Gravity and Information:
Heterogeneous Firms, Exporter Networks and the ‘Distance Puzzle’

Sebastian Krautheim
Gravity and Information: 
*Heterogeneous Firms, Exporter Networks and the ‘Distance Puzzle’*

Sebastian Krautheim
Gravity and Information:
Heterogeneous Firms, Exporter Networks and the ‘Distance Puzzle’*

Sebastian Krautheim**

November 2007

Abstract
Distance effects in empirical gravity equations appear to be too high to be explained by transport costs alone. Moreover, despite the strong and ongoing reduction of transport costs, the estimated coefficients are rather increasing than decreasing over the last six decades. To address the two dimensions of this ‘distance puzzle’, this paper proposes a model of international trade in which heterogeneous firms create informational networks to reduce their fixed costs of exporting. Since the variable trade cost (distance) affects the number of exporters, which in turn affects the available information, the fixed cost of exporting is endogenously increasing in distance. The model thus delivers higher predictions for the level of distance effects and, in addition, a quality improvement of the networks over time implies increasing distance elasticities. Major implications of the model regarding the effects of firm heterogeneity and market structure on distance effects are supported by existing empirical evidence. An empirical gravity equation is used to estimate the effect of exporter networks on distance coefficients. In the light of the results the empirical findings on the role of distance on international trade appear considerably less puzzling.

JEL: F10, F12, D85
Keywords: gravity, heterogeneous firms, networks, trade costs, distance, information

---
* I am grateful to Giancarlo Corsetti and Morten Ravn for constant advice and encouragement. I would also like to thank Itai Agur, Daniel Bernhoft, Thomas Chaney, Alan Deardorff, Ethanan Helpman, Benjamin Jung, Wilhelm Kohler, Gianmarco Ottaviano, Philip Sauvè, Eva Schneider, Daria Taglioni, Yves Zenou, participants of the EUI working group on international trade, the 7th RIEF Doctoral Meeting in Rennes 2007, the 6th GEP Postgraduate Conference in Nottingham 2007, the RIEF Conference in Rome 2007, the Globalization Workshop of the IAW in Tübingen 2007, the 9th ETSG annual conference in Athens 2007 and the VIIS annual meeting in Munich 2007 for fruitful discussions and comments. Any remaining errors are mine. First version: October 2006
** 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 the leading tools for the empirical analysis of international trade flows. Despite their remarkably simple structure, they fit the data quite 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.

Grossman (1998), however, argues that estimated distance effects are far too high to be explained by transport costs alone. According to Anderson and van Wincoop (2004), 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. As outlined by Désider and Head (2007), 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. Perhaps 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 periods of technological improvements, tariff reductions and decreasing transport costs the effect of geographical distance on trade flows is so strong, and how can this effect be even increasing over time? It has been argued (see for example Anderson (2000)) that this discrepancy between theoretical predictions and empirical findings points at a miss-specification of the empirical gravity equations: an omitted variable (correlated with geographical distance) might lead to an overestimation of the distance effect. The problem with this argument is, as pointed out among others by Aviat and Coeurdacier (2006), that such a variable has not been found yet.

The novel approach of this paper is not to question the empirical estimates of distance effects but to propose a theoretical model which rationalizes the strong effect of distance and which can account (for given technological improvements) for an increase of this effect over time.

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.

In the model, 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 magnifies the effect of variable trade costs on the extensive margin.
I will speak of the basic extensive margin and the magnified extensive margin in order to distinguish the effect present in standard models with heterogeneous firms and the effect which is introduced by the exporter networks.

The magnified 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 implies high estimated distance coefficients. The basic intuition is that variable trade costs influence the number of exporters which in turn affects the level of fixed costs faced by firms. This opens an additional channel for variable trade costs to affect aggregate trade flows. 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. The intuition behind this result is that a country facing an improvement in the network technology trades more with all its trading partners but disproportionately more with closer countries. In elasticity terms, this translates into a stronger elasticity of trade flows with respect to distance.

Existing empirical evidence strongly supports crucial implications of the model, in particular the role the degree of firm heterogeneity and the elasticity of substitution play for 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 yields an estimate of the magnitude of the magnified extensive margin. The results suggest that the model can also account quantitatively for the level of estimated distance coefficients.

The paper contributes to the recent literature on international trade with heterogeneous firms². The basic structure of the model is borrowed from Chaney (2006). Exporter networks are modelled drawing on basic concepts and methods of the recent microeconomic literature on networks pioneered by the work of Jackson and Wolinsky (1996).

There is a large literature considering (mostly empirically) the role business and social networks play

¹Note 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 is 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 costs, which is in line with the empirical evidence provided by Chaney (2006).

in international trade. There are 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 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.

Some authors have conjectured that different levels of information about export destinations (which are by assumption) decreasing in distance play a role for understanding distance effects. To the best of my knowledge, however, no theoretical model has been spelled out linking distance to available information. In this paper - based on a complete general equilibrium model - 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 which in turn lowers the fixed cost of exporting and leads even more firms to export.

The link between available information and fixed cost of exporting is crucial in the model. The empirical evidence on 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. Some direct evidence comes from survey studies. Roberts and Tybout (1997) report results of a survey of 186 Colombian firms interviewed by the World Bank and the Colombian government’s export promotion agency in 1990. 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 and had used private and public services to overcome these informational obstacles.

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 pro-

---


4It is important to note that in the proposed model networks are destination specific. If this were not the case, better networks would be observationally equivalent to important export specific infrastructure like e.g. ports (for the empirical relevance of ports in international trade see Wilson, Mann, and Otsuki (2003) and Blonigen and Wilson (2007)).


6I consider it an advantage of the model that, because of their dependence on available information, the fixed costs of exporting endogenously increase in distance. This intuitive relation has been assumed in empirical studies (see Kleinert and Tenhaef (2006) and Chahey (2006)), but to the best of my knowledge this model is the first to provide a theoretical mechanism in general equilibrium linking the fixed costs to geographical distance.
sion of information, consulting and exchange of experience organized by the chamber of commerce). The exporter networks in the model can be thought of as representing this indirect channel. 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.

The model suggests that empirically one should see a positive relationship between the number of exporting firms in country \( i \) and the probability of a firm in country \( i \) to start exporting. The evidence on such spillovers is mixed: some studies find evidence for positive spillovers and others do not.\(^7\) An important shortcoming of most studies is that no information about the export destination is available. But the model suggests that the spillovers observed should in fact be destination specific. Using a very detailed French firm level dataset which includes export destinations, Koenig (2005) closes this gap in the literature. In line with the predictions of the model, she finds strong evidence for destination specific spillovers on the regional level.

The remainder of the paper is structured as follows. The next section presents a model with heterogeneous firms and exporter networks. Section 3 discusses implications for the intensive and extensive margins focusing on the magnifying effect of exporter networks on the extensive margin. Section 4 shows how the model with heterogeneous firms and exporter networks can be used to explain the two dimensions of the ‘distance puzzle’. Section 5 derives some other implications of the model regarding the degree of firm heterogeneity, substitutability between varieties, the ‘border puzzle’ and the role of country size. Section 6 carries out an empirical exercise yielding an estimate of the magnitude of the network effects. Section 7 concludes.

2 The Model

This section introduces a model of trade with heterogeneous firms and exporter networks. The structure of the model, which is based on the work of Melitz (2003) and Clancy (2006), is outlined in the first subsection. Using a methodology based on Calvo-Armengol and Kilic (2007), exporter networks are introduced in the second subsection. In the third subsection all conditions for the world general equilibrium and the main theoretical result of the paper, a theoretically derived gravity equation, are derived taking into account the equilibrium network structure.

2.1 The Economy

**Basic structure:** The world economy consists of $N$ countries with $L_n$ denoting the population in country $n$. 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 implying 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:

$$U = \frac{1}{\mu_0} \prod_{h=1}^{H} \left( \int_{\Omega_h} q_h(\omega) \frac{\sigma_h^{-1}}{\sigma_h \omega} d\omega \right)^{\frac{\sigma_h}{\sigma_h - 1} \mu_h}$$

where $\Omega_h$ is the set of varieties from sector $h$ available to the workers, $q_h(\omega)$ is the quantity consumed of each individual variety $\omega$ from sector $h$, $q_0$ denotes the quantity of the homogeneous good consumed, $\mu_0 + \sum_{h=1}^{H} \mu_h = 1$ and $\sigma_h$ is the elasticity of substitution between varieties of sector $h$.

**Firms:** The number of firms (varieties) 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 competition case of the generalized framework proposed by Eaton, Kortum, and Kramarz (2005).\(^8\)

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 $\tau_{ij}^h$. 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 $q/\tau_{ij}^h$ units in $j$ is given by $c(q) = \frac{q}{x} + C_{ij}^h$. Facing isoelastic demand curves, the price charged in $j$ by a

\(^8\)Their empirical analysis using the above-mentioned French firm-level dataset strongly supports the monopolistic competition structure.
firm from \( i \) with productivity \( x \) is thus given by

\[
p_{ij}^h(x) = \frac{\sigma_h \pi_{ij}^h}{\sigma_h - 1} x.
\]

(1)

Firms differ in their productivity levels. As standard in the literature (see e.g. Helpman, Melitz, and Yeaple (2004) and references therein) an individual firm’s productivity is assumed to be drawn from a Pareto distribution with parameter \( \gamma_h \) so that \( P(X < x) = F_h(x) = 1 - x^{-\gamma_h} \). Without loss of generality the minimum productivity level \( x_{\min} \) is normalized to one. In order to have a finite second moment it is standard to assume \( \gamma_h > 2 \). Furthermore, it is assumed that \( \gamma_h > (\sigma_h - 1) \).

**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 consists of 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 \( \Pi \) are world profits and \( L \) stands for world population. Demand in \( j \) for a variety of good \( h \) imported from \( i \) is given by

\[
d_{ij}^h = \mu_h \left( 1 + \frac{\Pi}{L} \right) L_j \left( \frac{p_{ij}^h(x)}{P_j^h} \right)^{-\gamma_h} \left( \frac{P_h^i}{P_j^h} \right)^{-1},
\]

where \( P_j^h \) is the welfare based price index in sector \( h \).

### 2.2 Exporter Networks

This subsection introduces the crucial novel feature of the model. In each country, firms in a given sector can form networks in order to exchange information which is relevant (cost-reducing) for exporting. Formally, the model specification is enriched drawing on the recent microeconomic literature on networks.

---

9To see the importance of the last condition note the following: the lower \( \gamma \), the higher is the mass of firms with high productivities. If \( \sigma \) is high, goods are close substitutes. The closer \( \gamma \) 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 breaks down.

10The welfare-based price index for the varieties of good \( h \) in country \( j \) reflects all prices set by firms (from all countries) selling varieties of good \( h \) in \( j \): \( P_j^h = \left( \sum_{k=1}^N L_k \int_{\pi^h_{k,j}}^\infty \left( \frac{\sigma_h \pi_{k,j}^h}{\pi^h_{k,j}} x \right)^{1-\gamma_h} dF_h(x) \right)^{\frac{1}{1-\gamma_h}} \). Where \( \pi_{k,j}^h \) denotes the cutoff productivity level in sector \( h \). For firms with a productivity above this level it is profitable to export from \( k \) to \( j \), while firms below this level choose not to serve market \( j \). World profits are defined as the sum of the profits any firm (in any sector) makes in any market: \( \Pi = \sum_{h=1}^H \sum_{k=1}^N L_k \int_{\pi^h_{k,j}}^\infty \pi^h_{k,j}(x) dF_h(x), \) where \( \pi^h_{k,j}(x) \) are net profits a firm with productivity \( x \) in sector \( h \) of country \( k \) makes by exporting to \( l \).
Departures from standard network theory: Standard 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 unique amount of relevant information related to its export activity, which is proportional to its measure in the continuum. This information is sector- and destination-specific. The original information sets are “equidistantly disjoint,” which means that when firm $a$ obtains the original information of one additional firm in the sector, the effect on the information set of firm $a$ is the same, independently of the identity of the firm from which the information originates. What counts is only the amount of information a firm has gathered, not its origin.

(a2) Let $\pi^h_{ij}$ be the vector of profits individual firms in sector $h$ make by exporting from country $i$ to $j$. In the network formation game $\Psi^h(\pi^h_{ij})$, firms in sector $h$ can offer the creation of a link to other firms in the sector in order to exchange information about exporting from $i$ to $j$ (thereby reducing the costs of exporting) in order to maximize profits. 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 $\delta < 1$ i.e., if firm $a$ and $b$ exchange information, $b$ only receives a fraction $\delta$ of the information set of firm $a$. Being transferred along a link, information loses some degree of complexity/detailedness “on the way.” The information lost only depends on the number of links it is transferred through, not on the actual path it takes through the network.\footnote{It would be simple to model the decay factor $\delta$ as country or sector specific, but this would not yield any additional insights for the analysis of distance effects.}

(a4) If firms are indifferent between creating and not creating a link, they do not propose link creation.\footnote{These assumptions (a3) and (a4) are not crucial for the final results but they assure uniqueness of the equilibrium network.}

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 its level of fixed cost of exporting from country $i$ to $j$. So the fixed cost level faced by firm $a$ in sector $h$ to export from country $i$ to $j$, $C_{ij}^h(a)$ would be determined by some function $\Phi(.)$ such that $C_{ij}^h(a) = \Phi(\ell_{ij}^h(a), \bar{C}^h)$, where $\ell_{ij}^h(a)$ is firm $a$’s stock of information about exporting a variety of sector $h$ from $i$ to $j$ and $\bar{C}^h$ is an exogenous cost...
factor.\textsuperscript{13}

There are some desirable properties this function should have. \textit{Ceteris paribus} availability of more relevant information leads to a lower level of fixed cost and the marginal effect of additional information should be decreasing.\textsuperscript{14} Among the many functional forms that satisfy these conditions, a power function where the fixed costs of exporting decrease with some constant elasticity in the amount of information has the great advantage of analytical tractability ensuring a closed form solution. In line with Koenig (2005), the following functional form is assumed: \( C^h(i_j) = [b^h_{ij}(a)]^{-\frac{1}{b^h_i}} \). Where \( 1/b^h_{ij} \) is the elasticity of fixed costs with respect to information.

**Equilibrium:** The following proposition characterizes the equilibrium network emerging from the network formation game of the profit maximizing firms:

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

Definitions and a formal proof are provided in the appendix. The intuition is the following. Since information is lost every time it is passed on from firm to firm, proposing direct links to all firms exporting from \( i \) to \( j \) maximizes available information of a firm. Firms that are not exporting from \( i \) to \( j \) are indifferent towards link creation. By assumption (a4) these firms do not propose link creation (any other assumption on their strategy of proposing links would leave the equilibrium amount of information available to an exporting firm unchanged and would thus not affect the results). Since the creation of links is costless, the firm does not face a cost-benefit trade-off for the number of links. It could be, however, that maximizing available information does not maximize profits of firm \( a \), because profits also depend on the cost structure faced by the competitors, which depends on the structure of the network, which in turn depends on the links proposed by the individual firm \( a \). In the complete network, however, there is no room for such strategic interaction. An individual firm is small with respect to the amount of information it contributes, depriving its competitors of this information would thus not affect their cost structure. In addition to that, in the complete network all competitors have direct links to all other firms, it is thus not possible for firm \( a \) to use link creation or destruction in order to influence the amount of information firm \( b \) receives from the other firms.

\textsuperscript{13}The exogenous cost factor could also be country-pair specific, here it is assumed to be symmetric across countries in order to focus on the effect of differences in variable trade costs (distance) between countries.

\textsuperscript{14}Le. \( \partial \Psi^h(i_j^h(a),\pi^h)/\partial i_j^h(a) < 0 \) and \( \partial^2 \Psi^h(i_j^h(a),\pi^h)/\partial (i_j^h(a))^2 > 0 \).
Fixed cost of exporting in a complete exporter network: In a complete network of firms in a given sector exporting from \( i \) to \( j \) all firms exchange information. The information firm \( a \) gets from firm \( b \) is twofold: (a) it gets a fraction \( \delta \) of the original information of firm \( b \) and (b) it also gets \( \delta \) times all the information firm \( b \) has obtained through the network. In the complete network \( a \) is also linked with 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, in equilibrium the additional information firm \( a \) gets from the link with firm \( b \) is just \( \delta \) times its measure. So the overall information firm \( a \) obtains in the network about exporting from \( i \) to \( j \) is given by \( \delta n_{ij}^b \), where \( n_{ij}^b \) is the measure of firms in sector \( h \) exporting from \( i \) to \( j \).

The level of information of firm \( a \) is thus determined by the quality of the information transmission \( \delta \) and the number of firms in a sector exporting from country \( i \) to country \( j \). Since in equilibrium all firms in a sector exporting from \( i \) to \( j \) are part of the complete network, they all have the same level of available information \( \delta n_{ij}^b \) and thus face the same fixed cost \( C_{ij}^h \). Without any loss of generality we can define \( b_n \equiv b_n^h(\sigma_n - 1) \), which will simplify the algebra. The equilibrium fixed cost faced by a firm in sector \( h \) exporting from country \( i \) to \( j \) is then given by:

\[
C_{ij}^h = \left[ \delta n_{ij}^b \right]^{-\frac{\sigma_n - 1}{\sigma}} C^h. \tag{3}
\]

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. As the analysis below is valid for any of the \( H \) differentiated goods sectors, sectoral subscripts will be dropped where this causes no confusion.

Export Market Entry: 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 \( \pi_{ij}(x) = q_{ij}(x) \left( p_{ij}(x) - \frac{r_{ij}}{x} \right) - C_{ij} \). Using equations (1) and (2) this can be rewritten as

\[
\pi_{ij}(x) = \frac{\mu}{\sigma} \left( \frac{\sigma - 1}{\sigma} \right)^{\sigma-1} P_{pj}^{\sigma-1} \left( 1 + \frac{\Pi_{ij}}{L_{ij}} \right) L_{ij} \left( \frac{x}{\tau_{ij}} \right)^{\sigma-1} - C_{ij}. \tag{4}
\]
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. The cutoff productivity level, \( \pi_{ij} \), is the productivity level of a firm being just indifferent between exporting and not exporting, i.e. for which \( \pi_{ij}(\pi_{ij}) = 0 \). All firms with a productivity level below the cutoff level \( \pi_{ij} \) will not find it profitable to start exporting to \( j \). The other firms will. Given the assumptions on the productivity distribution, the measure of firms from country \( i \) that start exporting to country \( j \) is given by \( n_{ij} = \int_{\pi_{ij}}^{\infty} dF_i(x) = L_i \pi_{ij}^{-\theta} \).

Cutoff Productivity Level: Using the condition \( \pi_{ij}(\pi_{ij}) = 0 \) and the function for the fixed cost of exporting (3), the cutoff productivity level is given by

\[
\pi_{ij} = \left( \lambda_1 \delta^{-\frac{1}{b}} L_j^{\frac{1}{b}} \sigma L_i^{-\frac{1}{b}} C^{-\frac{1}{\sigma-1}} \right)^{\frac{1}{\theta}}. \tag{5}
\]

The exact expression for \( \lambda_1 \) (along with the expression for \( \lambda_2 - \lambda_6 \) and \( \lambda_T \)) is provided in the appendix. All these terms only depend on exogenous parameters and on equilibrium aggregate world profits, which remain to be determined but are, as will be shown in the appendix, constant in equilibrium. We can thus think of the \( \lambda \)-terms as constants. At this point the following parameter restriction is imposed: \( b > \gamma > (\sigma - 1) \).

**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

\[
P_j = \lambda_2 L_j^{\frac{1}{1-\gamma}} L^{-\frac{1}{\gamma}} \theta_j, \tag{6}
\]

where

\[
\theta_j = \sum_{k=1}^{N} \delta \frac{\gamma - (\sigma - 1)}{b - \gamma} \left( \frac{L_k}{L} \right)^{\frac{b - (\sigma - 1)}{b - \gamma}} \sigma \frac{1}{\gamma} \frac{1}{\gamma} C \frac{b - (\sigma - 1)}{b - \gamma}. \tag{4}
\]

For notational convenience I have defined \( \beta \equiv \frac{b - (\sigma - 1)}{b - \gamma} \) and \( \zeta \equiv \frac{b - (\sigma - 1)}{b - (\sigma - 1)} \) and \( \eta \equiv \frac{b - (\sigma - 1)}{b - (\sigma - 1)} \). \( \theta_j \) is an analog of the ‘multilateral resistance’ term of the importing country in Anderson and van Wincoop.

\(^{15}\)The first inequality ensures that the effect of the informational networks on the fixed costs of exporting is sufficiently small to assure that firm 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 9).
(2003) and can be interpreted as index of aggregate remoteness.

Given (6) it is now possible to derive equilibrium firm profits and equilibrium world profits. The derivations are presented in the appendix. It turns out that equilibrium world profits only depend on exogenous parameters of the model, which justifies treating the \( \lambda \)-terms as constants.

**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

\[
t_{ij}(x) = \lambda_3 L_j^{(\sigma-1)(\zeta-\gamma)} L_i^{-\frac{\gamma-1}{\zeta-\gamma}} \left( \frac{\tau_{ij}}{\theta_j} \right)^{1-\sigma} x^{\sigma-1}.
\]

(7)

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

\[
\tau_{ij} = \lambda_4 \delta^{-\frac{1}{\sigma-1}} L_j^{-\zeta} L_i^{\frac{1}{\sigma-1}} \left( \frac{\tau_{ij}}{\theta_j} \right)^{\frac{1}{\sigma-1}} \bar{C}^{\frac{1}{\sigma-1}}.
\]

(8)

**Equilibrium aggregate exports:** Aggregate trade flows between \( i \) and \( j \) can be obtained by aggregating all exports of individual firms exporting from \( i \) to \( j \): \( T_{ij} = \int_{\tau_{ij}}^{\infty} t_{ij}(x)L_i \ dF(x) \). 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:

\[
T_{ij} = \lambda \delta^\nu L_j \left( \frac{L_i}{L} \right)^{(1+\nu)} \left( \frac{\tau_{ij}}{\theta_j} \right)^{-(1+\nu) \bar{C}^{-b^*}},
\]

(9)

where I have defined \( \nu = \left[ \frac{\nu}{(\sigma - 1)} \right]^{-1} \). From the imposed parameter restrictions, it follows that \( \nu > 0 \). This expression for aggregate trade flows has a simple gravity structure: apart from constant terms, trade flows are determined by the economic masses \( L_j \) and \( L_i \), variable trade costs \( \tau_{ij} \) and remoteness of the destination market \( \theta_j \). The most important novel feature of this expression is that the elasticities of trade flows with respect to the exporters' country size \( L_i \) and with respect to variable trade costs \( \tau_{ij} \) depend on the parameter \( \nu \) which in turn depends on the degree of firm heterogeneity \( \gamma \), the elasticity of substitution between varieties \( \sigma \) and, most importantly, the quality of the exporter networks (which is manifested in the elasticity of fixed costs with respect to the amount of information, \( 1/b^* \)). The introduction of firm heterogeneity and exporter networks thus changes the distance (variable trade cost) effects predicted by the model compared to existing theories. The different margins of adjustment of aggregate trade flows to changes in variable trade costs and the role exporter networks play for these
margins will be analyzed in detail in the following section.

3 Intensive, basic and magnified extensive margin

Equation (9) pins down the impact of variable trade costs on aggregate trade flows. Most traditional models of international trade explicitly or implicitly assume homogeneous firms. Thus the impact of variable trade costs on individual firm exports will translate one-to-one into the aggregate flows (intensive margin). The crucial feature of models with heterogeneous firms is that variable trade costs also affect the cutoff productivity level and thus the number of exporting firms (extensive margin).

In the model with exporter networks the overall effect of variable trade costs on trade flows is given by \( \frac{\partial \ln T_{ij}}{\partial \ln \tau_{ij}} = -\gamma (1 + \nu). \)\(^{16}\) This effect can be decomposed into the intensive margin, into what I label the basic extensive margin which is standard in models with heterogeneous firms and the magnified extensive margin which is new in this model.

**Intensive margin:** The intensive margin can be isolated by ignoring changes in the number of exporters, i.e. setting the number of exporters constant. Then the adjustments of individual firms to changes in variable trade costs are just aggregated across firms in order to obtain the adjustments in aggregate trade flows, it is sufficient to derive the elasticity of an individual firms' exports with respect to variable costs using equation (7): \( \frac{\partial \ln T_{ij}}{\partial \ln \tau_{ij}} = (\sigma - 1). \) Note that a larger elasticity of substitution implies a stronger effect of variable trade costs on aggregate trade flows which is a standard feature of models of the Krugman (1980)-type. The intuition is that a larger substitutability between goods implies relatively strong demand shifts when marginal costs (and thus prices) change.

**Basic extensive margin:** In models with firm heterogeneity the cutoff level \( \tau_{ij} \) is endogenous and affected by variable trade costs \( \tau_{ij}. \) Thus variable trade costs will also influence the number of exporters, which opens a second channel for variable trade costs to affect aggregate trade flows. The basic extensive margin can be determined by considering a version of the model without exporter networks (i.e. \( 1/b = 0 \)). In this case, the overall effect of variable trade costs on aggregate exports is given by \( \frac{\partial \ln T_{ij}}{\partial \ln \tau_{ij}} = -\sigma (1) - \gamma + (\sigma - 1). \) The first term on the r.h.s. represents the intensive margin discussed above and the second and third term represent the basic extensive margin. Chaney (2006) provides a detailed analysis of the relation of these two margins. Note that the two terms including \( \sigma \) cancel out and the effect of variable trade costs on aggregate trade flows is independent of the elasticity of substitution.

\(^{16}\)Note that here and in following related derivations it is assumed that no country is large compared to the rest of the world. This implies that a change in variable trade costs towards one single trading partner \( i \) does not affect the overall remoteness of the importing country i.e. \( \frac{\partial \ln T_{ij}}{\partial \ln \tau_{ij}} \approx 0. \)
Magnified extensive margin: The introduction of exporter networks increases the effect of variable trade costs $\tau_{ij}$ on the cutoff productivity level $\bar{z}_{ij}$. This can be directly seen from equation (8) noting that the elasticity $\left(\frac{\theta_k}{\delta_{kj}}\right)$ can be rewritten as $(1+\nu)$. This magnified effect of variable trade costs on the cutoff productivity level carry through the aggregation over all individual firms’ exports, which finally implies: $\frac{\partial \ln T_{ij}}{\partial \ln \tau_{ij}} = -(\sigma - 1) - \gamma + (\sigma - 1) - \gamma \nu$. Simplifying gives the exponent of variable trade costs reported in equation (9). The last term on the r.h.s. reflects the effect of the exporter networks. It magnifies the (negative) effect of variable trade costs on aggregate trade flows by making the number of exporting firms more sensitive to variable trade costs. This will be referred to as the magnified extensive margin.

To better understand the different margins, consider two potential destinations $j$ and $k$ such that $\tau_{ij} < \tau_{ik}$. Models with homogeneous firms 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 $\tau_{ij}$ 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 (magnified extensive margin). This effect is the stronger the larger the number of exporters. By adding the magnified extensive margin, the introduction of exporter networks provides an endogenous mechanism magnifying the effect of variable trade costs on aggregate trade flows.

4 Exporter Networks and The ‘Distance Puzzle’

This section shows that, by introducing the magnified extensive margin, the model with heterogeneous firms and exporter networks delivers a possible explanation for both, the apparently high level and the increase over time of estimated distance coefficients.

The high level: First, consider the reasoning along the lines of Grossman (1998) outlined in Anderson and van Wincoop (2004). Based on their theoretical model (which - being a model with homogeneous firms and a CES expenditure system - only features the intensive margin), the theoretical prediction for the distance effect is given by $\frac{\partial \ln T_{ij}}{\partial \ln d_{ij}} = -(\sigma - 1) \frac{\partial \ln \tau_{ij}}{\partial \ln d_{ij}}$.

Defining $\tau'_{ij}$ 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 Wincoop (2004) use an estimate of Hummels (2001) for the elasticity of (tax equivalent) transport costs with respect to distance of about 0.3. Their preferred choice for transport costs 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 back-of-the-envelope calculation delivers a distance effect in the interval of \( \frac{\partial \ln T_{ij}}{\partial \ln d_{ij}} \in [-0.12, -0.46] \). It is common wisdom that estimated distance elasticities are around unity. To provide some illustration for how miserably the above interval matches empirical estimates, I reproduce figure 1 in Disdier and Head (2007) in the appendix. This figure reports 1467 estimated distance effects from 103 studies and shows clearly that most of the estimated distance coefficients are between 0.5 and 1.5.\(^{17}\)

It follows directly from (9) that the model with exporter networks implies an elasticity of trade flows with respect to distance of \( \frac{\partial \ln T_{ij}}{\partial \ln d_{ij}} = -\gamma (1 + \nu) \frac{\partial \ln T_{ij}}{\partial \ln d_{ij}} \).

Based on the findings of Chaney (2006) one can calibrate the degree of firm heterogeneity setting \( \gamma \approx 2(\sigma - 1) \). As regards an estimate for the parameter \( \nu \), section 6 discusses how an indicative estimate can be obtained using cross-sectional trade data. In the estimation I find \( \nu \approx 0.4 \).\(^{18}\) Using these parameter values, the same back-of-the-envelope calculation as above shows that the theoretical model with heterogeneous firms and exporter networks would suggest a distance effect of \( \frac{\partial \ln T_{ij}}{\partial \ln d_{ij}} \in [-0.34, -1.29] \).

It can be seen from figure 1 in Disdier and Head (2007) (reproduced in the appendix) that this interval includes the majority of the values of the distance elasticity commonly estimated.\(^{19}\)

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 rise over time:** As outlined above, the meta analysis of Disdier and Head (2007) provides evidence that empirical studies using data for different decades systematically find increasing distance effects from the 1950s to the 1980s and constant distance effects since then.

\(^{17}\)Obviously the theoretical predictions are computed under the assumption that distance only proxies for transportation costs. A missing variable correlated with distance affecting trade flows would be a possible explanation for the discrepancy between theory and empirical findings. But as outlined in the introduction such a variable has not been found yet.

\(^{18}\)This value implies an elasticity of the fixed costs of exporting with respect to the number of exporters of 0.22. Let an increase of the number of exporters of 10% decreases the fixed cost of exporting by 2.2%. This value is about twice as high as the value implied by the findings of Koenig (2005) for the regional level and does not seem to be outside plausible bounds.

\(^{19}\)Note that the basic effect comes from the introduction of firm heterogeneity. However, exporter networks are important because they magnify 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 \([-0.24, -0.32] \).
While Buch, Kleinert, and Toubal (2004) provide a very informative discussion of why a proportional fall in trade cost does not necessarily lead to falling distance elasticities of trade flows, the question remains why the distance effects are increasing over periods marked by considerable improvements in transport technologies and reductions of trade barriers.

The model suggests a possible explanation which is nicely in line with the notion that, generally speaking, technologies have improved over the last decades. Equation (9) implies that an increase in the quality of exporter networks (an increase in $1/b^*$ i.e., an increase in $\nu$) magnifies the impact of distance on aggregate trade flows. The intuition behind this result is straightforward. When variable trade costs fall, the number of exporters increases (basic extensive margin). Because of the informational networks the increase in the number of exporters reduces the fixed cost of exporting leading more firms to start exporting (magnified extensive margin). With a higher network quality this effect on the fixed costs is stronger, which leads more firms to enter the export market and thereby implies a stronger reaction of aggregate trade flows to a fall in variable trade costs. Thus an increase in the quality of exporter networks over time delivers a very natural explanation for increasing distance effects observed in the data.

It is important to note that an increasing distance elasticity does not mean that countries are trading less with more distant countries than three or four decades ago. In fact, an increase in the quality of the networks would imply that countries trade more with distant countries because better networks reduce the fixed cost of exporting. But the trade promoting effect would be disproportionately stronger for partners that are nearby, because the number of exporters (which is higher for close destinations) has become more important. So ceteris paribus a country with better networks trades more with all its partners but disproportionately more with closer countries. In elasticity terms, this translates into a stronger elasticity of trade flows with respect to distance.

5 Substitutability, Heterogeneity, Borders and Market Size

While the main contribution of the paper concerns the role of exporter networks for the two dimensions of the ‘distance puzzle’, there are other aspects of the model that deserve some attention. These are the role of the elasticity of substitution between goods, the role of firm heterogeneity, the importance of market size of the exporting country, as well as the implications of the model for the ‘border puzzle’. 
5.1 The Role of Market Structure and Firm Heterogeneity

As discussed in section 3, by taking into account the intensive margin only, standard theory with homogeneous firms implies that a high value of $\sigma$ should strengthen the (negative) effect of variable trade costs. As shown by Chaney (2006), the introduction of the basic extensive margin implies that the effect of variable trade costs should be independent of the elasticity of substitution: the degree of firm heterogeneity $\gamma$ alone determines the role of $\tau_{ij}$. In the model with exporter networks instead, the elasticity of aggregate trade flows with respect to variable trade costs is increasing in $\frac{\gamma}{\sigma - 1}$; thus the elasticity of substitution dampens the negative effect of variable trade costs on trade flows.\(^{20}\)

The impact of the term $\frac{\gamma}{\sigma - 1}$ on the distance effects in gravity equations is an interesting testable implication of the model. And indeed, Chaney (2006) himself provides evidence supporting the model's prediction on this effect. Based on his model (and consistent with the model presented above), Chaney obtains an estimate for the sectoral firm heterogeneity relative to the elasticity of substitution $\frac{\gamma}{\sigma - 1}$ from the firm size distributions of US manufacturing sectors. When this estimate is interacted with distance (and some other trade cost proxies) in a standard sectoral gravity equation, the coefficient of this interaction term is negative and highly significant (see table 2, p. 25 in his paper), which implies that a higher $\frac{\gamma}{\sigma - 1}$ implies stronger distance effects. Chaney also provides evidence on the role of the elasticity of substitution, using the elasticities of substitution estimated by Broda and Weinstein (2004) for 3-digit SITC (revision 3) sectors. The interaction of these sectoral elasticities with distance is both statistically and economically significant implying a dampening impact of the elasticity of substitution on the distance effect (table 3, p. 27).

Both, the model with exporter networks and the model proposed by Chaney (2006), have implications for the role of the term $\frac{\gamma}{\sigma - 1}$. The crucial difference is, however, that in Chaney's model the term $\frac{\gamma}{\sigma - 1}$ determines the role of the fixed costs of exporting while the role of variable trade costs is independent of the elasticity of substitution. While it is standard to use distance as a proxy for variable trade costs, the empirical analysis is then carried out under the maintained assumption that the fixed costs of exporting are also increasing in distance.

This assumption is not necessary in the model with exporter networks because the fixed cost of exporting is endogenously affected by variable trade costs. Due to this channel, the distance elasticity of trade flows becomes a function of $\frac{\gamma}{\sigma - 1}$. So even under the standard assumption that distance proxies for variable trade costs only, the model delivers a direct explanation for the empirical relevance of firm heterogeneity and elasticity of substitution for the level of the distance elasticity.

\(^{20}\)Recall that $\nu = \frac{(\gamma)(\sigma - 1)}{(\gamma)(\sigma - 1) - 1}$. 

16
5.2 The ‘Border Puzzle’

Anderson and van Wincoop (2003) show that when ‘multilateral resistance’ terms are taken into account, estimated border effects are reduced from the incredibly high values suggested by McCallum (1995) to about 46% (based on a value of the elasticity of substitution of 8). But the level of this tax equivalent still remains quite high. Chaney (2006) shows that the introduction of firm heterogeneity leads to a considerable reduction in the implied tax equivalents of border barriers. His back-of-the-envelope calculation shows that the tax equivalent of a border is reduced from 46% to 21% in a model with heterogeneous firms which includes the basic extensive margin only.

Exporter networks further strengthen this effect. Using the estimate of the parameter $\nu$ on the national level (which is derived in section 6) of 0.4 it turns out that the tax equivalent of the US-Canada border based on the model with exporter networks is further reduced to 14.4% which brings the estimate closer to what one would expect.\(^{21}\)

5.3 The Role of Country size

To the best of my knowledge all existing theoretical derivations of gravity equations imply a unit elasticity of aggregate trade flows with respect to the economic mass of the exporting and importing country. As can be seen from equation (9), the introduction of exporter networks strengthens the effect of the exporting country’s size on aggregate trade flows. The reason is straightforward: the amount of information in the network is larger the more firms export. Everything else equal, the number of exporters is larger in larger countries. This means that the relative size of the country of origin affects aggregate trade flows not only via the standard gravity type $\left(\frac{L_x}{L_y}\right)$ channel, but also by reducing the level of fixed cost for exporters i.e it has an additional positive effect of $L_y^{\nu}$.\(^{22}\) A casual review of some of the empirical gravity 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 is often ignored in the discussion of empirical results; or - with reference to existing theory - empirical models are constrained by setting the exponents equal to unity to start with.

The model with exporter networks provides a possible explanation for differences in the elasticities with respect to importer and exporter country size.

\(^{21}\)Note that taking the value of $\nu = 0.4$ is purely illustrative. To make a proper statement about the role of exporter networks for the effects of the US-Canada border, the network effects at work there would have to be estimated. Independent of the level of $\nu$, however, the effect of exporter networks will always go into the same direction, namely towards reducing the tax equivalent of the border. This is because the border is modeled as a variable trade cost. Just like the effect of distance, the effect of borders is magnified in the model with exporter networks (which implies that a lower tax equivalent is necessary in order to rationalize the observed effect).

\(^{22}\)Note that the gravity equation refers to the absolute level of trade flows, not to the exports per capita, which are known to be high in small countries.
6 Estimating the effect of networks

In order to get an idea of the magnitude of the parameter \( \nu \) which determines the strength of the effect of exporter networks in the model, one could calibrate its components. As discussed in section 4, the literature provides estimates for the elasticity of substitution between goods and the degree of firm heterogeneity. Estimates for the quality of exporter networks, however, are much more difficult to find. To my knowledge the only study estimating destination-specific network effects is Koenig (2005). Her findings imply an elasticity of the fixed cost of exporting with respect to the number of exporters of \(-11.3\%\) (which implies a \( \nu \) of about 0.15). While her results show nicely that the effect of the number of exporters on the fixed cost of exporting is statistically and economically significant, they refer to the regional level only, while the networks in the model act on the national level.

It is, however, possible to obtain an indicative estimate for the parameter \( \nu \) by estimating the aggregate trade flow equation (9). One can use the fact that \( \nu \) 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 \( \nu \), (9) can be rewritten as \( T_{ij} = \lambda_i \cdot L_{ij} \cdot L_{i}^{1+\nu} \cdot \tau_{ij}^{-\gamma (1+\nu)} \cdot \theta_j^{\gamma (1+\nu)} \), where \( \lambda_i \) collects constants. Rewriting (9) in this way nicely shows the gravity structure of the bilateral trade equation. Just like in standard gravity models, exports from \( i \) to \( j \) depend on a constant, the economic mass (here: the country sizes) and the variable trade costs. In addition, there is the index of aggregate remoteness of the destination country \( \theta_j \).

As standard in the gravity literature, variable trade costs are proxied by a function of the form \( \tau_{ij} = d_{ij}^\rho \) where \( d_{ij} \) is bilateral geographical distance between two countries and \( \rho \) is the elasticity of variable trade costs with respect to distance. The country sizes \( L_i \) and \( L_j \) will both be proxied by the respective country’s GDP. An importer country dummy is added to to account for the remoteness term \( \theta_j \).

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

\[
\ln T_{ij} = \beta_1 \ln d_{ij} + \beta_2 \ln GDP_i + \beta_3 \ln \varphi_m + \varepsilon_{ij}
\]

(10)

where \( \varphi_m \) 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 \( \beta_2 \) has a direct interpretation as an estimate of \( 1+\nu \).
Data: Apart from the introduction of the importer fixed effects, the r.h.s. of equation (10) comprises only the most basic ingredients of gravity equations.\textsuperscript{23} To test the model, data on unidirectional trade flows is needed. A detailed description of the dataset can be found in Feenstra, Markusen, and Rose (2001). 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 dataset 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 (9) represents the aggregate trade flows in the differentiated good sector, only. So, the empirical analysis uses the data on exports in differentiated goods only.\textsuperscript{24}

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

Estimation results: The results of the estimation with OLS of equation (10) are reported in table 1 in the appendix. 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\textsubscript{i} is also highly significant and is around 1.4, which according to the model implies a value of the parameter \( \nu \) of \( \nu \approx 0.4 \).\textsuperscript{25}

A value of \( \nu = 0.4 \) implies an elasticity of the fixed cost of exporting with respect to the number of exporters, \( 1/b^* \) of 0.22. An increase of 10\% in the number of firms exporting from \( i \) to \( j \) would then decrease the fixed cost of exporting from \( i \) 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 operate mainly on the national level, the estimated value of 0.22 seems quite reasonable. The value is economically significant.

\textsuperscript{23}Robustness checks include the addition of other standard gravity variables. This leaves the results unaffected.

\textsuperscript{24}Differentiated goods represent about two thirds of the flows. Robustness checks show that using all three goods classifications together the estimates for \( \nu \) decrease a bit but stay statistically and economically significant. Under the standard assumption that information matters more for differentiated products, the model would imply that networks should affect trade in differentiated products more than trade in reference-priced or homogeneous goods. An analysis using the data on reference priced and homogeneous goods only, confirms this prediction. The estimates of \( \nu \) for differentiated goods are highest, followed by the estimates of reference priced goods. The lowest estimates are obtained for homogeneous goods. In fact, in this case the estimate of \( \nu \) is even negative pointing at a competition effect.

\textsuperscript{25}Using the data on all three goods classifications reduces the absolute value of the distance coefficients by about 0.2 and delivers a very average estimate of \( \nu \) 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.
but at the same time does not appear unrealistically large. And it delivers, as we have seen in section 4, predictions for distance coefficients that are nicely in line with the empirical findings.

7 Conclusions

The main aim of this paper is to address the ‘distance puzzle’ along its two dimensions: the high level and the increase over time of estimated distance coefficients. Unlike most of the literature the puzzle is addressed from the theoretical side, constructing a model of international trade with heterogeneous firms and exporter networks. The inclusion of exporter networks introduces an endogenous feedback effect of variable trade costs on the level of fixed cost and thus on the number of exporters and aggregate trade flows (the magnified extensive margin). The model with exporter networks implies a higher elasticity of trade flows with respect to variable trade costs (distance) than existing theories. When more firms export because of low variable trade costs, more information is available and thus the fixed costs of exporting is lower, which further increases the number of exporters. This effect is the stronger the higher the quality of the network i.e., the faster the fixed cost of exporting decreases in the number of exporters. So not only the level of the distance effect predicted by the model is higher than in the existing literature but also an improvement in the network quality implies increasing distance elasticities. The model thus delivers a possible explanation for the increasing distance coefficients found in many empirical studies and is at the same time nicely in line with the general notion that technologies have improved over the last decades.

The model with heterogeneous firms and exporter networks also provides an explanation for the empirical finding that the elasticity of substitution dampens the effect of distance on trade flows. This is the case even if (as standard in the literature) distance is used as a proxy for variable trade costs only. As regards the so-called ‘border puzzle’, as shown in Chaney (2006) the effect of national borders found by Anderson and van Wincoop (2003) is reduced by the introduction of firm heterogeneity. In the model with exporter networks, the same mechanism that leads to stronger distance effects also implies a magnified reduction of the predicted border effect.
References


A Appendix

A.1 Network Theory and Proof of Proposition 1

In order to formally prove Proposition 1 some additional notation and definitions need to be introduced. The presentation follows Zenou (2006).

A network is a set of agents (firms) \( M = \{1, \ldots, m\} \) and a set \( A \) of links between them. A network is thus a list of unordered pairs of players \( \{a, b\} \) and can be denoted by \( g \equiv (M, A) \). 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 \( g_{ba} = g_{ab} \). 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 (agents) which have to be passed to get from a to b in a given network. The payoffs the different players get from the network \( g \) is summarized by some payoff function \( u(g) = (u_1(g), \ldots, u_n(g)) \) which assigns a payoff to everyone of the \( M \) agents as a function of the network \( g \) connecting them.

Consider the following network formation game \( \Psi(u) \): All agents in \( M \) individually announce the links they wish to form. For all \( a, b \in M \), \( s_{ab} = 1 \) if a wants to form a link with b and zero otherwise. A strategy of agent a is \( s_a = (s_{a1}, \ldots, s_{a\cdot-1}, s_{a\cdot+1}, \ldots, s_{am}) \in S_a \), and \( S_i = \{0, 1\}^{m-1} \) is the set of pure strategies available to a. The link \( ab \) is created iff \( s_{ab} \times s_{ba} = 1 \), i.e. if both players announce that they wish to form a link. Let \( S = S_1 \times \cdots \times S_m \). A pure strategy profile \( s = (s_1, \ldots, s_m) \in S \) induces a non-directed network \( g(s) \) and a vector of payoffs \( u(g(s)) \).

Definition 1: A network is a pairwise-Nash equilibrium of the network formation game \( \Psi(u) \) if and only if there exists a pure strategy profile \( s^* = (s^*_1, \ldots, s^*_m) \) that supports \( g \), that is \( g = g(s^*) \), and \( u_a(g(s^*)) \geq u_a(g(s_a, s^*_{-a})) \) for all \( s_a \in S_a \) and all \( a \in M \), and when for all \( a \in M \) and all \( ab \notin g \), if \( u_a(g) < u_a(g + ab) \) then \( u_b(g) > u_b(g + ab) \).

Here, \( s^*_{-a} \) is the pure strategy profile \( s^* \) without the strategy \( s_a \) and \( ab \) stands for a link between player a and player b and \( (g - ab) \) is the network \( g \) without the link \( ab \). In the following notation s is omitted where it causes no confusion.

In order to prove Proposition 1 it is convenient to first consider the equilibrium of a network formation game \( \Psi^h(i_{ij}, h) \) where firms maximize available information (and not necessarily profits). The equilibrium is characterized by the following Lemma:
Lemma 1. Under assumptions (a1)-(a4) the unique pairwise-Nash equilibrium of the network formation game $\Psi^h(\tilde{t}_{ij,h})$ is a complete network of firms in sector $h$ exporting from $i$ to $j$.

Proof: 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) and they will thus not create links in order to obtain this information.

Let's turn to a firm $a$ maximizing its available information about exporting from $i$ to $j$. In the network formation game $\Psi^h(\tilde{t}_{ij,h})$ the payoff function of firm $a$ in country $i$ is then given by: $u_a(g) = l_{ij}^h(a) = \sum_{d(a)=0} n_{ij,d(a)}^h g_{d(a)}^h$ where $u_a(g)$ is the payoff firm $a$ gets in the network $g$ and $n_{ij,d(a)}^h$ is the measure of firms in sector $h$ exporting from $i$ to $j$ that have a geodesic distance to firm $a$ of $d(a)$. This function adds the information obtained from the direct partners (geodesic distance of unity) of $a$ times their measure ($n_{ij,1}$), the information from indirect links with geodesic distance of two which is $s^2$ times their measure ($n_{ij,2}$) and so on. Since every exporter in sector $h$ has some geodesic distance to firm $a$ (zero for firm $a$ itself, and infinity if there is no connection at all), the overall measure of exporters in the network is given by: $n_{ij}^h = \sum_{d(a)=0} n_{ij,d(a)}^h$.

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 with. It is obvious that under the assumption of costless creation $l_{ij}^h(a)[g(s_a)]$ is maximized if and only if $n_{ij,1}^h = n_{ij}^h$, 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 $l_{ij}^h(a)[g(s_a, s_{-a}^*)] > l_{ij}^h(a)[g(s^*)]$.

In addition, $g(s^*)$ is pairwise-stable since for any firm in sector $h$ (regardless of the export status) for all possible pairs of firms we have (i) for all $ab \in g(s^*)$, $l_{ij}^h(a)[g(s^*)] \geq l_{ij}^h(a)[g(s^*)-ab]$ and $l_{ij}^h(b)[g(s^*)] \geq l_{ij}^h(b)[g(s^*)-ab]$ and (ii) for all $ab \not\in g(s^*)$, if $l_{ij}^h(a)[g(s^*)] < l_{ij}^h(a)[g(s^*)+ab]$ then $l_{ij}^h(b)[g(s^*)] > l_{ij}^h(b)[g(s^*)+ab]$. For pairs of firms in sector $h$ 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 (i) is also satisfied since for any firm $a$ in sector $h$ of country $i$ (regardless of export status) we have $l_{ij}^h(a)[g(s^*)] \geq l_{ij}^h(a)[g(s^*)+ab]$. Finally, the equilibrium is unique 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 firms not connected could always gain by creating an additional link. End of Proof.
Based on Lemma 1 what remains to be shown in order to prove Proposition 1 is that the equilibrium of the network formation game \( \Psi^h(\ell_{ij,k}) \) characterized by Lemma 1 is identical to the unique pairwise-Nash equilibrium of the network formation game \( \Psi^h(\ell_{ij,k}) \) in which firms choose to offer links in order to maximize profits.

**Proof of Proposition 1** It is straight forward to see that none of the firms can influence the aggregate effects of the complete network of exporters on the fixed costs (and thus prices and profits) of the other firms. 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 unchanged.\(^{26}\)

There are two ways how a firm \( a \) exporting from \( i \) to \( j \) might influence the effect of the network on its competitors. Firm \( a \) could refuse to share its information (i.e., not proposing link creation) or choose not to create some links in order to isolate other firms from information flowing in the network. Not creating any link would reduce the information available to the other firms by \( \delta \) times the measure of the firm which is zero. So individual non-participation does not affect the cost structure of the competitors. Isolating competitors from information flowing in the network is (as opposed to other standard networks like the star of circle network) not possible in the complete network, because in the complete network all exporting firms have direct links to the other exporters. **End of proof.**

### A.2 Firm and World Profits

In order to be able to derive the equilibrium world profits \( \Pi \), the expression for individual firm profits \( \pi_{ij} \) 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:

\[
\pi_{ij}(x) = \lambda_i L_j x^{1-\sigma}(\zeta-\eta) L^{-\frac{\sigma \alpha}{\theta}} \left( \frac{\tau_{ij}}{\theta_j} \right)^{1-\sigma} x^{\frac{\sigma \alpha}{\theta}} L_j^{\frac{\sigma \gamma - \alpha - 1}{\theta}} L_j \left( \frac{L_i}{L_j} \right)^{-\frac{\sigma \alpha}{\theta}} \left( \frac{\tau_{ij}}{\theta_j} \right)^{\frac{\alpha (\zeta-\eta)}{\theta_j}},
\]

where again the \( \lambda \)-terms are collecting parameters and \( \Pi \). 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 networks induce in equilibrium.

To close the model an expression for equilibrium world profits is needed. This can be obtained using the definition of \( \Pi \) together with equation (11). Evaluating the integral using (5), an expression emerges which can be simplified significantly using the definition of \( \theta_j \). This finally leads to an expression which

---

\(^{26}\)Which justifies the convenient assumption (a4).
can be solved for $\Pi$:

$$\Pi = \frac{\sum_{h=1}^{H} \left( \frac{a_{h} - 1}{\gamma_{h}} \right) \frac{\mu_{h}}{\sigma_{h}}}{1 - \sum_{h=1}^{H} \left( \frac{a_{h} - 1}{\gamma_{h}} \right) \frac{\mu_{h}}{\sigma_{h}}} L.$$

(12)

It can be directly seen that in equilibrium world profits only depend on parameters of the model and world population. $\Pi$ is thus constant and exogenous.

A.3 Constant Terms:

$$\lambda_1 = \left( \frac{\sigma}{\mu} \right)^{\frac{1}{\sigma - 1}} \frac{\sigma}{\sigma - 1} \left( 1 + \frac{\Pi}{L} \right)^{\frac{1}{\sigma - 1}}$$

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

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

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

$$\lambda_5 = \left( \frac{\gamma}{\gamma - (\sigma - 1)} \right)^{(\sigma - 1)(\zeta - \eta)} \left( \frac{\sigma}{\mu} \right)^{(\sigma - 1)(\zeta - \eta)} \left( 1 + \frac{\Pi}{L} \right)^{(1 - \sigma)(\zeta - \eta)}$$

$$\lambda_6 = \left( \frac{\gamma}{\gamma - (\sigma - 1)} \right)^{-\frac{(\sigma - 1)(\zeta - \eta)}{\zeta}} \left( \frac{\sigma}{\mu} \right)^{-\frac{(\sigma - 1)(\zeta - \eta)}{\zeta}} \left( 1 + \frac{\Pi}{L} \right)^{-\frac{(\sigma - 1)(\zeta - \eta)}{\zeta}}$$

$$\lambda_T = \mu \left( 1 + \frac{\sum_{h=1}^{H} \left( \frac{a_{h} - 1}{\gamma_{h}} \right) \frac{\mu_{h}}{\sigma_{h}}}{1 - \sum_{h=1}^{H} \left( \frac{a_{h} - 1}{\gamma_{h}} \right) \frac{\mu_{h}}{\sigma_{h}}} \right)$$

$$\lambda_0 = \lambda_T L^{(1 + \nu)} \bar{C} \left( \frac{\zeta - 1}{\zeta - 2} \right)^{\frac{\mu}{\sigma}} \delta^{(1 + \nu)}$$

$$\Pi = \frac{\sum_{h=1}^{H} \left( \frac{a_{h} - 1}{\gamma_{h}} \right) \frac{\mu_{h}}{\sigma_{h}}}{1 - \sum_{h=1}^{H} \left( \frac{a_{h} - 1}{\gamma_{h}} \right) \frac{\mu_{h}}{\sigma_{h}}} L$$
A.4 Estimation Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln d_{ij}$</td>
<td>-1.63</td>
<td>-1.47</td>
<td>-1.53</td>
<td>-1.54</td>
<td>-1.53</td>
</tr>
<tr>
<td>$\ln GDP_i$</td>
<td>1.47</td>
<td>1.35</td>
<td>1.45</td>
<td>1.38</td>
<td>1.29</td>
</tr>
<tr>
<td>importer dummy</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>no. of countries</td>
<td>7604</td>
<td>10004</td>
<td>10042</td>
<td>8648</td>
<td>7900</td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

A.5 Estimated distance coefficients reported in Disdier and Head (2007)

![Graph showing density of $R^2$ estimates](image)

Figure 1: This graph is taken from Disdier and Head (2007). It plots the frequency of estimates for distance coefficients from 1467 studies in the literature.