INTENSIVE AND EXTENSIVE MARGINS OF TRADE ON TRADE NETWORKS

RAYMOND T. WALTER

1. Rationale and Approaches for Studying Bilateral Trade

1.1. **Bilateral Trade: Extensive and Intensive Margins.** The evolution of bilateral trade relationships is among the less studied aspects of international trade. This is explained in part by the greater significance of a country's volume of total international trade (imports and exports) for the economic welfare of its population than of its volume of total trade with any particular country. Nonetheless some questions require systematic investigation of bilateral trade between pairs of countries.

A bilateral trade relationship is said to exist between two countries at a given period if there are positive imports (respectively exports) from one country to another during that period. Alternatively, the criterion for existence of such relationships is that imports (respectively exports) between countries exceed some threshold during that period. The evolution of these relationships is on two margins. Evolution on the extensive margin refers to the establishment or abandonment of trade relationships between pairs of countries; evolution on the intensive margin refers to changes in the volume (imports and exports) of existing trade relationships between pairs of countries. The intensive margin can be considered a generalization of the extensive margin, if the former considers instead the volume of potential rather than existing trade relationships. Direction can also be imputed to each margin; both the extensive and intensive margins could be considered into some country (imports into that country by others) or out of some country (exports out of that country to others).

1.2. **Bilateral Trade: Trade Networks.** Another way to study bilateral trade relationships uses complex network analysis. Such analysis has been applied, for instance, to

international economic integration and financial contagion (Kali and Reyes, 2007; Kali and Reyes, 2010). A network is a collection of nodes or vertices with the collection of all links or edges between pairs of these nodes or vertices, and is said to be complex when is necessary or most useful to characterize the network by various summary statistics.

More formally, consider a collection of nodes $N = \{1, 2, ..., n\}$ associated with a real-valued nxn matrix g such that g_{ij} is the relation between node i and node j. Then (N, g) is a graph or network with adjacency matrix g, and we will often refer to g itself as a network. By convention, $g_{ij} = 0$ always. If $g_{ij} = g_{ji}$ for all i and j in N (that is, g is symmetric), then g is said to be an undirected network; otherwise, g is said to be a directed network. If for all i and j in N, either $g_{ij} = 0$ or $g_{ij} = 1$, then g is an unweighed or binary network. If the entries of g can take more than these two values, then g is said to be a weighted network. A link or edge is a nonzero relation between two nodes. Further discussion is given in the references (Jackson, 2008).

In international trade, the most natural collection of nodes is the set of countries among which bilateral trade relationships are to be studied over some period. Considerable freedom remains in choosing the relevant relation between countries in a trade network. There is a choice of direction: the value g_{ij} can be defined with respect to the imports (respectively exports) into (out of) country i from (into) country j, and the network is directed; or defined with respect to the total trade volume (a sum or average of imports and exports) between countries i and j, and the network is undirected. There is also a choice of weight: g_{ij} is either 0 or 1 in accordance with whether or not these trade flows are zero (respectively less than some positive threshold) or positive (greater than some positive threshold), and the network is binary; or g_{ij} can take other values, say, the US dollar values of the relevant trade flows, and the network is weighted.

Our analysis defines g_{ij} in terms of imports into country i from country j, and thus we consider directed networks throughout. In most of our analysis, in which we consider weighted networks, the value of g_{ij} is the USD value of this flow of imports. Where binary networks are used, the value of 0 or 1 is in accordance with whether these imports are zero or positive respectively.

1.3. **Bilateral Trade: Margins of Trade in Trade Networks.** The analysis we are setting forth allows for each margin of trade to be associated with particular network structures. Insofar as one is measuring trade volumes by imports, the evolution of the intensive margin of trade (in the generalized sense of a *potential* trading relationship, mentioned above) from country i into country j is given by the evolution of g_{ij} in the adjacency matrix of a weighted directed network. It is similar for the extensive margin, but using a binary directed network.

Ideally our analysis would consider how the evolution of network properties of individual links affect the margins of trade, intensive or extensive, represented by these links. However, the summary statistics for complex networks have been much further developed for nodes than for links. Moreover, node statistics in general are much easier to compute. Thus we will instead consider an approach intermediate between strictly pairwise link analysis and strictly aggregated total trade volumes. We will consider the relation over time between the total volume of trade engaged in by one country with all other countries in the network and the position of that country in the trade network, where this position is characterized by various network measures. As we now argue, this total trade volume can be readily couched in terms of networks.

The position of a node in a network is measured by its *centrality*. The simplest measure of centrality is *degree centrality*, which considers how connected a node is to other nodes in the network. In a directed network, as we will be considering, has two measures of centrality for each node: node in-degree and out-degree.[1] The in-degree of node i is $d_i^{in}(g) = \sum_j g_{ji}$, the sum of all entries in the i-th column of the adjacency matrix g. The out-degree of node i is $d_i^{out}(g) = \sum_j g_{ij}$, the sum of all entries in the i-th row of the adjacency matrix g. Now recall in our representation of bilateral trade as a weighted directed network that g_{ij} (the relation from i to j) is the volume of imports into country i by country j. Thus to obtain the total imports into country i from all the countries in the network, it suffices

---

[1]The degree centrality of a node is typically defined as the degree of a node divided by the number of nodes other than that node in the network (or $n - 1$). In our subsequent analysis, we will always have the same number of nodes in our network, and thus degree is equivalent to degree centrality up to a fixed scaling factor.

to calculate $\sum_j g_{ij}$ – the out-degree. Similarly, the total exports out of country i to all the countries in the network are $\sum_j g_{ij}$ – the in-degree. A measure of total trade volume of a country is given by the sum of the out-degree and in-degree of the country - we call this the total degree of that country.

The definitions we have provided for weighted directed networks also apply to binary directed networks. In this case, the interpretation of in- and out-degree centrality is not trade volume of imports or export, but the number of countries from or to which the country is respectively buying imports or selling exports. The total degree would be sum of these degrees.[2]

We have now seen two ways to analyze bilateral trade relationships. One way considers intensive and extensive margins of trade, while the second way uses complex networks. Complex networks express both margins bilateral trade between pairs of countries. Furthermore, simple measures of degree centrality in networks express important aggregative measures of international trade volumes engaged by individual countries. We can now proceed to our data and empirical analysis.

2. DATA

Our source of bilateral trade data is the NBER-UN Trade Data, 1962-2000, available from the Center for International Data at UC Davis. A full discussion of these data is provided in the references (Feenstra et al., 2005). We consider a sample of 53 countries from 1992 to 2000. The procedure by which these countries were sampled is sufficiently involved that it requires a detailed examination.

In these data, for each importing country are given the USD values of total imports from each exporting country for each year in the data. Some pairs of importing and exporting countries are not included in the data for any of the years considered, but these missing pairs are not explicitly identified in the NBER-UN data. For our preliminary investigation,

[2]Both in the weighted and binary cases, it may seem that total degree is double-counting in some sense by adding imports *and* exports. But this misses the point, as we are trying to characterize the centrality of trade by some country relative to others in the network. The total degree need not necessarily be a real quantity for a country, but a useful quantity in analyzing the centrality of that particular country.

rather than making any assumptions about the values of these missing observations, we consider a subset of the data for which the USD values of imports are given for all importer-exporter country pairs in that subset.

The NBER-UN Trade Data is composed of a revised old dataset from 1962 through 1983. For 1984 through 2000, there is a subset of 72 countries for which the data is obtained from the UN, up to various adjustments (when changes in the names or political structures of these countries, this subset consists of only 69 countries); from 1996-2000, these countries accounted for 98 percent of world exports. We will only consider countries in this subset for the years 1962-1983. In some instances, a country was listed as its own importer; these were removed. However, even these data have missing pairs of countries. We narrowed down this set of 69 countries to its largest subset of countries with no missing pairs. (A) We identified first those countries for which imports from and exports to the 68 other countries were available. (B) The remaining countries were then sorted (1) in descending order of the number of the remaining 68 countries for which import data were available and then (2) in descending order of the number of the remaining 68 countries for which export data were available. (C) For countries with 68 exporters and 67 importers, the country was removed if data were missing for any one of the countries identified in (A) and kept otherwise. Then we consider countries with 68 exporters and 66 importers, and remove those countries without export data to the countries remaining in (C). We proceeded in this way until no countries remained, except when the number of countries for which import data was available decreased, in which case it was also necessary to consider whether import data was available for the countries already included.

Our final adjustments concerned years for which trade data were available for particular countries and how various countries were unified or separated into others. By starting from 1992, the Russian Federation could be used without the complications due to the dissolution of the Former USSR. This also avoids complications from the reunification of Germany in 1990. The Czech Republic was dropped, eliminating problems with Former Czechoslovakia and availability of full data for the Czech Republic being available only

since 1993, as were Former Yugoslavia, Yugoslavia, Slovakia, and Slovenia (eliminating problems with break-up and unification).

Our procedure is likely subject to selection bias, for reasons especially relevant to our network analysis. We have only used countries between which values of imports are available. Though it needs to be checked, it is likely that, for pairs of countries for which no import data is reported, the trade between these countries is zero or very small. Such small volumes of trade are not a significant source of revenue for governments of these countries, and thus there is little incentive to track these volumes. Thus the networks we are considering may be much more connected or denser than is representative of the world trade network. Our suggestion concerning the incentives of governments to record small trade volumes is more plausible when one considers a related issue for the accuracy of reported import and export data. As imports constitute a tax base and exports generally do not, there is incentive for governments to obtain accurate import data, but no such incentive exists for exports. For this reason, the NBER-UN Trade Data uses import data where and whenever possible.

A distinct but related selection bias may arise by restricting the analysis to a subset of the 72 countries obtained through the UN. Though it is not immediately evident, these countries are likely different from other countries not included; their trade is likely larger in volume and involves more countries, while trade of excluded countries is likely smaller in volume and involves fewer countries. Then we may have similar mischaracterization of the of the connectedness or density of the world trade network as above.

From our reduced NBER-UN data, we built adjacency matrices for the network characterizing these bilateral trade relationships for each year (1992 through 2000). Using these matrices we calculated various network statistics, such as different kinds of degree centrality, clustering coefficients, etc. Various dynamical variables and transformed variables were calculated from these measures. GDP data for each country during each year were also used, obtained from the USDA dataset of Historical Real GDP Values.

Having discussed these data and its potential issues, we now proceed to examine the relation over time between the total volume of trade for each country with all other countries

in the network and the position of that country in the trade network, where this position is characterized by various network measures.

3. ANALYSIS

In this preliminary investigation, we consider the relation between the evolution of the total trade volume of a country and a recently introduced measure of betweenness centrality, *random walk betweenness centrality (RWBC)*, to characterize the position of that country on the international trade network. Betweenness measures for a node seek to measure the importance of a node for connecting other nodes in a network. Exploring the rationale and computation of RWBC is beyond our present scope. In the context of communication networks, the intuitive idea of RWBC is that it measures the potential of a particular node for receiving a signal traveling through a network from an arbitrary node via a random walk process. Our MATLAB code for computing RWBC is included below.

Consider the trade network of our data as a weighted directed matrix. Let $wdtotdeg_{i,t}$ be the total degree of country i in the adjacency matrix of the weighted directed (trade) network at time t. Let $wdrwbc_{i,t}$ be the RWBC of country i in the adjacency matrix of the weighted directed (trade) network at time t. We consider a fixed effects regression of the following form:

$$\Delta wdtotdeg_{i,t} = wdtotdeg_{i,t} - wdtotdeg_{i,t-1} = wdrwbc_{i,t-1}\gamma + x_{i,t}\beta + \alpha_i + \epsilon_{i,t}, \text{ where}$$

$x_{i,t}$ is a vector of control variables, α_i is a fixed effect, and $\epsilon_{i,t}$ is an error term.

Our hypothesis is that changes in total trade volume (total degree) of particular country (node) depend on the RWBC for that country. Any such effect of *wdrwbc* on *wdtotdeg* should not be instantaneous, and thus we have lagged *wdrwbc* by one period. More formally, we are testing the hypothesis $\gamma \neq 0$ against the null $\gamma = 0$.

In Regression (1), we naively include no control variables. The coefficient is negative and significant at the 1-percent level. According to the p-value for our F-statistic, we reject the null hypothesis. Our overall R-squared is .3721. In these results, the standard errors initially seem alarmingly large, but these are only about 6 (11) percent the magnitude of our slope (constant) coefficient. By removing the lag on *wdrwbc*, as in (2), we can no

longer reject the null. Neither the slope coefficient or the constant remain significant for even 10-percent significance.

To ensure that our results for Regression (1) are some trivial consequence of a relation between the variables from which $\Delta wdtotdeg_{i,t}$ can be defined and $wdrwbc$, we did several regressions of each such variable on $wdrwbc$ without controls. We regressed $wdtotdeg_{i,t}$ on $wdrwbc_{i,t}$ in (3) and on $wdrwbc_{i,t-1}$ in (4). In both cases, the slope coefficients were significant at the 5-percent level or better. Both slope coefficients are positive, but the one in (3) is larger; this is consistent with our negative coefficient in (1). Evidently there is some relationship here. However, the improvements in within and overall R-squared in (1) compared to (3) and (4) are sufficient encourage us in asserting $wdrwbc$ is more important for changes in levels of $wdtotdeg_{i,t}$ rather than levels of $wdtotdeg_{i,t}$ themselves.

Repeating (1) but using changes in in-degree in (5) or out-degree in (6), the coefficients for $wdrwbc$ are not significant. Moreover, the various R-squared are low for both regressions and the signs of these coefficients are inconsistent with those in (1). That $wdrwbc$ only significantly affects a measure of total trade volume rather than a volume with direction, as it were, such as total imports or exports, makes sense, as $wdrwbc$ should have not be preferential to any particular direction. This relates to properties of a random walk.

So far, we have not considered any control variables. It is not clear what are the most relevant control variables, and this will require a further examination of the literature, especially in gravity-based models of international trade. It seems reasonable to believe $\Delta wdtotdeg_{i,t}$ to be depend on changes in GDP ($\Delta GDP_{i,t} = GDP_{i,t} - GDP_{i,t-1}$) and the number of countries with which a country is trading (say, as measured by the total degree of a country i considered as a node in a binary directed network at time t, $bdtotdeg_{i,t}$). These are included in (7). $delta_G DP$ is not significant and does not have the expected sign (increases in productivity are expected to be reflected in increases in international trade), while $bdtotal$ is significant with the expected sign. Otherwise there is little change from the results in (1). This suggests that $bdtotal$ is an important control, and motivates us to seek better controls in subsequent research.

We conclude that *wdrwbc* is likely to be negatively related to changes in *wdtotdeg*. To identify what mechanism explains this relationship will require further study. As *wdtotdeg* is usually increasing over time, we will need to better control for factors accounting for positive changes in *wdtotdeg*, those positive changes which *wdrwbc* is working against.

4. Conclusion

We have developed a general framework for studying the intensive and extensive margins of international trade, both on the disaggregated level of pairs of countries and on the more aggregated level of total trade by individual countries, using complex network analysis. We have identified the need for measures on network links to study the pairwise interactions of countries on the international trade network. In the absence of such measures on links, we have considered the effect of betweenness centrality measures on countries in the international trade network – in particular, the recently developed notion of random walk betweenness centrality (RWBC) – on total trade volume by these individual countries. RWBC appears to have negative and significant impact on changes in total trade volume, but explaining this relationship requires us to better account for the positive changes that total trade volume usually undergoes.

5. Bibliography

Felbermayr, Gabriel J. and Wilhelm Kohler, 2006. "Exploring the Intensive and Extensive Margins of World Trade," Review of World Economics (Weltwirtschaftliches Archiv), Springer, vol. 142(4), pages 642-674, December.

Feenstra, R. C., R.E. Lipsey, Deng, H., Ma, A.C. and Hengyong Mo, (2005),World Trade Flows, 1962-2000,NBER Working Paper 11040.

Jackson, Matthew O. Social and Economic Networks. Princeton, NJ: Princeton UP, 2008.

Kali, R. and Reyes, J., The architecture of globalization: A network approach to international economic integration, *J. Int. Bus. Stud.* 38 (2007a) 595620.

Kali, Raja and Reyes, Javier A., Financial Contagion on the International Trade Network. Economic Inquiry, Vol. 48, No. 4, pp. 1072-1101, October 2010.

Reyes, Javier A., Schiavo, Stefano and Fagiolo, Giorgio, Assessing the Evolution of International Economic Integration Using Random Walk Betweenness Centrality: The Cases of East Asia and Latin America (October 2008). Advances in Complex Systems, Vol. 11, No. 5, pp. 685-702, 2008.