Multilateral Gravity – A network approach

Karolien De Bruyne\textsuperscript{1,2,3}, Glenn Magerman\textsuperscript{1,2,3} and Jan Van Hove\textsuperscript{1,2,3}

Extended abstract

Summary

This paper extends the empirical gravity literature by including, and focusing on, the impact of networks on trade. Traditional gravity equations model bilateral trade flows as a function of income and a vector of distance variables. While the explanatory power of these models has proven to be very high, all measures of distance captured in these models are prone to some ‘ad hoc-ness’ that is related to international networks, but fail to capture more specific network elements. At the same time, the literature has mainly focused on bilateral effects between trading partners, neglecting the multilateral nature of trade in general. Network analysis presents the opportunity to address both these issues. First, network analysis allows to naturally include multilateral measures such as the number and intensity of trading partners, the effect of social networks in trade etc. Secondly, this paper estimates different weighted network measures and adds those to the vector of distance measures of the gravity models, introducing a more natural way to incorporate multilateral resistance into the bilateral models. Estimation techniques discussed are Poisson Pseudo Maximum Likelihood (PPML), Negative Binomial (NB) and zero-inflated versions of both models.

Introduction and literature

The first gravity model proposed by Tinbergen (1962) included only geographical distance in the vector of economic distance variables. Modern economies however, are not only characterized by these measures alone. Of course, physical proximity is something that is impossible to change,

\textsuperscript{1} KULeuven – Center for Economic Studies (CES), Naamsestraat 69, 3000 Leuven, Belgium.

\textsuperscript{2} University College Brussels – Center for Research on Economic Markets and their Environments (CREME), research group Globalization, Innovation and Competition (GIC), Warmoesberg 26, 1000 Brussels, Belgium. Email: glenn.magerman@hubrussel.be.

\textsuperscript{3} KULeuven - Leuven Center for Irish Studies (LCIS), Janseniussstraat 1, 3000 Leuven, Belgium.
but virtual networks can reduce distance in many ways impossible before. Other economic distance measures have since then been included, such as common language (Rose and Spiegel, 2002), common religion (De Groot et al., 2004), colonial ties (Marquez-Ramos & Martinez-Zarzoso, 2005; Lewer & Van den Berg, 2007), common currency areas and institutional influences like WTO and IMF Membership (Rose, 2000, 2004; Zarroz and Lehmann, 2003), migration flows (Karemera et al., 2000; Rauch and Trindade, 2002; Bandyopadhyay, Coughlin and Wall, 2008; Grogger and Hanson, 2011), political allies and zones of conflict (Mansfeld, Milner and Rosendorff, 2000), FDI (Bergstrand and Egger, 2007), multilateral trade barriers (Anderson and Van Wincoop, 2003), patent data (Picci, 2010) etc. These virtual networks can also include social and professional network ties (Jovanovic, 2009), path dependence and other measures not yet included in the vector of distance variables in the gravity model.

Some multilateral measures have been considered to expand the mainly bilateral gravity measures up until recently. These include Anderson and Van Wincoop (2003), who point out that trade intensities are governed by relative trade costs rather than absolute trade costs. It is not the goal of this paper to present a uniform solution, which points to the most significant universal measures to include in a gravity model, nor can we claim that we present the best multilateral approach to the gravity model, but we aim to include some new significant measures and indicators into the vector of distance for the gravity model by drawing from the inherent multilateral nature of network measures that are present in trade flows in general.

Empirical gravity has been criticized by a lack of theoretical foundation. Structural models have been characterized by Anderson (1979), using the Armington assumption (differentiated goods by country of origin) and a CES demand side; Bergstrand (1985) and Anderson and Van Wincoop (2003), adding monopolistic competition; Eaton and Kortum (2002) and Chaney (2008) model heterogeneous productivity draws for countries. Baldwin and Harrigan (2006) famously dedicate their paper on the proper estimation of these recent models with large portions of zeros in the bilateral trade matrix. Especially the more recent models are useful for the analysis presented here for various reasons: (i) account for zero-trade observations\(^4\) (ii) allow for asymmetric trade flows, and (iii) evaluation of the extensive margin of trade (i.e. the number of countries or varieties that trade with respect to a given partner) and the intensive margin of trade (i.e. the volume of trade per country).

Recent research has included some descriptive statistics of the World Trade Web (WTW) (Fagiolo, Reyes and Schiavo, 2008; Bhattacharya, Mukherjee, Saramäki, Manna, 2008), but only serve as a standalone observation of the network itself. De Benedictis and Tajoli (2011) give a first pass at incorporating network analysis into gravity models, using data from Subramanian

\(^4\) in 2005, the World Trade Web (WTW) had a network density of .55, measured as the ratio of connections with respect to the total possible number of connections or \(\frac{2m}{n(n-1)}\) in an undirected network, increasing from a density of .3 in 1980.
and Wei (2007) and IMF DoT statistics, over a shorter time frame and with less estimation techniques than used in this paper.\(^5\)

**Motivation & Research Question**

The WTW is a complex entity. Understanding its underlying structures and its natural rules is of fundamental importance for firms, policy makers and consumers in general, as it affects everyone’s wellbeing. This explains the burgeoning literature in trade, focusing on gravity models that aim to explain the effects of this network structure and the effects of changes therein. These include for example, the impact of membership of the WTO and IMF, the effect of common currency areas, the effect of lowering tariff and non-tariff barriers, the effect of migration etc. on total trade flows. However, the bilateral notion of gravity models fails to take into account the omnipresent diversity of this WTW. Many countries do not trade at all, or only trade in closed formations with each other. Some countries are trading more inter-industry, others trade more intra-industry. This heterogeneity of the WTW is largely discarded in the bilateral approach, while it is naturally inherent in the concept of network theory and that is the direction we take in this paper.

This paper contributes to both the theoretical and the empirical literature surrounding gravity models. The main innovations are the inclusion of multilateral measures into the gravity model, taking into account effects that are not merely bilateral. This multilateral environment is natural in the setting of trade but mainly absent from the literature. A first pass has been made by Anderson and Van Wincoop (2003) with the inclusion of multilateral trade resistance (MTR), stressing the importance of relative trading costs rather than absolute trading costs. We refine the nature of these MTR and their effect on trade flows in this paper. The multilateral innovation presented in this paper is possible due to the recent developments in network theory, its measures and its econometrics. This leads directly to the underlying research question: “What is the effect of networks on the world trading web?”

**Datasets**

Bilateral worldwide trade data on the national level, from the publicly available UN database of COMTRADE has been used for the first set of estimations. Data is available for the period 1962-2011, and in current US dollars. In a parallel estimation procedure, estimates are derived from the CoW (Correlates of War) dataset. This dataset contains dyadic imports and exports in current US

---

\(^5\) The authors also tried using lagged exports as an instrumental variable, but in their sample, this did not pass the weak instrument test. In our opinion, lagged variables of exports are not strong instruments in general, as the trade flow variations over one year are not that big. The effect could be more pronounced over a larger span of 25 years for example.
dollars for the unbalanced panel of 1870-2009. For the economic mass of countries, measured as GDP per capita in current US dollars, we use the Penn World Tables, version 7.0.

Methodology – part 1: Networks

To represent the WTW as a network, countries represent nodes or ‘vertices’, and the bilateral trade flow between countries is given by links between those vertices, or ‘edges’. If we contemplate two countries merely whether they trade or not, this is given by an element 0 or 1 in the ‘adjacency matrix’, a matrix which describes the edges between all vertices present in the network, and the network is then called an non-weighted network. If we attach weights to the links however, for example by using ‘trade volume’ measured by the level of exports and imports between a pair of vertices, we obtain a weighted network.

When looking at the total trade intensity between two countries, i.e. the total level of exports and imports between those countries combined, we call this an undirected edge, and if all edges are undirected in a network, we call this an undirected network. This is a first approach to modeling the WTW as a network. However, we lose information by doing so, as we can distinguish between exports and imports separately and the volume of trade between two trading partners can be vastly different in this bilateral composition. Therefore, we also measure a directed network, only using export or import intensities. If we model all export patterns only, this coincides with all import pattern values, only reversed, and one of both is redundant.

We will discuss the descriptive statistics and measures and metrics of the network analysis next. For a formal definition and rigorous explanation of these measures and related issues, we strongly recommend Newman (2010).

Descriptive statistics and network measures for the cross-sections in the undirected network are given by the elements of the adjacency matrix \( A \), which represent:

- \( n \) is the number of vertices \( i \) inside the WTW (i.e. countries that have non-zero trade values) at a given time, giving the size of the square matrix,
- \( m \) is the number of edges connecting the vertices inside the WTW at a given time, representing elements in that matrix.

---

6 Note that we do not need to go to a balanced panel, as is explained by the estimation procedures below.
7 In their 2008 paper, Bhattacharya et al., only model an undirected network, and due to slight differences in reporting, export and import patterns do not coincide completely (also reported by Anderson, 2010). They overcome this by taking the average of imports and exports and remain in the undirected realm. This fallacy is not present in our COMTRADE and CoW datasets, and we can safely discard one of two models, while addressing the more interesting case of the directed network.
From this, the network density in an undirected network is calculated as a fraction of the total network: $\rho = \frac{2m}{n(n-1)}$. From the adjacency matrix, we can calculate some more measures:

- $k_i = \sum_{j=1}^{N} A_{ij}$ is the degree of a vertex $i$, i.e. the number of undirected links it has, or the number of trading partners in our study. This follows a distribution over the network that can be estimated.
- $c = \frac{1}{n} \sum_{i=1}^{N} k_i$ is the mean degree of a vertex in the undirected network.

Descriptives can be given for both the non-weighted as the weighted network. For the directed network measures we add the following measures:

- $k_i^{in} = \sum_{j=1}^{N} A_{ij}$ is the in-degree, or the number of incoming edges for a particular vertex $i$.
- $k_j^{out} = \sum_{i=1}^{N} A_{ij}$ is the out-degree, or the number of outgoing edges for a particular vertex $j$. Both measures also follow a distribution over the network, and we can also estimate the joint distribution of $(k_i^{in}, k_j^{out})$.
- $c_{in} = \frac{1}{n} \sum_{i=1}^{N} k_i^{in} = \frac{1}{n} \sum_{j=1}^{N} k_j^{out} = c_{out}$ is the mean degree of the directed network.

With regard to the degree distributions in both the undirected and the directed network, we can estimate a range of the distribution that closely follows a power law, as is indicated by Newman (2010), by calculating the relationship directly from the data, which is preferred over log-log transformations that lead to linear regression analysis. The probability distribution function of a power law is given by $p_k = Ck^{-\alpha}$, where $k$ is the degree of the vertex, and $\alpha$ is an exponent to be estimated. The higher $\alpha$, the more convex the power law will be. The calculation of $\alpha$ is then given by:

$$\alpha = 1 + N \left( \sum_{i} \ln \frac{k_i}{k_{min} - \frac{1}{2}} \right)^{-1}$$

where $k_{min}$ is the minimum degree for which the power law holds, and $N$ is the number of vertices for which $k \geq k_{min}$ holds. This can easily be estimated with ML (Clauset, Shalizi and Newman, 2009). From a more applied point of view, it is nice to note that the pdf of a power law is the cdf of Pareto’s law. This measure can then be used to test income or trade distributions according to Pareto’s law.

Network measures and metrics allow for typical network characterization. For example, how ‘central’ a vertex is in the WTW can be measured by centrality. This kind of analysis is a first fundamental departure from the bilateral analysis in traditional gravity models, as this looks at the
effect of a particular trading partner inside the whole network.\textsuperscript{8} Centrality measures come in some flavors, with some basic varieties: degree centrality, eigenvector centrality, Katz centrality and the PageRank. Degree centrality is merely the degree of the vertex, and the more edges are connected to a vertex, the more important the vertex is assumed to be in the network. Degree centrality exists in both an undirected version and in-degree centrality and out-degree centrality in the directed network.

Eigenvector centrality extends the degree centrality, by measuring the centrality of the vertices that have edges to the measured vertex. In other words: not all neighbors are equally important, trading with a well-connected partner can have strong positive effects on the centrality of the contemplated vertex. Due to the chicken-and-egg problem present, eigenvector centrality is estimated with iterative methods.

In a directed network, eigenvector centrality can cause a particular problem: suppose a particular country is a very strong exporter, with respect to his import behavior. This country is arguably an important vertex in the network, however, his eigenvector centrality will be close to zero, as (almost) only outgoing edges are present. Especially when zero-inflation is present (see later), this can be an important issue. This is remedied by the concept of Katz centrality, which in essence “gives away” some small amount of centrality for free, from which the centrality measure is then calculated. A nice extension to this Katz centrality is to assign individual specific values for this free amount of centrality, based on non-network factors such as income (Newman, 2010). It might be very interesting to ‘invert’ the gravity equation, by including economic mass inside this Katz centrality directly, and measure trade from the Katz centrality as a function of, inter alia, national income. This would represent a true network approach to the gravity model, PPML with multilateral distance as the main vector on the right hand side, with income defined implicitly inside the centrality measure. What the differences in results are will be evaluated in the next stage of this project. A further extension to the concept of centrality is the PageRank, named after Google co-founder Larry Page. An undesirable effect of the Katz centrality is namely that vertices with high centrality values give high values to all connected vertices, regardless of the number of vertices that attach to it. For example, if the US, supposedly having a high value of Katz centrality, trades with many different partners, all these partners receive a high centrality value too. PageRank corrects for this by dividing centrality proportionally over the number of outgoing edges from that vertex.

We also present some clustering measures, including transitivity and local clustering. Transitivity means in loose terms: “a friend of a friend is also my friend”, or in mathematical terms, if $a \circ b$ and $b \circ c$ together imply $a \circ c$. We can measure the degree of transitivity in the network as the clustering coefficient, which counts the amount of transitive relationships in the network. Related

\textsuperscript{8} This can serve as a starting point for policy analysis, for example, what is the effect of a trade embargo posed upon a single trading partner over the whole network? Traditional analysis will only account for bilateral effects, and can underestimate the total effect of the policy on the network.
to trade, this can be seen as the probability that country A trades with country C on a bilateral basis, given that A trades with B and B trades with C. The clustering coefficient gives insight in the cliquishness of trading countries, countries that tend to trade heavily between themselves, rather than with outside partners, and can also give an indication for the horizontal dispersion of trade in general: if trade is more horizontal (from A to B to C but not to A or back to B), this can indicate that B is more a transit country, or that B uses goods from A as inputs and exports outputs to country C. This is close to the theoretical work of Ahn, Khandelwal and Wei (2011). It must be noted that using the directed network for the clustering coefficient can add great value and insights to both theoretical network literature, as to empirical models, as the directed variant of the clustering coefficient is hardly used for some reason (Newman, 2010).

Methodology – part 2: Gravity models

Traditional gravity models have been estimated using OLS. This technique however can be severely biased in the case of trade data, as many zeros exist in the trade matrix. The log-transform of the gravity model drops all zero observations, and the regression is only run on positive values. OLS will then lead to upward-biased intercepts and downward-biased coefficients. We estimate the new gravity model using Poisson Pseudo-Maximum Likelihood (PPML), Negative Binomial (NB), Zero-inflated Poisson (ZIP) and Zero-Inflated Negative Binomial (ZINB). We are aware of the controversy surrounding the last two estimation techniques (see for example Santos Silva and Tenreyo (2006) and Burger, Van Oort and Linders (2009) for a discussion), but we will only take a final stance after the successful completion of this project.

The PPML model

The Poisson estimation is always consistent, even when the underlying distribution is not truly Poisson. However, when the variance is not equal to the mean in this model (equidispersion), we can control for overdispersion by using the PPML, which is just a Poisson model with robust standard errors (Gourieroux, Montfort & Trognon, 1984). We estimate the PPML from the econometric model given by Burger et al. (2009). The novelty with respect to the paper of Burger et al. (2009) is that we include the new network measures instead of the traditional bilateral gravity regressors. We continue this practice across all the estimated econometric models. In the panel dimension, we estimate the PPML with fixed effects for individual heterogeneity that is constant over time. The advantage of this method is that estimates are always consistent as in the cross-section case. The Poisson model is then given by:
Where:

- $I_{ijt}$ is the estimated exports from country $i$ to country $j$ at time $t$
- $\mu_{ijt}$ is the conditional mean of the distribution, linked to an exponential function of a set of variables, $X_{ijt}$

The conditional mean is given by elements of the traditional gravity model:

$$\mu_{ijt} = \exp(\alpha_0 + \beta'X_{ijt} + \gamma_t)$$

Where:

- $\alpha_0$ is a constant,
- $X_{ijt}$ is the vector of distance related explanatory variables, including our network measures,
- $\gamma_t$ is a year-specific fixed effect.

Note that we do not have to include country-specific effects to control for unobserved heterogeneity, exactly due to using these network variables.

**The NB model**

The PPML econometric model is omnipresent in the contemporary literature surrounding the estimation of gravity models. There are however some alternatives to the PPML. A reason for overdispersion in the gravity models can be found in unobserved heterogeneity not taken into account by the Poisson model (Greene, 1994). The NB model explicitly models the variance (assumed to be equal to the mean in the Poisson model) and relates it to the mean. In this model, the variance is a function of bot the mean $\mu_{ijt}$ and a dispersion parameter $\alpha$, so that unobserved heterogeneity is incorporated into the model (Burger et al., 2009). The larger $\alpha$ is, the larger is the degree of overdispersion. The NB model is given by:

$$Pr(I_{ijt}) = \frac{\Gamma(I_{ijt} + \alpha^{-1})}{I_{ijt}! \Gamma(\alpha^{-1})} \left( \frac{\alpha^{-1}}{\alpha^{-1} + \mu_{ijt}} \right)^{\alpha^{-1}} \left( \frac{\mu_{ijt}}{\alpha^{-1} + \mu_{ijt}} \right)^{I_{ijt}}$$

where $\Gamma$ is the gamma function. A drawback of this model is that estimates are no longer consistent when the underlying distribution is not truly NB. A simple likelihood ratio test can be used ($H_0 = \alpha = 0$) to see whether the NB is preferred over the PPML or not.
Another issue is that both PPML and NB can underestimate the true number of zeros in the model, which lead to ZIP or ZINB models. These models start from the consideration that we have two latent groups in the data: one group has value zero, because there is no trade between countries at all in general, and one group has value zero because (i) in a cross section, there was no trade at that specific time point, but it can be trade exists for that country pair at other time points, or (ii) the data was truncated at a cutoff level of reporting. These kinds of observations are clearly visible in both the COMTRADE dataset as the CoW dataset, and deserve proper attention. The ZIP and ZINB models then lead to a conditional Poisson model. The first part of the econometric model is a logit regression that represents the probability of having no trade at all (Burger et al., 2009). The second part then models the Poisson model, conditional on the probability of having non-zero trade. This ZIP is represented by:

\[
\begin{align*}
\text{Pr}(I_{ijt} = 0) &= \psi_{ijt} + (1 - \psi_{ijt}) \exp(-\mu_{ijt}) \\
\text{Pr}(I_{ijt}) &= (1 - \psi_{ijt}) \frac{\exp(-\mu_{ijt}) \mu_{ijt}^{I_{ijt}}}{I_{ijt}!}
\end{align*}
\]

Now, \( \psi_{ijt} \) is the probability of having a strictly zero count (i.e. the first latent group in the above example).

The ZINB is presented in a similar way:

\[
\begin{align*}
\text{Pr}(I_{ijt} = 0) &= \psi_{ijt} + (1 - \psi_{ijt}) \left( \frac{\alpha^{-1}}{\alpha^{-1} + \mu_{ijt}} \right)^{\alpha^{-1}} \\
\text{Pr}(I_{ijt}) &= (1 - \psi_{ijt}) \Gamma(I_{ijt} + \alpha^{-1}) \left( \frac{\alpha^{-1}}{\alpha^{-1} + \mu_{ijt}} \right)^{\alpha^{-1}} \frac{\mu_{ijt}^{I_{ijt}}}{I_{ijt}! \Gamma(\alpha^{-1})}
\end{align*}
\]

To see whether the zero-inflated or the original models are more appropriate, the Vuong LR test (Vuong, 1989) can be used.
References