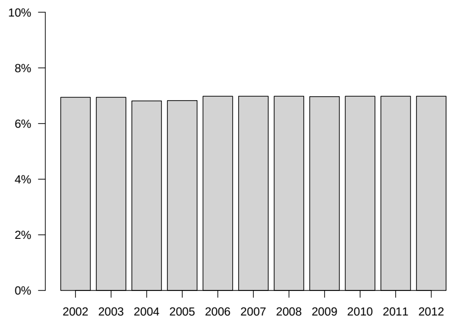
Supporting Information for:
Complex Dependence in Foreign Direct Investment: Network Theory and Empirical Analysis

Abstract

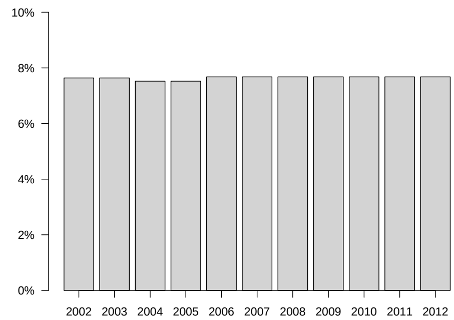
We develop theory that accounts for complex dependence in foreign direct investment (FDI) relationships. Conventional theories of FDI focus on firm-, country-, or dyad-level characteristics to account for cross-border capital movements. Yet, today's globalization is characterized by the increasing fragmentation and dispersion of production processes, which gives rise to complex dependence among production relationships. Consequently, FDI flows should be represented and theorized as a network. Specifically, we argue that FDI flows are reciprocal and transitive. We test these hypotheses along with conventional covariate determinants of FDI using an exponential random graph model (ERGM) for weighted networks. We find that FDI networks exhibit both reciprocity and transitivity. Our network approach to studying FDI provides new insights into global investment flows and their political and economic consequences. In addition to our substantive findings, we offer a broad methodological contribution by introducing the ERGM for count-weighted networks in political science. (150 words)

A Summary Statistics

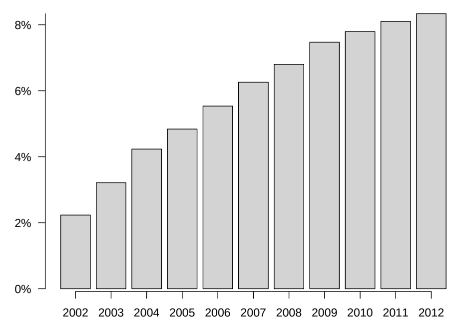
Alliance Treaty



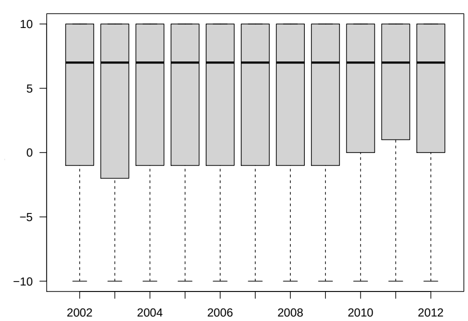
Defense Treaty



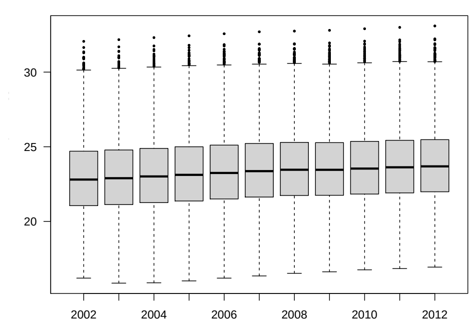
Bilateral Investment Treaty



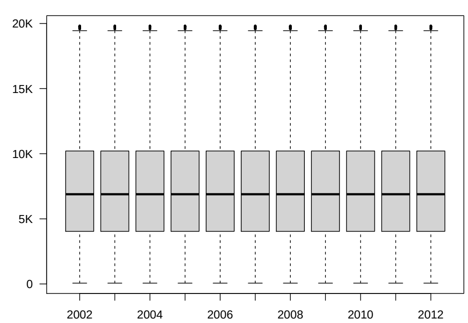
Polity



Dyad GDP Product



Distance



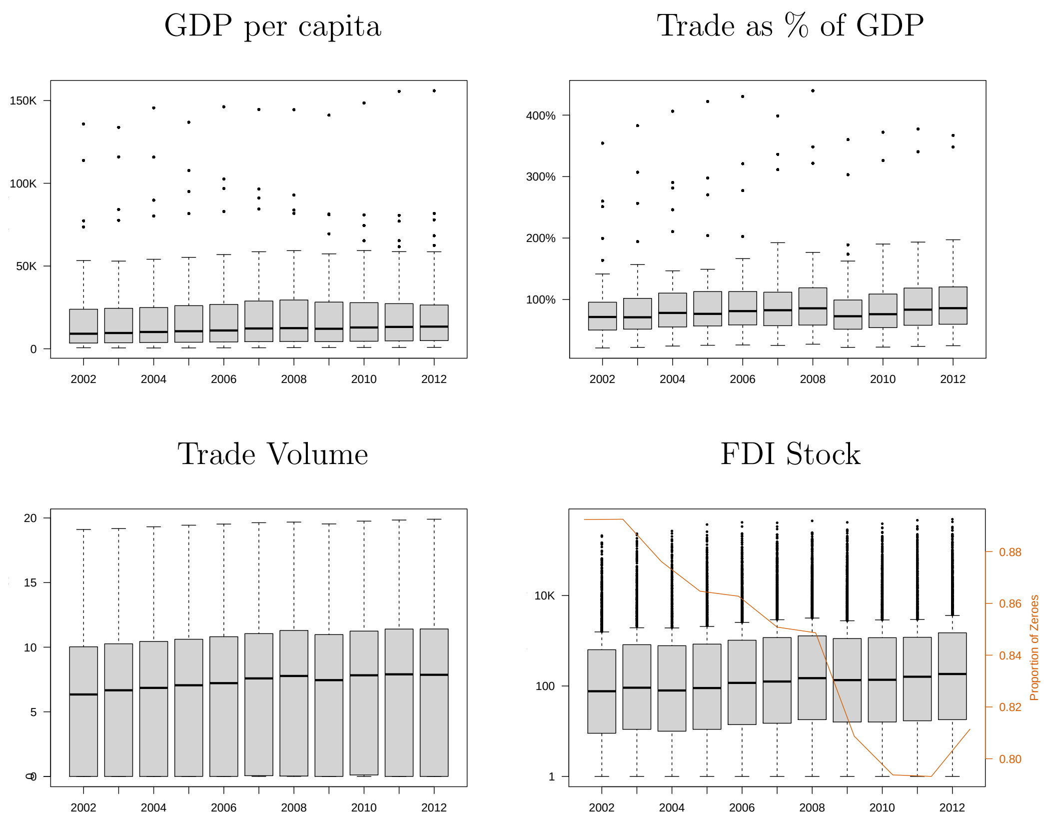


Table A: Summary Statistics.

Table B: Correlation Matrix

	Mass	Distance	Polity
Mass	1	-0.003	0.091
Distance (logged)	-0.003	1	0.008
Polity	0.091	0.008	1
Trade Openness	-0.166	-0.057	-0.078
BITs	0.141	-0.085	0.018
Trade Volume	0.714	-0.215	0.215
GDP per capita (logged)	0.392	-0.084	0.166
Alliance Treaty	0.133	-0.348	0.073
Defense Treaty	0.065	-0.391	0.065

	Trade Openness	BITs	Trade Volume
Mass	-0.166	0.141	0.714
Distance (logged)	-0.057	-0.085	-0.215
Polity	-0.078	0.018	0.215
Trade Openness	1	0.032	-0.055
BITs	0.032	1	0.143
Trade Volume	-0.055	0.143	1
GDP per capita (logged)	0.225	0.093	0.330
Alliance Treaty	-0.044	0.021	0.216
Defense Treaty	-0.046	0.010	0.177

	GDP per capita	Alliance Treaty	Defense Treaty
Mass	0.392	0.133	0.065
Distance	-0.084	-0.348	-0.391
Polity	0.166	0.073	0.065
Trade Openness	0.225	-0.044	-0.046
BITs	0.093	0.021	0.010
Trade Volume	0.330	0.216	0.177
GDP per capita (logged)	1	0.098	0.038
Alliance Treaty	0.098	1	0.850
Defense Treaty	0.038	0.850	1

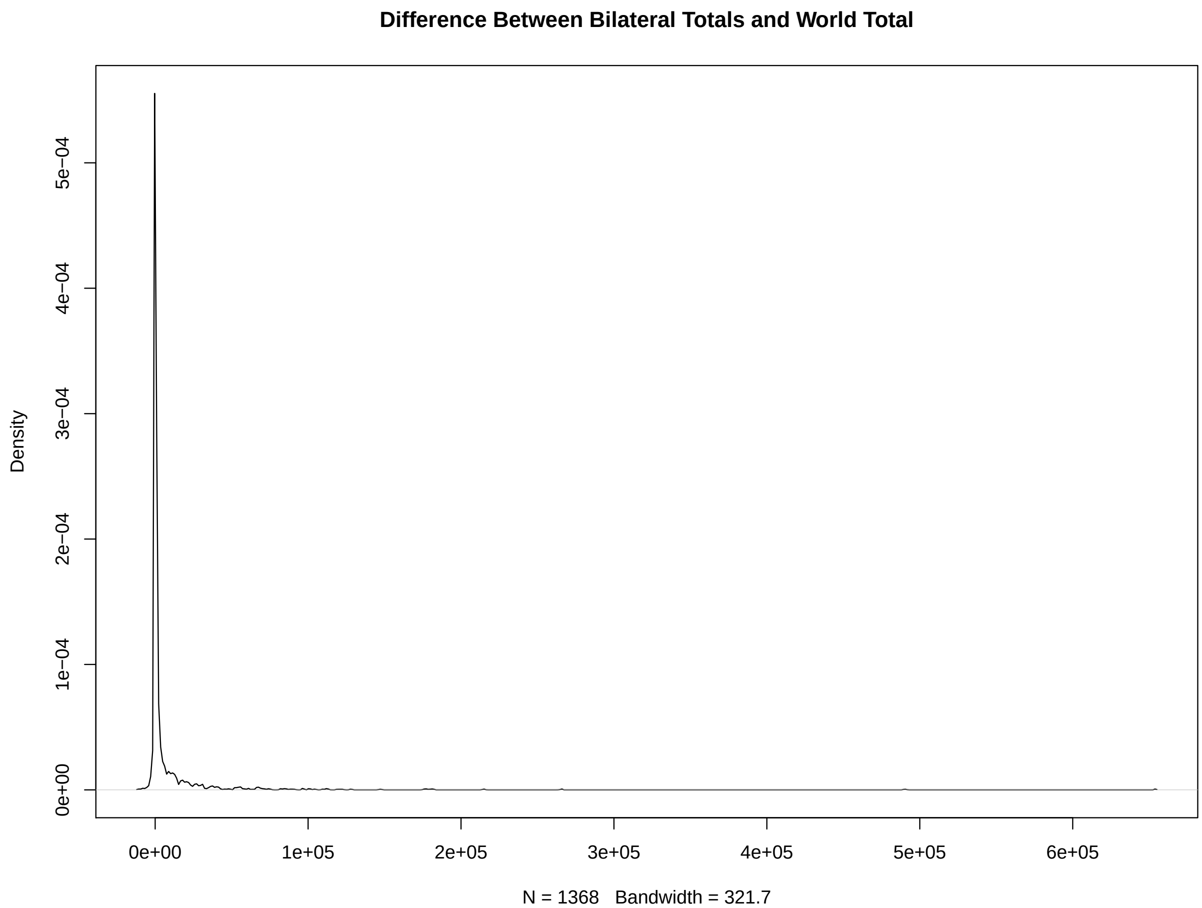


Figure A: Density Plot of the Difference between Total FDI stocks and Summing Bilateral FDI stocks.

B Time-Pooled Model Results

For robustness checks, we re-estimate the count ERGM by pooling the data, which is common in the literature for regression based models. Figure B shows that after pooling, network terms remain positive and statistically significant, supporting our hypothesis that reciprocity and transitivity characterize FDI flows. The exogenous covariates from the pooled model are presented in Table C. The estimates are similar to yearly results in terms of direction and statistical significance.

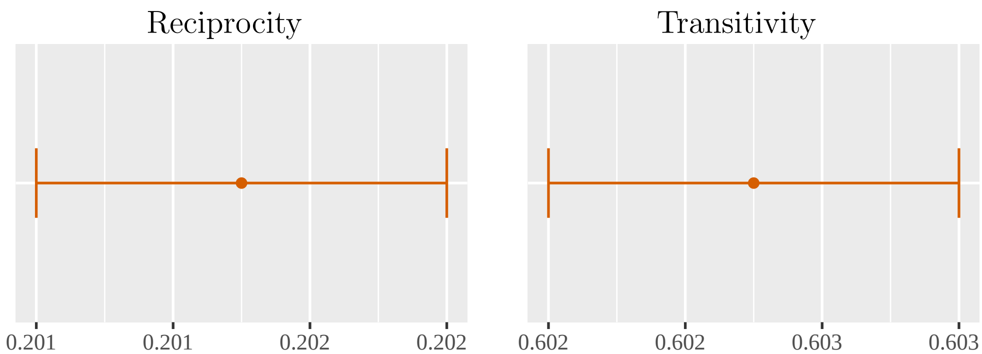
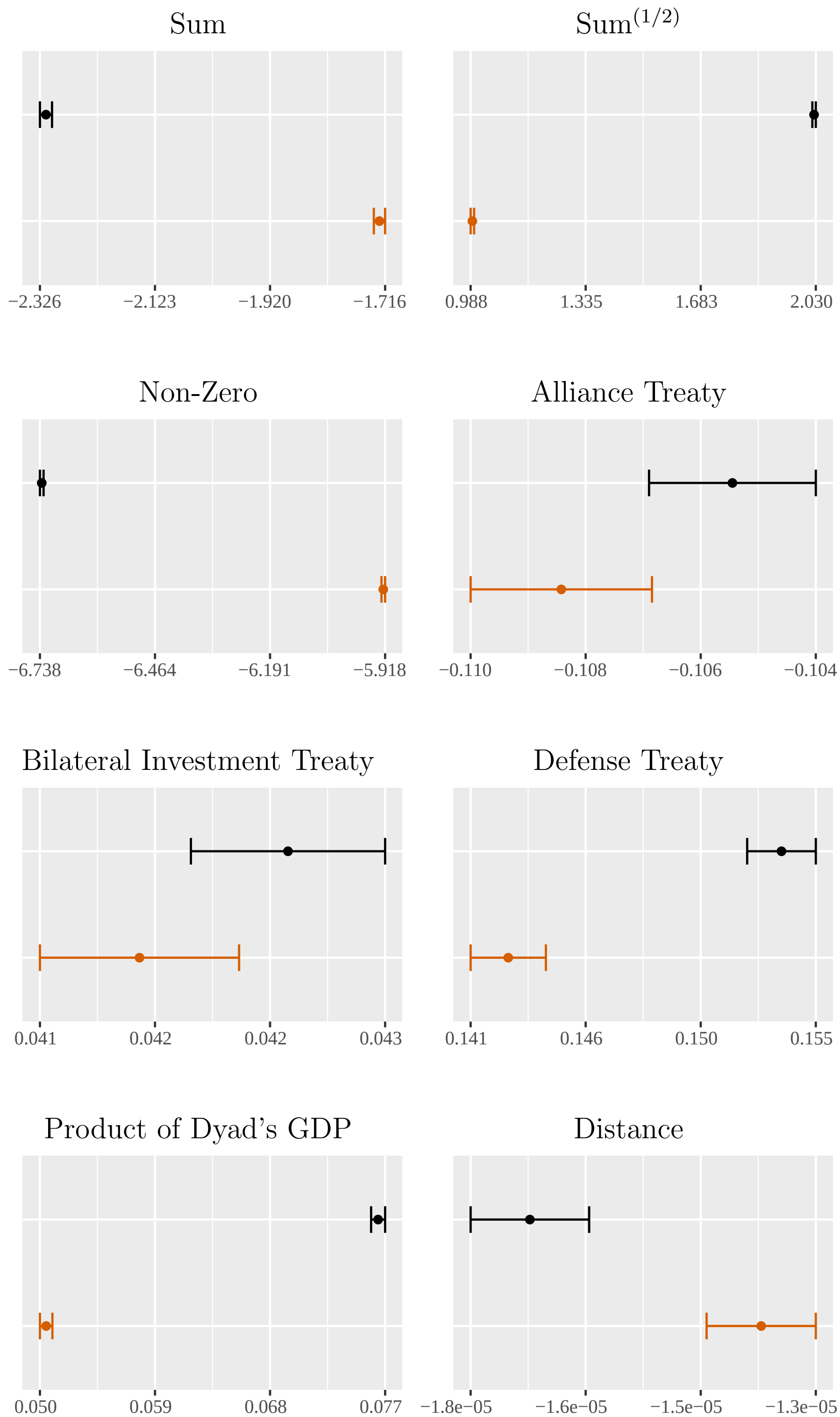


Figure B: Estimates of Dependence terms in time-pooled ERGMs. Bars span 95% confidence intervals.

The results also show that ignoring network structure lead to biased estimates in several covariates. We see significant differences in the coefficients for distance, the product of dyad's GDP, the three treaty variables, as well as origin and destination's GDP per capita, Polity, and trade openness. These findings are consistent with those from the yearly models. It illustrates that failure to include network structure results in biased estimates.


Sum
−2.326
−2.123
−1.920
−1.716
Sum^(1/2)
0.988
1.335
1.683
2.030
Non-Zero
−6.738
−6.464
−6.191
−5.918
Alliance Treaty
−0.110
−0.108
−0.106
−0.104
Bilateral Investment Treaty
0.041
0.042
0.042
0.043
Defense Treaty
0.141
0.146
0.150
0.155
Product of Dyad's GDP
0.050
0.059
0.068
0.077
Distance
−1.8e−05
−1.6e−05
−1.5e−05
−1.3e−05

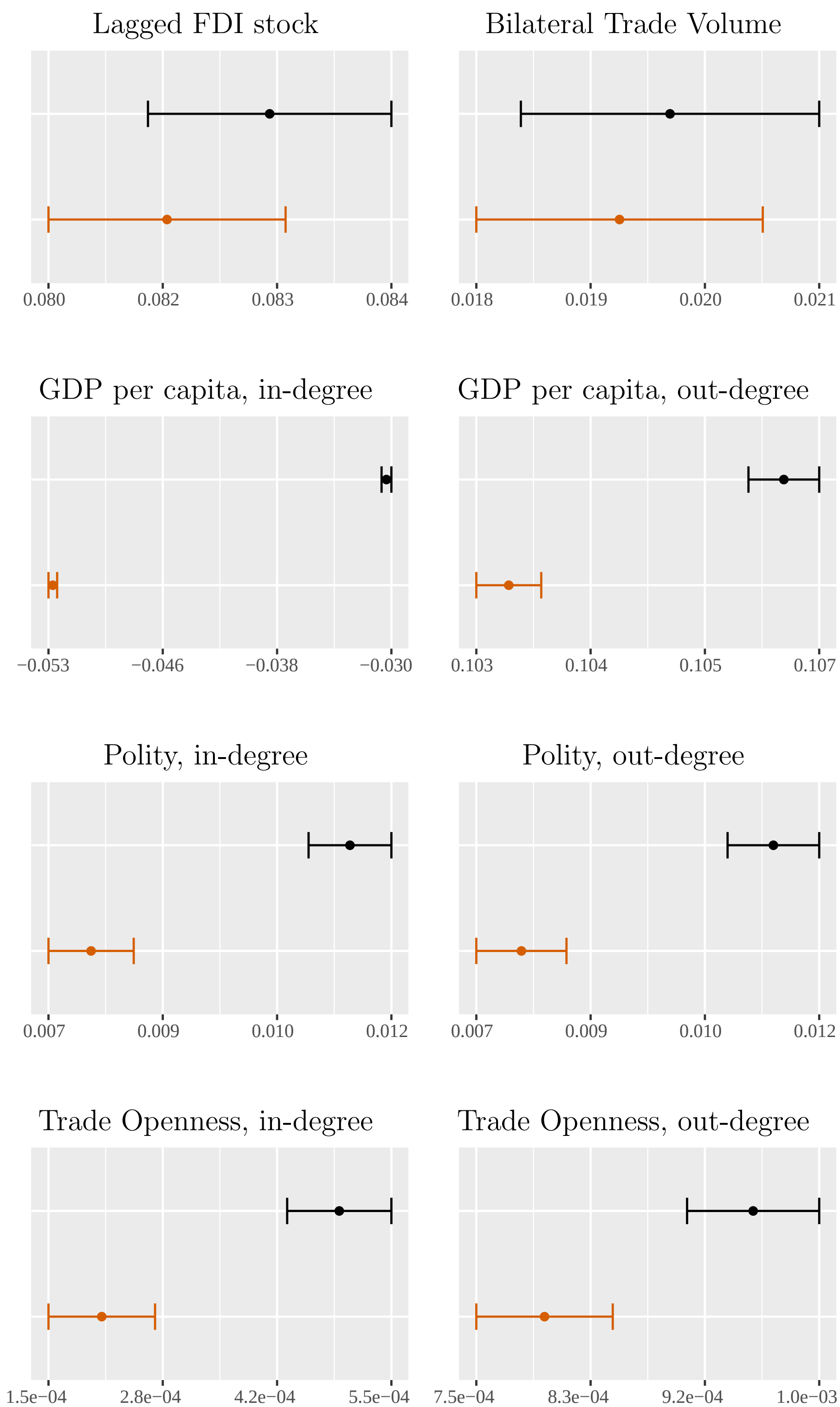


Table C: Estimates of terms in time-pooled ERGMs. Bars span 95% confidence intervals. Black coefficient representations are from models excluding dependence terms (i.e., transitivity and reciprocity).

C Subset by Missingness Results

In the paper, we imputed missing values with zeros. In this section, we check whether our results are robust if we analyze a subset of the data set based on the level of missingness. To subset the data, we approximate total level of missingness in the adjacency matrices (q) by using the proportion of missing values for each node (p). We conduct two robustness checks: (1) when $p = 0.86$, $q \approx 0.50$ and $n = 70$; and (2) when $p = 0.72$, $q \approx 0.25$ and $n = 28$. In the first case, we only include nodes with missing values that are 86% or less of the possible edges for the entire data set, which leaves us with an adjacency matrix that is only missing 50% of the values (70 countries in total). Similarly, the second set only includes nodes with missing values that are 72% or less of the possible edges for the entire dataset, which leaves us with an adjacency matrix that is only missing 25% of the values (28 countries in total). Following our approach in the paper, we impute missing values in the two subsets of the data with zeros.

Figures C and D present the results for the two robustness checks, respectively. We see that FDI networks show strong transitivity for all years, but reciprocity effects become weak and insignificant in some years. This may be because most nodes (i.e. states) in the subsets are developed countries that have substantial two-way FDI flows between them and thus there is little variation in the level of reciprocity.

C.1 q ≈ 0.50

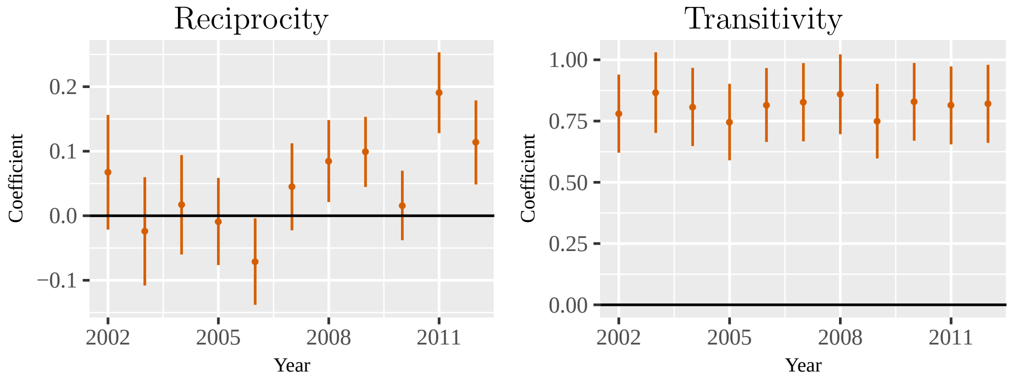


Figure C: Estimates of Dependence terms. Bars span 95% confidence intervals.

C.2 q ≈ 0.25

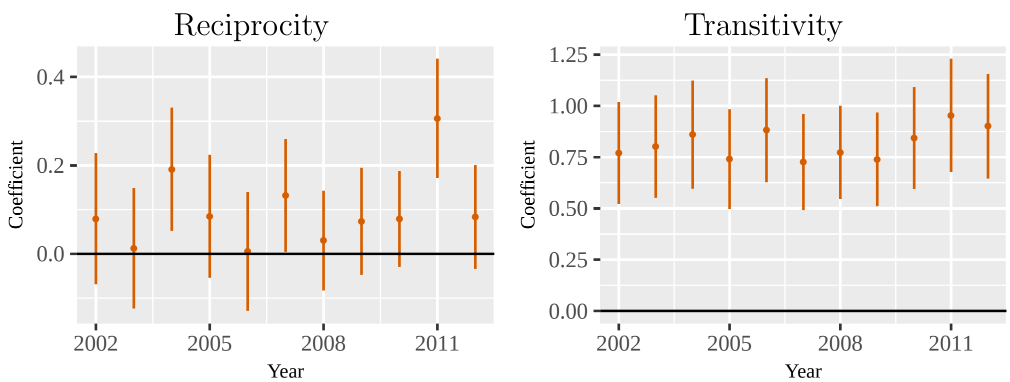


Figure D: Estimates of Dependence terms. Bars span 95% confidence intervals.

D Multiple Imputations with Amelia Results

In this section, we utilize the R package Amelia to impute the missing values in the full data set, when $q \approx 0.50$, and when $q \approx 0.25$ (Honaker, King, Blackwell et al., 2011; King, Honaker, Joseph and Scheve, 2002). Figures E and F show the results. We see that transitivity effects are significant in all years and reciprocity effects are also significant in most years.

Together with the results in Sections C and D, these results give us confidence that our findings regarding the reciprocity and transitivity of FDI are not a result of the pattern of missingness in the data set.

D.1 Full

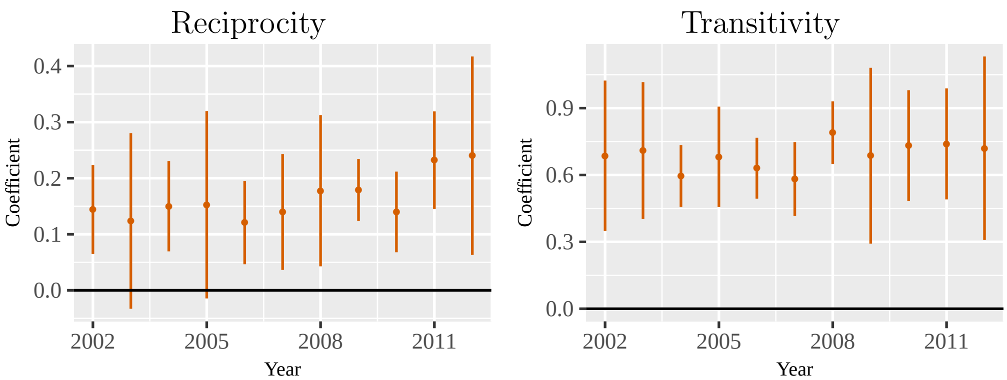


Figure E: Estimates of Dependence terms. Bars span 95% confidence intervals.

D.2 q ≈ 0.50

Figure F: Estimates of Dependence terms. Bars span 95% confidence intervals.

D.3 q ≈ 0.25

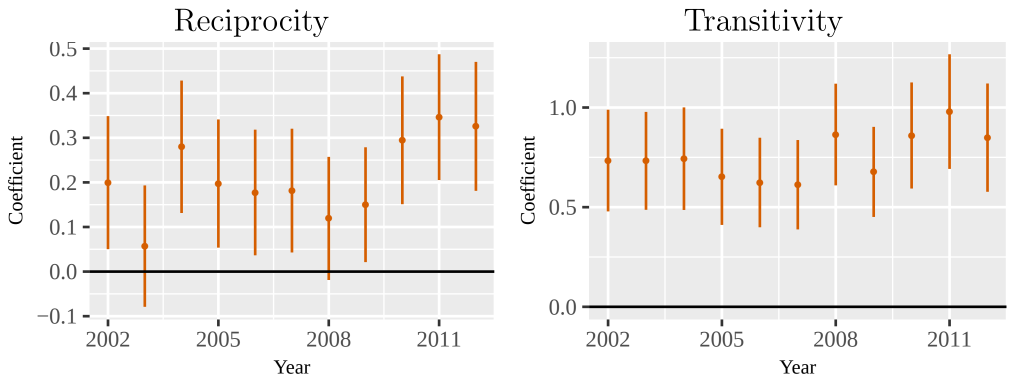


Figure G: Estimates of Dependence terms. Bars span 95% confidence intervals.

E Independent Covariate Interpretation

For the Poisson-reference ERGM these covariate estimates are usually interpreted by exponentiating Euler's constant to the power of the coefficient times the number of unit changes in the covariate to get the expected change in the tie weight (Krivitsky and Butts, 2013). Taking Polity, in-degree for example, if the FDI destination had a Polity score of 10 in 2002, we would expect the value of logged FDI being sent to be 1.27 times more than a destination that had a Polity score of -10. In the model with network terms this expected increase is only 1.17 times higher. Another method for interpreting independent terms in the model is to simulate networks using the estimated coefficients while fixing all other independent terms at the mean value and then comparing changes in the average edge value to the range of values of the covariate. For node-level covariates the method is similar. The only change is that you use the column (receiver) averages in the adjacency matrix to compare against covariate values. For illustrative purpose, we present a plot of this for Polity, in-degree below in Figure H. Here the plot shows that as the destination state's Polity increases from the score -10 to the score of 6 there is a slow increase in the average level of FDI received, with a sharper increase between Polity Scores 7 and 10.

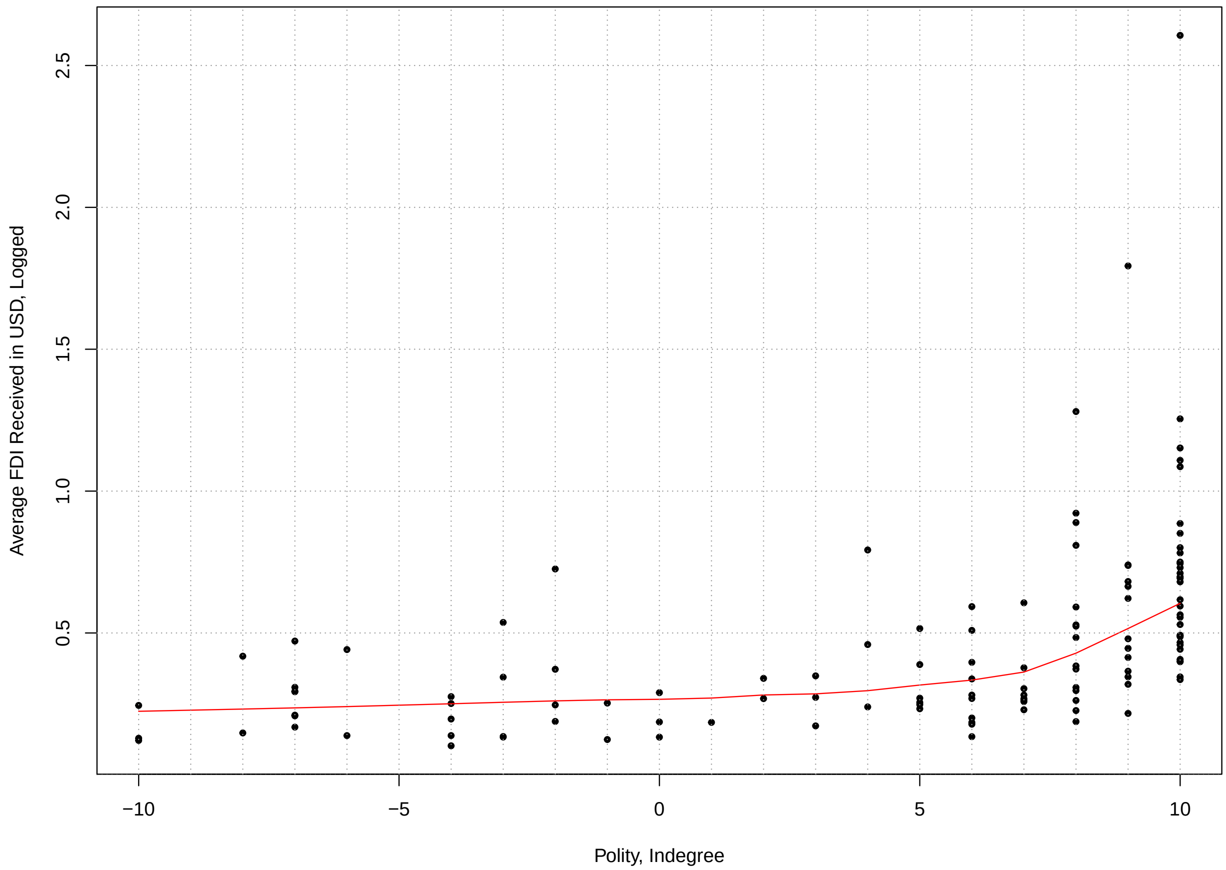


Figure H: Results from the simulation exercise investigating the effects of Polity, in-degree for year 2002. The points are calculated as the the receiver node averages in the adjacency matrix. These results are from 500 simulations. The line in red is the Loess curve.

References

Honaker, James, Gary King, Matthew Blackwell et al. 2011. "Amelia II: A program for missing data." *Journal of statistical software* 45(7):1–47.

King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2002. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95(1):49–69.

Krivitsky, Pavel N and Carter T Butts. 2013. "Modeling valued networks with statnet." *The Statnet Development Team* p. 2013.