrically, i.e. such that the weights are decreasing for higher numbers of edgewise shared partners. This weighting allows the distribution to be represented as a single term in the model, which will account for some amount of transitivity in the network. The decay factor of 1.3 that I use was estimated in a previous ERGM, not shown.

Table 5: ERGM model fit excluding unstable terms

	Estimate	Std. Error	MCMC %	p-value
gwesp.fixed.1.3	3.320e-01	1.126e-03	9	$<$ 1e-04
nodecov.GDP	1.820e-04	1.754e-05	0	$<$ 1e-04
nodecov.GDPgrowth	-9.329e-02	7.067e-03	0	$<$ 1e-04
nodecov.GDPpercap	-2.630e-05	3.840e-06	0	$<$ 1e-04
nodecov.GDPpercapgrowth	-3.879e-02	5.995e-03	1	$<$ 1e-04
nodecov.ImportsPct	-4.114e-02	7.569e-03	0	$<$ 1e-04
nodecov.ExportsPct	1.993e-02	7.404e-03	0	0.00714

Notably, the p-values suggest that all or essentially all of the coefficients shown in Table 5 are significantly different from zero (depending on the desired threshold for significance). Furthermore, a comparison with Table 4 shows that the coefficients of the included covariates seem to generally maintain the same size and magnitude. This is an encouraging sign, as it suggests that these parameter estimates are robust to different model specifications. The exception is the coefficient for GDP growth, which went from being large, positive, and significant to being negative (and still significant), abeit with a slightly smaller magnitude. This is cause for concern; the two model specifications show statistically significant effects of appreciable magnitude in opposite directions. While I may give preference to the model which appears to have a more stable fit, any interpretations must be cautious. Thus, while it is encouraging to see that essentially all of the parameters are statistically significant, the result for GDP growth demonstrates cause for reasonable doubt, at least with respect to that term and possibly for the other terms as well.

Taking the results at face value, Table 5 suggests that disputes involving countries with a high

GDP and a high volume of exports as a percent of GDP increase the likelihood of a network. On the other hand, disputes between countries with high levels of GDP per capita, high levels of growth in GDP and GDP per capita, and high levels of imports all make a network less likely. It is worth noting that the values of the GDP and GDP per capita variables (which are in billions of dollars and dollars, respsectively) are potentially several orders of magnitude larger than the variables for the growth terms and the variables for imports and exports as a percentage of GDP. Summaries of the terms are given in Table 6 below. With this in mind, all of the coefficients in Table 5 appear to be of appreciable magnitude.

Table 6: Summary of covariate values

	GDP	GDPgrowth	GDPpercap	GDPpercapgrowth	ImportsPct	ExportsPct
min	0.9	-0.3	553	-0.3	12.0	11.0
1st quartile	43.6	4.7	2303	4.2	26.4	24.6
median	168.3	7.4	6978	5.8	36.9	32.3
3rd quartile	429.3	8.4	25,860	7.0	55.6	53.2
max	13,388.4	15.7	63,313	14.9	172.7	194.6

It should come as little surprise that a large GDP would tend to increase the potential for a trade dispute. As was noted previously, the two most central countries to the network are the European Union and the United States, which are also the largest economies. On the other hand, the finding that countries with high levels of GDP per capita are less likely to be involved in disputes is a refutation of my previous hypothesis, which was that these countries would be more active on the international stage and more likely to end up in disputes.

The terms for GDP growth and GDP per capita growth are also unexpected, and perhaps more difficult to intrepret. They suggest that countries with high levels of growth, both per capita and in nominal terms, are less likely to be involved in disputes. However unexpected, the effect seems accurate: a list of the top countries in terms of GDP and GDP per capita growth is full of countries such as Armenia, Costa Rica, Romania, and Sri Lanka, which are involved in very few disputes - alongside China and India, who are involved in more disputes but who also have a much higher GDP.

The coefficients for the statistics derived from imports and exports as percentages of GDP suggest that disputes between countries with high levels of imports are rare; on the other hand, disputes between countries with high levels of exports increase the likelihood of a network. For these variables, it may be sensible to decompose the network statistics into two components, an indegree component and an outdegree component. For example, a country with high imports may be more likely to be accused of unfair trade restrictions, and a country with high exports may be more likely to make such an accusation. Therefore I fit another ERGM which splits each covariate into indegree and outdegree components.

The results, given in Table 7, show that this hypothesis is not correct. Disputes that accuse countries with high levels of imports significantly decrease the likelihood of a given network, as do disputes initiated by countries with high levels of imports. Similarly, ties both to and from countries with high levels of exports make a given network more likely. The model does however produce an unexpected result: the influence of GDP growth and GDP per capita growth appears to be bifurcated. High levels of GDP growth tend to increase outgoing ties and increase incoming ties; the reverse is true for GDP per capita growth. It is unclear to me how this should be interpreted.

Table 7: Decomposition into indegree and outdegree components

	Estimate	Std. Error	MCMC %	p-value
gwesp.fixed.1.3	3.375e-01	2.194e-03	2	$< 1e-04$
nodeicov.GDP	1.998e-04	2.384e-05	0	$< 1e-04$
nodeocov.GDP	1.656e-04	2.357e-05	0	$< 1e-04$
nodeicov.GDPgrowth	-2.241e-01	7.275e-03	1	$< 1e-04$
nodeocov.GDPgrowth	4.031e-02	6.685e-03	1	$< 1e-04$
nodeicov.GDPpercap	-3.382e-05	6.249e-06	0	$< 1e-04$
nodeocov.GDPpercap	-1.916e-05	5.440e-06	0	0.000434
nodeicov.GDPpercapgrowth	1.191e-01	6.208e-03	1	$< 1e-04$
nodeocov.GDPpercapgrowth	-2.011e-01	5.601e-03	1	$< 1e-04$
nodeicov.ImportsPct	-4.887e-02	1.137e-02	0	$< 1e-04$
nodeocov.ImportsPct	-3.192e-02	1.091e-02	0	0.003452
nodeicov.ExportsPct	2.105e-02	1.092e-02	0	0.053957
nodeocov.ExportsPct	1.631e-02	1.016e-02	0	0.108431

One important question left unanswered by the above analysis is the goodness of fit of the ERGM itself. It is all good and well to find factors that are statistically significant. However, if the model does a poor job of explaining the observed network, then the statistical significance loses much of its lustre. For better or for worse, there is no single measure for assessing the goodness of fit of an ERGM model. Instead, the goodness of fit is measured by simulating from the fitted model and judging the similarity between the simulated networks and the observed network. In particular, if the simulations consistently produce networks that are substantially different from the observed network, then we should be concerned that the model fit is poor.

Figure 5 shows a plot of the observed network along with 8 sampled networks from the distribution specified by the model in Table 7. The simulated networks seem reasonably similar to the observed network. In general they seem a little more dense than the observed network; that is, the nodes tend to be more strongly interconnected than in the observed network. The distributions of the indegrees and outdegrees, edgewise shared partners, and minimum distances between nodes, shown in Figure 6, also match the observed network relatively well. This provides a reasonable degree of confidence in the model estimates, suggesting that the included factors likely do play a significant role in determining the structure of the WTO network.

Figure 5: Simulations from the ERGM

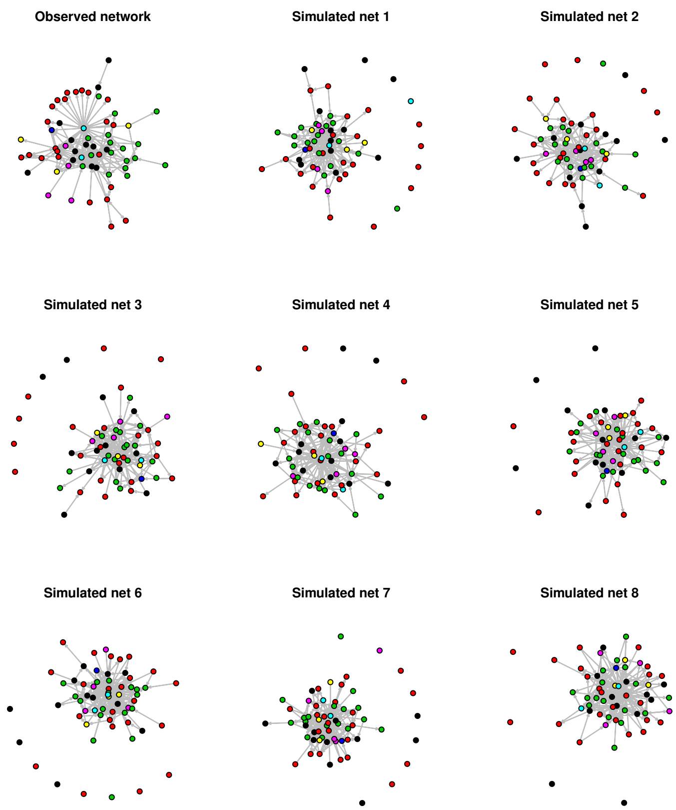
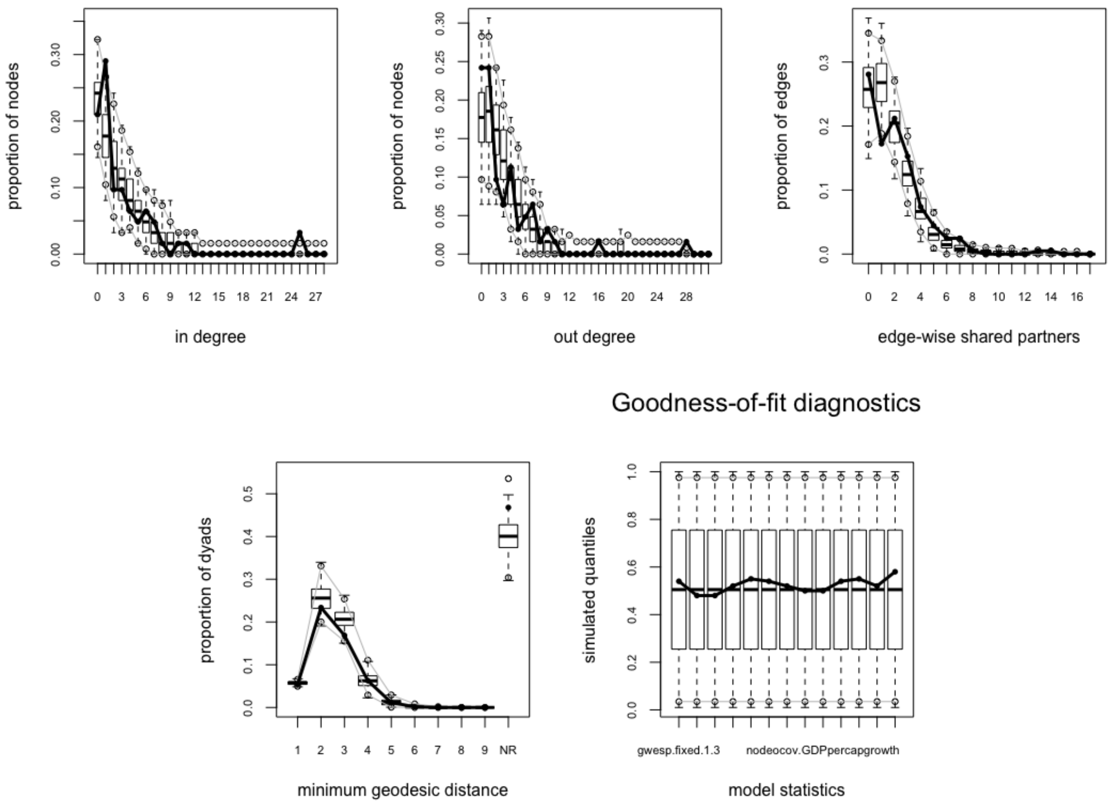


Figure 6: ERGM goodness of fit plots



One thing that these models ignore, however, is the fact that the edges in the WTO trade network are weighted. Recall that amongst the 501 disputes from 1995 to 2015, there are only 203 unique edges; any given country may file multiple disputes against another country. By ignoring these repeats, as I did above, I am to some extent losing the full richness of the data. To account for this, I fit a valued ERGM, which follows roughly the same framework as the ERGMs described above, with the main distinction that it fits the number of disputes between countries, and not merely the existence of disputes. This requires specifying a reference distribution for the network; I choose a Poisson distribution, which matches the property of my network that larger edge weights are less common. A geometric distribution would likely be a better match for this property; however, fitting the WTO network to a geometric distribution proved to be numerically unstable.

The results of my valued ERGM are shown in Table 8. Comparing the results to those found using the regular ERGM models shows a large degree of consistency. The coefficients are all statistically significant, and with only one exception, the signs and magnitures are similar to those found previously. The exception is GDP growth, which now has a positive coefficient, compared to the negative coefficient in Table 5. This result is somewhat disconcerting. However, GDP growth was already shown to be problematic; recall that the coefficient changed significantly between Tables 4 and 5. Therefore, fitting this model did not appreciably change the interpretations of my previous results.

Table 8: Coefficients of the valued ERGM

	Estimate	Std. Error	MCMC %	p-value
nodecov.sum..GDP	2.626e-04	7.603e-06	0	<1e-04
nodecov.sum..GDPgrowth	7.720e-02	4.855e-03	2	<1e-04
nodecov.sum..GDPpercap	-1.792e-05	2.528e-06	0	<1e-04
nodecov.sum..GDPpercapgrowth	-1.296e-01	4.283e-03	4	<1e-04
nodecov.sum..ImportsPct	-6.210e-02	6.534e-04	2	<1e-04
nodecov.sum..ExportsPct	2.672e-02	1.847e-03	0	<1e-04

MCMC diagnostics for the valued ERGM are shown in Figure 7. Figures 8 and 9 show simulations from the valued ERGM and a summary of the indegree and outdegree distributions across these samples. The dots in Figure 9 represent the values in the observed network. The MCMC diagnostics show that the fit appears to be stable; the sampled statistics remain close to the value observed in the actual network, and the variance appears fairly constant. The simulations suggest that the model likely does a poor job modeling the transitivity in the network. The simulated networks appear to be quite densely packed compared to the observed network. The indegree and outdegree plots show that the model does an acceptable, but not stellar, job of representing the network. The observed network generally lies within the range of simulated values, although in some cases it is quite close to the margin, or just outside it. With more time, it likely would not be difficult to find a valued ERGM model with a better fit; however, this model does lend some credibility to the results, due to its consistency with the unvalued ERGM.

Figure 7: MCMC diagnostics for the valued ERGM

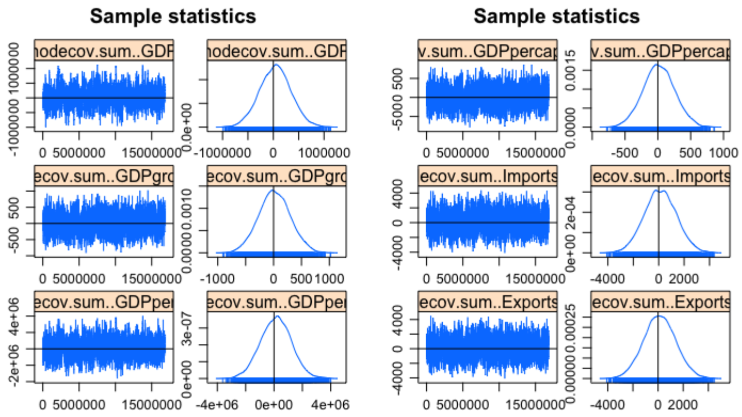


Figure 8: Simulations from the valued ERGM

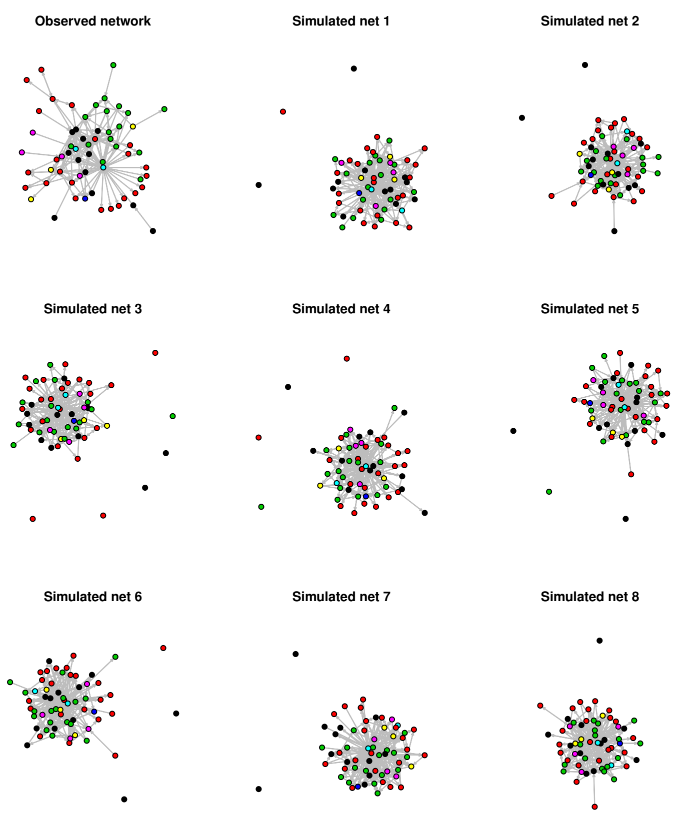
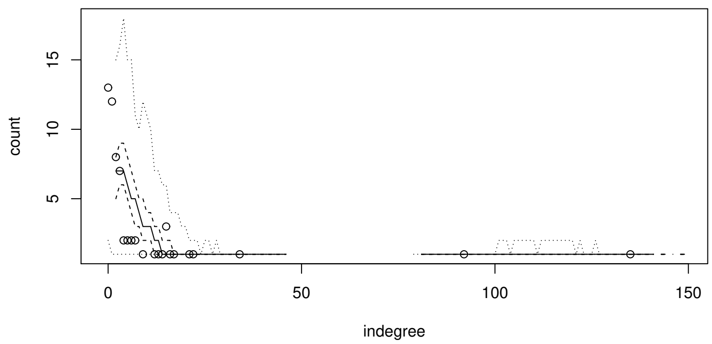
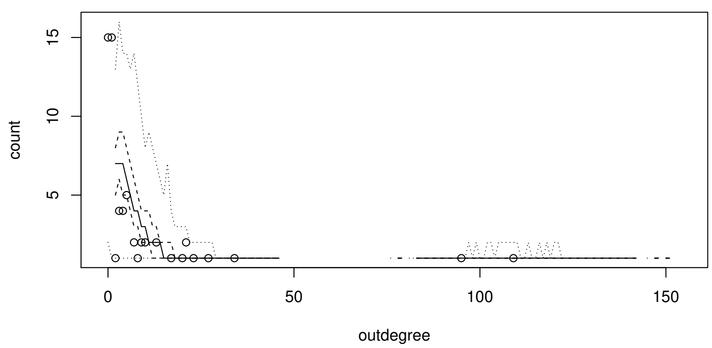


Figure 9: Valued ERGM goodness of fit plots





Discussion

The goodness of fit metrics in the previous section show that the ERGM models provide an imperfect but reasonably good fit for the network of WTO trade disputes. However, a number of caveats are warranted. First, the models were run on a subset of the full network, including only those countries which were involved in disputes, rather than the full set of WTO member countries. This means that the ERGM models assign a probability of 0 to any disputes involving a country that was not involved in any disputes in the observed data. This is an incorrect assumption; in 2017, for example, numerous disputes were filed by Qatar against other Arab nations, none of which are featured in the observed network from 1995 to 2015. While this effectivly demonstrates that the model is wrong, it does not necessarily mean that the model is bad. However some caution is warranted.

Additionally, there is also cause for concern due to the fact that the coefficient for GDP growth changed sign depending on the model specification. The other coefficients seemed relatively robust, but this is evidence that the fit might nonetheless be unstable and the estimates should only be regarded as suggestive, and not as the whole truth. Further study may be necessary to lend more confidence to the results.

There are a number of ways in which this analysis could be extended. For example, a key feature of the network which was ignored by the models in this analysis is the temporal nature of the ties. The disputes are filed over time, and the network could potentially be modeled as such. The nodal covariates such as GDP are constantly changing over time; it would seem more sensible to use covariates that are contemporaneous with the filing of the disputes, rather than using averages over the entire period as I did in this analysis (as explained in the Appendix). Furthermore, analyzing the state of the network over time could allow for the testing of a retaliatory effect, as countries that are targeted in one time period file disputes against their accusers in the next.

One last point which is relevant for all of these models is that the data which I used are by no means the best of all possible worlds. In particular, many of the variables serve, at least to some extent, as proxy variables for characterizing the trade relationship between pairs of countries. If possible, it would be better to include more detailed data about the volume of trade itself between each pair of countries. This could allow for a much more detailed analysis. For example, if the data were available, it could be of interest to examine whether countries whose main export is agricultural products would differ from countries whose main export is manufactured goods. This and many other particular questions could be answered about the effects of certain kinds of trade relationships between countries on the number of disputes. Therefore there remain many potential improvements to the analysis described in this paper.

References

[1] Kaggle. *World Development Indicators.* https://www.kaggle.com/worldbank/world-development-indicators. Updated June 2017. Accessed December 3, 2017.

[2] World Bank. *Countries and Economies.* https://data.worldbank.org/country. Accessed December 3, 2017.

[3] World Trade Organization, a. *Chronological List of Disputes Cases.* https://www.wto.org/english/tratop_e/dispu_e/dispu_status_e.htm. Accessed November 17, 2017.

[4] World Trade Organization, b. *Regional Trade Agreements Database.* http://rtais.wto.org/UI/PublicAllRTAList.aspx. Accessed December 5, 2017.

[5] World Trade Organization, c. *The WTO.* https://www.wto.org/english/thewto_e/thewto_e.htm. Accessed December 3, 2017.

Appendix - A note regarding the World Development Indicators

The covariates that I use in this analysis come from the World Development Indicators dataset, collected by the World Bank and provided by Kaggle. The dataset includes a large number of indicators, over varying time periods, at an annual frequency. I filtered this large dataset down to the relevant set of countries (WTO members) and picked out indicators which were both available for most countries in most time periods, and which seemed particularly relevant to international trade. Nonetheless, some values remained missing. Moreover, because the network which I examine is the collection of all trade disputes filed with the WTO from 1995 to 2015, the issue arises of how to use the annual-frequency data for this larger time period. I used the following procedure:

- For GDP, dollar-valued exports, and dollar-valued imports, for each country I took the average of all available data from 1995 through 2015. I then rescaled the data to be in units of $1 billion.
- For GDP per capita and exports and imports as a percentage of GDP, for each country I took the average of all available data from 1995 through 2015.
- For GDP growth and GDP per capita growth, for each country I took the first observation and the last observation of the dollar-valued variable (GDP and GDP per capita) within the range of 1995 through 2015. Using these two values, and the distance (in years) between them, I calculated the geometric average growth rate, and then converted the number to a percent. For example, if GDP data were available for 1995 and 2015, I calculated the GDP growth as $\left[\left(\frac{\text{GDP}_{2015}}{\text{GDP}_{1995}}\right)^{\frac{1}{20}} - 1\right] * 100$.