Gravity and Information: Heterogeneous Firms, Exporter Networks and the 'Distance Puzzle'*

Sebastian Krautheim**

First version: Oct. 2006, this version: Feb. 2007

preliminary draft - comments appreciated

do not quote without permission

Abstract

One of the most robust features of estimated gravity equations is that - based on standard theory - the effect of distance on trade flows appears too strong to be explained by trade costs alone. Maybe even more surprisingly, the effect seems to increase rather than decrease over time. In this paper a simple model of international trade is presented in which heterogeneous firms can form informational networks in order to decrease their fixed costs of exporting. Compared to existing theories the effect of distance on aggregate trade flows is endogenously enforced and an amelioration of the (exogenous) network technology turns out to be a possible explanation for the increase of estimated distance effects over time. Existing empirical evidence is shown to support some crucial implications of the model and an empirical gravity equation is estimated and used to determine the magnitude of the effect. It turns out that when the proposed model is used to interpret the empirical results, the 'distance puzzle' appears considerably less puzzling.

JEL: *F10, F12, D85*

Keywords: *gravity, heterogeneous firms, exporter networks, fixed costs, distance*

---

* I am grateful to Giancarlo Corsetti and Morten Ravn for constant advice. I would also like to thank Itai Agur, Thomas Chaney, Philip Sauré, Eva Schneider, Daria Taglioni, Yves Zenou, the participants of the EUI working group on international trade and the VIIth RIEF Doctoral Meeting in Rennes 2007 for fruitful discussions and comments. All remaining errors are mine.

** Department of Economics, European University Institute, Villa San Paolo, Via della Piazzuola 43, 50133 Firenze, Italy. E-mail: sebastian.krautheim@eui.eu.

1 Introduction

Gravity equations are clearly the dominating tool for the empirical analysis of international trade flows. Despite their remarkably simple structure, they fit the data surprisingly well. Recent advances in the theoretical foundation of this relationship (e.g. Anderson and van Wincoop (2003)) have made the use of gravity equations even more popular.

However, Grossman (1998) argues that estimated distance effects are far too high to be explained by trade costs alone. Anderson and van Wincoop (2004) show that standard trade theory would suggest that estimated elasticity of trade flows with respect to distance should be in a range of 0.1 to 0.2. But empirical estimates are much higher. Disdier and Head (2007) show that surprisingly high coefficients on distance are one of the most robust features found in estimated gravity equations. They compute a mean of about 0.9 across 1467 estimations in 103 papers in the literature. Maybe even more surprisingly, they find that estimated distance coefficients tend to rise for data between the 1950s and the 1980s and have remained high since then.

How can it be that in the ongoing process of globalization the effect of geographical distance on trade flows is so much higher than theory suggests and how can this effect even be increasing over time?

A possible explanation could be a mis-specification of the gravity model. For example Aviat and Coeurdacier (2006) write: "One can argue that the 'gravity model' might be misspecified: an omitted variable (correlated with physical distance) might lead to an over-estimation of β [the distance coefficient] but the difficulty consists in finding the missing variable. The right one has not been found yet." (p. 6). Anderson (2000) reasons that the "inescapable conclusion is - *there must be some other transaction costs*. Moreover, *the missing transactions costs must be very sensitive to distance* [...]" (p. 118, emph. in original) and: "The objective [of future work] is to build consistent general equilibrium models in which the volume of trade interacts endogenously with the size of the transaction cost." (p. 125).

Instead of searching for missing transaction costs or a mis-specification of the empirical gravity equations, the approach of this paper is to take the empirical results serious and to develop a theoretical model which rationalizes the strong effect of distance on aggregate trade flows. The basic idea is to reconsider two aspects that have played an important role in the discussion of the determinants of international trade: information and networks.

A general equilibrium model with heterogeneous firms and exporter networks is constructed in which firms within a country have the possibility to form informational networks. A larger amount of information about exporting from country i to country j decreases the level of fixed costs of exporting to j and thus encourages more firms to enter the export market than theories with exogenous (possibly: zero) fixed cost would suggest.

It is shown that the model provides a mechanism which magnifies the effect of variable trade costs (distance) on the trade volume between countries compared to the level predicted by standard theory. Trade flows are affected by distance via two margins: lower distance between country i and country j raises the volume of exports of each individual firm exporting from i to j (*intensive margin*). But lower distance also increases the number of firms exporting from i to j (*extensive margin*). The introduction of exporter networks introduces a feedback effect on the level of fixed costs and thus enforces the effect of trade barriers on the number of exporting firms. I will speak of the two components as the *basic extensive margin* and the *enforced extensive margin*.

The enforced extensive margin is the most important novel feature of the model. I find that depending on the quality of the networks (reflected by the elasticity of fixed costs with respect to available information) the model provides a mechanism which potentially accounts for the high estimated distance coefficients. Moreover, it is shown that an increase in the quality of exporter networks over time leads to an increase in the elasticity of trade flows with respect to distance. Thus exporter networks deliver a possible explanation for both dimensions of the 'distance puzzle' at the same time: the *high level* and the *increase* of the distance coefficient over time.

In an empirical section, testable implications of the model are discussed. It is shown that existing empirical evidence strongly supports crucial implications of the model. Namely how the degree of firm heterogeneity and the elasticity of substitution affect the effect of *variable* trade costs on trade flows. In contrast to existing theory but in line with the empirical evidence the suggested model predicts that the elasticity of substitution between varieties *dampens* the effect of variable trade costs on aggregate trade flows.[1] In an empirical exercise a gravity-type trade equation derived from the theoretical model is used to estimate the magnitude of the enforced extensive margin. The results suggest that when the proposed model is used to interpret the estimated distance elasticities, the high estimates of the distance coefficient in the literature appear far less puzzling.

The modeling is related to different strands of literature. The model presented in this paper is based on the recent literature on international trade with heterogeneous firms.[2] The basic structure is borrowed from Chaney (2006). He provides a version of the model of Melitz (2003) which is simplified along some dimensions but additionally allows to consider different sectors and countries of different size preserving analytical tractability. The second related field is the recent microeconomic literature on networks pioneered by the work of Jackson and Wolinsky (1996). Its basic concepts are used to introduce an informational network of exporting firms *within* a country.

There is a large literature that (mostly empirically) considers the role networks in international trade. Starting with Rauch (1999) a literature has evolved which argues that business, social or ethnic networks reduce informational frictions between countries (or between regions within a country) and thus increase trade (see for example Rauch (2001), Rauch and Casella (2003), Rauch and Trindade (2002) and Combes, Lafourcade, and Mayer (2005)). There are, however, two crucial differences between the way this literature considers networks in international trade and the approach taken in this paper. Firstly, the existing literature focuses on networks *across* countries or regions, while this paper considers networks *within* countries or regions. Secondly, while in the literature networks

[1] Note that standard theory (e.g. Krugman (1980)) implies that the elasticity of substitution *strengthens* the effect of variable trade costs on trade flows. Chaney (2006) provides a model in which the elasticity of substitution dampens the effect of *fixed* trade costs on aggregate trade flows, while the effect of variable trade costs in *independent* of the elasticity of substitution. To the best of my knowledge my model is the first to imply the elasticity of substitution to *dampen* the effect of *variable* trade flows, which is in line with the empirical evidence.

[2] Pioneered by Melitz (2003), some of the major contributions of the literature based on monopolistic competition framework à la Krugman (1980) are Helpman, Melitz, and Yeaple (2004), Melitz and Ottaviano (2005), Chaney (2006) and Helpman, Melitz, and Rubinstein (2006). To the best of my knowledge, the only model in the literature introducing endogenous fixed costs of exporting is Arkolakis (2006) who introduces marketing costs in order to endogenize market size for individual firms. Eaton and Kortum (2002) propose a Ricardian model of international trade, Bernard, Eaton, Jensen, and Kortum (2003) propose a model with Bertrand competition and Eaton, Kortum, and Kramarz (2005) provide an integrated framework which also includes monopolistic competition.

are assumed to be exogenously given, in the model suggested here the network emerges endogenously as a result of profit maximization of firms. I view these two ways of thinking of informational networks in international trade as complementary.

The main mechanism in the model goes via the impact of information on the fixed costs of exporting. The empirical evidence on these fixed costs is limited. Bugamelli and Infante (2003) provide some evidence that the cost of information acquisition is an important part of the fixed costs. In a survey of about 4,400 German firms by the German chamber of commerce (DIHK (2005)), the acquisition of information about the destination market and the development of a sound business plan based on this information are found to be crucial for successful exporting. The relevant information is in most cases obtained via a *direct channel* (personal travel, participation in trade fairs and own market analysis) and via an *indirect channel* (business partners, acquaintances, personal networks and provision of information, consulting and exchange of experience organized by the chamber of commerce). I think of the exporter networks as representing this indirect channel.

Similar results are reported in Roberts and Tybout (1997). In 1990, 186 Colombian firms were interviewed by the World Bank and the Colombian government's export promotion agency. Most firms considered information acquisition on buyer identification, foreign prices, market selection and standards and testing requirements as a major entry cost into foreign markets. Most firms had used private and public services to overcome these informational obstacles.

An indication that the fixed costs of starting to export are of considerable magnitude (and potentially different depending on the destination markets) can be found in DIHK (2005): the time German firms report to be spending for preparation before they actually start exporting ranges from an average of 1.7 years for exporting to EU15 over 6 years for US/Canada to 7.1 for Africa.

In the model a larger number of exporters to a particular market decreases the fixed costs of exporting to this market. One should thus find positive spillover effects between exporters in the data. The evidence on spillovers is mixed. Using firm level data Aitken, Hanson, and Harrison (1997) (for Mexico) and Greenaway, Sousa, and Wakelin (2004) (for the UK) find evidence for exporting spillovers between firms, Barrios, Görg, and Strobl (2003) (for Spain) and Bernard and Jensen (2004) (for the U.S.) do not find such evidence. An important shortcoming of these studies in the light of the proposed model is that they have no information about the export destination. But the model suggests that the spillovers observed should in fact be *destination specific*. Using a very detailed French firm level data set which includes export destinations, Koenig (2005) closes this gap in the literature. In line with the predictions of the model, she finds strong evidence that spillovers (on the regional level) are of considerable magnitude and indeed *destination specific*.

Information has long been conjectured to play a role in understanding the 'distance puzzle'. Grossman (1998) conjectured that the high distance elasticities found in the empirical literature could be related to 'unfamiliarity' which is by assumption increasing in distance. Huang (2006) uses this assumption and provides empirical evidence that countries with a high degree of risk aversion trade disproportionately less with more distant countries. Portes and Rey (2005) use telephone traffic between two countries as a proxy for available information. However, in all these cases no clear theoretical mechanism is spelled out and the reasoning is that (a) information matters for trade, (b) information decreases in distance (*by assumption*), so (c) trade decreases in distance. In this paper - based on a complete general equilibrium model - the argument is reversed: information matters for trade, but trade also *generates* information: lower variable trade costs (lower distance) encourage trade, *ceteris paribus* increase the level of available information, increase the number of exporting firms and thus further increase aggregate exports.

2 The Model

In this section a model of trade with heterogeneous firms and exporter networks will be presented. In the first subsection the basic structure of the model is outlined. In the second subsection exporter networks are introduced: exporting firms have the possibility to form networks and to exchange information which reduces their cost related to exporting. In the third subsection all conditions for the world general equilibrium are derived taking into account the equilibrium network structure. The last subsection provides a discussion and interpretation of the theoretical results.

2.1 The Economy

Basic structure: The world economy consists of N countries with L_n denoting the population in country n.[3] There are $H+1$ sectors, H of which are producing differentiated products, while sector zero produces a homogeneous good with a constant returns to scale technology. The homogeneous good is freely traded and is used as the numeraire with its price normalized to one. As is standard in such a setting (see for example Helpman, Melitz, and Yeaple (2004)) only those equilibria are considered where all countries produce the homogeneous good which implies that wages are equalized across countries and can also be normalized to one. Labor is the only input in the production process. Each worker holds a share of a perfectly diversified portfolio of all firms in the world.

Preferences: The workers are all identical. They share the same preferences over consumption of the goods produced in the $H+1$ sectors:

$$U = q_0^{\mu_0} \prod_{h=1}^{H} \left(\int_{X_h} (q_h^x)^{\frac{\sigma_h - 1}{\sigma_h}} dx \right)^{\frac{\sigma_h}{\sigma_h - 1}\mu_h}$$

where q_h^x is the quantity of variety x of good (sector) h, q_0 is the quantity of the homogeneous good consumed, $\mu_0 + \sum_{h=1}^{H} \mu_h = 1$ and σ_h is the elasticity of substitution between varieties of sector h.

Note that here we have a Cobb-Douglas structure between goods (sectors) and a CES structure between varieties of the same good (within sectors). It is a well known property of Cobb-Douglas preferences that agents spend a constant fraction of their income on varieties of each sector. This allows us to analyze each sector separately.

Taking prices as given, consumers choose a quantity of each variety of each good available in their country.

Firms: Different to Melitz (2003) the number of firms in each sector is assumed to be fixed and proportional to country size. No firm entry and exit takes place on the national level. This feature makes the structure of the model similar to the monopolistic competi-

[3] Note that the model is not restricted to the analysis of trade flows between countries. It could as well be used for the analysis of regions withing a country.

tion case of the generalized framework proposed by Eaton, Kortum, and Kramarz (2005).[4] Two types of costs emerge when a firm exports. In order to be able to start exporting a variety of good h from country i to country j a firm has to pay a fixed cost of exporting C_{ij}^h. In addition to this the firm has to pay variable trade costs of the "melting iceberg" type τ_{ij}^h. This means that of each unit produced in i and shipped to j only a fraction $1/\tau_{ij}^h$ arrives.

Production in the differentiated good sectors takes place according to a standard increasing returns to scale technology. So the costs for a firm with productivity x in country i of producing output to sell q/τ_{ij}^h units in j is given by

$$c(q) = \frac{q}{x} + C_{ij}^h.$$

Facing isoelastic demand curves, firms charge a constant mark-up over marginal costs. Note that marginal costs include the variable cost of shipping the good from i to j. The expression for the price charged in j by a firm from i with productivity x is thus given by

$$p_{ij}^h(x) = \frac{\sigma_h}{\sigma_h - 1} \frac{\tau_{ij}^h}{x}. \tag{1}$$

Firms differ in their productivity levels. Asstandard in the literature (see e.g. Helpman, Melitz, and Yeaple (2004) and references there in) an individual firm's productivity is assumed to be drawn from a Pareto distribution with parameter γ_h

$$P(X < x) = F_h(x) = 1 - x^{-\gamma_h}.$$

Without loss of generality the minimum productivity level has been normalized to one ($x_{min} \equiv 1$) which implies $x \geq 1$. In order to have a finite second moment it is standard to assume $\gamma_h > 2$. Furthermore we have to impose $\gamma_h > (\sigma_h - 1)$. This assumption assures that the mean of the productivity distribution is finite.[5] Taking the optimal choices of

[4] Note that their empirical analysis using the above mentioned French firm-level data set strongly supports the monopolistic competition structure.

[5] To see the importance of the last condition note the following: the lower γ, the higher is the mass of firms with high productivities. If σ is high, goods are close substitutes. The closer γ_h and $(\sigma_h - 1)$ get, the larger the mass of firms with a very high productivity with some of them being so productive that they sell at a price close to zero. But if then substitutability was too high (the above condition is violated) these firms would take over the whole market and the equilibrium brakes down.

the other firms as given, each firm chooses a subset of countries to export to and prices to charge in these countries.

Demand: With the wages in all countries normalized to one, the total labor income in j is given by L_j. Since firms make positive profits the second component of income are dividends paid on the shares of the global fund holding all firms. Dividends received by workers in country j are given by $(L_j/L)\Pi$ where Π are world profits and L stands for world population. Demand in j for a variety of good h imported from i is given by

$$q_{ij}^h = \mu_h \left(1 + \frac{\Pi}{L}\right) L_j \left(\frac{p_{ij}^h(x)}{P_j^h}\right)^{-\sigma_h} P_j^h. \tag{2}$$

Where P_j^h is the welfare based price index in sector h.

Welfare based price index and global profits: Denote with $\bar{x}_{ij}^h$ the cutoff productivity level in sector h. For firms with a productivity above this level it is profitable to export from i to j, while firms below this level choose not to serve market j, i.e. they remain inactive in market j and make zero profits there.

The welfare based price index for the varieties of good h in country j has to take into account all prices set by all firms (from all the N countries) selling varieties of good h in j. It is given by:

$$P_j^h = \left(\sum_{k=1}^{N} L_k \int_{\bar{x}_{kj}^h}^{\infty} \left(\frac{\sigma_h}{\sigma_h - 1}\frac{\tau_{ij}^h}{x}\right)^{1-\sigma_h} dF_h(x)\right)^{\frac{1}{1-\sigma_h}}.$$

World profits are defined as the sum of the profits any firm (of any sector) makes in any market

$$\Pi = \sum_{h=1}^{H} \sum_{k,l=1}^{N} L_k \int_{\bar{x}_{kj}^h}^{\infty} \pi_{kl}^h(x)\ dF_h(x)$$

where $\pi_{kl}^h(x)$ are net profits a firm with productivity x in sector h of country k makes by exporting to l.

2.2 Exporter Networks

In this subsection the most important novel feature of the model is introduced by assuming that firms within a country have the possibility to form a network and exchange information which is relevant for exporting. This will be done formally using standard concepts of network theory.[6] The subsequent analysis is valid for any of the H differentiated goods sectors. Sectoral subscripts will thus be dropped where this causes no confusion.

Departures from traditional network theory: Traditional network theory has focused on networks between a finite number of players. Different assumptions about the cost structure can lead to different network structures which in many cases are neither stable nor unique equilibria. However, we will see that under appropriate simplifying assumptions some of the essential mechanisms of network theory can still be applied to the case of a continuum of firms and that this can provide some useful insights.

Assumptions: (a1) Each firm that decides to export from country i to country j has a given amount of relevant information related to its export activity, given by Υ_a for firm a. This information is specific to the firm. The information sets are 'equidistantly disjoint' which means that given that firm a obtains the information set from one other additional firm, the effect of the additional information set is the same independently of the identity of the firm from which the information set originates. What counts is thus only the *amount* of information, not its *origin*. Normalize $\Upsilon_a \equiv 1 \; \forall \; a$.

(a2) In the network formation game $\Psi(u)$ firms can create links to other firms in order to exchange information. Links can only be formed with mutual consent. Link creation and the exchange of information along a link is costless for the firms.

(a3) Information can be exchanged along the links with a decay factor $\delta_i < 1$ i.e. if firm a and b exchange information, b only receives a fraction δ_i of the information set of firm a. Being transferred along a link, information looses some degree of complexity/detailedness 'on the way'. In fact, one can think of the original information set as consisting of information with different degrees of 'fragility'. The most fragile (complex/detailed) information is lost when it is first passed on, the information of the next degree of fragility is lost

[6] Since network theory has not yet become part of the theoretical tools commonly used in trade theory, the main concepts and definitions used are outlined in appendix A1.

when the information is passed on via two links and so on. It is important to note that the information lost only depends on the number of links it is transferred through *not* on the actual path it is taking through the network.

(a4) If firms are indifferent between creating and not creating a link, they don't propose link creation.[7]

Information and fixed cost of exporting: An important element of the model is that the set of information about exporting from i to j available to a firm influences the level of fixed cost of exporting to j. So the fixed cost level faced by firm a in country i to export to j, $C_{ij}(a)$, would be determined by a function $\Phi(.)$ such that $C_{ij}(a) = \Phi(\iota_{ij}(a), \bar{C})$. Where $\iota_{ij}(a)$ is firm a's stock of information about exporting from i to j and $\bar{C}$ is an exogenous cost factor.

There are some properties this function should have. It should be decreasing monotonically in $\iota_{ij}(a)$ i.e. $\partial\Phi(\iota_{ij}(a), \bar{C})/(\partial\iota_{ij}(a)) < 0$ implying that *ceteris paribus* availability of more relevant information leads to a lower level of fixed cost. At the same time the marginal effect of additional information should be decreasing: $\partial^2\Phi(\iota_{ij}(a), \bar{C})/(\partial\iota_{ij}(a)^2) > 0$.[8]

The simplest functional form which ensures analytical tractability and is in line with the assumptions on the fixed costs in Koenig (2005) is a power function of the form: $C_{ij}(a) = [\iota_{ij}(a)]^{-\frac{1}{b^*}}\ \bar{C}$. Where $1/b^*$ is the elasticity of fixed costs with respect to information and $\bar{C}$ is a scaling parameter. The minimum level of information an exporting firm can possibly have, $\bar{\iota}_{ij}(a)$, is the amount of information it individually acquires by exporting: $\bar{\iota}_{ij}(a) = \Upsilon_a = 1$ thus the maximum level of fixed costs is given by $\bar{C}$.

Equilibrium: Firms are maximizing profits. In order to analyze the pairwise-Nash equilibrium of the network formation game outlined in section two, we will proceed as follows. First we will consider the outcomes of a network formation game, $\Psi(\iota_{ij})$, in which all firms maximize their available information on exporting activities relevant to them.[9] This

[7]This assumption rules out equilibria where firms are in the network which neither contribute nor benefit from the network (the non-exporters). Since link creation is costless and they don't contribute information to the network, the equilibrium effects of the network would be just the same if they (or some of them) joined.

[8]We will see later on that this property is in line with the empirical findings of Koenig (2005).

[9]Note that this implies that a firm not exporting from i to j will not maximize its available information for doing so.

behavior will reduce their fixed cost of exporting, *ceteris paribus* increasing their profits. But it could be that in this 'information maximizing network', firms have the power to influence the network structure and thus affect the cost structure and price setting of their competitors. This would for example be the case of a circle- or star-shaped network. We will thus determine in a second step whether the information maximizing network gives rise to strategic interaction or not.

Consider first the firms not exporting from i to j. They are completely indifferent regarding the information about exporting from i to j by assumption (a4) they will thus not create links in order to obtain this information.

Let's turn to a firm a maximizing it's available information about exporting from i to j. In the network formation game $\Psi(\iota_{ij})$ the payoff function of firm a in country i is then given by: $u_a(g) = \iota_{ij}(a) = \int_0^{n_{ij}} \delta_i^{d(a,b)}\, db$ where $u_a(g)$ is the payoff firm a gets in the network g and n_{ij} is the measure of firms exporting from i to j. This function adds the information obtained from the direct partners (geodesic distance of unity) of δ_i times their measure ($n_{ij,1}$), the information from indirect links with geodesic distance of two δ_i^2 times their measure ($n_{ij,2}$) and so on. Defining the measure of exporters from i to j with geodesic distance to firm a of φ with $\varphi \in \{1, 2, 3, ...\}$, then $\sum_{\varphi=1}^{\infty} n_{ij,\varphi} = n_{ij}$.[10]

Define s_a as a pure strategy profile of firm a i.e. the whole list of firms firm a proposes to form a link to. It is obvious that under the assumption of costless link creation $\iota_{ij}(a)[g(s_a)]$ is maximized if and only if $n_{ij,1} = n_{ij}$ i.e. firm a has a direct link to all firms exporting from i to j. Denote the corresponding pure strategy profile of firm a by s_a^*. Thus, under the assumptions (a1)-(a4) the complete network of exporters, $g(s^*)$, maximizes the relevant available information of each individual firm.

The network $g(s^*)$ is a Nash equilibrium because given the behavior of the other firms no firm could choose a strategy profile $s_a \neq s_a^*$ such that $\iota_{ij}(a)g(s_a, s_{-a}^*) > \iota_{ij}(a)[g(s^*)]$.

In addition, $g(s^*)$ is pairwise-stable since for any firm in country i (regardless of the export status) for all possible pairs of firms we have (i) for all $ab \in g(s^*)$, $\iota_{ij}(a)[g(s^*)] \geq \iota_{ij}(a)[g(s^*) - ab]$ and $\iota_{ij}(b)[g(s^*)] \geq \iota_{ij}(b)[g(s^*) - ab]$ and (ii) for all $ab \notin g(s^*)$, if $\iota_{ij}(a)[g(s^*)]) < \iota_{ij}(a)[g(s^*) + ab]$ then $\iota_{ij}(b)[g(s^*)] > \iota_{ij}(b)[g(s^*) + ab]$. For pairs of firms

[10]It is standard to define a geodesic distance of ∞ as the case where two players are parts of separated networks. So a network where all players have a geodesic distance of ∞ is the empty network with no links.

exporting from i to j the strict inequalities in (i) hold whereas the equalities apply when a firm not exporting from i to j is involved. Condition (ii) is also satisfied since for any firm a in country i (regardless of export status) we have $\iota_{ij}(a)[g(s^*)] \geq \iota_{ij}(a)[g(s^*) + ab]$. Finally, the equilibrium is the unique pairwise-Nash equilibrium because by (a4) firms not exporting from i to j stay out of the network and for firms exporting from i to j it always holds that any two not connected firms could always gain by creating an additional link. We can thus state (proof in the text):

Lemma 1 *Under assumptions (a1)-(a4) the unique pairwise-Nash equilibrium of the network formation game $\Psi(\iota_{ij})$ is a complete network of firms exporting from i to j.*

We shall now turn to the question of whether the complete network among firms exporting from i to j is also the unique pairwise-Nash equilibrium for a network formation game, $\Psi(\pi_{ij})$, where firms maximize profits from exporting from i to j, π_{ij}.

The structure of the model outlined above is such that firms maximize profits of exporting from i to j separately for all j. Firm a exports if $\pi_{ij}(a) > 0$ and does not export to j if $\pi_{ij}(a) \leq 0$. *Ceteris paribus* an exporting firm maximizes profits by minimizing $C_{ij}(a)$ i.e. by maximizing $\iota_{ij}(a)$. Since we are in a context of monopolistic competition also prices charged by the competitors play a role for $\pi_{ij}(a)$. If the firm could thus influence the cost-structure faced by its competitors by deviating from the cost-minimizing strategy profile s^*, the unique pairwise-Nash equilibrium of $\Psi(\iota_{ij})$ might not be the unique pairwise-Nash equilibrium of the network formation game $\Psi(\pi_{ij})$.

It is easy to see that none of the firms can influence the aggregate effects of the complete network of exporters on their fixed costs of exporting. Consider first the firms not exporting from i to j. Whether one or many of these firms join or leave the network leaves its effects on the fixed costs of the exporters unaffected.[11] The same holds true for an individual firm exporting from i to j. The strongest impact it can have on the aggregate effect of the network is by cutting all its links. This would reduce the information available to the other firms by δ_i. But since firms are small leaving the network does only marginally affect the information about exporting from i to j available to the competing firms within

[11] Which justifies the convenient assumption (a4).

the network. It follows that no individual firm has the possibility to affect the aggregate outcomes of the network by deviating from the pure strategy profile s^*. We can thus state (proof in the text):

Proposition 1 *Under assumptions (a1)-(a4) the unique pairwise-Nash equilibrium of the network formation game $\Psi(\pi_{ij})$ is identical to the unique pairwise-Nash equilibrium of the network formation game $\Psi(\iota_{ij})$, namely a complete network of all firms exporting from i to j.*

Fixed cost of exporting in a complete exporter network: In a complete network of firms exporting from i to j all firms exchange information. The information firm a gets from firm b is twofold: (a) it gets a fraction δ_i of the original information of firm b and (b) it also gets δ_i times all the information firm b has obtained through the network. In the complete network a is also linked to all the partners of b so the information a gets from the other firms via b is a subset of the information a gets directly from the other firms. Thus, the additional information firm a gets from the link to firm b is just δ_i.

It follows that the overall information firm a obtains in the network about exporting from i to j is given by[12]

$$\iota_{ij}(a) = \delta_i n_{ij} = \delta_i L_i \bar{x}_{ij}^{-\gamma}.$$

Three factors determine the level of information of firm a: the quality of the information transmission δ_i, the overall number of firms in i, L_i (which is by assumption proportional to country size) and the fraction of these firms exporting from i to j.

Since in equilibrium all firms exporting from i to j are part of the complete network, one has $\iota_{ij}(a_i) = \iota_{ij} \ \forall \ a_i$ and thus $C_{ij}(a_i) = C_{ij} \ \forall \ a_i$. Without any loss of generality we can define $b \equiv b^*(\sigma - 1)$, which will simplify the algebra.

Using this definition together with the equilibrium expression for ι_{ij}, in equilibrium the

[12] Note that the measure of firms exporting from i to j is given by $n_{ij} = L_i \int_{\bar{x}_{ij}}^{\infty} dF_h(x) = L_i \bar{x}_{ij}^{-\gamma_h}$. Where $\bar{x}_{ij}$ is the productivity level above which firms export from i to j. Given the distributional assumptions of the previous subsection, $\bar{x}_{ij}^{-\gamma}$ is the fraction of firms with a productivity above this level.

fixed cost of exporting from i to j is given by

$$C_{ij} = \left[\delta_i L_i \bar{x}_{ij}^{-\gamma}\right]^{-\frac{(\sigma-1)}{b}} \bar{C}. \tag{3}$$

In line with the empirical findings in Koenig (2005), the fixed cost for firms in country i to export to country j is decreasing in the number of other firms exporting to the same market and the *marginal* effect of additional exporters decreases as the number of exporters grows larger.

2.3 Equilibrium with Exporter Networks

Given the equilibrium outcome of the network formation game, we can now determine the general equilibrium expressions for all variables in the model.

Cutoff Productivity Level: Firms decide to enter an export market as long as profits from doing so are non-negative the profits a firm with productivity x makes by exporting from i to j are

$$\pi_{ij}(x) = q_{ij}(x)\left(p_{ij}(x) - \frac{\tau_{ij}}{x}\right) - C_{ij}.$$

Using equations (1) and (2) this can be rewritten as

$$\pi_{ij}(x) = \frac{\mu}{\sigma}\left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1} P_j^{\sigma-1}\left(1+\frac{\Pi}{L}\right) L_j \left(\frac{x}{\tau_{ij}}\right)^{\sigma-1} - C_{ij}. \tag{4}$$

All the elements determining a firm's profits made by exporting from i to j are identical for all firms in i except the productivity of the firm, x. As long as $C_{ij} > 0$, it depends on the productivity level of the individual firm whether or not it makes positive profits from exporting. Define the cutoff productivity level, $\bar{x}_{ij}$, as the productivity level of a firm being just indifferent between exporting and not exporting i.e. for which $\pi_{ij}(\bar{x}_{ij}) = 0$. All firms with a productivity level below the cutoff level $\bar{x}_{ij}$ will not find it profitable to start exporting to j. The other firms will.

Using the condition $\pi_{ij}(\bar{x}_{ij}) = 0$ and the function for the fixed cost of exporting (3), the cutoff productivity level is given by

$$\bar{x}_{ij} = \left(\lambda_1 \delta^{-\frac{1}{b}} P_j^{-1} L_j^{\frac{1}{1-\sigma}} \tau_{ij} L_i^{-\frac{1}{b}} \bar{C}^{\frac{1}{\sigma-1}} \right)^{\frac{b}{b-\gamma}}. \tag{5}$$

This expression still depends on P_j and, via λ_1, also on aggregate world profits Π which remain to be determined in equilibrium.[13]

At this point we have to impose the following parameter restriction: $b > \gamma > (\sigma - 1)$. Where the first inequality ensures that the effect of the informational networks on the fixed costs of exporting is sufficiently small to assure that firms entry does not drive down the fixed costs too fast. If this condition was violated, this would drive the value of the cutoff productivity below its lower bound, implying that all firms would enter all export markets which would be equivalent to a model without fixed costs. The second inequality is a standard assumption which assures that the mean of the productivity distribution is finite (see footnote 5).

Equilibrium price index: Equation (5) together with the expression for P_j derived above can be used to determine the equilibrium price index which is given by

$$P_j = \lambda_2 L_j^{\beta \frac{1}{1-\sigma}} L^{-\frac{1}{\gamma}} \theta_j \tag{6}$$

where

$$\theta_j^{-\frac{1}{\zeta-\eta}} \equiv \sum_{k=1}^{N} \delta^{\frac{\gamma-(\sigma-1)}{b-\gamma}} \left(\frac{L_k}{L} \right)^{\frac{b-(\sigma-1)}{b-\gamma}} \tau_{kj}^{-\frac{1}{\zeta-\eta}} \bar{C}^{\frac{b(\sigma-1)-b\gamma}{b(\sigma-1)-\gamma(\sigma-1)}}$$

and where, for notational convenience, I have defined $\beta \equiv \frac{b\gamma - b(\sigma-1)}{b\gamma - \gamma(\sigma-1)}$ and $\zeta \equiv \frac{b}{\gamma[b-(\sigma-1)]}$ and $\eta \equiv \frac{\gamma}{\gamma[b-(\sigma-1)]}$.[14] θ_j is an analog of the 'multilateral resistance' term of the importing country in Anderson and van Wincoop (2003) and can be interpreted as index of aggregate remoteness.

---

[13] $\lambda_1 = \left(\frac{\sigma}{\mu}\right)^{\frac{1}{\sigma-1}} \frac{\sigma}{\sigma-1} \left(1 + \frac{\Pi}{L}\right)^{\frac{1}{1-\sigma}}$. This is just the same expression as in Chaney (2006) except for the fact that in his paper the exponent of the last term is missing. Note at this point that we will see later on that λ_1 to λ_4 all consist of parameters and aggregate world profits only, which also only depend on parameters of the model. We can thus think of λ_1 to λ_4 as constants.

[14] $\lambda_2 = \frac{\sigma}{\sigma-1} \left(\frac{\gamma-(\sigma-1)}{\gamma}\right)^{\zeta-\eta} \left(\frac{\sigma}{\mu}\right)^{\beta\frac{1}{\sigma-1}} \left(1 + \frac{\Pi}{L}\right)^{-\beta\frac{1}{\sigma-1}}$.

Equilibrium firm exports: Exports of a firm with productivity x from i to j are given by $t_{ij}(x) = p_{ij}(x)q_{ij}(x)$. Using the expressions for $p_{ij}(x)$ and $q_{ij}(x)$ from above and equation (6), equilibrium exports of an individual firm are given by

$$t_{ij}(x) = \lambda_3 L_j^{-(\sigma-1)(\zeta-\eta)} L^{\frac{\sigma-1}{\gamma}} \left(\frac{\tau_{ij}}{\theta_j}\right)^{1-\sigma} x^{\sigma-1}. \tag{7}$$

Where λ_3 is a collection of constants and Π.[15]

Equilibrium cutoff productivity: The equilibrium expression for the cutoff productivity level can be derived combining (5) and (6) to get

$$\bar{x}_{ij} = \lambda_4 \delta^{-\frac{1}{b-\gamma}} L_j^{-\zeta} L_i^{-\frac{1}{b-\gamma}} L^{\frac{b}{b-\gamma}\frac{1}{\gamma}} \left(\frac{\tau_{ij}}{\theta_j}\right)^{\frac{b}{b-\gamma}} \bar{C}^{\frac{b}{b-\gamma}\frac{1}{\sigma-1}}. \tag{8}$$

Where once again λ_4 collects constants and Π.[16]

Equilibrium firm profits: In order to be able to derive the equilibrium world profits Π the expression for individual firm profits π_{ij} is needed. To get this we can start from equation (4). Using the expression for the fixed costs (3) and the equilibrium price level given by equation (6) leads to the following expression:

$$\pi_{ij}(x) = \lambda_5 L_j^{(1-\sigma)(\zeta-\eta)} L^{-\frac{\sigma-1}{\gamma}} \left(\frac{\tau_{ij}}{\theta_j}\right)^{1-\sigma} x^{\sigma-1} - \lambda_6 \delta^{-\frac{\sigma-1}{b-\gamma}} L_j^{\frac{\gamma(\sigma-1)}{b-\gamma}(\zeta-\eta)} \left(\frac{L_i}{L}\right)^{-\frac{\sigma-1}{b-\gamma}} \left(\frac{\tau_{ij}}{\theta_j}\right)^{\frac{\gamma(1-\sigma)}{b-\gamma}} \tag{9}$$

where again the λ-terms are collecting parameters and Π.[17] Note that the first part of the expression equals revenue minus variable costs, while the second part equals the fixed cost taking into account the cost reduction the spillovers induce in equilibrium.

Equilibrium world profits: We have now derived equilibrium expressions for all important variables which only depend on fundamentals of the model and Π. To close the

[15] $\lambda_3 = \mu \left(\frac{\sigma}{\mu}\right)^{\beta} \left(\frac{\gamma-(\sigma-1)}{\gamma}\right)^{(\sigma-1)(\zeta-\eta)} \left(1+\frac{\Pi}{L}\right)^{-(\sigma-1)(\zeta-\eta)}$.

[16] $\lambda_4 = \left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{\zeta} \left(\frac{\sigma}{\mu}\right)^{\zeta} \left(1+\frac{\Pi}{L}\right)^{-\zeta}$.

[17] $\lambda_5 = \left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{(\sigma-1)(\zeta-\eta)} \left(\frac{\sigma}{\mu}\right)^{(\sigma-1)(\zeta-\eta)} \left(1+\frac{\Pi}{L}\right)^{(1-\sigma)(\zeta-\eta)}$ and $\lambda_6 = \left(\frac{\gamma}{\gamma-(\sigma-1)}\right)^{-\frac{\gamma(\sigma-1)}{b-\gamma}(\zeta-\eta)} \left(\frac{\sigma}{\mu}\right)^{-\frac{\gamma(\sigma-1)}{b-\gamma}(\zeta-\eta)} \left(1+\frac{\Pi}{L}\right)^{\frac{\gamma(\sigma-1)}{b-\gamma}(\zeta-\eta)}$.

model an expression for equilibrium world profits is needed. This can be obtained using the definition of Π above together with equation (9). Evaluating the integral using (5), an expression emerges which can be simplified significantly using the definition of θ_j. This finally leads to an expression which can be solved for Π:

$$\Pi = \frac{\sum_{h=1}^{H}\left(\frac{\sigma_h - 1}{\gamma_h}\right)\frac{\mu_h}{\sigma_h}}{1 - \sum_{h=1}^{H}\left(\frac{\sigma_h - 1}{\gamma_h}\right)\frac{\mu_h}{\sigma_h}} L. \tag{10}$$

It can be directly seen that world profits only depend on parameters of the model and world population. Π is thus constant and exogenous.[18] Using this expression in equations (6) to (9) delivers the equilibrium values for all variables which now only depend on parameters and exogenous variables.

Equilibrium aggregate exports: Aggregate trade flows between i and j are given by $T_{ij} = \int_{\bar{x}_{ij}}^{\infty} t_{ij}(x) L_i \; dF(x)$. Using the equilibrium expressions for the cutoff level and individual firm exports from equations (8) and (7) together with $b^* = (b/(\sigma - 1))$ the following expression can be derived:

$$T_{ij} = \lambda_T \, \delta^{\frac{[\gamma/(\sigma-1)]-1}{b^* - [\gamma/(\sigma-1)]}} L_j \left(\frac{L_i}{L}\right)^{\left(1 + \frac{[\gamma/(\sigma-1)]-1}{b^* - [\gamma/(\sigma-1)]}\right)} \left(\frac{\tau_{ij}}{\theta_j}\right)^{-\gamma\left(1 + \frac{[\gamma/(\sigma-1)]-1}{b^* - [\gamma/(\sigma-1)]}\right)} \bar{C}^{-b^*\left(\frac{[\gamma/(\sigma-1)]-1}{b^* - [\gamma/(\sigma-1)]}\right)}. \tag{11}$$

Where λ_T is a constant.[19] This is a theoretically derived gravity-type trade equation and constitutes the main theoretical outcome of the analysis. Apart from constant terms trade flows are determined by the economic masses L_j and L_i, variable trade costs τ_{ij} and remoteness θ_{ij}.[20] The central feature of this equation are the exponents of these variables in which the impact of firm heterogeneity and exporter networks are reflected. For notational convenience, I define $\nu \equiv \frac{[\gamma/(\sigma-1)]-1}{b^* - [\gamma/(\sigma-1)]} > 0$.

[18] There are several aspects which are noteworthy. (a) The expression contains sums over all differentiated sectors (profits in sector 0 are zero by construction). (b) Only parameters enter the expression so it is justified to think of λ_1 to λ_6 as constants. (c) There is a small discrepancy between this expression and the one in Chaney (2006) which is (note that the superscript "*ch*" stands for Chaney): $\Pi^{ch} = \sum_{h=1}^{H} \frac{\mu_h(\sigma_h - 1)}{\gamma_h \sigma_h - \mu_h(\sigma_h - 1)} L$. (d) In the case of one sector ($H = 1$) (but only then) the two expression are equivalent. (e) However, this discrepancy does not affect the central findings and conclusions of the models.

[19] The constant is given by $\lambda_T = \mu\left(1 + \frac{\sum_{h=1}^{H}\left(\frac{\sigma_h - 1}{\gamma_h}\right)\frac{\mu_h}{\sigma_h}}{1 - \sum_{h=1}^{H}\left(\frac{\sigma_h - 1}{\gamma_h}\right)\frac{\mu_h}{\sigma_h}}\right)$.

[20] Assuming identical network technologies (just like production technologies) across countries, neither the elasticity b^* nor the decay factor δ are taken to be country specific.

2.4 Interpretation of the theoretical Results

In the context of this paper the most important question to be addressed using equation (11) is the impact of variable trade costs (i.e. empirically: distance) on aggregate trade flows. Traditional models of international trade explicitly or implicitly assume homogeneous firms. Thus the impact of variable trade costs on individual firm exports (the *intensive margin*) will translate one to one into the aggregate flows. A central feature of the introduction of heterogeneous firms is that variable trade costs also affect the cutoff productivity level and thus the number of exporting firms (*extensive margin*).

The overall effect of variable trade costs in the model is given by: $\frac{\partial \ln T_{ij}}{\partial \ln \tau_{ij}} = -\gamma\,(1+\nu)$. In the following it will be outlined how this overall effect can be decomposed into intensive and extensive margin and how the extensive margin can further be decomposed into a *basic extensive margin* (independent of network effects) and an *enforced extensive margin* (caused by the introduction of exporter networks).

Intensive margin: In order to isolate the intensive margin in the model, it is sufficient to assume that the number of exporting firms is exogenously fixed.[21] From equation (7) it follows that the elasticity of aggregate trade flows with respect to variable trade flows would then be given by: $\frac{\partial \ln T_{ij}}{\partial \ln \tau_{ij}} = \frac{\partial \ln t_{ij}}{\partial \ln \tau_{ij}} = -(\sigma-1)$. Note that here, a larger elasticity of substitution implies a *stronger* effect of variable trade costs on aggregate trade flows.

Basic extensive margin: In models with firm heterogeneity the cutoff level $\bar{x}_{ij}$ is endogenous and affected by variable trade costs τ_{ij}. Closing down the exporter networks in the model (taking C_{ij} to be exogenous), the cutoff level is given by $\bar{x}_{ij}^{nn} = const\left(\frac{L}{L_j}\right)^{\frac{1}{\gamma}}\left(\frac{\tau_{ij}}{\theta_j}\right)C_{ij}^{\frac{1}{1-\sigma}}$ where the superscript '*nn*' stands for 'no network' and *const* is a constant similar to λ_4. When this expression is used to derive T_{ij}, it is easy to check that $\bar{x}_{ij}$ (and thus τ_{ij}) is raised to the power $-\gamma+(\sigma-1)$ so, by reducing the number of exporters, variable trade costs further reduce trade (recall that $\gamma > (\sigma-1)$). This margin will be called the *basic extensive margin*. The overall effect of variable trade costs on aggregate exports would then be given by $\frac{\partial \ln T_{ij}}{\partial \ln \tau_{ij}} = -(\sigma-1)-\gamma+(\sigma-1)$. Where the first term on the r.h.s. represents the intensive margin discussed above and the second

[21] This is equivalent to Krugman (1980) in the sense that there the number of exporters is also exogenously given by the number of firms (all firms export).

and third term represent the basic extensive margin. Chaney (2006) provides a extensive analysis of the relation of these two margins. Note that the two terms including σ cancel and the effect of variable trade costs on aggregate trade flows is *independent* of the elasticity of substitution.

Enforced extensive margin: The crucial novel feature of the model is the introduction of exporter networks. Given that in equilibrium firms share information about exporting and thereby reduce their fixed cost of exporting, there is a feedback effect of the number of exporters on the cutoff level $\bar{x}_{ij}$. The impact of variable trade costs on the cutoff level is described in equation (8). Note that the elasticity $\left(\frac{b}{b-\gamma}\right)$ can be rewritten as $\left(1+\frac{[\gamma/(\sigma-1)]}{b^{*}-[\gamma/(\sigma-1)]}\right)$, where the latter part represents the impact of exporter networks (without the networks - i.e. $b^{*}=\infty$ - one would just have an exponent of unity). When computing T_{ij} by adding all individual firm exports, $\bar{x}_{ij}$ (and thus τ_{ij}) is additionally raised to the power $-\gamma+(\sigma-1)$ which implies: $\frac{\partial \ln T_{ij}}{\partial \ln \tau_{ij}}=-(\sigma-1)-\gamma+(\sigma-1)-\gamma\left(\frac{[\gamma/(\sigma-1)]-1}{b^{*}-[\gamma/(\sigma-1)]}\right)$. Simplifying gives the exponent of variable trade costs reported in equation (11). The last term on the r.h.s. reflects the effect of the exporter networks. It adds an additional (negative) effect of variable trade costs on aggregate trade flows on top of the intensive and basic extensive margin. This will be called the *enforced extensive margin.*

Some intuition: To better understand the three margins, consider two potential destinations j and k such that $\tau_{ij}<\tau_{ik}$. Traditional models would *ceteris paribus* predict that the number of exporters is the same but each individual firm's exports to j are higher than exports to k (intensive margin). In addition to this intensive margin, models with heterogeneous firms would imply a larger number of firms exporting from i to j than from i to k (basic extensive margin). The introduction of the exporter network finally implies that the positive effect of lower variable trade costs τ_{ij} on the number of firms (basic extensive margin) feeds back on the fixed cost of exporting (and more so for country j), further lowers the cutoff productivity level and thus leads even more firms to start exporting (enforced extensive margin). This effect is the stronger the larger the number of exporters. By adding the enforced extensive margin, *the introduction of exporter networks provides an endogenous mechanism enforcing the effect of variable trade costs on aggregate trade flows.*

The role of market structure: An interesting feature of the model is the role of the elasticity of substitution between varieties σ. Taking into account the intensive margin only, traditional theory implies that the a high value of σ should strengthen the (negative) effect of variable trade costs, the introduction of the basic intensive margin implies that the effect of variable trade costs should be *independent* of the elasticity of substitution: the degree of firm heterogeneity alone determines the role of τ_{ij}.

An interesting and novel feature of the model with exporter networks is that the elasticity of substitution actually does affect the impact of variable trade costs on trade flows but *dampening* the negative effect. This testable implication will be discussed in the next section.[22]

Country size: To the best of my knowledge all theoretical derivations of gravity equations imply elasticities with respect to the economic mass of countries of unity. It is noteworthy that the introduction of exporter networks strengthens the effect of the exporting country's size on aggregate trade flows. The reason is simple: the amount of information in the network is larger the more firms export. For a given cutoff-level the number of exporters is the larger the larger the overall number of firms in the country. This means that the relative size of the country of origin affects aggregate trade flows not only positively via the standard gravity type $\left(\frac{L_i L_j}{L}\right)$ channel but has an additional positive effect of L_i^{ν} on trade by reducing the level of fixed cost for exporters. *Ceteris paribus* larger countries export more because they acquire more relevant information and thus their firms face lower fixed costs.[23] A casual review of some of the empirical gravity

[22] There is a subtle but important difference to the results in Chaney (2006) which merits some attention. One of Chaneys main points is that in his model the elasticity of substitution *dampens* the negative effect of *trade costs* on aggregate exports. Note that this is true if one considers variable *and* fixed costs of trade *together*. In his model only the basic extensive margin is at work so the effect of *variable* trade costs on aggregate trade is *independent* of the elasticity of substitution. However, the negative impact of the *fixed* costs of exporting is *dampened* by the elasticity of substitution. So his claim that compared to Krugman (1980) in his model the role the elasticity of substitution for the effect of *trade costs* is reversed is correct. However, when he tests this prediction he uses bilateral distance as a proxy for these trade costs which is usually used as a proxy for *variable costs*. This is only a valid test of his model under the *assumption* that the (exogenous) fixed costs of exporting are *increasing* in distance. Note however that in the model with exporter networks there is an *endogenous* mechanism linking the fixed costs of exporting to the variable trade costs and thus directly to distance. So the effect of the elasticity of substitution goes directly via the variable trade costs and (as will be argued in the following section) Chaneys empirical results give direct support to the model.

[23] Note that the gravity equation refers to the absolute value of trade flows, not to the exports per capita which are known to be higher in small countries.

literature suggests that in many cases the coefficients estimated for the economic mass of the exporting country tend to be larger than the estimates for the economic mass of the importer. This difference appears to be mainly ignored or - with reference to existing theory - the exponents are imposed to be unity to start with. The model with exporter networks provides a possible explanations for differences in the exponents and might thus point at an interesting subject of future investigation.

3 Empirical Evidence

From the empirical side two issues are of particular interest. One is to identify (and if possible test) testable implications of the model in order to find general support for it. The second is to find an estimate of the parameter ν which is governing the strength of the network effect and which determines how much of the 'distance puzzle' can actually be explained by the model.

3.1 Testable Implications

The impact of exporter networks on the effect of variable trade costs on bilateral trade flows delivers different testable implications. Recall that in the model the effect of distance on aggregate trade flows is determined by $\nu \equiv \frac{[\gamma/(\sigma-1)]-1}{b^{*}-[\gamma/(\sigma-1)]}$. This term is increasing in the elasticity of fixed costs with respect to information $1/b^{*}$ and also in the degree of firm heterogeneity (or better: homogeneity - recall that a high γ stands for a homogeneous sector) γ and decreasing in the elasticity of substitution σ. These implications can potentially be tested.

Effect of network quality: We have seen in the analysis of the theoretical results above that a larger elasticity of fixed costs with respect to information $1/b^{*}$ enforces the distance effect on aggregate trade flows. This elasticity can be interpreted as the quality of the network: in a network of higher quality, the level of fixed costs decreases faster in the number of exporters participating in the network. Based on the model one would thus expect distance effects to be stronger in countries (or sectors) with higher network quality. If there was a good proxy for network quality available one could estimate a standard gravity equation including an interaction term of distance and this proxy. Based on the model one would expect the coefficient of this interaction term to be negative and significant. Unfortunately, by now I have not found an appropriate proxy for network quality in countries or sectors. However, the model has some other testable implications regarding the effect of distance on aggregate trade flows which are not directly related to network quality.

Effects of heterogeneity and market structure: The model predicts that for a given b^* sectors with a low degree of heterogeneity (high γ) and a low elasticity of substitution σ should be affected more strongly by variable trade costs than other sectors. This implies that a sectoral analysis using data on sectoral elasticities of substitution and sectoral firm heterogeneity could deliver support for the model.

Based on the model one would expect the effect of distance on trade flows in sector h to be stronger the the larger γ_h and the lower σ_h.[24] For sectoral trade flows this could be tested by adding an interaction term of distance and γ_h (σ_h) which would be expected to be of negative (positive) sign.

Interestingly, this analysis has already been carried out by Chaney (2006). From the firm size distributions of US manufacturing sectors he cannot derive a direct estimate of γ but, based on his model (and consistent with the model presented above), he obtains an estimate for the sectoral firm heterogeneity relative to the elasticity of substitution $\frac{\gamma_h}{\sigma_h - 1}$ (which is precisely the term governing the impact of exporter networks in my model). He uses this estimate interacted with distance (and some other trade cost proxies) in a standard sectoral gravity equation. He finds a positive and highly significant coefficient of this interaction term (see table 2, p. 25 in his paper).

He also provides evidence on the role of the elasticity of substitution. He uses the elasticities of substitution estimated by Broda and Weinstein (2004) for 3-digit SITC (revision 3) sectors. He interacts these sectoral elasticities with distance and finds a both statistically and economically significant *dampening* impact of the elasticity of substitution on the distance effect (table 3, p. 27).[25]

[24] Note that this implies that the elasticity of substitution σ_h plays exactly the opposite role compared to standard theory: it is dampening instead of magnifying the negative effect of variable trade costs.

[25] Since in Chaney's model only the basic extensive margin is present, the impact of variable trade costs on trade flows is determined by the degree of firm heterogeneity γ_h only. However, he uses the fact that his model implies that the elasticity of trade flows with respect to the (exogenous) *fixed costs* of exporting

Both results are nicely in line with the prediction of my model: without relying on any assumptions regarding the relation of exogenous parameters to distance, the model predicts the effects of firm heterogeneity and market structure on the distance effect present in Chaney's sectoral gravity equations.

3.2 Estimation of ν

In order to get an idea of the magnitude of the parameter ν which determines the strength of the effect of exporter networks in the model, one could calibrate its components. Consensual estimates of the elasticity of substitution between varieties σ are between 5 and 10 (see for example Anderson and van Wincoop (2004)). The degree of firm heterogeneity γ can be calibrated following Chaney (2006) who finds $\gamma \approx 2(\sigma - 1)$. However, because of a lack of empirical literature in the field, a value for the remaining elasticity of fixed cost of exporting with respect to the number of exporters cannot be easily picked.[26]

It is however possible to obtain an estimate for the full parameter ν by estimating the aggregate trade flow equation (11). One can use the fact that ν does not only enter the exponent of variable trade flows but also determines the elasticity of trade flows with respect to the size of the exporting country.

Bringing the aggregate trade flow equation to the data: Using the definition of ν, (11) can be rewritten as

$$T_{ij} = \lambda_\nu \; L_j \; L_i^{1+\nu} \; \tau_{ij}^{-\gamma(1+\nu)} \; \theta_j^{\;\gamma(1+\nu)}$$

where $\lambda_\nu = \lambda_h \; L^{-(1+\nu)} \; \bar{C}^{-\left(\frac{\gamma}{\sigma-1}-1\right)\frac{b}{b-\gamma}} \; \delta^{(1+\nu)}$ is a constant. Rewriting (11) in this way nicely shows the gravity structure of the bilateral trade equation. Just like in standard gravity

is determined by $\frac{\gamma_h}{\sigma_h - 1}$. He *assumes* that these fixed costs are increasing in distance and then carries out the analysis outlined above. Since in both cases the elasticity of substitution is present, this assumption is crucial in order to interpret the empirical results as support for his model. The main difference in my model is that the level of fixed costs is *endogenously* affected by variable trade costs, so taking distance as a proxy for variable trade costs, the model can explain the influence of $\frac{\gamma_h}{\sigma_h - 1}$ and σ_h on the distance effect.

[26] The findings of Koenig (2005) imply an elasticity of the fixed cost of exporting with respect to the number of exporters of -11.3% (which implies a ν of about 0.15). So according to her findings a firm's fixed cost of exporting to a particular market decrease by 1.13% when the number of exporters to that market in the same region increases by 10%. While her results show nicely that the effect of the number of exporters on the fixed cost of exporting is economically significant, her results are for the regional level and can thus not be used to calibrate the elasticity on the national level.

models, exports from i to j depend on a constant, the economic mass (here: the country sizes) and the variable trade costs. The parameter θ_j is an index of aggregate remoteness of the destination country. This index corresponds the 'multilateral resistance term' in Anderson and van Wincoop (2003).[27]

After Anderson and van Wincoop (2003) had pointed out that based on their theoretical model one would expect traditional gravity estimations to be suffering from an omitted variable bias, it has become standard in the literature to account for the presence of 'multilateral resistance terms' by adding country fixed effects (see e.g. Feenstra (2002) or Baldwin and Taglioni (2006)). As standard in the gravity literature (see e.g. Anderson and van Wincoop (2003)), variable trade costs are proxied by a function of the form $\tau_{ij} = d_{ij}^{\rho}$ where d_{ij} is bilateral geographical distance between (in most cases the capitals of) two countries and ρ is the elasticity of variable trade costs with respect to distance.[28]

L_j enters the model through the CES expenditure system and is one determinant of aggregate demand in country j. L_i enters the model as the number of firms. As standard in the literature they will both be proxied by the respective country's GDP.

Taking the theoretical model serious, in addition to distance and the $GDPs$ one needs an importer country fixed effect in order to account for the index of remoteness of the importing country, θ_j.

The following specification would thus deliver a direct estimate of $(1+\nu)$:

$$\ln T_{ij} = \beta_1 \ln d_{ij} + \beta_2 \ln GDP_i + \beta_m^j \varphi_m^j + \varepsilon_{ij} \tag{12}$$

where φ_m^j is a dummy variable which takes the value of one if in a given observation country j is the importer. On the grounds of the theory presented above, the estimate of β_2 has a direct interpretation as an estimate of $(1+\nu)$.

[27]Note that in their theoretical model Anderson and van Wincoop (2003) have two similar terms, one for the importer and one for the exporter. While the former corresponds to the θ_j, the latter has no correspondence in the model presented above. This is due to the simplifying assumption of the existence of the sector 'zero' which equalizes wages (and thus variable cost of production) across countries. Allowing for different cost of production Anderson and van Wincoop (2003) get the price index in the exporter country P_i as an additional term.

[28]If border effects are being investigated, a function of the type $\tau_{ij} = b_{ij}\, d_{ij}^{\rho}$ is used where $b_{ij} = 1$ if the trade flow takes place between two regions of the same country and one plus the tariff-equivalent of the border barrier if the flow crosses a border. Other determinants of trade cost such as common language or membership in the same free trade agreement can be introduced in a similar manner.

The data: Apart from the introduction of the importer fixed effects, the r.h.s. of equation (12) comprises only the most basic ingredients of gravity equations.[29] To test the model, data on unidirectional trade flows is needed. A detailed description of the data set can be found in Feenstra, Markusen, and Rose (2001).[30] Here, cross-sectional data for 1990, 1985, 1980, 1975 and 1970 is used. The GDP measures are drawn from the Penn World Table 5.6, the distance variable represents Great Circle distance between capital cities. An interesting feature of the data set is that trade flows (exports) are not only reported unidirectional but also sorted by the classification scheme introduced by Rauch (1999) into homogeneous, reference-priced and differentiated goods. This is very convenient because in the theoretical model equation (11) represents the aggregate trade flows in the differentiated good sector, only. So, in the empirical analysis the data on exports in differentiated goods will be used.[31]

In the analysis all countries for which the control variables are available are used. Due to data availability issues, the sample size varies between 108 and 137 countries for the different years.

Estimation results: The results of the estimation with OLS of equation (12) are reported in table (1). The coefficients of the dummy variables are omitted.

For all the available years, the coefficient of distance has the expected sign and is highly significant. It is relatively high in absolute value compared to other estimates in the literature, but it is still within the conventional range. The coefficient of GDP_i is also highly significant and is around 1.4 which implies according to our theory a value of the parameter ν of $\nu \approx 0.4$.[32] Despite its simplicity, the model fits the data quite well, explaining about 85% of the overall variation in unidirectional exports of differentiated goods.[33] This

[29] Robustness checks include the addition of other standard gravity variables. This leaves the results unaffected.

[30] I'm indebted to Andrew Rose who provides a large full gravity data set on his home page at http://faculty.haas.berkeley.edu/arose/RecRes.htm#Trade

[31] Differentiated goods represent about two thirds of the goods. Robustness checks show that using all three goods classifications together the estimates for ν decrease a bit but stay statistically and economically significant

[32] Using the data on all three goods classifications reduces the absolute value of the distance coefficients by about 0.2 and delivers an average estimate of ν of about 0.3. Adding dummies for common language, common border and membership in the same regional free trade agreement does not alter the results.

[33] ⋄ Note that the results should be the same independent of whether N dummies are used or whether $(N-1)$ dummies are used and a constant is added. But e.g. for 1990, for the former specification one finds $R^2 = 0.89$ (adj$R^2 = 0.89$), the latter specification implies $R^2 = 0.61$ (adj$R^2 = 0.60$) while the estimates

Estimations of Equation (12) with max. sample size					
	1990	1985	1980	1975	1970
$\ln d_{ij}$	-1.63	-1.47	-1.53	-1.54	-1.53
	(.04)	(.04)	(.04)	(.04)	(.03)
$\ln GDP_i$	1.47	1.35	1.45	1.38	1.29
	(.02)	(.01)	(.01)	(.01)	(.01)
importer dummy	yes	yes	yes	yes	yes
no. of countries	108	137	133	115	113
no. of obs.	7604	10004	10042	8648	7900
R^2	0.89	0.86	0.86	0.85	0.85

Table 1: Dependent variable: exports of differentiated goods from country i to country j. Robust standard errors in parenthesis, coefficients of dummies are not reported.

is a typical feature of gravity equations.

A value of $\nu = 0.4$ would imply an elasticity of the fixed cost of exporting with respect to the number of exporters, $1/b^*$ of 0.22. This means that an increase of 10% in the number of firms exporting from i to j decreases the fixed cost of exporting form i to j by 2.2%. This value is about twice as high as the value implied by the findings of Koenig (2005) for the regional level. Given that firms from the same country share many relevant characteristics and export promoting institutions like the chambers of commerce operate on the national level, the estimated value of 0.22 seems quite reasonable. The value is economically significant but at the same time does not appear unrealistically large.

Using this first-pass estimate of ν, we can make an attempt to put some numbers to the model and to get an idea of how much of the 'distance puzzle' the model can explain.

4 Exporter Networks and The 'Distance Puzzle'

In this section the results of the theoretical model and the empirical exercise will be combined and their implications for the 'distance puzzle' will be presented and discussed.[34]

of the coefficients for distance and GDP_i remain unchanged. This discrepancy might merit some further investigation.⋄

[34] It is of interest to note that the model can also be used to shed some more light on the 'border puzzle'. Chaney (2006) shows that the introduction of firm heterogeneity leads to a considerable reduction in the implied tax equivalents of border barriers. The introduction of exporter networks enforces this effect in a similar manner as in the case of the 'distance puzzle'.

The level: Equipped with an estimate of the parameter ν, we can determine the level of the distance effect predicted by the model with heterogeneous firms and exporter networks to the predictions of standard theory.

To do so, first consider the reasoning along the lines of Grossman (1998) outlined in Anderson and van Wincoop (2004). Based on their theoretical model (which - just like most traditional models with homogeneous firms and a CES expenditure system - only features the intensive margin) one obtains as theoretical prediction for the distance effect

$$\frac{\partial \ln T_{ij}}{\partial \ln d_{ij}} = -(\sigma - 1)\ \frac{\partial \ln \tau_{ij}}{\partial \ln d_{ij}}.$$

Defining τ_{ij}' as the tax equivalent of transport costs i.e. $\tau_{ij} = 1 + \tau_{ij}'$ it is easy to show that $\frac{\partial \ln \tau_{ij}}{\partial \ln d_{ij}} = \frac{\tau_{ij}'}{1+\tau_{ij}'}\ \frac{\partial \ln \tau_{ij}'}{\partial \ln d_{ij}}$. Like this, all the different determinants of the distance effect can be calibrated. Anderson and van Winccop use an estimate of Hummels (2001) for the elasticity of (tax equivalent) trade costs with respect to distance of about 0.3. Their preferred choice for the trade cost is 11% but they argue that including time costs one could go up to 21%. For the elasticity of substitution between varieties σ they use consensus estimates between 5 and 10. This calibration implies a distance effect in the interval of $\frac{\partial \ln T_{ij}}{\partial \ln d_{ij}} \in [-0.12, -0.46]$. Where the lower bound represents $\sigma = 5$ and $\tau_{ij}' = 11\%$ and the upper bound is computed using $\sigma = 10$ and $\tau_{ij}' = 21\%$. It is obvious that this interval is far away from capturing the distance effects commonly found in empirical estimations. To provide some illustration for how miserably the standard theory fails to explain empirical distance effects, I present figure 1 in Disdier and Head (2007) in the appendix. It reports 1467 estimated distance effects from 103 studies. It is obvious that most of the estimates are far above the interval proposed by standard theory.

It follows directly from the analysis of the theoretical results above that by adding the basic and enforced extensive margin on top of the intensive margin the model with exporter networks implies an elasticity of trade flows with respect to distance of

$$\frac{\partial \ln T_{ij}}{\partial \ln d_{ij}} = \ -\gamma\ (1+\nu)\ \frac{\partial \ln \tau_{ij}}{\partial \ln d_{ij}}.$$

Based on the findings of Chaney (2006) one can calibrate the degree of firm heterogeneity according to $\gamma \approx 2(\sigma - 1)$. Doing so and using the same values as above for the other parameters and the estimate of $\nu \approx 0.4$ from the last section, one finds that the theoretical model with heterogeneous firms and exporter networks would suggest a distance effect of $\frac{\partial \ln T_{ij}}{\partial \ln d_{ij}} \in [-0.34, -1.29]$. A look at the figure 1 in Disdier and Head (2007) (reported in the appendix) reveals that this interval includes the major part of the values of the distance effect commonly estimated.[35] [36]

This suggests that the introduction of firm heterogeneity and exporter networks into a standard monopolistic competition trade model can account for the otherwise puzzlingly high values of estimated distance effects. But the model can also be used to address the second dimension of the 'distance puzzle': the increase over time of the distance effects on international trade.

The dynamics: The model can also be used as a formal framework to think about the increase of the distance effects many studies find over time. A very informative analysis of this finding is provided in the meta analysis of Disdier and Head (2007).

To address this dimension of the 'distance puzzle' it is useful to first understand why a distance effect increasing over time should be puzzling. It is commonly accepted that the past fifty years have been marked by a strong improvement in transportation and information technologies. In addition, variable trade costs have been reduced by cuts in tariffs and other trade barriers. So how can it be that the negative impact of geographical distance (commonly used as a proxy for variable trade costs) on trade flows has been increasing over long time periods?

The model suggests a possible explanation which is nicely in line with the notion that technologies have improved over the last decades. In fact, the model implies that an *increase* in the quality of exporter networks actually *enforces* the impact of distance on

[35] Note that the basic effect comes from the introduction of firm heterogeneity. However, exporter networks are important in order to enforce the effect such that the intervals include a large part of the estimates. Closing down the exporter networks ($\nu = 0$) would imply an interval for the expected distance effect of $[-0.24, -0.92]$.

[36] The estimate on $\nu \approx 0.4$ does not fully solve the 'distance puzzle' in the empirical analysis in the previous section: here we find quite high distance effects around 1.5 in order to get the intervals up to this magnitude one would need a ν of about 0.6. this would imply an elasticity of fixed costs with respect to the number of exporters of about 0.38 which is relatively high but does not seem to be out of plausible bounds.

aggregate trade flows. Note that this does not mean that countries are trading *less* with more distant countries than three of four decades ago. In fact an increase in the quality of the networks would imply that countries trade *more* with distant countries but the trade promoting effect would even be *stronger* for partners that are nearby. So in a world in which an improvement of the network technology takes place, a country's trade flows to all trading partners increase but the increase in the trade flows to partners that are closer is relatively stronger than the increase in the flows to distant partners. One would thus expect the elasticity of trade flows with respect to distance to increase.[37]

While Disdier and Head (2007) find significant increases in estimated gravity equations of studies using data of the time periods between the 1950s and the 1980s, the increase in the 1990s is not significant which means that the distance effect might have remained constant at a high level since the end of the 1980s. Does this mean that in a period of huge improvement in information technology the network technology has stopped improving? Not necessarily. A possible interpretation which goes beyond the framework of the model is that while from the 1950s to the 1980s networks on the *national* level (and their improvement) might have played the dominant role, in the 1990s networks beyond national borders might have become more important. This intuition is in line with the amelioration of information technologies, is supported by the strong regional integration in many parts of the world and also by the emergence of more and more multinational firms. Obviously, models and empirical studies considering trade between counties cannot address this dimension of globalization. So the amelioration of exporter networks might have continued in the 1990s but one would have to find a different way of looking for its effects in the data.

5 Conclusions

The main aim of this paper has been to address the 'distance puzzle' along its two dimensions: the high level and the increase over time of estimated distance coefficients.

[37]A very informative discussion of why estimated distance effects are not decreasing over time can be found in Buch, Kleinert, and Toubal (2004) the main intuition is that as long as the decrease in variable trade costs is proportional across destinations, the estimated distance effect remains constant because it represents an estimate of an *elasticity*. My model adds on top of this intuition that a better network technology may actually *increase* this elasticity while trade among all partners increases.

Different to the rest of the literature the puzzle was addressed from the theoretical side, constructing a model of international trade with heterogeneous firms and exporter networks. The model turned out to be appropriate for this purpose. Based on the model the strength of the effect of exporter networks on international trade could be estimated. The results suggest that in the light of the model level and dynamics of the estimated distance coefficients loose much of their mystery. However, the model has some other interesting implications that merit some further investigation. Ongoing work addresses implications of exporter networks for the 'border puzzle', the role of popular gravity variables such as common language, common borders, colonial ties and FTA membership in gravity equations and how the model might serve as a theoretical foundation for studies addressing the impact of familiarity on international trade.

References

Aitken, Brian, Gordon H. Hanson, and Ann E. Harrison, 1997, "Spillovers, Foreign Investment, and Export Behavior", *Journal of International Economics* 43, 103–132.

Anderson, James E., 2000, "Why do Nations Trade (So Little)?", *Pacific Economic Review* 5, 115–135.

Anderson, James E., and Eric van Wincoop, 2003, "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review* 93, 170–192.

Anderson, James E., and Eric van Wincoop, 2004, "Trade Costs", *Journal of Economic Literature* 42, 691–751.

Arkolakis, Costas, 2006, "Market Access Costs and the New Consumers Margin in International Trade", *University of Minnesota, mimeo.*

Aviat, A., and N. Coeurdacier, 2006, "The geography of trade in goods and asset holdings", *Journal of International Economics, in press.*

Baldwin, R., and D. Taglioni, 2006, "Gravity for Dummies and Dummies for Gravity Equations", *NBER Working Paper No. 12516.*

Barrios, Salvador, Holger Görg, and Eric Strobl, 2003, "Explaining Firms' Export Behaviour: The Role of R& D and Spillovers", *Oxford Bulletin of Economics and Statistics* 65, 475–496.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum, 2003, "Plants and Productivity in International Trade", *American Economic Review* 93, 1268–1290.

Bernard, Andrew B., and J. Bradford Jensen, 2004, "Why Some Firms Export", *The Review of Economics and Statistics* 86, 561–569.

Broda, C., and D. Weinstein, 2004, "Determinants and Effects of Foreign Direct Investment: Evidence from German Firm-Level Data", *Federal Reserve Bank of New York, mimeo* pp. 51–110.

Buch, Claudia M., Jörn Kleinert, and Farid Toubal, 2004, The Distance Puzzle: On the Interpretation of the Distance Coefficient in Gravity Equations, *Economics Letters* 83, 293–298.

Bugamelli, M., and L. Infante, 2003, "Sunk Costs of Exports", *Bank of Italy Research Papers.*

Chaney, Thomas, 2006, "Distorted Gravity: Heterogeneous Firms, Market Structure and the Geography of International Trade", *mimeo: University of Chicago.*

Combes, Pierre-Philippe, Miren Lafourcade, and Thierry Mayer, 2005, "The trade-creating effects of business and social networks: evidence from France", *Journal of International Economics* 66, 1–29.

DIHK, 2005, *"Going International: Erfolgsfaktoren im Auslandsgeschäft"*. (Deutscher Industrie- und Handelskammertag - German Chamber of Commerce, Berlin).

Disdier, Anne-Célia, and Keith Head, 2007, "The Puzzling Persistence of the Distance Effect on Bilateral Trade", *Review of Economics and Statistics (forthcoming).*

Eaton, Jonathan, and Samuel Kortum, 2002, Technology, Geography, and Trade, *Econometrica* 70, 1741–1779.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz, 2005, "An Anatomy of International Trade: Evidence from French Firms", mimeo: University of Minnesota.

Feenstra, R., 2002, "Border Effects and the Gravity Equation: Consistent Methods for Estimation", *Scottish Journal of Political Economy* 49, 491–506.

Feenstra, R., J. Markusen, and A. Rose, 2001, "Using the gravity equation to differentiate among alternative theories of trade", *Canadian Journal of Economics* 34, 430–447.

Greenaway, D., N. Sousa, and K. Wakelin, 2004, "Do Domestic Firms Learn to Export from Multinationals?", *European Journal of Political Economy* 20, 1027–1043.

Grossman, G., 1998, "Comment on Deardorf", in *The Regionalization of the World Economy* J.A. Frankel, ed. U. Chicago Press, pp.33-57.

Helpman, Elhanan, Marc J. Melitz, and Yona Rubinstein, 2006, "Trading Partners and Trading Volumes", mimeo: Harvard university.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple, 2004, "Export versus FDI with Heterogeneous Firms", *American Economic Review* 94, 300–316.

Huang, Rocco R., 2006, "Distance and trade: Disentangling unfamiliarity effects and transport cost effects ", *European Economic Review, In Press.*

Hummels, 2001, "Towards a Geography of Trade Costs", Working paper Purdue University.

Jackson, Matthew O., and Asher Wolinsky, 1996, "A Strategic Model of Social and Economic Networks", *Journal of Economic Theory* 71, 44–74.

Koenig, P., 2005, "Agglomeration and the Export Decision of French Firms", *INSEE Working Paper no. 2005-02.*

Krugman, Paul, 1980, Scale Economies, Product Differentiation, and the Pattern of Trade, *American Economic Review* 70, 950–59.

Melitz, M., 2003, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica* 71, 1695–1725.

Melitz, Marc J., and Gianmarco I.P. Ottaviano, 2005, "Market Size, Trade, and Productivity", NBER Working Papers 11393 National Bureau of Economic Research, Inc.

Portes, R., and H. Rey, 2005, The Determinants of Cross-Border Equity Flows, *Journal of International Economics* 65, 269–296.

Rauch, James E., 1999, Networks versus markets in international trade, *Journal of International Economics* 48, 7–35.

Rauch, James E., 2001, "Business and Social Networks in International Trade", *Journal of Economic Literature* 39, 1177–1203.

Rauch, James E., and Alessandra Casella, 2003, "Overcoming Informational Barriers to International Resource Allocation: Prices and Ties", *Economic Journal* 113, 21–42.

Rauch, James E., and Vitor Trindade, 2002, Ethnic Chinese Networks in International Trade, *Review of Economics and St* 84, 116–130.

Roberts, M., and J. Tybout, 1997, "What Markets Exports Boom?", The International Bank for Reconstruction/The World Bank, Washington D.C.

Zenou, Yves, 2006, "Networks: Theory and Applications", *Lecture Notes*, Working paper, IUI, GAINS and CEPR.

A Appendix

A.1 Basic concepts of Network Theory

In this appendix some of the crucial concepts and definitions of network theory are described. The presentation follows Zenou (2006) closely.

A *network* is a set of agents, called nodes, $M = \{1, ..., m\}$ and a set Λ of links between them. A network is thus a list of unordered pairs of players $\{a, b\}$ and can be denoted by $g \equiv g(M, \Lambda)$. The links are represented by a network $g \in G$, where $g_{ab} = 1$ if a has a link to b and $g_{ab} = 0$ otherwise. Links are taken to be reciprocal, so that $g_{ab} = g_{ba}$. The *geodesic distance* between players a and b, $d(a, b)$, is the length of any shortest path between a and b i.e. it is unity in the case of a direct link and, for indirect connections, one plus the minimum number of nodes which have to be passed to get from a to b in a given network.[38]
Consider some payoff function $u(g) = (u_1(g), ..., u_n(g))$ which assigns a payoff to every one of the M agents as a function of the network g connecting them. Denote by ab the link between a and b. Then, following Jackson and Wolinsky (1996), a *pairwise-stable* network is defined as follows:

Definition 1: A network g is *pairwise-stable* for the payoff function u if and only if
(i) for all $ab \in g$, $u_a(g) \geq u_a(g - ab)$ and $u_b(g) \geq u_b(g - ab)$ and
(ii) for all $ab \notin g$, if $u_a(g) < u_a(g + ab)$ then $u_b(g) > u_b(g + ab)$.

Here, ab stands for a link between player a and player b and $(g - ab)$ is the network g without the link ab. In words the definition implies that, a network is pairwise-stable if (i) no player gains by cutting an existing link, and (ii) no two unconnected players both gain by creating a direct link between each other.
Since pairwise stability only considers the creation or destruction of single links it is useful to also have an equilibrium definition which allows for creation (destruction) of a larger number of links. Consider the following network game: All agents in M individually

[38] Take for example a network of four players A, B, C and D with every player having two links (a square). Then the geodesic distance of A to its two neighbors is one while the geodesic distance to the fourth player is two.

announce the links they wish to form. For all $a, b \in M$, $s_{a,b} = 1$ if a wants to form a link with b and zero otherwise. A strategy of agent a is $s_a = (s_{a,1}, ..., s_{a,a-1}, s_{a,a+1}, ..., s_{a,m}) \in S_i$, and $S_i = \{0,1\}^{n-1}$ is the set of pure strategies available to a. The link ab is created iff $s_{a,b} \times s_{b,a} = 1$ i.e. if both players announce that they wish to form a link. Let $S = S_1 \times ... \times S_m$. A pure strategy profile $s = (s_1, ...s_n) \in S$ induces a non-directed network $g(s)$ and a vector of payoffs $u(g(s))$. In the following notation s is omitted where it causes no confusion. Denote this network formation game by $\Psi(u)$.

A pure strategy profile $s^* = (s_1^*, ..., s_m^*)$ is a Nash equilibrium of the network formation game $\Psi(u)$ if and only if: $u_a(g(s^*)) \geq u_a(g(s_a, s_{-a}^*))$ for all $s_a \in S_a$ and all $a \in M$. Here, s_{-a}^* is the pure strategy profile s^* without the strategy s_a. The Nash equilibrium is is too weak an equilibrium concept to deliver meaningfull results.[39] It is thus combined with the concept of pairwise stability:

Definition 2: A network is a *pairwise-Nash equilibrium* if and only if there exists a Nash equilibrium strategy profile profile s^* that supports g, that is $g = g(s^*)$, and, for all $a \in M$ and all $ab \notin g$, if $u_a(g) < u_a(g + ab)$ then $u_b(g) > u_b(g + ab)$.

In words, g is a pairwise-Nash equilibrium if it is both pairwise-stable and a Nash equilibrium. This concept adds to pairwise stability that it allows for multiple link creation and destruction.

A.2 Empirical distance effects:

The following graph taken from Disdier and Head (2007) shows 1467 estimates of distance effects found in 103 studies. The median is around 0.9 and the largest number of estimates is between 0.5 and 1.4.

[PUT FIG 1 OF DISDIER AND HEAD HERE]

[39] One could for example have a Nash equilibrium where two players could both increase their utility by forming a link. But *given* the partner does not announce a link creation, not announcing neither is a Nash equilibrium which also includes the empty network.