ENDOGENOUS HUB FORMATIONS IN INTERNATIONAL TRADE

Nicholas-Joseph Lazarou

Paul Krugman in 1993 explored the role of transportation hubs in the pattern of trade. He asks when and why will a transportation hub emerge. In this paper, I incorporate network theory in a simple trade model of monopolistic competition with representative firms in order to yield conditions governing the optimal network formation - a hub or a direct links structure. Firms take into consideration the optimal network for their trading strategy. Reductions in variable and fixed costs affect trade flows in the predictive way. But trade volumes also have an impact on the optimal transportation network structure because of increasing returns in the transport sector, herein modelled via the manufacturing firm.

I provide empirical evidence for the existence of hub formations and direct connections in international trade flows. I employ HS2 bilateral trade flows for a panel of 63 countries between 2003 and 2007. I compare two measures of distance, direct capital and indirect capital distance, under two different regimes: countries trading directly, and countries trading indirectly. Indirect distance under hub networks reduces exports by about 5 per cent less relative to direct distance. The latter reduces export flows 24 per cent less than indirect distance when direct trading is allowed. I show that the network formation is driven through sectoral heterogeneity. The decision of the firm to choose network formations results in the endogeneity of distance. I consequently utilise sector fixed effects or time difference as an instrument to isolate the impact of the network structure in trade flows.

Keywords: International Trade, Network Formations, Transport Costs.

JEL Classification Numbers: F12, L91, C71.

DRAFT NO: 07/0312-THIS WORK IS STILL IN PROGRESS.

1. INTRODUCTION

Paul Krugman in 1993 delivered a paper exploring the role of transportation hubs in the pattern of interregional and international trade. He tackled a problem that is widely ignored in international trade theory: when and why will a transportation hub emerge. He shows that production of commodities will be concentrated in a location from which all arrivals and departures have the lowest transportation costs. The interaction between increasing returns to scale in manufacturing and transport leads to the emergence of a hub region.

Transportation hubs are not a fundamental element of international trade study because not enough attention is cast upon the structure of the transport industry. For transportation costs are just perceived as a subset of trade costs and specifically variable trade costs. They are modelled using the iceberg assumption first developed 60 years ago. This simple modelling concept allows for a wedge that serves to inhibit allocation of production and consumption.

In this paper I perturb Krugman's model of transportation hubs. I do this in order to raise awareness of another dimension of the distortive impact of transportation costs on trade flows. I model the

[a]PhD Student in Economics, University of Southampton, N.Lazarou@soton.ac.uk

[b]I am grateful to my Supervisors Dr. Hector Calvo-Pardo and Dr. Jackline Wahba for their faithful support and advice. I thank Dr. Luca David Opromolla for his constructive suggestions.

significance of the network organisation of the transport sector and display its pervasive effect on trade flows. Specifically I show that when transportation costs are high, manufacturing firms instead of lowering their output or exiting the market, can choose to trade via a central hub. When transportation costs are low, these firms will always trade directly with foreign countries or otherwise face losses. I empirically test for the existence of hub formations using sectoral aggregate export flows and find supporting evidence for such structures: Manufacturing firms tend to trade indirectly via a central hub as well as directly. Whilst firms producing agricultural goods only trade directly to their destinations. I also show that network selection is defined endogenously by the firm as a form of a hedge against adverse transport costs. Such a decision induces simultaneous causality between export flows, distance and fixed costs and is attributed to the network effect. I then implement appropriate instrumentation to isolate the effect created by the network. In this way I am able recover the true magnitude of trade costs on trade flows whilst controlling for network induced information.

According to Hummels (2007), 99% of the world's trade by weight is carried by sea. The introduction of containerisation in the latter half of the last century led to hub and spoke configuration of maritime trade flows. By endogenously creating hub networks, the shipping industry minimises exposure to transport costs by virtue of higher traffic density and when compared to direct trading.

A hub and spoke configuration is a particular form of a network. It is this network structure that I incorporate in the standard trade model of monopolistic competition with representative firms. I conveniently abstain from modelling a separate transport sector[1] but keep the network structure as the product of the sector's industrial organisation. I then study the impact of the network on trade flows through decisions of manufacturing firms.

I do this by substituting the notion of iceberg costs with benefits and costs associated with network formations. I merge the symmetric connections framework of Jackson & Wolinsky (1996) in a simple Krugman monopolistic competition model of trade for three countries. Using sufficiently less assumptions than Krugman (1993), I show how firms decide whether to export or not. If they do, I study which conditions will determine their organisation under a hub formation or direct links. The reason for this approach is to have an alternative methodological tool than the Krugman (1993) model. In the latter one needs to postulate concentration of production, centrality together with a defecting firm as sufficient conditions to yield equilibria. Further, it is not possible to develop a testable prediction for the existence of hub formations.

In my simple extension of this model, locational advantage and the level of transportation costs are the only prerequisites after controlling for country size. The economic intuition behind this framework is related to the very nature of maritime trade. Whilst concentration of production jointly with centrality may be of importance under air transport, the same need not apply for maritime transport. Some of the less developed regions of the world obtain hub status conditional only on centrality amongst the distribution of trade routes. Concentration of production may then take place but is certainly not a condition precedent.

Empirically I show, under different distance and fixed cost estimations, when hub networks appear to be more beneficial for trading partners than direct trading. When I allow indirect trading, distances and fixed costs associated with hub formations appear to affect exporters less. Distances and fixed costs associated with direct trading appear to affect exporters more. The reverse holds true when I allow for direct trading. This application holds for manufacturing firms, but is violated for firms pro-

[1] I am concurrently extending this model to cater for: a) network structure, b) increasing returns in the transportation sector and c) firm heterogeneity in manufacturing.

ducing agricultural goods. My estimators take into consideration the network effect and are proved to be endogenous. This is because the firm always decides the formation that will allow it to increase its trade volume. I show that this effect is driven through sectoral heterogeneity and I control for it using sector fixed effects or time difference as an instrument for distance. Then I recover the true impact of distance costs on trade flows.

The implications of this paper are directed towards developing nations. Exporters there have to absorb higher transport costs so as to be in the position to penetrate markets abroad. This situation prevents export-led development, reducing workers' wages and inducing a welfare impact. Higher transport costs are attributed to geographical disadvantage and lack of proximity. Prevention of market access for developing nations translates to losses from trade of about 68% lower GDPpc on average (Redding & Venables 2004, p.77). Therefore improvement of own and transit country infrastructure together with hub formations could make possible the amelioration of excessive transport costs (Limao & Venables 2001, p.471).

My work balances between two strands of theoretical and empirical trade literature. It also incorporates a branch of the separate field of social and economic networks. On the one hand I try to yield predictive inference through the standard trade model of representative firms developed in Krugman (1980) and Krugman (1991) by distorting the iceberg assumption. This is done after examining more closely the nature of transportation costs and the industrial organisation of the transport sector. The importance of transportation costs and maritime costs has been highlighted in the works of Hummels (2007) and Hummels et al. (2009). The relative importance of the validity of the iceberg cost assumption for transportation costs is contested in Irarrazabal et al. (2011). Insight about the structure of the transportation industry is provided in Hummels et al. (2009) and Hendricks et al. (1995) whilst the pervasive impacts of the modes of transport on trade flows are presented in Lux (2011). The importance of the network literature in international trade is currently being highlighted by Chaney (2011).

The remainder of this paper is organised as follows: Section 2 provides some basic notions of economic and social networks that will be used in the model. Section 3 incorporates network structure in a simple trade model of monopolistic competition. Equilibria for two countries and three countries are studied in section 4. Subsequently I develop the conditions for endogenous hub formations. Section 5 derives the gravity equation for aggregate trade flows and develops the empirical prediction for the existence of hub formations. Calibration of and endogeneity in the gravity equation are discussed in section 6. I then present evidence of hub network formations and provide treatment of the endogeneity of distance. The last section concludes.

2. SETUP OF THE NETWORK

Consider a set of countries $K = \{1, 2, ..., k\}$ that exist in the world and engage or not in international trade through manufacturing firms. I will state as *directly connected* or *directly linked*, the countries that actually have a direct trading relationship using no other intermediary country. Thus, I can define a network G as a list of pairs of countries $\{i, j\}$ that are linked to each other. Each link can be represented as a *graph* $g \in G$. And hence if country i and country j are directly connected g takes a value of 1 or if they are not directly connected g becomes 0. I denote the existence of a *direct link* as $g_{ij} = 1$ and $g_{ij} = 0$ if countries i and j are *not directly connected.*

I postulate that there are costs and benefits associated with *links*. These costs and benefits affect the firms that choose to enter the export market in each particular country. There are two types of links that can exist, a *direct* and an *indirect* link. If a direct link is formed by a firm it must incur a cost c. There is also a benefit $0 \leq \delta \leq 1$ associated with proximity or distance between i and j in the

sense that the firm will prefer to trade to closer trading partners rather than more distant ones. The firm has the additional option to form an indirect link. Implicitly, there must already exist a direct link to another country for the indirect link to be feasible. The indirect link is formed without cost and receives only a pure benefit $\delta^{t_{ij}}$, where $t_{ij} \geq 0$ is the integer number of links between countries i,j. This construction allows a firm in country i entering the export market to take into consideration the distance δ and the cost c associated with this distance. It can also consider whether to connect directly to the destination country and incur this cost. Or alternatively, it can consider connecting indirectly. In the latter case it avoids the cost but receives a discounted benefit as the proximity decreases. The difference between the benefit of forming a link and the cost of a type of link is thus defined as:

$$v_{ij} = \delta^{t_{ij}} - c_{ij|ij \in G} \tag{2.1}$$

By convention we have $g_{ii} = 0$ if a country remains autarkic. Further, $t_{ij} = 0$ if there is no path that connects directly or indirectly countries i and j. An exposition of this construct is as follows. For $N = 3$ symmetrically placed countries assume that countries 1 and 2 are at the edges and 3 is in the middle. There can be two types of available networks. One network formation is direct links between all participants. Then the network is defined as $G = \{12, 21, 13, 31, 23, 32\}$. The second formation is an indirect link between 1 and 2 and direct links from and to country 3, such that $G = \{13, 31, 23, 32\}$. The net benefit term between countries 1 and 2 becomes for the case of direct links $v_{12} = \delta - c = v_{21}$. For the case of an indirect link between 1 and 2 we have: $v_{12} = \delta^2 - 0 = v_{21}$. The latter indirect link implies the existence of two direct links forming this particular connection: the link between 1 and 3 and the link between 3 and 2.

I develop also the participation share of each country in the network depending on the types of links that are formed. The network participation share will be perceived as the fixed cost associated with the network. While the participation share is not employed in the theoretical model, it will assist in the gravity equation specification in lieu of the unmeasurable benefit of forming an indirect link. In the empirical part I will substitute the parameter δ with the fixed costs associated with the network for each country. Denote the set of country i's direct connections in a network as $N_i(G) = \{i \neq j | g_{ij} = 1\}$. The cardinality of this set is $n_i(G)$. The size of the network is $n(G) = \sum_i^N n_i(G)$. The participation share of i in the network is simply $F_i = \frac{n_i(G)}{n(G)}$. To provide an example, consider the direct links network for the 3 countries. Country 1's set of direct connections is $N_1 = \{12, 13\}$ and the cardinality of the set is $n_1 = 2$. The total number of direct connections is 6, and country 1's fixed costs associated with the network are $F_1 = 1/3$. Equivalently for the case of indirect links between 1 and 2 we have $F_1 = 1/4$, since country 3 in the middle is burdened by the additional share $F_3 = 1/2$.

For the remainder of this paper, countries are symmetrically spaced on the surface of the plane or a sphere: $ij = ji$ therefore $c_{ij} = c_{ji} = c, t_{ij} = t_{ji} = t$. I now proceed to incorporate this structure in a simple trade model of monopolistic competition with representative firms. I provide the general model of trade and exemplify equilibria between two trading countries and three trading countries. I then recover information on the impact of the network formation on trade flows for three trading countries and how this becomes an endogenous decision for the firm in each country.

3. SETUP OF THE TRADE MODEL

Symmetric countries produce goods using only labour. Country n has a population L_n and two sectors. One sector is responsible for the production of a single homogeneous good that can be traded freely. This good is the numeraire. The other sector produces a continuum of differentiated goods that can be traded at a cost. Each specific variety is produced by a single monopolist. In both sectors firms can freely enter or exit production. The population works in the sectors, moves freely across sectors

but not across countries and consumes goods. Each consumer is endowed with one unit of labour.

Demand — A representative consumer receives utility U from consuming q_0 units of the numeraire and q units of the differentiated variety ω which may be produced domestically or may be imported. Her preferences are given by a C.E.S. utility function over the continuum of differentiated varieties:

$$U = q_0^{1-\mu} \left[\int_{\omega \in \Omega} q_{ij}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}\mu} \tag{3.1}$$

where $\sigma > 1$ is the elasticity of substitution between pairs of varieties and Ω is the mass of available goods. Maximising utility subject to exhausting her labour income share, the representative consumer in country j has demand for differentiated goods:

$$q_{ij} = \frac{\mu L_j p_{ij}^{-\sigma}}{\sum_{j:i=1}^{N} \int_{\Omega} p_{ij}^{1-\sigma}(\omega) d\omega}, \tag{3.2}$$

where the denominator represents an aggregate price if the set of differentiated goods was consumed as an aggregate good.

Production and Trade Costs — Good 0 is the numeraire homogeneous good. One unit of labour produces w and its price is normalised to 1, so that the wage is equal to the price of the good. In this respect I set the wage equal to 1 across countries due to free trade, and across the two sectors within each country.

One firm can produce one variety of the differentiated good. Labour costs for differentiated goods are split between a marginal and a fixed cost and thus the sector is characterised by increasing returns. The fixed cost for the remainder of the theoretical model will be constant. In the aggregate trade (gravity) equation I shall relax this assumption by incorporating the share of each country's participation in the network. The marginal cost consists of a constant parameter $\gamma > 0$ and a variable trade cost. The variable trade cost is the net benefit term that stems from forming a link to another country. In this way I introduce network structure in the trade model. For domestic consumption the net benefit becomes by construction 1.

To produce and sell a variety ω either domestically or abroad, the firm in country i employs Labour input:

$$L(q) = \gamma \frac{q_{ij}}{v_{ij}} + F_i = \gamma \frac{q_{ij}}{\delta^{t_{ij}} - c_{ij|ij \in G}} + F_i \tag{3.3}$$

As such, if a direct link ij is formed the firm receives a benefit δ in the sense that $\frac{\partial \Pi}{\partial \delta} > 0$. In addition it incurs a transport cost c as $\frac{\partial \Pi}{\partial c} < 0$. In case of an indirect link, the firm receives a decayed benefit $\delta^{t_{ij}}, t_{ji} > 1$ and incurs no cost at all.

The firm solves its maximisation problem constrained by the quantity demanded. It sets its optimal price equal to a constant markup over the unit cost $p_{ij} = \frac{\sigma}{\sigma-1} \frac{\gamma}{v_{ij}}$. Positive profits require firms to enter the sector exhausting any profit margin. When profits are zero, the firm produces quantity $q_{ij} = \frac{F_i v_{ij}}{\gamma}(\sigma - 1)$.

Over all varieties ω produced in each country, the total labour input must equal the labour share in the increasing returns sector: $\int_{\omega \in \Omega} L_{q,i}(\omega) d\omega = \mu L_i$. Since each firm produces one variety, the number

of firms becomes finite and equal to $n_i = \frac{\mu L_i}{\sigma F_i}$.

This enables me to write the aggregate price index as $\sum_{j:i=1}^{N} \int_{\Omega} p_{ij}^{1-\sigma}(\omega) d\omega = \sum_{j:i=1}^{N} n_i p_{ij}^{1-\sigma}$.

4. EQUILIBRIA AND ENDOGENOUS FORMATIONS

In this section I characterise equilibria based on network structure, symmetric geographical placement of countries, optimal prices given trade costs and traded quantities. I analyse equilibria for a two country world and a three country world. The two country case does not differ from the Krugman (1980) trade model of monopolistic competition. It yields the home market effect including also network. In order to proceed to a higher number of trading countries I must make a strong assumption that the home market effect is normalised to 1. In this respect, all countries and number of firms become identical. I then show in equilibrium, that when firms are exposed to an exogenous cost shock they decide to switch in the short term their formation of links. Otherwise they will exit the export market. Other network-specific notions such as stability and efficiency are relegated to the appendix. These notions are not intrinsic to the inferential properties of trade costs on trade flows. They serve to check whether the network will be stable depending on the type of links formed.

Two Country Equilibrium — The equilibrium is characterised by the zero profit condition across two countries due to free entry and exit of firms within each country. The net benefit term associated with the two countries becomes $v_{12} = \delta - c = v_{21}$ because of symmetry of the two direct links g_{12} and g_{21}. Hence I impose $\pi_1 = \pi_2 = 0$:

$$p_{12}q_{12} - \frac{\gamma}{v_{12}}q_{12} - F_1 = p_{21}q_{21} - \frac{\gamma}{v_{21}}q_{21} - F_2 = 0 \tag{4.1}$$

Given that fixed costs of production are equal and countries differ only in their size, this relationship can be written compactly as:

$$q_{12}(p_{12} - \frac{\gamma}{v_{12}}) = q_{21}(p_{21} - \frac{\gamma}{v_{21}}) \tag{4.2}$$

Domestic prices are equal across countries as γ is a common constant. Prices abroad differ only by the net benefit term which is symmetric. By expanding the price indices in country 1 and 2 given domestic prices and prices abroad, in equilibrium I yield the home market effect:

$$\frac{n_1}{n_2} = \frac{\frac{L_1}{L_2} - (\delta - c)^{\sigma-1}}{1 - \frac{L_1}{L_2}(\delta - c)^{\sigma-1}} > 0 \tag{4.3}$$

Introducing a network leaves things unchanged in the standard model of trade with two countries. Yet a notable remark is that the decision to export entails an additional inherent condition for the firm. Provided the cost of transport will never exceed the benefit and as long as profits cover the fixed costs, the firm will always be favourable towards establishing a link.

The link is beneficial for society as utility increased due to the greater number of varieties available to consumers. Clearly there is autarky when $c > \delta$ and the benefit is greater the more proximal countries 1 and 2 are. The equilibrium outcome is if the relative size of country 1 increases there is a more than proportional increase in the relative number of domestic firms given the net benefit of forming a direct link. The condition holds as long as $(\delta - c)^{\sigma-1} < \frac{L_1}{L_2} < (\delta - c)^{1-\sigma}$ and n_1 and n_2 are non-zero.

Three Country Equilibrium — Similarly to Krugman (1993), in order to proceed to the three country example I need to make a strong simplifying assumption that will enable me to construct the equilibrium: That is, all countries have the same size $L_1, L_2, L_3 = L$ so that the number of firms is also equalised: $n_1, n_2, n_3 = n$. Essentially, I am assuming the home market effect between any two trading partners in a three country world is normalised to 1. In relationship (4.3) this would be equivalent to stating that the ratio $L_1/L_2 = 1$. Then it must be that $n_1/n_2 = 1$.

The impact of the network structure becomes apparent when a firm has to consider whether it will form a direct link or an indirect link to a trading partner. In order to form an indirect link with one partner it must have implicitly formed a direct link with another. Exploiting the symmetry assumption any decision that a firm in country 1 may make is an equivalent decision for a firm that could be in the complement countries. I focus on the decision of a firm in country 1 that has the option to trade directly with country 2 or indirectly with country 2 via country 3.

Similarly as in the two country case, the equilibrium with three countries is characterised by the zero profit condition $\pi_1 = \pi_2 = \pi_3 = 0$ irrespective of the types of links formed. The equilibrium condition is expressed for any country pair as:

$$q_{12}(p_{12} - \frac{\gamma}{v_{12}}) + q_{13}(p_{13} - \frac{\gamma}{v_{13}}) = q_{21}(p_{21} - \frac{\gamma}{v_{21}}) + q_{23}(p_{23} - \frac{\gamma}{v_{23}}) \Leftrightarrow q_{12} = q_{21},\ q_{13} = q_{23} \tag{4.4}$$

While this equilibrium may not be of particular interest, it helps us extract inferences between the two available network formations. This will occur through the differences in the price indices when alternate formations occur. I consider first the case of a network consisting only of direct links.

Direct Links Network — The net benefit term becomes $v_{ij} = \delta - c$ for all pairs $\{i, j\}$ in the network. The zero profit condition and the assumption of no home market effect naturally equalises the traded quantities across countries $q_{12} = q_{21}$. The price index that any country faces is an aggregate measure of domestic prices and imported prices given the types of links established. For country 2 for example it can be expressed as:

$$P_2 = n\left(\frac{\sigma}{\sigma - 1}\gamma\right)^{1-\sigma}\left(1 + 2(\delta - c)^{\sigma-1}\right) \tag{4.5}$$

Indirect Links Network — In the case of indirect links between countries 1 and 2 and direct links with country 3, the same equilibrium condition holds, $q_{12} = q_{21},\ q_{13} = q_{23}$. The price index with one indirect connection and one direct is written:

$$P_2 = n\left(\frac{\sigma}{\sigma - 1}\gamma\right)^{1-\sigma}\left(1 + (\delta - c)^{\sigma-1} + (\delta^2)^{\sigma-1}\right) \tag{4.6}$$

where the term δ^2 indicates the benefit from having a hub location intervening between countries 1 and 2. The two equilibria must be identical by the zero profit condition and the assumption of no home market effect. Due to the latter assumption the equilibria are points not ranges. They must collapse from the range of the two country case $\left((\delta - c)^{\sigma-1} < \frac{L_1}{L_2} < (\delta - c)^{1-\sigma}\right)$ to a single point. This single point accommodates indifference across network formations; otherwise a specific network formation would prevail and the zero profit condition would be violated. Equalising the two types of price indices that occur in equilibrium yields the condition:

$$\delta - c = \delta^2 \tag{4.7}$$

The cost level eliminates any benefit from choosing one particular formation such that the firm becomes indifferent between network formations.

4.1. *Endogenous Formations*

I now study what happens when an exogenous transportation cost shock occurs. Here I am required to make another strong simplifying assumption, that if a shock occurs, the number of firms or county sizes do not adjust simultaneously. They do so but with a lag. What a firm can do in the short run to cover its losses is to switch the formation or leave the export market.

Suppose that a high cost shock appears between country 1 and 2. The two countries could continue trading directly. The profits for a firm in country 1 trading with 2 and 3 are (notation D denotes a direct links network):

$$\pi_1^D = q_{12}^D(p_{12}^D - \frac{\gamma}{\delta - c}) + q_{13}^D(p_{13}^D - \frac{\gamma}{\delta - c}) \tag{4.8}$$

Whilst the profits for the same firm if it chose to bypass the direct connection and trade indirectly with country 2 using country 3 as a hub (notation I denotes a network with one indirect link):

$$\pi_1^I = q_{12}^I(p_{12}^I - \frac{\gamma}{\delta^2}) + q_{13}^I(p_{13}^I - \frac{\gamma}{\delta - c}) \tag{4.9}$$

I derive the endogenous decision of the firm to change network formation by minimising its losses. The indirect network formation will prevail if the cost from forming a direct link is very high. Then the firm may decide to sever the direct link and begin trading indirectly. In this way it has the chance to remain in the market otherwise see its profits decrease or even exit the market. I set the equilibrium condition to $0 = \pi_1^I > \pi_1^D$ to test when will the indirect links network formation prevail. The effect works through the price index of country 2 and the difference across export prices given the two network types. Solving this inequality yields the simple relationship $\delta - c \leq \delta^2$. Given a high exogenous transportation cost, it is more sensible for the firm to choose a hub network formation. It minimises each country's exposure to transport costs: the indirect links network has the minimal number of links that connects all countries. The average path length is minimised given the minimal number of links.

Alternatively, given a very low exogenous transportation cost shock, the direct network formation will prevail. A firm will then suffer losses if it is trading indirectly. I now set the equilibrium condition to $0 = \pi_1^D > \pi_1^I$ to test when will the direct links network formation prevail. As expected, it gives the simple solution $\delta - c \geq \delta^2$. When link costs are low enough it becomes even beneficial to form all direct links. The cost of adding a link is less than the benefit the firm gains from shortening the link of length two (δ^2) into a link of length one.

When the exogenous cost is extremely high, none of these formations should prevail. The countries become autarkic. The equilibrium condition then requires that firm profits are negative for both formations simultaneously. The condition $c > \delta + \delta^2$ prevents indirect trading. Similarly the condition $c > \delta$ prevents direct trading. For the equilibrium condition to be autarky for all partners (due to symmetry) it must be that simultaneously both of the above statements are true. This holds when $c > \delta + \frac{\delta^2}{2}$.

These three conditions coincide with Proposition 1 of Jackson and Wolinsky (1996). I summarise them as follows. The network formation decisions for a representative firm in the symmetric trade

model with increasing returns are:

i. A direct links network, if $c \leq \delta - \delta^2$.
ii. A hub with indirect links, if $\delta - \delta^2 \leq c \leq \delta + \frac{\delta^2}{2}$.
iii. Autarky if $\delta + \frac{\delta^2}{2} < c$.

These formations are *uniquely efficient* in the sense that each case is a prevailing case and no other network can accommodate higher profits. While I develop the concepts of network efficiency and stability in the appendix, I analyse here some simple outcomes of this proposition. If costs are forbidding it does not make sense for a firm (or for a consumer at the receiving end) to proceed with trading (consuming) a specific variety. The empty network, or autarky is the only efficient outcome of the three country problem. If costs are high but less than the autarkic level, the only efficient network is the hub network. Autarky would have lower utility levels and direct links would give lower profits for the firm and lower consumption. For sufficiently low transportation costs, the cost of adding an extra link is less than the firm's gain from replacing an indirect link to a direct link. And so it will always prefer to have a direct link at these costs. The same applies for the consumer.

The hub network is *stable* for cost values consistent in the range of $\delta - \delta^2 \leq c$: country 3 being in the center, becomes worse off if a link is severed since utility for consumers there decreases. A firm in country 1 similarly is adversely affected by this choice. The indirect link is severed and the varieties traded decrease. Profits for the firm decrease. Therefore a firm will never choose to sever the link with country 3. Suppose also that a firm in country 1 forms a direct link at this cost level with country 2 instead of the indirect link via country 3. Profits from this configuration become less and thus the firm will never choose to do so. If it actually did, a firm in country 2 would have to sever another direct link with country 3 due to the high cost of maintenance. Thus the hub formation is *pairwise stable* but not necessarily unique as it can also rotate between countries. For lower transportation costs all direct connections are pairwise stable as no country would be willing to sever a link. Therefore any two countries which are not directly connected benefit from forming a link (Jackson 2003).

This approach develops a very simple economic concept. Contrary to the Krugman (1993) three country trade model, countries which are not necessarily benefited from concentration of production, possibly created by historical incidence, can yield hub network formations as well. This arises by incidence merely of geographical placement and as a form of a hedge. In the Krugman model of trade with three countries the equilibrium arises by postulating concentration of production and a defecting firm to survey other countries' production possibilities. Instead herein I can simply postulate an excessively large transportation cost and start to decrease it. At some autarkic liberating level, the profit of a firm producing only domestic goods can be increased. This happens when there is a benefit from entering the export market. It decides to export, due to the positive profit margin. But it also decides the formation that minimises exposure to the exceedingly high costs it faces. The network formation will be a hub configuration with indirect links.

Then more firms enter production, after labour migrating from the numeraire sector, due to the positive profit margin in manufactures. More firms entering decrease the sectoral profits and through decreasing transportation costs, the power of a hub formation gradually dwindles. The net benefit of an indirect links network weakens. When the number of firms in the world is such that profits are driven to zero, for a given cost, the firm is indifferent whether to choose a particular formation. But if an exogenous very low cost shock is imposed it will decide the formation that is minimising its exposure to the adversity. Direct links are formed. This is the short term effect. Any deviation from the equilibrium given a direct links network leaves a positive profit margin. More firms enter and they

restore the equilibrium condition to the zero profit. Indifference between network formations is also restored. The reverse holds true when there is an exogenous high transportation cost shock. The firm forms indirect links to hedge itself. The positive profit margin allows firms to enter driving global profits to zero and restoring the equilibrium once more.

5. AGGREGATE TRADE AND THE GRAVITY EQUATION

In this section I derive the general case for aggregate trade in the manufacturing sector of a country. This represents the traditional trade model of representative firms. However, in the place of the iceberg transportation costs that the model commands, I will present the net benefit term and implicitly the network structure. In order to yield a testable prediction however I need to replace the unmeasurable benefit of forming a link given the network structure. Instead, I implement the changes in fixed costs associated with the network and distance. I then develop the trade-off between the two network types. If a hub network prevails, fixed costs are reduced because the network participation share for a trading partner is decreased. But distance, the proxy for transportation costs increases. This setting will be preferred against a network formation with direct links if exports under a hub are higher. Otherwise the opposite should hold. For direct links distance is less but fixed costs are higher. I then proceed to test empirically whether there is evidence of a hub network formation in aggregate exports.

For a particular sector, the value of exports from i to j of all firms belonging in this sector is $x_{ij} = n_i p_{ij} q_{ij}$ or,

$$x_{ij} = L_i \mu_j L_j \frac{v_{ij}^{\sigma-1}}{\theta_j} \frac{1}{\sigma F_i}, \text{ where } \theta_j = \sum_{j:l=1}^{N} n_l v_{lj}^{\sigma-1} \tag{5.1}$$

and θ_j is an aggregate index of network costs in j, derived from the price index. Since I cannot measure the net benefit from forming a link, I decompose the problem into two parts. First I replace the transportation cost proxy with the distance between i and j. For this I assume that transportation costs are of the form $c_{ij} = d_{ij}^{\beta} * exp(\beta_0)$. Second for every network formation there are fixed costs associated with the network. These fixed costs burden all firms in a specific country. The fixed costs are proportional to the share of each country's participation in the network.

Therefore, the gravity equation of exports between two countries involving also a hub will incorporate an increase in distance. It will also involve a reduction in fixed costs by lowering country $i's$ network participation share. But if there is no hub involved, the gravity equation is just the standard outcome of the trade model with representative firms. If there is a hub formation present for a bilateral trading relationship then it must be because aggregate exports are higher. Otherwise bilateral trade must be conducted directly, without the presence of a hub.

If there is a hub involved between two countries trading, it must also imply that transportation costs were very high thus preventing direct connections. If there is no hub involved it must be that transportation costs are very low between these trading partners. Suppose that the hub country is k. Country i trades with j via country k. I denote x_{ij}^I the aggregate exports using indirect links, or the hub country k. If country k is not involved I denote x_{ij}^D the aggregate exports using direct links. Taking logs the gravity equation becomes:

$$\ln x_{ij}^{D} = \beta_0 + \ln L_i + \ln \mu L_j - \beta_1(\sigma - 1)\ln d_{ij} - \beta_2 \ln F_i^{D} - \ln(\sigma\theta_j) \tag{5.2}$$

and,

$$\ln x_{ij}^{I} = \beta_0 + \ln L_i + \mu \ln L_j - \tilde{\beta}_1(\sigma - 1)\ln(d_{ik} + d_{kj}) - \tilde{\beta}_2 \ln F_i^{I} - \ln(\sigma\theta_j) \tag{5.3}$$

I can control for country size and the impact of relative prices, herein the aggregate network cost indices θ_j, by using country fixed effects (Chaney 2005). Subsequently, a hub formation will exist when $\ln x_{ij}^{I} \geq \ln x_{ij}^{D}$. Two conditions must then hold:

i. Ceteris paribus, $-\tilde{\beta}_1 \ln(d_{ik} + d_{kj}) \geq -\beta_1 \ln d_{ij}$ for the variable trade costs across formations.
ii. Ceteris paribus, $-\tilde{\beta}_2 \ln F_i^{I} \geq -\beta_2 \ln F_i^{D}$ for the fixed costs associated with the network.

The first condition is derived given that the elasticity of substitution is constant and greater than 1 for a particular sector. The condition can be reorganised as:

$$\frac{\beta_1}{\tilde{\beta}_1} \geq \frac{\ln(d_{ik} + d_{ij})}{\ln d_{ij}} \tag{5.4}$$

The right hand side can be at least 1. Thus the coefficient of indirect distance must be at least less than the coefficient of direct distance, $\tilde{\beta}_1 \leq \beta_1$.

The reduction of fixed costs ceteris paribus, must also yield similar results:

$$\frac{\beta_2}{\tilde{\beta}_2} \geq \frac{\ln F_i^{I}}{\ln F_i^{D}} \tag{5.5}$$

The right hand side of this relationship can be at most 1 as there is a percentage *reduction* in the network participation share for the country trading under a hub. This implies that the coefficient of fixed costs under a network involving a hub must be less than the coefficient under direct links and constrained by the right hand percentage difference can obtain some values greater than β_2.

Conversely, a direct connection will exist if $\ln x_{ij}^{I} < \ln x_{ij}^{D}$. Subject to a constant and greater than unity elasticity of substitution, the two conditions are hence reversed:

i. Ceteris paribus, $-\tilde{\beta}_1 \ln(d_{ik} + d_{kj}) < -\beta_1 \ln d_{ij}$ for the variable trade costs across formations.
ii. Ceteris paribus, $-\tilde{\beta}_2 \ln F_i^{I} < -\beta_2 \ln F_i^{D}$ for the fixed costs associated with the network.

Reorganising the first condition yields:

$$\frac{\beta_1}{\tilde{\beta}_1} < \frac{\ln(d_{ik} + d_{kj})}{\ln d_{ij}} \tag{5.6}$$

As the right hand side is greater than 1, the coefficient of indirect distance must be greater than the coefficient of direct distance, $\tilde{\beta}_1 > \beta_1$.

Straightforwardly, the same applies for fixed costs, where the right hand side is less than unity and so $\tilde{\beta}_2$ must be sufficiently larger than β_2:

$$\frac{\beta_2}{\tilde{\beta}_2} < \frac{\ln F_i^I}{\ln F_i^D} \tag{5.7}$$

I govern the decision to trade via hub country k with a binary indicator function.

$$H(k) = \begin{cases} 1 & \text{trade via country } k \\ 0 & \text{trade directly} \end{cases} \tag{5.8}$$

Therefore trading via a hub implies the indicator is one and subsequently it must be that $\ln x_{ij}^I \geq \ln x_{ij}^D$ and vice versa. I can then summarise the conditions that must hold in order to deduce evidence of the existence of hub formations:

$$\text{if } H(k) = \begin{cases} k = 1 & \text{then } \dfrac{\beta_1}{\tilde{\beta}_1} \geq \dfrac{\ln(d_{ik} + d_{kj})}{\ln d_{ij}},\ \dfrac{\beta_2}{\tilde{\beta}_2} \geq \dfrac{\ln F_i^I}{\ln F_i^D} \\ k = 0 & \text{then } \dfrac{\beta_1}{\tilde{\beta}_1} < \dfrac{\ln(d_{ik} + d_{kj})}{\ln d_{ij}},\ \dfrac{\beta_2}{\tilde{\beta}_2} < \dfrac{\ln F_i^I}{\ln F_i^D} \end{cases} \tag{5.9}$$

These two conditions constitute the specification for empirically testing whether hub formations exist in international trade. If a hub formation exists then the coefficient on indirect distance via the hub is sufficiently smaller than the coefficient of direct distance. In addition, the coefficient of fixed costs associated with a hub network must be smaller (and for a small range range greater) than the coefficient of fixed costs associated with a network with direct links. If both these conditions hold, then I conjecture that the finding is attributed to the existence of a network structure in international trade flows.

6. EMPIRICAL STRATEGY AND ENDOGENEITY

To apply the model to the data, I calibrate the two gravity equations to yield the empirical prediction. My aim, dictated by the model, is to show that the ratio of the direct to indirect distance coefficients is weakly greater than the increase in distance when the hub indicator becomes one. For fixed costs, I am required to show that the ratio of direct to indirect fixed costs coefficients is greater than the reduction in fixed costs induced by a hub network formation. Conversely, the ratio of direct to indirect distance coefficients is less than the increase in distance when the hub indicator is zero, with the same applying for reductions in fixed costs. If this prediction holds, I can make inference about the existence of hub formations, or network formations in international trade flows.

However, a subtle point of the analysis so far is that distance and fixed costs are in fact, endogenous. The causal relationship runs from the regressors to the independent variable in the sense that longer distance, higher fixed costs dampen aggregate trade flows. This is what standard trade theory predicts. Except for this channel of causality I have shown theoretically another channel of causality. That the firm decides in the interest of increasing export levels to switch the network formation and thus distance and fixed costs. If $\ln x_{ij}^I > \ln x_{ij}^D$ then it is because a hub network formation was preferred and vice versa. If all firms in the sector decide to trade via a hub, they choose implicitly to change the trading route and fixed costs to export. Firms always choose the route that will guarantee higher

export flows. So the causality in this case runs from the dependent variable to the regressors of distance and fixed costs. This is a case of simultaneous causality.

The reverse causal relationship is solely attributed to the network. So my initial estimators for distance and fixed costs will be biased and inconsistent and thus require appropriate instrumentation. For this reason I proceed as follows. I define the empirical prediction and test my hypothesis for the existence of hub network formations in international trade flows. I then isolate the effect of the network by instrumenting distance and fixed costs. Then I re-test my hypothesis. If the instruments are capturing the network effect, the hypothesis should fail.

I write the gravity equation for aggregated goods in a particular sector as:

$$\ln x_{ij}^D = \beta_0 + \beta_1 \ln L_i + \beta_2 \ln \mu_j L_j + \beta_3(1-\sigma)\ln d_{ij} + \beta_4 \ln F_i^D + \beta_5 \ln(\sigma\theta_j) + \zeta_{ij} \tag{6.1}$$

and,

$$\ln x_{ij}^I = \beta_0 + \beta_1 \ln L_i + \beta_2 \ln \mu_j L_j + \tilde{\beta}_3(1-\sigma)\ln(d_{ik}+d_{kj}) + \tilde{\beta}_4 \ln F_i^I + \beta_5 \ln(\sigma\theta_j) + \xi_{ij} \tag{6.2}$$

I estimate both these specifications with OLS. For the moment I assume that ζ_{ij}, ξ_{ij} are both orthogonal to the independent variables and normally distributed. I will later relax this assumption due to the presence of endogeneity. I retain throughout the assumption that the shocks ζ_{ij}, ξ_{ij} affect trade flows within each country pair and so all observations are clustered at this level.

In the data I can only observe one value of export volumes. This value may correspond to trade flows using direct connections and/or some network structure involving a hub location. I cannot observe the differences between $\ln x_{ij}^D$ and $\ln x_{ij}^I$. So my dependent variable is the same in both estimations. Therefore I just test how the change in the indicator affects the ratio of coefficients of distance and fixed costs. I also check if the fit deteriorates or improves in each estimation. If a hub network formation prevails then the coefficient of indirect distance should be decreased compared to the coefficient of direct distance. The same applies for fixed costs with the exception that the coefficient of indirect fixed costs cannot be much larger than the coefficient of direct fixed costs. And the reverse must hold true. Given that the negative signs of trade costs coefficients, the prediction for existence of hub formations becomes:

Empirical Prediction

$$\text{if } H(k) = \begin{cases} k=1 & \text{then } \dfrac{\beta_3}{\tilde{\beta}_3} \geq \dfrac{\ln(d_{ik}+d_{kj})}{\ln d_{ij}},\ \dfrac{\beta_4}{\tilde{\beta}_4} \geq \dfrac{\ln F_i^I}{\ln F_i^D} \\ k=0 & \text{then } \dfrac{\beta_3}{\tilde{\beta}_3} < \dfrac{\ln(d_{ik}+d_{kj})}{\ln d_{ij}},\ \dfrac{\beta_4}{\tilde{\beta}_4} < \dfrac{\ln F_i^I}{\ln F_i^D} \end{cases} \tag{6.3}$$

For distance, I expect $\beta_3, \tilde{\beta}_3 < 0$. When $h = 1$, I also expect $\tilde{\beta}_3 < \beta_3$. When $h = 0$, I expect $\tilde{\beta}_3 > \beta_3$. The same application should hold for fixed costs. When $h = 1$, I anticipate $\tilde{\beta}_4 < \beta_4$ or $\tilde{\beta}_4 \geq \beta_4$ by a small amount. When the hub indicator is 1 I should observe $\tilde{\beta}_4 >> \beta_4$.

If both these findings occur, I can infer that the dependent variable, export flows of a particular sector, corresponds to trading using a hub formation or direct trading and then this is proof of a network structure in international trade flows. I can also infer that distance and fixed costs are endogenous due

to the decision of the firms to export in that sector using a particular network structure.

Lastly, I will control for heterogeneity and for the price index with fixed effects. I introduce country dummies, year dummies, industry, sector and transport mode dummies.

7. CONSTRUCTION OF THE DATASET

The dataset is constructed by merging data from various available sources. Exports from 2003 to 2007 are obtained through the World Bank WITS interface and the classification level is 2 digit HS 1988/1992 for all possible trading partners. I perceive that goods under the umbrella of each 2 digit aggregation forms a distinct sector. These observations are matched with corresponding data on ad-valorem maritime transport costs of the same classification level and year. Observations for maritime transport costs as well as the mode of transport are available under subscription to the OECD Maritime Transport Costs Database. The justification for such an operation is twofold. First, according to Hummels (2007), 99% of the world's trade by weight is carried by sea and based on the OECD (2008) 90% of the volume of world trade is transported by sea. Second with the existence of hub and spoke networks induced by the advent of containerisation I postulate that the observed price of shipping services, i.e. transport costs, is a function of the network organisation of the transportation sector. By conducting this operation I am able to capture some of the global export volume that has been transported by sea and therefore by some form of network.

All observations not having matching exports and transport costs are removed. Each surviving trade partnership is assigned nominal GDP values, two measures of distance, border and language characteristics. GDP values are obtained through the World Bank Databank. Information on capital distances, border and languages come from the Trade, Production, and Protection Database compiled by Nicita and Olarreaga and the CEPII GeoDist dataset compiled by Mayer and Zignago.

I construct a measure of indirect distance in addition to capital distance. For a subset of trading partners I construct observations by measuring the distance from the capital of the exporter to the closest major exporting port. The exporting port is located within the country of export and for the cases of landlocked countries I consider the closest major foreign port through which it may proceed to export. I then measure trading partner port distances using the U.S. National Imagery and Mapping Agency Distances Between Ports publication as well as the online resource Port World Distance Calculator. For each particular partnership, listed in table B-IV, the main country ports of origin and destination are assessed and the shortest shipping route is calculated. If the shipping route requires passage through any of the below exogenously imposed "*hub*" areas, the route is assigned an indicator value equal to one or otherwise zero. No such distinction is made for direct capital distance. These areas are the Panama Canal, the port of Gibraltar, Port Said, Singapore, Cape Town, Istanbul and the port of Arica. Finally I measure the distance from the major importing port to the capital. To provide an example of this construction, the direct distance between Beijing and Brasilia is 16,937 km. I then measure the distance from Beijing to Shanghai (1,267 km), and from Shanghai to Singapore (3,934 km), where the indicator is assigned a value of 1. I then add the distance from Singapore to Rio (16,366 km) and Rio to Brasilia (1,160 km). The observation for indirect distance is 22,727 km.

The data come with a weakness in the sense that transport costs for individual E.U. 15 countries are not observed. For this reason, and when otherwise not available, all other units are aggregated to provide approximations at the 15 country level. The capital distance of the E.U. 15 entity with the rest of the world is then measured from Brussels and its main export/import port becomes Rotterdam.

8. EXISTENCE OF HUB NETWORKS IN INTERNATIONAL TRADE

The theoretical approach herein asserts that geographical characteristics now become endogenous elements as countries open to trade: in order to protect themselves from adverse transport costs they create a hub network and can determine hub location either by concentration of production and historical precedence or simply by geographical advantage. This helps to explain why neighbouring countries have more integrated networks than countries which are not, and infrastructure crucially determines country access and transit facilitation (Limao & Venables 2001), (Krugman 1993).

I estimate the empirical specifications robust to heteroscedastic clustered standard errors in the presence of country, year and sectoral, industrial and transport mode fixed effects. I find that (6.3) holds for the distance variables, when controlling for all effects except sectoral. When I control for heterogeneity across sectors the effect is reversed and (6.3) fails for distance.

I cannot make any conclusive inference about the coefficients for fixed costs as the border and language variables act in the same manner as the religion variable in Helpman et al. (2008). They assume that common religion affects the fixed cost of export but not the variable cost of export. This assumption must therefore hold for common language and border, however Anderson (2010) mentions the potential identification issue that this approach may entail.

Most importantly, the reason why these are not good proxies to the theoretical model is that they do not capture the change in costs due to the change in the network. Suitable variables to extract inference concerning fixed costs would be customs data such as exporting costs per good type per destination but are widely unavailable. While country specific export and import costs are available from the World Bank, they are absorbed by the country fixed effects. The lack of suitable proxies for fixed costs remains a challenge for the validation of the second half of the empirical prediction (6.3) and may necessitate structural estimation. Hence, for the remainder of this version of the paper only the variable cost prediction is validated and treated for endogeneity: half of the empirical prediction is proved due to lack of suitable data.

I now proceed to show the evidence of hub formations in international trade. I first show the negative relationship stemming from the two measures of distance. I then present the results that lead to the construction of (6.3).

The first set of results in Table I is the outcome of (6.1) and (6.2) which are represented in columns (1) and (2) respectively. I control for year, sectors within an industry, the industry itself and the transport mode, and apply no control on country size. An industry is defined in the OECD dataset as an aggregate of sectors that fall under the umbrella of Agriculture, Manufacturing, Crude Oils and Raw Materials[2]. Given these controls, I then assume that aggregate prices are incorporated in the orthogonal error term (Anderson & van Wincoop 2004). For these estimations I make no assumption about the presence of a hub indicator. The results exhibit the typical adversity caused by direct distance. They also exhibit an increased adversity caused by indirect distance. Since on average indirect distance is 3,773 km longer than direct distance, I expect, and observe, a higher coefficient for indirect distance than direct distance for the same sample. This result is straightforward: given two measures of distance, one indirect and one direct, with the former being on average greater than the latter, an exporter will choose the shorter distance. Both coefficients of distance fall within the range of the observed estimates in the trade literature (Overman et al. 2001) and the same applies for the elasticities of GDP being ~ 1. The common border variable enters also with a typical observed elasticity (see Limao & Venables

[2]For the detailed industrial composition of sectors see Appendix Table B-III.

TABLE I

EXPORTS: DIRECT AND INDIRECT DISTANCE

Variable	(1) Exports	(2) Exports
Direct Distance	-0.957***	
	(0.143)	
Indirect Distance		-1.081***
		(0.198)
Common Language	1.130***	1.040***
	(0.199)	(0.217)
Common Border	0.958**	1.185***
	(0.461)	(0.388)
Exporter GDP	1.058***	1.093***
	(0.0341)	(0.0358)
Importer GDP	0.836***	0.905***
	(0.0366)	(0.0366)
Fixed Effects		
Exporter	-	-
Importer	-	-
Sector	YES	YES
Year	YES	YES
Mode	YES	YES
Industry	YES	YES
Observations	73,237	73,237
R-squared	0.521	0.517

Pooled OLS estimation for years 2003 – 2007, monetary units are real U.S. Dollars. Standard errors are allowed to be correlated within country pair clusters. Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Constants included but not reported.

(2001) and Helpman et al. (2008)) but the language coefficient is perceived as very high.

For these two specifications I now proceed to incorporate the hub indicator. That is, I test relationship (6.1) employing direct distance when the indicator is zero and one and I do exactly the same for (6.2) which employs indirect distance. In this case I control for country size and the relative prices by including country dummies (Chaney 2005). I control for the transport mode and the industry but allow for sectoral changes.

Tables II and III outline the results for this experiment. Let us assume that we are in a world with direct connections. That is the hub indicator is zero. Therefore if everybody is trading directly it is because transportation costs are lower than indirect trading according to our theoretical treatise. The first two columns of Table II indicate this effect. Direct distance impacts trade flows 24% less than trading via a hub country. That is for trade flows in the data whose indirect distance does not involve transit via a hub, exporters should (*and appear to*), utilise the shortest possible distance to the importer. The situation is reversed when the hub indicator is turned to one. Now we are assuming that

partners are trading indirectly because trade costs are less in this regime compared to the alternative. Columns (3) and (4) reveal the reversed impact of the two measures of distance. For observed trade flows that pass through a hub location, utilising direct routing is considered 5% more costly than trading indirectly. This holds despite the fact that indirect distance is greater than direct distance by approximately 36%. This somewhat counter intuitive result seems to confirm the theoretical prediction of this study. Whenever a geographically advantageous country exists that helps to link two relatively remoter countries, partners should (*and appear to*) avoid direct trading due to the higher costs involved.

TABLE II

EXPORTS: IMPACT OF THE HUB INDICATOR

Variable	(1) Exports Hub=0	(2) Exports Hub=0	(3) Exports Hub=1	(4) Exports Hub=1
Direct Distance	-1.430***		-0.925***	
	(0.113)		(0.268)	
Indirect Distance		-1.877***		-0.881***
		(0.249)		(0.278)
Common Language	0.971***	0.815***	0.696***	0.671***
	(0.198)	(0.276)	(0.147)	(0.139)
Common Border	1.329***	1.154*	-1.104**	-0.874*
	(0.469)	(0.690)	(0.513)	(0.512)
Fixed Effects				
Exporter	YES	YES	YES	YES
Importer	YES	YES	YES	YES
Sector	-	-	-	-
Year	YES	YES	YES	YES
Mode	YES	YES	YES	YES
Industry	YES	YES	YES	YES
Observations	41,481	41,481	32,055	32,055
R-squared	0.413	0.402	0.465	0.465

Pooled OLS estimation for years 2003 − 2007, monetary units are real U.S. Dollars. Standard errors are allowed to be correlated within country pair clusters. Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Constants included but not reported.

TABLE III

EMPIRICAL PREDICTION

	$\frac{\beta_3}{\bar{\beta}_3}$		$\frac{\ln(d_{ik}+d_{kj})}{\ln d_{ij}}$
h=1	1.05	>	1.03 (mean)
h=0	0.76	<	1.03 (mean)

Taking the ratios of the distance coefficients and comparing them to the mean of the ratio of indirect to direct distance we confirm the empirical prediction (6.3) as shown in Table III. This result seems to hold when I allow for sectoral variation in the data. However when I control for sectors by incorporating a dummy for each of the 96+1 HS 2 digit aggregates the result breaks down as is shown in Table IV.

When adding sectoral fixed effects, indirect distance always seems to be costlier to exporters compared to direct distance irrelevant of the value of the hub indicator. This interesting finding implies that the impact of the hub indicator is driven through heterogeneity existing across sectors, given the mode of transport for each commodity produced. It is equivalent to stating that when controlling for sectoral differences by virtue of the fixed effect, any unobserved information implying that some sectors within each country pair trade by using hub formations is subsumed by the fixed effect.

TABLE IV

EXPORTS: IMPACT OF THE HUB INDICATOR

Variable	(1) Exports Hub=0	(2) Exports Hub=0	(3) Exports Hub=1	(4) Exports Hub=1
Direct Distance	-1.482***		-1.031***	
	(0.117)		(0.241)	
Indirect Distance		-2.080***		-1.101***
		(0.242)		(0.229)
Common Language	1.329***	1.106***	0.478***	0.454***
	(0.229)	(0.278)	(0.132)	(0.114)
Common Border	1.403***	1.123	-1.017**	-0.855*
	(0.506)	(0.700)	(0.456)	(0.461)
Fixed Effects				
Exporter	YES	YES	YES	YES
Importer	YES	YES	YES	YES
Sector	YES	YES	YES	YES
Year	YES	YES	YES	YES
Mode	YES	YES	YES	YES
Industry	YES	YES	YES	YES
Observations	41,481	41,481	32,055	32,055
R-squared	0.615	0.607	0.627	0.627

Pooled OLS estimation for years 2003 − 2007, monetary units are real U.S. Dollars. Standard errors are allowed to be correlated within country pair clusters. Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Constants included but not reported.

TABLE V

EMPIRICAL PREDICTION

	$\frac{\beta_3}{\widetilde{\beta}_3}$		$\frac{\ln(d_{ik}+d_{kj})}{\ln d_{ij}}$
h=1	0.93	<	1.03 (mean)
h=0	0.71	<	1.03 (mean)

Grouping observations within each industrial aggregate, we can observe how sectoral differences affect the impact of the hub indicator from the viewpoint of agricultural/perishable goods, or differentiated manufactured products. While the former goods would accommodate modes of transport or transport networks to be dispatched to the destination directly, or in the least possible time, the same may not

apply for the latter. When I relax the sector fixed effects and repeat the estimation at the industry level for agriculture and manufactures, I observe the empirical prediction in manufactures but not in agriculture. I list below the summaries of the empirical prediction for these two broad aggregates. I relegate the output tables together with the sectors contained within each industry to the appendix.

	$\frac{\beta_3}{\tilde{\beta}_3}$		$\frac{\ln(d_{ik}+d_{kj})}{\ln d_{ij}}$
h=1	0.93	$<$	1.03 (mean)
h=0	0.69	$<$	1.03 (mean)

(a) Agriculture - Empirical Prediction

	$\frac{\beta_3}{\tilde{\beta}_3}$		$\frac{\ln(d_{ik}+d_{kj})}{\ln d_{ij}}$
h=1	1.14	$>$	1.03 (mean)
h=0	0.79	$<$	1.03 (mean)

(b) Manufactures - Empirical Prediction

Summarising these results, I confirm, at least for variable trade costs, the existence of hub network formations in international trade flows. These formations appear to be driven mainly by the characteristic nature of the types of goods produced by each sector. The reason for existence of these formations according to my theoretical framework is solely attributed to cost reductions associated with network formations. When transportation costs are high (longer distances), partners are more inclined to form hub network formations in order to increase their export volume rather than trade directly. When transportation costs are low (shorter distances) partners are increasing their export volume by trading directly rather than involving a beneficially located trading partner.

When controlling for differences across sectors the result is treated; information on the network appears to be subsumed by the fixed effect. To validate this result further, more reliable measures of fixed costs associated with export flows are required. Specifically measurements of fixed costs related to the network structure of transportation be it direct trading or transit costs at hub locations are desirable, however these costs are unavailable at present.

8.1. ***Addressing Endogeneity***

The decision to trade using a hub network or a direct connection is made by the exporting firm based on its observations of transportation costs. It will prefer the formation that allows for the greatest exporting volume to be produced. For this reason exists a reverse channel of causality which dictates my estimates for the two distance variables are endogenous. The endogeneity is caused by the effect of the network formation. And so if a candidate instrument that eradicates the network effect but retains the standard causal relationship exists, it will also invalidate straightforwardly the empirical prediction (6.3). I am selecting an instrument that should be impervious to the fluctuations caused by the hub indicator but retain the information of the impact of variable trade costs on trade flows.

A candidate instrument satisfying relevance and exogeneity criteria is time difference. I assume that firms in a country export to other countries which have small time differences rather than large time differences. This has not so much to do with the actual distance that goods have to travel to reach the destination, but is related with the relative ease of building trade partnerships through channels of communication.[3] I can deduce that the time barrier of trade should contain information that is relevant to distance but then it should not contain information pertaining to the network structure of the transportation sector. The current model is parsimonious to explain the effect of the network

[3]An interesting validation of this thought experiment was realised on 29 December 2011, when Samoa and Tokelau skipped a day in order to ease trade relationships with Australia and New Zealand. According to the Prime Minister of Samoa, "While it's Friday here, it's Saturday in New Zealand, and when we're at church on Sunday, they're already conducting business in Sydney and Brisbane." Source: (BBC News, 30 December 2011).

through the construction of a separate transportation sector. In reality, the firm would most probably produce specific export volumes of a good by observing a transportation cost. It would not be able to deduce or exert any power upon the route that the good will take other than commanding an arrival date to the destination. This reasoning leads to the satisfaction of the relevance and exogeneity criteria for the instrument.

Yet the coefficient of time difference would be inefficient as it would fail to capture variation on a vertical trade axis which it interprets as minimal. The notion of time difference can then be exploited on a two axis assumption. Since time difference on a scale from zero to 24 hours can be captured at 15 degree intervals of longitudinal lines, the same can be applied at 7.5 degree intervals of latitudinal lines. I construct measures of time difference by subtracting the longitude of the importer's capital from the longitude of the exporter's capital divisible by 15 degrees to obtain the 24 hour measure of standard time difference. I subtract the latitude of the importer's capital from the latitude of exporter's capital divisible by 7.5 degrees and obtain the 24 hour measure of vertical time difference. I then take the absolute value of each and obtain a measure of time difference expressed on a decimal scale rather than hours and minutes in order to generate sufficient variation across country pairs. I then estimate the gravity equations (6.1) and (6.2) using one overidentifying restriction for both direct and indirect distances. If both instruments are valid, I test for relevance and exogeneity. If exogeneity fails, I report the 2SLS estimate otherwise I report the OLS equivalent of Table II. If one instrument is valid I remove the vertical measure of time difference and repeat relevance and exogeneity tests using standard time difference. If exogeneity fails, I report the 2SLS equivalent otherwise the OLS distance equivalent of Table II.

The results of this experiment are contained in Tables VI and VII. Two interesting outcomes are observed.

First, I observe all estimations except (3) are exactly identified and (2) has no trace of endogeneity. The desired network configuration, which is the realisation of the firms endogenous decision, contains the endogenous variable of distance. In that sense when the hub indicator is zero, firms according to the theoretical expansion prefer direct trading compared to indirect trading. Then column (1) contains the treated coefficient for direct distance as the null of exogeneity was rejected. This does not seem to apply for indirect distance as the Hausman statistic supports exogeneity and so column (2) listing the alternative decision of the firm is unchanged. When the hub indicator is one I have shown that firms are preferring indirect trading to direct trading. Then column (4) lists the treated coefficient for indirect distance since the latter was rejected under the null of exogeneity. Column (3) however shows the alternative decision for direct trading when the hub is one, and lists the overidentified treated coefficient of direct distance. The endogenous decisions of firms in columns (1) and (4) are rejected under the null of exogeneity when allowing for variation across sectors. The alternative decisions of firms are somewhat misleading: when the hub indicator is zero, the alternative decision is exogenous; when the hub indicator is one, the alternative decision is endogenous. One possible reasoning for this finding is that perhaps the instrument is also selecting sectoral variation (since all other variation according to the model is controlled for) that is not related to distance in either or both of the two samples. Another reasoning could be identification, in the sense that the exogenous variables in the first stage may not be identified correctly or are very strongly correlated with each other.

Second, I observe that by instrumenting distance with time difference I am able to invalidate the empirical prediction (6.3). In Table VII, ratios of distance coefficients under the one and zero outcomes of the hub indicator are less than the ratios of indirect to direct distance. The implication is that the endogenous variable, when identified and treated, should contain no information about the network

TABLE VI

EXPORTS: IMPACT OF THE HUB INDICATOR

Variable	(1) 2SLS Exports Hub=0	(2) OLS Exports Hub=0	(3) 2SLS Exports Hub=1	(4) 2SLS Exports Hub=1
Direct Distance	-1.175***		-1.164***	
	(0.156)		(0.259)	
Indirect Distance		-1.877***		-1.179***
		(0.249)		(0.277)
Common Language	1.104***	0.815***	0.710***	0.681***
	(0.224)	(0.276)	(0.152)	(0.152)
Common Border	1.543***	1.154*	-1.344**	-1.112**
	(0.496)	(0.690)	(0.518)	(0.533)
Fixed Effects				
Exporter	YES	YES	YES	YES
Importer	YES	YES	YES	YES
Sector	-	-	-	-
Year	YES	YES	YES	YES
Mode	YES	YES	YES	YES
Industry	YES	YES	YES	YES
Observations	41,481	41,481	32,055	32,055
Overidentifying Restrictions	0	0	1	0
Robust Hausman Statistic	9.63***	0.02	4.95**	4.04**
R-squared	0.412	0.402	0.464	0.446

Pooled 2SLS/OLS estimations for years 2003 − 2007, monetary units are real U.S. Dollars. Standard errors are allowed to be correlated within country pair clusters. Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Constants included but not reported.

TABLE VII

EMPIRICAL PREDICTION

	$\frac{\beta_3}{\tilde{\beta}_3}$		$\frac{\ln(d_{ik}+d_{kj})}{\ln d_{ij}}$
h=1	0.98	<	1.03 (mean)
h=0	0.62	<	1.03 (mean)

structure, be it direct connections or a hub network since it violates the testable prediction. I am led to the conclusion that time difference is a valid instrument for harnessing the impact of network structures that might be present, but unobserved, in variable trade costs suggesting caution on the exogeneity assumptions for trade costs.

I conclude that the endogenous decision of firms to choose the network formation is evident in international trade flows. I show this by validating the empirical prediction of the theoretical model: firms choose the formation that is minimising their exposure to transportation costs, herein proxied by distances that correspond to actual measurements of transport costs. The endogeneity can be harnessed

using two tools. One can either incorporate sector fixed effects as it appears that the information on the network structure based on the types of goods produced is alleviated by the fixed effect. As an alternative, one can choose to employ an instrument for variable trade costs but only if these variable costs are suspect of containing network effects.

9. CONCLUSION

I study the reasons for formation and the impacts of transportation hub networks in international trade. I do this by merging the symmetric connections model of Jackson & Wolinsky (1996) in a trade model of monopolistic competition with representative firms. Under reasonable assumptions, I show that firms begin to export and choose the network formation that is minimising their exposure to transport costs. As such when transport costs are extremely high countries remain closed. Conditional on their gradual reduction, firms commence exporting and create hub networks that are associated with higher levels of costs. Further reduction of variable trade costs leads eventually to indifference between network formations in equilibrium. When a low cost shock occurs, firms favour direct trading until the equilibrium is restored. When a high cost shock occurs, firms turn to indirect trading until once again the equilibrium is restored.

I then yield the gravity equation of trade and develop the empirical prediction that will provide evidence on the existence of hub formations in international trade flows. Firms choose the formation that maximises the volume of output given two measures of distance, one direct and one indirect. I find that when direct connections are preferred the coefficient for direct distance is less than the coefficient of indirect distance to the same destination. I also observe that when indirect connections are preferred the coefficient of indirect distance is lower than the coefficient of direct distance. These findings carry the justification of endogeneity insofar as firms, depending on the type of goods they produce, prefer a direct or an indirect route. As such, agricultural goods seem to be always directly traded. Manufactured goods tend to be indirectly traded. This endogeneity can be treated in two ways, either by incorporating sectoral fixed effects or by instrumentation when there is suspicion that the variable trade costs may contain information on the unobserved or unknown network formation.

The reasons for these formations are currently being researched in a forthcoming theoretical and empirical model. The driving force appears to be economies of scale in transportation due to sectoral export volumes and their interaction with distance. Thus transportation costs on high-volume trading routes, irrespective of distance tend to be low. This does not affect much the productivity or number of firms. Transportation costs on low-volume trading routes, tend to be higher and distance here plays a crucial role. This should affect the productivity of exporting firms and the number of firms more severely. Additional factors are directional imbalances penalising countries which cannot provide return cargoes, costs for importing and exporting commodities and exercise of market power.

It is concluded that geographically disadvantaged countries absorbing high transport costs achieve a more beneficial trading position when forming a transportation network by utilising a hub. The link with at least one proximate geographically advantaged partner improves market access, ameliorates exposures to these costs and leads to improvements in own and transit infrastructure.

APPENDIX A: PROOF OF EFFICIENCY AND PAIRWISE STABILITY OF NETWORK FORMATIONS

A.1. *Efficiency*

Based on the definitions of Jackson & Wolinsky (1996), each graph g has a value $y : \{g|g \subset g^N\} \to \mathbb{R}$, where $y \in Y$ is the set of functions. The value is an aggregate of individual exporting firm values $y_j(g) = \sum_{i:ij \in g} v_{ij}(g)$, where $v_{ij} : \{g|g \subset g^N\} \to \mathbb{R}$ is a net benefit value. The graph $g \subset g^N$ is *efficient* if $y_j(g) \geq y_j(g')$ for all $g' \subset g^N$.

Case 1, Direct Trading vs Autarky: $G = \{g_{ij} = \{12, 21\}, g'_{ij} = \{\emptyset\}\}$— The simplest case shows how forming a direct link is more efficient than not forming a link. Firms profits and utility increases as long as the benefit of forming a link exceeds its cost for countries 1 and 2. Due to symmetry, global utility and firm profits (until the trade equilibrium restores profits to zero) increase:

$$y_1(g) = (\delta - c_{12}) = (\delta - c_{21}) > 0, \text{when } \delta > c_{ij} \tag{A-1}$$

Case 2, Direct Trading vs Indirect Trading: $G = \{g_{ij} = \{12, 21, 13, 31, 23, 32\}, g'_{ij} = \{13, 31, 23, 32\}\}$— In the case of the direct links network with three countries trading, direct connections are more efficient than indirect connections and no trading at all. I denote the graphs originating from direct trading between any integer pair $\{i, j\} \in [1, 3]$ as g^D. For pairs that are indirectly trading I denote their graph as g^I. Therefore in case (i) of Proposition 1, country 1 obtains from trading to 2 and 3:

$$y_1(g^D) = (\delta - c_{13}) + (\delta - c_{12}) \geq \delta - c_{13} + \delta^2 = y_1(g^I) \tag{A-2}$$

Country 3 similarly yields:

$$y_3(g^D) = (\delta - c_{31}) + (\delta - c_{32}) = (\delta - c_{31}) + (\delta - c_{32}) = y_3(g^I) \tag{A-3}$$

Because in the two alternate formations, country 3, being in the middle must always form direct connections. The symmetric case for country 1 applies for country 2.

Hence, total profits (or utilities in the receiving countries) have increased by $y = 6(\delta - c) > 0$.

Case 3: Indirect Trading vs Direct trading: $G = \{g_{ij} = \{13, 31, 23, 32\}, g'_{ij} = \{12, 21, 13, 31, 23, 32\}\}$—
Countries 1 and 2 are symmetric, therefore the following is obtained also for country 2.

$$y_1(g^I) = (\delta^2 + \delta - c_{13}) \geq (\delta - c_{13}) + (\delta - c_{12}) = y_1(g^D) \tag{A-4}$$

For country 3:

$$y_3(g^D) = (\delta - c_{31}) + (\delta - c_{32}) = (\delta - c_{31}) + (\delta - c_{32}) = y_3(g^I) \tag{A-5}$$

Hence, total profits (or utilities in the receiving countries) have increased by $y = 2\delta^2 + 4(\delta - c) > 0$.

A.2. *Pairwise Stability*

There exists an allocation rule $K : \{g|g \subset g^N\} \times Y \to \mathbb{R}^{\mathbb{N}}$ and $K_j(g, y_j)$ is the distribution of each network value to individual firms or representative agents. The graph is pairwise stable w.r.t. y and K if:

a. For all $ij \in g, K_j(g, y_j) \geq K_j(g - ij, y_j)$ and $K_i(g, y_i) \geq K_i(g - ij, y_i)$.
b. For all $ij \notin g, K_j(g, y_j) < K_j(g + ij, y_j)$ then $K_i(g, y_i) > K_i(g + ij, y_i)$.

The implication is that if j strictly prefers to form link ij and i is indifferent, the link is formed.

Case 1, Direct Trading $c \leq \delta - \delta^2$***:*** —

a. For $\{ij\} \in [1, 3]$: If any one link is severed, utility (profits) decrease for the trading pair: $\delta - c_{ij} > 0$ as long as $\delta > c_{ij}$.
b. For $\{ij\}$, $i, j \in \{1, 3\}$ and country 2 remaining autarkic: utility (profits) of 2 connecting to 1 or 3 are lowered: $\delta - c_{2j} < 0$. But if $\delta - c_{2j} > 0$, then utility (profits) increase and all links are formed.

Case 2, Indirect Trading $\delta - \delta^2 \leq c$***:*** —

a. For $\{ij\} \in [1, 3]$: If any one link is severed, utility (profits) decrease for the trading pair: $\delta^2 + \delta - c_{ij} > \delta - c_{ij} > 0$ for countries, 1 and 2. For country 3 connected to 1 and 2, $2(\delta - c_{3j}) > (\delta - c_{3j}) > 0$ as long as $\delta > c_{3j}$.

b. For $\{ij\},\ i,j \in \{1,3\}$ trading directly and country 2 trading directly to 3. It considers connecting directly or indirectly to 1. Utility (profits) of 2 connecting directly to 1 are lowered: $2(\delta - c_{2j}) < 0 < (\delta - c_{23}) + \delta^2$. But if it actually connected to 1 directly, country 1 has utility (profits) decreased: $2(\delta - c_{1j}) < 0 < (\delta - c_{12}) + \delta^2$ and therefore country 1 will break the link to country 3 and start trading indirectly.

It is deduced that when costs are low, $c \leq \delta - \delta^2$, the direct links network is uniquely efficient and pairwise stable. When costs are consistent with the range $\delta - \delta^2 \leq c$, the indirect links network is uniquely efficient. It is also pairwise stable but not necessarily unique as the system can rotate between partners.

APPENDIX B: RESULTS FOR BROADLY DEFINED AGRICULTURAL AND MANUFACTURING INDUSTRIES

TABLE B-I
AGRICULTURAL EXPORTS

Variable	(1) Exports Hub=0	(2) Exports Hub=0	(3) Exports Hub=1	(4) Exports Hub=1
Direct Distance	-1.333***		-0.993**	
	(0.0981)		(0.436)	
Indirect Distance		-1.909***		-1.060**
		(0.235)		(0.444)
Common Language	0.713***	0.491*	0.769**	0.713**
	(0.225)	(0.276)	(0.360)	(0.341)
Common Border	1.191***	0.913*	0.0493	0.253
	(0.387)	(0.547)	(0.677)	(0.622)
Fixed Effects				
Exporter	YES	YES	YES	YES
Importer	YES	YES	YES	YES
Sector	-	-	-	-
Year	YES	YES	YES	YES
Mode	YES	YES	YES	YES
Observations	11,148	11,148	6,985	6,985
R-squared	0.463	0.455	0.531	0.531

Pooled OLS estimations for years 2003 – 2007, monetary units are real U.S. Dollars. Standard errors are allowed to be correlated within country pair clusters. Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Constants included but not reported.

TABLE B-II

MANUFACTURED EXPORTS

Variable	(1) Exports Hub=0	(2) Exports Hub=0	(3) Exports Hub=1	(4) Exports Hub=1
Direct Distance	-1.479***		-0.992***	
	(0.130)		(0.201)	
Indirect Distance		-1.857***		-0.869***
		(0.280)		(0.243)
Common Language	1.227***	1.097***	0.710***	0.697***
	(0.207)	(0.303)	(0.182)	(0.196)
Common Border	1.416***	1.261*	-1.610***	-1.308***
	(0.485)	(0.745)	(0.451)	(0.477)
Fixed Effects				
Exporter	YES	YES	YES	YES
Importer	YES	YES	YES	YES
Sector	-	-	-	-
Year	YES	YES	YES	YES
Mode	YES	YES	YES	YES
Observations	28,238	28,238	23,875	23,875
R-squared	0.415	0.403	0.473	0.473

Pooled OLS estimations for years 2003 – 2007, monetary units are real U.S. Dollars. Standard errors are allowed to be correlated within country pair clusters. Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Constants included but not reported.

Table B-III: Industrial Composition of Commodities.

Agriculture		Crude oils		Raw Materials		Manufactures	
Commodity	Transport Mode	Commodity	Transport Mode	Commodity	Transport Mode	Commodity	Transport Mode
01: Live animals	Containers	27: Mineral fuels, oils, distillation products, etc	Dirty bulk	25: Salt, sulphur, earth, stone, plaster, lime and cement	Dirty bulk	15: Animal,vegetable fats and oils, cleavage products, et	Dirty bulk
02: Meat and edible meat offal	Containers	27: Mineral fuels, oils, distillation products, etc	Tankers	26: Ores, slag and ash	Dirty bulk	15: Animal,vegetable fats and oils, cleavage products, et	Containers
03: Fish, crustaceans, molluscs, aquatic invertebrates ne	Containers			72: Iron and steel	Dirty bulk	28: Inorganic chemicals, precious metal compound, isotope	Dirty bulk
04: Dairy products, eggs, honey, edible animal product ne	Containers					29: Organic chemicals	Dirty bulk
05: Products of animal origin, nes	Containers					30: Pharmaceutical products	Containers
06: Live trees, plants, bulbs, roots, cut flowers etc	Containers					31: Fertilizers	Dirty bulk
07: Edible vegetables and certain roots and tubers	Containers					32: Tanning, dyeing extracts, tannins, derivs,pigments et	Containers
08: Edible fruit, nuts, peel of citrus fruit, melons	Containers					33: Essential oils, perfumes, cosmetics, toiletries	Containers
09: Coffee, tea, mate and spices	Containers					34: Soaps, lubricants, waxes, candles, modelling pastes	Containers
10: Cereals	Clean bulk					35: Albuminoids, modified starches, glues, enzymes	Containers
11: Milling products, malt, starches, inulin, wheat glute	Containers					36: Explosives, pyrotechnics, matches, pyrophorics, etc	Containers
12: Oil seed, oleagic fruits, grain, seed, fruit, etc, ne	Clean bulk					37: Photographic or cinematographic goods	Containers
12: Oil seed, oleagic fruits, grain, seed, fruit, etc, ne	Containers					38: Miscellaneous chemical products	Containers
13: Lac, gums, resins, vegetable saps and extracts nes	Containers					39: Plastics and articles thereof	Containers
14: Vegetable plaiting materials, vegetable products nes	Containers					40: Rubber and articles thereof	Containers
15: Animal,vegetable fats and oils, cleavage products, et	Tankers					41: Raw hides and skins (other than furskins) and leather	Containers
15: Animal,vegetable fats and oils, cleavage products, et	Containers					42: Articles of leather, animal gut, harness, travel good	Containers
16: Meat, fish and seafood food preparations nes	Containers					43: Furskins and artificial fur, manufactures thereof	Containers
17: Sugars and sugar confectionery	Containers					44: Wood and articles of wood, wood charcoal	Containers
18: Cocoa and cocoa preparations	Containers					45: Cork and articles of cork	Containers
19: Cereal, flour, starch, milk preparations and products	Containers					46: Manufactures of plaiting material, basketwork, etc.	Containers
20: Vegetable, fruit, nut, etc food preparations	Containers					47: Pulp of wood, fibrous cellulosic material, waste etc	Containers
21: Miscellaneous edible preparations	Containers					48: Paper & paperboard, articles of pulp, paper and board	Containers
22: Beverages, spirits and vinegar	Containers					49: Printed books, newspapers, pictures etc	Containers
23: Residues, wastes of food industry, animal fodder	Containers					50: Silk	Containers
24: Tobacco and manufactured tobacco substitutes	Containers					51: Wool, animal hair, horsehair yarn and fabric thereof	Containers
						52: Cotton	Containers
						53: Vegetable textile fibres nes, paper yarn, woven fabri	Containers
						54: manmade filaments	Containers
						55: manmade staple fibres	Containers
						56: Wadding, felt, nonwovens, yarns, twine, cordage, etc	Containers
						57: Carpets and other textile floor coverings	Containers
						58: Special woven or tufted fabric, lace, tapestry etc	Containers
						59: Impregnated, coated or laminated textile fabric	Containers
						60: Knitted or crocheted fabric	Containers
						61: Articles of apparel, accessories, knit or crochet	Containers
						62: Articles of apparel, accessories, not knit or crochet	Containers
						63: Other made textile articles, sets, worn clothing etc	Containers
						64: Footwear, gaiters and the like, parts thereof	Containers
						65: Headgear and parts thereof	Containers
						66: Umbrellas, walking-sticks, seat-sticks, whips, etc	Containers
						67: Bird skin, feathers, artificial flowers, human hair	Containers
						68: Stone, plaster, cement, asbestos, mica, etc articles	Containers
						69: Ceramic products	Containers
						70: Glass and glassware	Containers
						71: Pearls, precious stones, metals, coins, etc	Containers
						73: Articles of iron or steel	Containers
						74: Copper and articles thereof	Containers
						75: Nickel and articles thereof	Containers
						76: Aluminium and articles thereof	Containers
						78: Lead and articles thereof	Containers
						79: Zinc and articles thereof	Containers
						80: Tin and articles thereof	Containers
						81: Other base metals, cermets, articles thereof	Containers
						82: Tools, implements, cutlery, etc of base metal	Containers
						83: Miscellaneous articles of base metal	Containers
						84: Nuclear reactors, boilers, machinery, etc	Containers
						85: Electrical, electronic equipment	Containers
						86: Railway, tramway locomotives, rolling stock, equipmen	Containers
						87: Vehicles other than railway, tramway	Containers
						88: Aircraft, spacecraft, and parts thereof	Containers
						89: Ships, boats and other floating structures	Containers
						90: Optical, photo, technical, medical, etc apparatus	Containers
						91: Clocks and watches and parts thereof	Containers
						92: Musical instruments, parts and accessories	Containers
						93: Arms and ammunition, parts and accessories thereof	Containers
						94: Furniture, lighting, signs, prefabricated buildings	Containers
						95: Toys, games, sports requisites	Containers
						96: Miscellaneous manufactured articles	Containers
						97: Works of art, collectors pieces and antiques	Containers

PRELIMINARY DRAFT, No. 07/0312

TABLE B-IV

TRADE PARTNERSHIPS WITH FINITE PORT DISTANCE.

Exporter	Importer	Exporter	Importer	Exporter	Importer	Exporter	Importer	Exporter	Importer	Exporter	Importer
ARE	AUS	CAN	MYS	ECU	USA	ISL	ARG	MYS	USA	THA	USA
ARE	EU15	CAN	NZL	EGY	AUS	ISL	AUS	NGA	AUS	TUN	AUS
ARE	IND	CAN	PAK	EGY	USA	ISL	BRA	NZL	ARG	TUN	USA
ARE	NZL	CAN	PER	EU15	ARE	ISL	COL	NZL	AUS	TUR	ARG
ARE	USA	CAN	RUS	EU15	ARG	ISL	ECU	NZL	BRA	TUR	AUS
ARG	AUS	CAN	SAU	EU15	AUS	ISL	NZL	NZL	COL	TUR	BRA
ARG	COL	CAN	TUN	EU15	BGD	ISL	PER	NZL	ECU	TUR	COL
ARG	ECU	CAN	TWN	EU15	BRA	ISL	URY	NZL	PER	TUR	ECU
ARG	EU15	CAN	URY	EU15	CHN	ISL	USA	NZL	URY	TUR	NZL
ARG	MEX	CAN	VEN	EU15	COL	ISR	ARG	NZL	USA	TUR	PER
ARG	NZL	CAN	ZAF	EU15	ECU	ISR	AUS	PER	AUS	TUR	URY
ARG	PER	CHE	ARG	EU15	EGY	ISR	BRA	PER	NZL	TUR	USA
ARG	USA	CHE	AUS	EU15	IDN	ISR	COL	PER	USA	TWN	AUS
AUS	ARG	CHE	BRA	EU15	IND	ISR	ECU	PHL	AUS	UKR	AUS
AUS	BGD	CHE	COL	EU15	JOR	ISR	NZL	PHL	EU15	UKR	USA
AUS	BRA	CHE	ECU	EU15	JPN	ISR	PER	PHL	USA	URY	AUS
AUS	CHN	CHE	NZL	EU15	KOR	ISR	URY	POL	ARG	URY	NZL
AUS	COL	CHE	PER	EU15	LBY	ISR	USA	POL	AUS	URY	USA
AUS	ECU	CHE	URY	EU15	MAR	JPN	ARG	POL	BRA	USA	ARG
AUS	EGY	CHE	USA	EU15	MYS	JPN	AUS	POL	COL	USA	AUS
AUS	IDN	CHL	AUS	EU15	NZL	JPN	BRA	POL	ECU	USA	BRA
AUS	JPN	CHL	BRA	EU15	PER	JPN	COL	POL	NZL	USA	CHN
AUS	KOR	CHL	COL	EU15	PHL	JPN	ECU	POL	PER	USA	COL
AUS	LKA	CHL	ECU	EU15	SAU	JPN	EU15	POL	URY	USA	ECU
AUS	NZL	CHL	NZL	EU15	SDN	JPN	NZL	POL	USA	USA	EGY
AUS	PAK	CHL	URY	EU15	SGP	JPN	PER	RUS	AUS	USA	EU15
AUS	PER	CHL	USA	EU15	THA	JPN	URY	RUS	USA	USA	IDN
AUS	SAU	CHN	ARG	EU15	TUN	JPN	USA	SAU	AUS	USA	IND
AUS	TWN	CHN	AUS	EU15	URY	KAZ	AUS	SAU	EU15	USA	JPN
AUS	URY	CHN	BRA	EU15	USA	KOR	ARG	SAU	USA	USA	KOR
AUS	USA	CHN	COL	EU15	VNM	KOR	AUS	SGP	AUS	USA	LKA
AUS	ZAF	CHN	ECU	EU15	YEM	KOR	BRA	SGP	EU15	USA	MAR
BGD	AUS	CHN	EU15	GHA	AUS	KOR	COL	SGP	NZL	USA	MYS
BGD	USA	CHN	IND	HUN	ARG	KOR	ECU	SGP	USA	USA	NZL
BGR	USA	CHN	NZL	HUN	AUS	KOR	NZL	SVK	ARG	USA	PAK
BHR	AUS	CHN	PER	HUN	BRA	KOR	PER	SVK	AUS	USA	PER
BOL	AUS	CHN	URY	HUN	COL	KOR	URY	SVK	BRA	USA	PHL
BOL	NZL	CHN	USA	HUN	ECU	KOR	USA	SVK	COL	USA	RUS
BOL	USA	COL	AUS	HUN	NZL	KWT	AUS	SVK	ECU	USA	SAU
BRA	EU15	COL	NZL	HUN	PER	LBN	AUS	SVK	NZL	USA	SGP
CAN	ARG	COL	USA	HUN	URY	LKA	AUS	SVK	PER	USA	THA
CAN	AUS	CRI	AUS	HUN	USA	LKA	USA	SVK	URY	USA	TUN
CAN	BRA	CRI	COL	IDN	AUS	LTU	AUS	SVK	USA	USA	URY
CAN	CHN	CRI	ECU	IDN	EU15	MAR	AUS	SVN	ARG	USA	VEN
CAN	COL	CZE	ARG	IDN	USA	MAR	USA	SVN	AUS	USA	VNM
CAN	DZA	CZE	AUS	IND	AUS	MEX	ARG	SVN	BRA	USA	YEM
CAN	ECU	CZE	BRA	IND	BRA	MEX	AUS	SVN	COL	USA	ZAF
CAN	EGY	CZE	COL	IND	CHN	MEX	BRA	SVN	ECU	VEN	AUS
CAN	EU15	CZE	ECU	IND	COL	MEX	COL	SVN	NZL	VEN	USA
CAN	IDN	CZE	NZL	IND	ECU	MEX	ECU	SVN	PER	VNM	AUS
CAN	IRN	CZE	PER	IND	EU15	MEX	NZL	SVN	URY	VNM	EU15
CAN	JOR	CZE	URY	IND	NZL	MEX	PER	SVN	USA	VNM	USA
CAN	JPN	CZE	USA	IND	PER	MEX	URY	THA	ARG	YEM	USA
CAN	KOR	DOM	USA	IND	SAU	MYS	AUS	THA	AUS	ZAF	ARG
CAN	LKA	DZA	USA	IND	URY	MYS	BRA	THA	BRA	ZAF	AUS
CAN	MAR	ECU	AUS	IND	USA	MYS	EU15	THA	EU15	ZAF	NZL
CAN	MEX	ECU	NZL	IRN	AUS	MYS	NZL	THA	NZL	ZAF	USA

REFERENCES

Anderson, J. E. (2010), The gravity model, NBER Working Papers 16576, National Bureau of Economic Research.

Anderson, J. E. & van Wincoop, E. (2004), 'Trade costs', *Journal of Economic Literature* **42**(3), 691–751.

Chaney, T. (2005), Distorted gravity: Heterogeneous firms, market structure and the geography of international trade. Job Market Paper.

Chaney, T. (2011), The network structure of international trade, NBER Working Papers 16753, National Bureau of Economic Research.

Helpman, E., Melitz, M. & Rubinstein, Y. (2008), 'Estimating trade flows: Trading partners and trading volumes', *The Quarterly Journal of Economics* **123**(2), 441–487.

Hendricks, K., Piccione, M. & Tan, G. (1995), 'The economics of hubs: The case of monopoly', *The Review of Economic Studies* **62**(No. 1), 83–99.

Hummels, D. (2007), 'Transportation costs and international trade in the second era of globalization', *Journal of Economic Perspectives* **21**(3), 131–154.

Hummels, D., Lugovskyy, V. & Skiba, A. (2009), 'The trade reducing effects of market power in international shipping', *Journal of Development Economics* **89**(1), 84–97.

Irarrazabal, A., Moxnes, A. & Opromolla, L. (2011), The tip of the iceberg: A quantitative framework for estimating trade costs. Banco de Portugal, Economics and Research Department, Working Paper, No. 25/2011.

Jackson, M. O. (2003), A survey of models of network formation: Stability and efficiency, Game theory and information, EconWPA.

Jackson, M. O. (2008), *Social and Economic Networks*, Princeton University Press, Princeton.

Jackson, M. O. & Wolinsky, A. (1996), 'A strategic model of social and economic networks', *Journal of Economic Theory* **71**(1), 44–74.

Krugman, P. (1980), 'Scale economies, product differentiation and the pattern of trade', *The American Economic Review* **70**(5), 950–959.

Krugman, P. (1991), 'Increasing returns and economic geography', *Journal of Political Economy* **99**(3), 483–499.

Krugman, P. (1993), The hub effect: or, threeness in interregional trade, *in* 'Theory, Policy and Dynamics in International Trade', Cambridge University Press, New York, pp. 29–37.

Limao, N. & Venables, A. (2001), 'Infrastructure, geographical disadvantage and transport costs.', *The World Bank Economic Review* **15**, 451–479.

Lux, M. (2011), Defying gravity: The substitutability of transportation in international trade, Mpra paper, University Library of Munich, Germany.

OECD (2008), Clarifying trade costs in maritime transport. OECD Trade Committee Working Paper.

Overman, H. G., Redding, S. J. & Venables, A. J. (2001), The economic geography of trade production and income: A survey of empirics. Discussion Paper, C.E.P.R., No. 2978.

Redding, S. & Venables, A. (2004), 'Economic geography and international inequality', *Journal of International Economics* **62**, 53–82.