Multi-Relational International Trade Networks, 1965-2000

Matthew C. Mahutga
Department of Sociology
University of California, Riverside

1. Overview

Beginning with the seminal work of Snyder and Kick (1979), social network analysts have utilized trade relations to quantify the social structure of the world-economic system (e.g. Breiger, 1981; Clark, 2010; Clark & Beckfield, 2008; Kim & Shin, 2001; Mahutga, 2006; Mahutga & Smith, 2011; Nemeth & Smith, 1985; Smith & Nemeth, 1988; Smith & White, 1992). Some of this work utilizes data from the International Monetary Fund's Direction of Trade Statistics (IMF 2012). These data provide total trade for each country dyad, and cover a large number of dyads over a relatively long period of time (1950 to the present) (see Lloyd et al., 2009 for a review). One advantage to the Direction of Trade Statistics is that they are easy to work with because there are only N(N-1) data points in each year, and there is little year on year missing data because it is relatively easy for state agencies to accurately record total imports and exports with each of their partners. While there is much to learn from these data, total trade masks a significant amount of inter-industry variation in the structure of international trade (e.g. Hidalgo et al., 2007).

The other major data source is the United Nations Commodity Trade Database (UNCOMTRADE), which covers a large number of country dyads over a relatively long period of time (1962 to the present) (United Nations, 2012). A major advantage to the UNCOMTRADE data is that it disaggregates dyadic trade flows into industry and sub-industry categories, and thereby allows users to analyze inter-industry variation in the structure of international trade. These advantages present two unique challenges. First, while both sources are publically available, users face a more complex task to collect, clean and organize the UNCOMTRADE data insofar as the data points scale with both N and the number of relations (each year now has N(N-1)R data points, where R is the number of relations). And, the detailed classificatory schemes provided by the UN make it much more difficult for (especially poor) state agencies to comply with the reporting requirements consistently from year to year, resulting in a substantial amount of missing data on a year-on-year basis.

The data recorded here overcome some of the obstacles to employing UNCOMTRADE data because they record dyadic trade among a constant set of 94 countries that together account for 96 to 99 % of world trade, cover multiple commodity relations and span a relatively long period of time. In particular, the trade matrices contain ordered dyadic trade flows reported in three time points (1965, 1980, and 2000). The 15 particular industries covered represent the 5 distinct commodity clusters identified by Smith and Nemeth (1985). The 45 matrices include a constant node-set of 94 countries in each year. Moreover, the data include roughly 33 percent more cases than the UNCOMTRADE database records for the three specific years, owing to a set of procedures that allowed me to infer missing trade between reporting and non-reporting countries and between non-reporting countries.

These data provide network analysts a rare opportunity to apply network methods to multi-relational international trade networks. These trade data are unique relative to other publicly available data insofar as they cover multiple commodity-trade relations, three time points spanning thirty five years, and include a large sample of countries representing all world-regions and levels of development. The data described in this article should facilitate the wider usage of multi-relational commodity-trade data because they require minimal processing prior to analysis, which has probably been the single largest obstacle to their usage thus far. In what follows I describe the industries covered, how the data were collected and reported, and discuss the procedures I followed to gather missing trade flows.

2. Data Collection

2.1 Industries

The UNCOMTRADE data base has nine industrial classification systems for categorizing the types of goods traded between countries. The data described here are classified according to the Standard International Trade Classification (SITC) Rev. 1. While each of the nine classificatory schemes has its own advantages, the major advantage to SITC Rev. 1 is that it extends back to the first year that the UN began collecting data. Contrarily, the newer alternative schemes cover fewer years because commodities cannot be categorized "backward" in time once new schemes are developed.

The SITC Rev. 1 system classifies commodities under a multi-digit scheme that varies from total trade to hundreds of unique five-digit codes. Shorter digits imply a higher level of aggregation. For example, the one-digit code "7" is "Machinery and Transport equipment", which subdivides into three unique two-digit codes: "71" is "Machinery, other than electric", "72" is "Electrical machinery, apparatus and appliances", and "73" is "Transport Equipment." Each of these two-digit codes subdivides into greater specificity. For example, "72996" is "Electrical carbons."

The data reported here were collected at the two-digit level, which capture recognizable industries. And, while the two-digit level certainly misses some of the finer disaggregation communicated by the longer codes, many countries do not report beyond the two-digit level for reasons related to administrative and resource burdens, or to preserve national secrets. Thus, the sum of smaller digit reports rarely equal the volume of trade reported at the two-digit level, but the sum of the two-digit level flows equals the one-digit level.

Table 1. UN Commodity Categories Classified in Relational Categories from Smith and Nemeth (1988).

1) High Tech/Heavy Manufacturing
58) Plastic Materials, Regenerated Cellulose and Artificial Resins
69) Manufactures of Metal
71) Machinery – nonelectrical
2) Sophisticated Extractive
25) Pulp and waste paper
34) Gas, natural and manufactured
64) Paper, paperboard, and manufactures thereof
3) Simple Extractive
04) Cereal and cereal preparations
22) Oil seeds, oil nuts and oil kernels
41) Animal oils and fats
4) Low Wage/Light Manufactures
83) Travel bags, handbags, and similar containers
84) Clothing
85) Footwear
5) Animal Products and Byproducts
01) Meat and meat preparations
02) Dairy products and bird's eggs
29) Crude animal and vegetable materials

There are fifty five two-digit codes in the SITC Rev. 1 system (United Nations 1963). Table 1 reports the fifteen industries covered by the data described in this article. The industries in Table 1 were selected on the basis of Smith and Nemeth's (1988) factor analysis, and were analyzed in part by Smith and White (1992), Mahutga (2006), and in full by Mahutga and Smith (2011). Smith and Nemeth's factor analysis reduced the 55 two-digit matrices to five unique factors within which commodity matrices were highly correlated. Substantively, Smith and Nemeth's factor analysis implied that the fifty five two-digit codes reduced to five broad categories, within which individual commodities were more or less interchangeable. A quick scan of the commodity clusters provides some intuition to their analysis. For example, the matrices for commodity codes 01 ("Meat and meat preparations"), 02 ("Dairy products and bird's eggs"), and 29 ("Crude animal and vegetable materials") were among a group of highly correlated trade matrices that clustered on a factor that Smith and Nemeth labeled "Animal Products and Byproducts". Clearly, countries that for whatever reason—climate, geography, factor abundance, etc.—excel at the production and export (or conversely, do not excel and therefore import) of one type of animal product and byproduct, also excel at others.

2.2 Imports, Exports and Units of Measurement

In order to compile UNCOMTRADE data, the UN asks countries to report both their exports to and imports from each other country, which makes it possible to rely on either reported imports or reported exports to assemble a trade matrix. Exports and imports are very highly but imperfectly correlated. For example, the correlation of the vector of the US's reported exports to its partners with the vector of the US' reported imports from its partners will approach 1, but the value of the US' reported import from Mexico on any given relation may not correspond exactly to the value of Mexico's reported export to the US on the same relation . However, reported imports tend to be more accurate because of the care taken by state agencies to record imports precisely for the purpose of tariffs (Durand 1953). In general, I therefore rely on reported imports to assemble the trade matrices here. Thus, the vast majority of the cell entries in each N x N commodity matrix represent country j's reported imports from country I, except as noted below. The dyadic trade flows in these matrices record the dollar amount of the given commodity group in thousands of current (i.e. not adjusted for inflation) US dollars.

2.3 Sample Selection and Missing Data

Dyadic trade flows on each of the fifteen commodity groups described above were collected for a constant panel of 94 countries in 1965, 1980 and 2000. The countries are reported in Table 2. However, only 63 of the 94 countries detailed in Table 2 reported trade (either imports or exports) in each of the three years. In order to increase the coverage above 63, I sampled as follows. I first included any country that reported in each year. I then included any country that reported trade flows in at least two of the three time periods, and used the following strategy to fill in missing data for each country that did not report in one of the years. I began by following StatCanada in utilizing "mirror flows" (i.e. reported exports to missing countries from non-missing countries), which left systematically missing data for the possible trade ties between non-reporting countries. In order to fill in the flows between countries that did not report in a given year, I used reported imports from a tem-

Table 2. Countries and Respective UN Codes.

UN Code	Country Name	UN Code	Country Name
012	Algeria	388	Jamaica
024	Angola	392	Japan
032	Argentina	400	Jordan
036	Australia	410	South Korea
040	Austria	414	Kuwait
048	Bahrain	434	Libya
052	Barbados	450	Madagascar
058	Belgium	454	Malawi
068	Bolivia	458	Malaysia
076	Brazil	466	Mali
096	Brunei Darussalam	470	Malta
120	Cameroon	480	Mauritius
124	Canada	484	Mexico
140	Central African Republic	504	Morocco
144	Sri Lanka	528	Netherlands
148	Chad	554	New Zealand
152	Chile	558	Nicaragua
156	China	562	Niger
170	Colombia	566	Nigeria
178	Congo	579	Norway
188	Costa Rica	586	Pakistan
196	Cyprus	590	Panama
200	Czechoslovakia	600	Paraguay
204	Benin	604	Peru
208	Denmark	608	Philippines
218	Ecuador	620	Portugal
222	El Salvador	634	Qatar
230	Ethiopia	682	Saudi Arabia
246	Finland	686	Senegal
251	France	702	Singapore
266	Gabon	724	Spain
270	Gambia	752	Sweden
276	Germany	757	Switzerland
288	Ghana	764	Thailand
300	Greece	768	Togo
320	Guatemala	780	Trinidad/Tobago
340	Honduras	788	Tunisia
344	Hong Kong	792	Turkey
348	Hungary	818	Egypt
352	Iceland	826	UK
356	India	841	USA
360	Indonesia	854	Burkina Faso
364	Iran	858	Uruguay
372	Ireland	862	Venezuela
376	Israel	882	Samoa
381	Italy	891	Yugoslavia
384	Côte d'Ivoire	894	Zambia

porally proximate year. The 31 countries for which I filled in missing data in this way are as follows:

- 1965: Algeria, Angola, Bahrain, Barbados, Czechoslovakia, China, Ethiopia, Gambia, Indonesia, Jamaica, Kuwait, Malawi, Mauritius, Poland, Qatar, Saudi Arabia, Trinidad/Tobago, Uruguay.
- 1980: Chad, Côte d'Ivoire, Iran, Nigeria, Romania, Zambia.
- 2000: Benin, Brunei, Cyprus, Central African Republic, Sri Lanka, Congo (Democratic Republic), Gabon.

Thus, 129,735 (or 33%) of the 393,390 dyads reported here were obtained with the procedure for handling missing data outlined above. Finally, users will note that Table 2 lists both Czechoslovakia and Yugoslavia even though neither existed as independent states in 2000. The trade flows reported in 2000 for Czechoslovakia and Yugoslavia were obtained by aggregating the imports reported by the former Czechoslovakian and Yugoslavian republics.

In sum, the 94 countries appearing in this sample appear if either they reported imports in every year, or I could rely on a combination of "mirror flows" and temporally proximate flows between non-reporting countries for no more than one missing year. The full sample is representative of all world regions and accounts for between 96 and 99 percent of world trade, between 92 and 98 percent of world GDP, and roughly 80 percent of world population through time.

3. Data Files and Formats

The data appear in two formats—excel and UCINET. In each, the file names correspond to the year and commodity code of each relation. For example, y196501 is commodity code 01 (Meat and Meat Preparations) in the year 1965. The excel files do not contain labels, but the accompanying excel file titled "labels" lists both the UN country code and country name in the same order as the countries appear in the rows/columns of the data files. The UCINET files include the UN country codes on the rows and columns.

4. Data Details

Table 3. Data details.

Response Rate	N/A
Non-Respondent Bias	N/A
Theoretical Grounding	These data are relevant to questions about the organizational structure of manufacturing industries worldwide, as well as change in these organizational structures over time
Publications Using These Data	These data appear in part in Boyd et al. (2010); Mahutga (2006); (forthcoming); Smith and White (1992) and in full in Mahutga and Smith (2011)
Data Context	N/A
Respondents	N/A
Longitudinal	Yes, 15 relations in 1965, 1980 and 2000
Temporality	The valued dyads are measured in current US dollars
Analytical Utility	Any analytic context calling for comparisons of network structure across relations and over time
Known Issues	See description for procedures employed to handle missing data

References

Boyd, John P., Fitzgerald, William J., Mahutga, Matthew C., & Smith, David A. (2010). Computing Continuous Core/ Periphery Structures for Social Relations Data with MINRES/SVD. *Social Networks, 32*, 125-137

Breiger, Ronald (1981). Structures of Economic Interdependence Among Nations. In Peter M Blaue, & Robert K. Merton (Eds.) *Continuities in Structural Inquiry (*pp. 353-79). London, England: Sage.

Clark, Rob (2010). World-System Mobility and Economic Growth, 1980 – 2000. *Social Forces, 88*, 1123-1152.

Clark, Rob & Beckfield, Jason (2009). A New Trichotomous Measure of World-System Position Using the International Trade Network. I*nternational Journal of Comparative Sociology, 50*, 5-38.

Durand, Edward D. (1953). Country Classification. In Roy George, Douglas Allen, & J.E. Ely (Eds.) *International Trade Statistics* (pp. 117-29). Hoboken, NJ: John Wiley.

Hidalgo, C.A, Klinger, B., Barabasi, A.L., & Hausmann, R. (2007, July). The Product Space Conditions the Development of Nations. *Science, 317*, 482-487.

IMF (2012). *Direction of Trade Statistics*. Washington D.C: International Monetary Fund.

Lloyd, Paulette, Mahutga, Matthew C., & de Leeuw, Jan (2009). Looking Back and Forging Ahead: Thirty Years of Social Network Research on the World-System. *Journal of World-System Research, 15*(1), 48-85.

Mahutga, Matthew C. (2006). The Persistence of Structural Inequality? A Network Analysis of International Trade, 1965-2000. *Social Forces, 84*(4), 1863-89.

_____. Forthcoming. "Global Models of Networked Organization, the Positional Power of Nations and Economic Development." Review of International Political Economy.

Mahutga, Matthew C., & Smith, David A. (2011). Globalization, The Structure of the World Economy and Economic Development. *Social Science Research, 40*(1), 257-72.

Nemeth, Roger J. & Smith, David A. (1985). International Trade and World System Structure: A Multiple Network Analysis. *Review, 8*(4), 517-160.

Smith, David A., & Nemeth, Roger J. (1988). Research Note: An Empirical Analysis of Commodity Exchange in the International Economy: 1965-1980. *International Studies Quarterly, 32*, 227-240.

Smith, David A., & White, Douglas R. (1992). Structure and Dynamics of the Global Economy: Network Analysis of International Trade 1965-1980. *Social Forces, 70*(4), 857-893.

Snyder, David, & Kick, Edward L. (1979). Structural Position in the World System and Economic Growth, 1955-1970: A Multiple-Network Analysis of Transnational Interactions. *American Journal of Sociology, 84*(5), 1096-1126.

United Nations (2012). COMTRADE. New York, NY: United Nations

_____. (1963). "Commodity Indexes for the Standard International Trade Classification, Revision 1." Statistical Papers M:38(II). New York, NY: United Nations.