A Note on Trade Costs and Distance

Martina Lawless*

Central Bank and Financial Services Authority of Ireland

Karl Whelan†

University College Dublin

December 2008

Abstract

One of the most famous and robust findings in international economics is that distance has a strong negative effect on trade. Bernard, Jensen, Redding, and Schott (2007) discuss how this can be decomposed into an effect due to the number of products and an effect due to average exports per product. Using US firm-level data, they show that distance has a strong negative effect on the number of products exported. However, they find that the intensive margin—average sales of individual products—is *increasing* with distance. We show that this apparently puzzling finding is consistent with models featuring firm heterogeneity in productivity and fixed costs associated with exporting to each market. We also show how evidence of this type can be used to derive new estimates of how distance affects fixed and variable trade costs and how these two costs combine to generate the distance effect on trade.

*E-mail: martina.lawless@centralbank.ie. We would like to thank Forfás for providing the data used in this paper. The views expressed in this paper are our own, and do not necessarily reflect the views of the Central Bank and Financial Services Authority of Ireland or the ESCB.

†E-mail: karl.whelan@ucd.ie.

1. Introduction

One of the most famous and robust findings in international economics is that distance has a strong negative effect on trade.[1] This pattern suggests that distance must have a substantial impact on the costs associated with trade. There is little evidence, however, on exactly how the distance effect operates and which types of trade costs are impacted.[2] From a theoretical perspective, perhaps the best known model of this so-called gravity relationship is due to Anderson and van Wincoop (2003).[3] Their model views all firms within each country as being identical with trade costs taking the form of tarriff-like "iceberg" costs that are proportional to export volumes. However, more recent contributions, such as Chaney (2008) and Helpman, Melitz and Rubinstein (2008) have emphasized that trade costs may also take the form of fixed costs that must be incurred independent of how much export revenue is generated. An empirical motivation for this approach is the widely documented fact that exporters are more productive than non-exporters, suggesting that only those productive enough get over a barrier due to fixed trade costs will decide to export.

In this paper, we provide new estimates of how distance affects fixed and variable trade costs and how these two costs combine to generate the distance effect on trade. Our starting point is some recent evidence presented by Bernard, Jensen, Redding, and Schott (2007, henceforth BJRS). Using US firm-level data on exports by destination, BJRS decompose the effect of distance on exports into two elements: An extensive margin due to variations in the number of products exported to each market, and an intensive margin due to variations in average sales per product in these markets. A priori, one might expect distance to have a negative effect on both of these margins, and indeed BJRS find that distance has a strong negative effect on the number of firms that sell to an export market as well as the number of products per firm exported. However, somewhat surprisingly, they find that average sales of individual products *increase* with distance. BJRS observe that this finding "is at first sight puzzling" and suggest one potential explanation is that variable trade costs may operate in a different manner than the "iceberg" formulation standard in the trade literature.

Our paper's first contribution is to show that the apparently puzzling finding of distance having a positive effect on average sales per product does not require a new formulation of variable trade costs. In fact, we show that this finding is consistent with the traditional iceberg approach, once it is combined with the assumptions first introduced by Melitz (2003), namely firm heterogeneity in productivity and fixed costs associated with exporting to each market.

[1] Disdier and Head (2006) reported that that the average elasticity of trade with respect to distance from 103 empirical papers was -0.9.

[2] One notable exception is the work of David Hummels (2001, 2007).

[3] According to the IDEAS/REPC site, as of December 2008, this was the sixth most-cited paper released over the previous five years

We illustrate this point by applying a model similar to Chaney (2008) to the product level, implying heterogeneity in productivity and fixed and variable trade costs associated with each product. We discuss the effects of both types of trade costs on the number of products exported to each market as well as the average sales per product in these markets. As would be expected, this model predicts that the number of products exported to a market depends negatively on both fixed and variable trade costs. More surprisingly, however, it also predicts that average sales per product does not depend on variable trade costs at all, and depends positively on fixed trade costs. This is because profitably selling a product in a foreign market requires covering fixed trade costs and this requires a minimum level of productivity and sales. Thus, to the extent that fixed trade costs increase with distance, one should expect to find individual product sales relating positively to distance.

The paper's second contribution is to demonstrate how data on numbers of products exported to each market and average sales per product can be used to estimate the effects of distance on fixed and variable trade costs. We show that the BJRS evidence implies that distance has a stronger effect on fixed trade costs than on variable costs. In addition, we show how these estimates can be used to decompose the elasticity of trade with respect to distance into a component due to fixed trade costs and a component due to variable trade costs. We find that despite fixed costs being more sensitive to distance than variable costs, the effect of distance on trade is largely due to its effect on variable trade costs. This is because reductions in fixed trade costs increase aggregate trade only by introducing new firms to exporting, but these are more marginal low-productivity firms and so have a weaker effect on total exports.

The rest of the paper is organized as follows. Section 2 reviews the US evidence presented by BJRS on how numbers of firms, numbers of products and exports per product vary across export markets according to their distance from the US and their level of GDP. Section 3 presents our model while Section 4 discusses and extends the results. Section 5 uses the model to estimate the effect of distance on fixed and variable trade costs and to decompose the contributions of these costs to the distance effect on trade. Section 6 concludes.

2. Evidence on Distance and Trade

Almost all of the previous research on the so-called gravity relationship in international trade has focused on aggregated data, which sum up bilateral exports over sectors or whole economies. One reason for this limited focus is that, until recently, researchers have not had access to firm-level data reporting both the quantity and the destination of each firm's exports. However, papers such as Eaton, Kortum, and Kramarz (2004) and Bernard, Jensen, Redding and Schott (2007) have shown how such data can generate substantial insights into the processes underlying international

trade.[4]

Eaton, Kortum, and Kramarz (2004) do not explicitly discuss the effect of distance on the pattern of trade, but they report results that indicate the traditional approach to the gravity relationship, based on homogenous firms within each country, is incorrect. Using a cross-sectional sample of French firms from 1986, they show that the so-called extensive margin of trade (variations in the number of firms that serve export markets) appears to be more important than the intensive margin (variations in average export sales per firm).

More recently, BJRS use transactions-based data from the US Census (the Linked-Longitudinal Firm Trade Transaction Database or LFTTD) to provide a detailed picture of US exporting firms. A unique aspect of the LFTTD data is that, in addition to specifying which markets firms sell to, it also specifies how many products they sell (as described by ten-digit product classifications), as well as the total sales of each product. BJRS use this dataset to estimate a standard log-linear gravity equation for US exports in 2000 and then decompose the elasticities with respect to distance and GDP into three components: Extensive components due to the number of firms and number of products that are exported and an intensive component due to the value of export sales per product. We report their estimated elasticities for the extensive and intensive margins in Table 1.

Focusing in particular on the coefficients on distance, it is striking that the negative distance elasticity of -1.36 obtained from this regressions is completely determined by the extensive margins: Distance has approximately equal negative effects on numbers of firms entering a market and on the average number of products they sell, so that these combine to give a negative elasticity of -2.2 for the total number of firm-product combinations being exported. In contrast, the effect of distance on the intensive margin—the sales of individual products sold by a firm—is a *positive* elasticity of 0.84.

BJRS note that this last finding is "at first sight puzzling." They suggest an explanation based upon variable trade costs that differ from the usual "iceberg" formulation. The iceberg approach assumes a certain fraction of goods produced for export "melt away" during the exporting process. Thus, the extra costs of producing a unit for export are proportional to the unit's production cost. BJRS argue that if these costs depended on quantity or weight then only high unit value products would be worth exporting: For instance, if variable trade costs depended on weight then diamonds and computer chips are more likely to be exported to a distant country than tins of baked beans. Baldwin and Harrigan (2007) provide some support for this, showing that distance has a positive effect on the average value of individual product units shipped to US export destinations.[5]

[4]Lawless (2007) also analyzes a data set of this type for Irish firms.

[5]Hummels and Skiba (2004) have also questioned the iceberg assumption using evidence on the effect of freight costs and tarriffs on export prices.

Table 1: Gravity Equation Coefficients for Aggregate US Exports in 2000

	Total Export Value	Number of Exporting Firms	Number of Products Per Firm	Export Sales Per Product Per Firm
GDP	0.98	0.71	0.52	-0.25
Distance	-1.36	-1.14	-1.06	0.84

Source: Figures are based on Table 6 in Bernard, Jensen, Redding and Schott (2007). All coefficients are significant at the 1% level.

Arguments about unit values of products, however, do not explain the positive 0.84 elasticity with respect to distance in Table 1. This figure relates to the effects of distance on the average value of export sales of firm-specific product lines and this is not necessarily related to high unit values. To again give a concrete example, if both diamonds and computer chips are exported to a particular country, then there is little reason to expect that the those who export diamonds will necessarily have higher total sales than those who export computer chips. In the next section, we show that one can in fact explain this finding with a model based upon the standard iceberg formulation of export costs.

3. Modelling Numbers of Products and Sales Per Product

In this section, we show how a model containing the features first introduced by Melitz (2003)—heterogeneity in productivity and both fixed and variable trade costs—can explain the finding of average sales per product increasing with distance. Specifically, we adapt a model with these features presented by Chaney (2008) to derive its predictions for the number of products exported and average sales per product.

We will assume that prices and quantities are set separately for each product that a firm produces. This means that, beyond noting that firms can potentially produce a range of products with varying productivity levels, we will not develop an explicit theory of how many products firms choose to sell. This is currently a very active area of research with important recent contributions from Bernard, Redding, and Schott (2006), Eckel and Neary (2008) and Nocke and Yeaple (2008)

all taking different approaches. However, as we discuss below, our arguments about the effects of distance on extensive and intensive margins are likely to be consistent with a range of different theories of the forces underlying multiproduct firms.

3.1. Assumptions

We assume that there are N countries and that each country can produce a continuum of separate differentiated products. Consumers in country j have derive utility from the products available to them according to a Dixit-Stiglitz aggregator:

$$Y_j = \left[\int_{k \in \Omega_j} q_j(k)^{\frac{\epsilon-1}{\epsilon}} dk \right]^{\frac{\epsilon}{\epsilon-1}} \tag{1}$$

where $q_j(k)$ is the quantity of good k consumed in country j and Ω_j is the set of goods available in country j. From this utility function, the demand for good k in country j is

$$q_j(k) = \frac{p_j(k)^{-\epsilon} Y_j}{P_j^{1-\epsilon}} \tag{2}$$

where $p_j(k)$ is the price charged in country j for good k, Y_j is real income in country j and P_j is the Dixit-Stiglitz price level defined by

$$P_j = \left[\int_{k \in \Omega_j} p_j(k)^{1-\epsilon} dk \right]^{\frac{1}{1-\epsilon}} \tag{3}$$

We will focus on the model's predictions for the exports from a single country (the home market) and, for convenience, we will normalise the number of products that a country can potentially produce to one.

We assume that all costs are incurred at the product level. Specifically, our assumptions about technology are as follows:

- Labour is the only input and must be paid a wage of w. A fixed amount of labour of mass one is supplied in our home market.
- Variable production costs are modelled using a Ricardian technology, so that the total number of workers required to produce q units of output is $\frac{q}{a}$, where a varies across products.
- Following Helpman, Melitz and Yeaple (2004), the product-specific productivity parameter a is assumed to be randomly drawn from a Pareto distribution with probability density function $g(a) = \gamma a^{-\gamma-1}$ on the support $[1, \infty]$.

- There are iceberg transport costs so that τ_j units have to be shipped from the home market to country j for one unit to arrive
- Producing for all markets, including the home market, also requires fixed costs, such that selling in market j requires F_j units of labour independent of how much is actually sold in market j.

The fixed costs, F_j, can be viewed as due to compliance with regulations, marketing, and running wholesale and retail distribution chains. These costs are likely to be lowest (though not zero) in the home market. It is also likely that many of these costs increase with the scale of exports to a particular market. However, at least some part of them will need to be incurred independent of the scale of export sales and we represent this part by F_j.

3.2. Effects of Trade Costs

Our assumptions imply that firms with identical productivity levels make the same decisions, so we will describe firm-level outcomes by indexing them with the productivity level a. The assumptions about production technology imply that the profits generated in market j by a good produced with technology level a (if it is sold there) will be

$$\pi_j(a) = p_j(a)\, q_j(a) - \frac{\tau_j w q_j(a)}{a} - wF_j \tag{4}$$

Factoring in the assumptions about demand, the optimal selling price in country j will be

$$p_j(a) = \frac{\epsilon}{\epsilon - 1}\frac{\tau_j w}{a} \tag{5}$$

Combining this with the demand curve (2), one can show that the profits generated by this product in country j would be

$$\pi_j(a) = \mu \left(\frac{P_j a}{\tau_j w}\right)^{\epsilon-1} Y_j - wF_j \tag{6}$$

where $\mu = (\epsilon - 1)^{\epsilon-1}\, \epsilon^{-\epsilon}$. These profits will be positive as long as

$$a > \left(\frac{F_j}{\mu Y_j}\right)^{\frac{1}{\epsilon-1}} \frac{\tau_j}{P_j} w^{\frac{\epsilon}{\epsilon-1}} = \bar{a}_j \tag{7}$$

This defines a cut-off level of productivity, $\bar{a}_j$, such that only firms with productivity above this level will sell in country j. As would be expected, the cut-off level of productivity is increasing in both types of trade costs and with the home country wage rate, while it is negatively affected by destination country GDP and the price level in country j. Because fixed and variable trading

costs will be lowest in the home market, it is likely that the home market will have the lowest productivity cutoff. However, because a cutoff exists for all markets, not all potential firms will choose to produce.

To interpret the evidence presented by BJRS, we derive the model's predictions for the extensive and intensive margins, that is, the number of products exported and the average exports per product. The extensive margin can be derived as follows using the formula for the cut-off level of productivity

$$N_j = \int_{\bar{a}_j}^{\infty} g(a)da = \bar{a}_j^{-\gamma} = \left(\frac{P_j}{\tau_j}\right)^{\gamma} \left(\frac{\mu Y_j}{F_j}\right)^{\frac{\gamma}{\epsilon-1}} w^{\frac{-\gamma\epsilon}{\epsilon-1}} \tag{8}$$

The number of firms from the home market that export to country j will depend positively on country j's price level and GDP and negatively on both types of trade costs.

To calculate the intensive margin, we start by calculating the total value of export sales to country j. Export sales for a good produced with technology level a are

$$s_j(a) = p_j(a)q_j(a) = \left(\frac{P_j}{p_j(a)}\right)^{\epsilon-1} Y_j \tag{9}$$

Inserting the formula for the optimal price, we get

$$s_j(a) = \left(\frac{\epsilon-1}{\epsilon}\frac{P_j a}{\tau_j w}\right)^{\epsilon-1} Y_j \tag{10}$$

Thus, sales of an individual good in country j depend positively on productivity, on the country's GDP and price level, and negatively on variable trade costs. Once it has been decided that a product will be sold in a particular market, its subsequent sales are independent of fixed trade cost. Total export sales to country j are obtained by integrating across all productivity levels above the threshold:

$$S_j = \left(\frac{\epsilon-1}{\epsilon}\frac{P_j}{\tau_j w}\right)^{\epsilon-1} Y_j \int_{\bar{a}_j}^{\infty} a^{\epsilon-1} g(a)da \tag{11}$$

$$= \frac{\gamma}{\gamma-\epsilon+1}\left(\frac{\epsilon-1}{\epsilon}\frac{P_j}{\tau_j w}\right)^{\epsilon-1} Y_j \bar{a}_j^{\epsilon-\gamma-1} \tag{12}$$

Note from this last calculation that it is necessary to assume $\gamma > \epsilon - 1$. Higher values for γ imply that the distribution of productivity levels falls off faster. As this value approaches $\epsilon - 1$, the cross-sectional distribution of firm sales would become more and more skewed towards large firms and, in the limit, does not converge. Given the well-known empirical regularities relating to

the size distributions for firms, it is reasonable to assume that this particular inequality holds.[6]

The average value of exports per product can now be calculated directly as

$$\frac{S_j}{N_j} = \frac{\gamma}{\gamma - \epsilon + 1} \left(\frac{\epsilon - 1}{\epsilon} \frac{P_j}{\tau_j w} \right)^{\epsilon - 1} Y_j \bar{a}_j^{\epsilon - 1} \tag{13}$$

This can be simplified considerably by inserting the formula for the cutoff value of productivity. In this case, all of the terms involving Y_j, P_j and τ_j cancel out, leaving the strikingly simple formula

$$\frac{S_j}{N_j} = \frac{\gamma \epsilon w}{\gamma - \epsilon + 1} F_j \tag{14}$$

Sales per product in market j will be directly proportional to fixed trade costs and variations in these trade costs will be the only factor determining cross-sectional differences. Thus, to the extent that fixed trade costs increase with distance, one should expect to obtain BJRS's finding of sales per product depending positively on distance.

Finally, combining equation (8) for number of firms and equation (14) for sales per firm produces a simplified formula for total exports from our model economy to market j

$$S_j = \left(\frac{\gamma \epsilon}{\gamma - \epsilon + 1} \right) \left(\frac{P_j}{\tau_j} \right)^{\gamma} (\mu Y_j)^{\frac{\gamma}{\epsilon - 1}} (F_j)^{\frac{\epsilon - 1 - \gamma}{\epsilon - 1}} w^{\frac{\epsilon(1-\gamma) - 1}{\epsilon - 1}} \tag{15}$$

which shows that total export sales across countries are determined by a log-linear specification, as in the standard gravity regression.

3.3. Intuition

These results show how the combination of fixed costs and firm heterogeneity can lead to somewhat counter-intuitive results for the effects of trade costs on the extensive and intensive margins seen in the data, i.e the number of products exported and average sales of these products. Equation (8)'s prediction that the number of products sold to a market is negatively related to both fixed and variable trade costs would be expected. However, equation (14)'s predictions that average export sales per product are independent of τ_j and depend *positively* on the fixed cost F_j are more surprising.

Intuitively, this result can be explained as follows. First consider the effects of variable trade costs. Equation (9) tells us that, for each individual product, an increase in τ_j reduces the exports of all firms that choose to continue sell to market j. However, this increase also eliminates some marginal low-sales products from the market and these two counteracting forces exactly offset

[6] See, for instance, Axtell (2001)

each other. As a result, variable trade costs (as well as foreign country GDP and price level) have no effect on average exports per product. In contrast, fixed trade costs have no effect on sales of individual products (once a firm has decided to supply the product) but an increase in these costs removes some marginal products with low sales from the market. For this reason, average exports per firm depend positively on fixed costs.

4. Discussion and Extensions

In this section, we discuss some technical aspects of the results just derived and present some extensions relating to multiproduct firms and general equilibrium considerations.

4.1. Pareto Distribution

One issue worth clarifying is the role played by the assumption of a Pareto distribution in generating these results. The Pareto assumption leads to clean analytical solution and also has significant empirical evidence in its favor: There is empirical evidence that important firm-level distributions, such as for firm size, follow a Pareto distribution.[7]

Beyond such evidence, however, the general nature of the results above—how trade costs affect the number of products exported and exports per product—will hold across a wide range of distributional assumptions for productivity. Equation (8) shows that the negative effect of both types of trade costs on numbers of products exported works solely through their positive effect on $\bar{a}_j$ and does not rely on distributional assumptions. Similarly, because an increase in fixed costs has no effects on sales of continuing products, the positive relationship between $\frac{S_j}{N_j}$ and F_j does not depend on distributional assumptions. Finally, because an increase in τ_j reduces sales of continuing products but also eliminates marginal products with lower sales, this effect is in general ambiguous: Only in the Pareto case do these effects exactly cancel.[8] Thus, the general idea that the negative effect of trade costs on exports should work mainly through the extensive margin, with the effect on the intensive margin potentially being positive, does not rely on special distributional assumptions.

[7] See Axtell (2001) for evidence on size distributions of US firms. In addition, Gabaix (1999) has shown that Pareto distributions can be generated from an aggregation of random micro-level exponential growth shocks to each of the individual units, while Kortum (1997) has shown that the upper tail of productivity distributions needs to be Pareto if steady-state growth paths are to be sustained.

[8] See Lawless (2008) for a demonstration of these results for the general case.

4.2. Definitions of Intensive and Extensive Margins

Our definitions of extensive and intensive margins of trade as relating to numbers of firms and export sales per firm are simple and intuitive definitions. We should probably note, however, that the related literature has focused on two different definitions of these margins.

Firstly, our definitions of these margins differs from those discussed by Chaney (2008). Equations such as (8) and (14) for number of firms and sales per firm are not discussed in his paper. Instead, he uses the Leibniz Integral Rule to divide changes in total exports due to shifts in an exogenous parameter x as follows

$$\frac{\partial S_j}{\partial x} = \int_{\bar{a}_j}^{\infty} \frac{\partial s_j(a)}{\partial x} g(a)da - s_j(\bar{a}_j)G(\bar{a}_j)\frac{\partial \bar{a}_j}{\partial x} \tag{16}$$

For instance, consider a change that leads to firms exiting market j. The first term in this decomposition (Chaney's intensive margin) describes the change in exports keeping the group of exporting firms unchanged: The exports of firms that have exited are measured based on the optimal (loss-making) level of sales that they would have obtained had they stayed in the market. The second term (his extensive margin) measures the loss in sales due to these firms having exited the market. This provides a useful theoretical decomposition of the effects of changes in exogenous parameters. It has the disadvantage, however, of being based on a thought experiment: How much would firms that have exited sell if they were still exporting, or how much would firms that have just arrived have sold if they were exporting last period? This means that, in practice, these two components cannot be observed over time in datasets such as the LFTTD.[9] In addition, unlike the decomposition of exports into number of products and exports per product, this decomposition has no analogue in cross-sectional data such as those that generated the BJRS results.

Secondly, our definition of these margins differs from that used in papers based on aggregate bilateral trade data such Baldwin and Harrigan (2007) and Helpman, Melitz and Rubinstein (2008). These papers focus on the fact that some countries do not trade with each other at all and then analyse the extensive margin as relating to whether countries trade and the intensive margin as relating to total bilateral export sales if they do trade. In focusing on the one-zero nature of whether we observe any trade data between countries, this approach misses the more general importance of an extensive margin such that numbers of firms participating in export markets tend to decline relatively rapidly as one moves away from the most popular markets. This has been documented both by BJRS for the US and by Eaton, Kortum and Kramarz (2004) for France.[10]

[9] An additional complication is that one is likely to see simultaneous entry and exit to each individual export market. Lawless (2007) documents this is a significant pattern using a sample of Irish exporting firms.

[10] A third definition looks at aggregate data on the number of different product types (as defined by detailed product codes) that a country exports and calls this the extensive margin, with average sales per product type defined as the

4.3. Multiproduct Firms

Our results can explain why the total number of firm-product combinations sold in a market will decline with distance while average sales for each individual product line may increase. Thus far, however, we have been silent on the results described in the third column of Table 1, which show that the average number of products sold by individual firms in a market tends to increase with that market's GDP and decline with its distance from the home market. This prediction is likely to fit with any one of a number of models of multiproduct firms.

For example, our approach is consistent with the following simple view of multiproduct firms. Each firm has the ability to produce M differentiated products and its efficiency in making these products is determined by M uncorrelated random draws $(a_1, a_2, ..., a_M)$ from the aggregate productivity distribution, $g(a)$. Whether a firm chooses to make a particular product at all will depend upon whether its productivity draw for that product exceeds one of the thresholds $(\bar{a}_1, \bar{a}_2, ..., \bar{a}_N)$ which exist for all markets, including the home market. Similarly, whether a firm chooses to export a product to market j will depend upon whether its productivity draw for that product exceeds $\bar{a}_j$. Firms will thus choose to sell more products in markets with low productivity thresholds. For this reason, when looking at an individual firm, we would expect that firms would sell more products in markets with high GDP and which are geographically close (implying lower trade costs) and this is exactly what is found.

This is, of course, a highly simplistic theory of a multiproduct firm but the same principle will apply to more sophisticated models. Models such those of Bernard, Redding and Schott (2006) and Eckel and Neary (2008) assume that firms have range of different potential productivities across the products they can make. Bernard, Redding and Schott assume a multiplicative approach in which a firm's efficiency in making an individual product is determined by a firm-specific value times a random value drawn for that specific product. Eckel and Neary instead assume that each firm attains its maximum productivity for a particular "core" product and that its efficiency gradually declines as one moves away from this product according to a defined space for differentiated products. Either of these assumptions would still be consistent with the analysis just outlined in which the average number of products exported per firm increases with GDP and declines with distance.

4.4. Labour Market Equilibrium

The exact value taken by the wage rate w does not matter for our discussion of variations across export destinations in numbers of products sold and sales per product. The wage features in both equation (8) for the number of products exported to each market and in equation (14) for sales per

intensive margine. This approach is taken by Hummels and Klenow (2005).

product. However, we have used these equations to explain coefficients on distance and GDP from cross-sectional regressions, and because we are looking at exports from a single country (the US) there is only one wage rate being considered and its empirical counterpart is the constant term.

That said, it is possible to solve for the implications of labour market equilibrium for the wage rate and we report these calculations in an appendix. Specifically, we show that equating the total amount of labour required to produce the output of this economy (adding up all the S_j terms) and equating this sum with the assumed fixed unit supply of labour produces the following condition for the equilibrium wage rate

$$w = \left(\sum_{j=1}^{N} \left(\frac{\gamma\epsilon - \epsilon + 1}{\gamma - \epsilon + 1} \right) \left(\frac{P_j}{\tau_j} \right)^{\gamma} (\mu Y_j)^{\frac{\gamma}{\epsilon-1}} F_j^{\frac{\epsilon-1-\gamma}{\epsilon-1}} \right)^{\frac{\epsilon-1}{\gamma\epsilon}} \tag{17}$$

This formula shows how wages depend negatively on the home country's fixed and variable trade barriers relative to all countries and positively on the GDP and price levels of all countries. Because threshold productivity barriers for all foreign markets depend positively on this wage rate, this formula provides an illustration of the re-allocation effects due to trade liberalization first discussed by Melitz (2003). A generalized reduction in trade barriers will raise wages and will raise average sales levels for those products that continue to be exported. However, those using less efficient technologies may not be competitive enough to export any more and may have to cease producing altogether.

One can take these calculations one step further and calculate endogenous price-levels, P_j, for each country under the assumption that this model of production applies to all countries. Calculations along these lines have been reported by Chaney (2008).[11] However, because P_j plays no role in our expression for (14) for sales per product, these calculations would be of limited relevance to our paper.

5. Distance, Trade Costs, and Gravity

Equations (8) and (14) describe how fixed and variable trade costs can affect exports when there is heterogeneity in productivity. These relationships can also be linked directly to the gravity regressions presented by BJRS by making assumptions about the form of the relationships between these trade costs and both distance and destination country GDP. Here, we provide a simple log-linear model of these relationships and derive estimates of its parameters. We then show how these estimates can be used to decompose the elasticity with respect to distance in an estimated

[11] Note, however that Chaney considers the wage to be exogenously determined in each country.

gravity relationship into a component due to fixed trade costs and a component due to variable trade costs.

5.1. How Does Distance Affect Trade Costs?

Because the gravity regression takes a log-linear form, it is natural to also assume this specification for the relationship between trade costs and distance and GDP. Omitting other factors, this leads to the following simple formulation:

$$F_j = d_j^{\theta_1} Y_j^{\theta_2} \tag{18}$$

$$\tau_j = d_j^{\theta_3} Y_j^{\theta_4} \tag{19}$$

Substituting these into equations (8) and (14), we now obtain expressions for the intensive and extensive margins directly in terms of the effects of distance and destination-country GDP, as in the BJRS regressions (reported here in Table 1):

$$N_j = w^{\frac{-\gamma\epsilon}{\epsilon-1}} P_j^{\gamma} d_j^{-\gamma\left(\theta_3+\frac{\theta_1}{\epsilon-1}\right)} Y_j^{\left(\frac{\gamma}{\epsilon-1}\right)(1-\theta_2-\theta_4(\epsilon-1))} \tag{20}$$

$$\frac{S_j}{N_j} = \frac{\gamma\epsilon w}{\gamma-\epsilon+1} d_j^{\theta_1} Y_j^{\theta_2} \tag{21}$$

The elasticities with respect to distance and GDP in these equations can now be related directly to the estimated values obtained by BJRS, so that

$$\begin{aligned} -\gamma\left(\theta_3+\tfrac{\theta_1}{\epsilon-1}\right) &= -2.2 & \left(\tfrac{\gamma}{\epsilon-1}\right)(1-\theta_2-\theta_4(\epsilon-1)) &= 1.23 \\ \theta_1 &= 0.84 & \theta_2 &= -0.25 \end{aligned} \tag{22}$$

Note that here we are using the combined extensive margin coefficients (for both numbers of firms and product per firms) as the relevant elasticities to match up with the N_j equation. The estimated elasticities for sales per product translate directly into the elasticities for fixed trade costs. Thus, the elasticity of fixed trade costs with respect to distance is $\theta_1 = 0.84$, while the elasticity with respect to GDP is $\theta_2 = -0.25$. This latter estimate is somewhat surprising because one might have expected larger markets to have higher set-up costs. One possible explanation is that richer countries tend to have better infrastructure and lower regulatory burdens and these advantages may be more important for fixed trade costs.

The parameters of the variable trade equation cannot be directly identified because the remaining two equations have four unknown parameters: θ_3, θ_4, ϵ and γ. Our approach is to use values for ϵ and γ that have been derived elsewhere on the basis of firm-level data. Specifically, we follow Bernard, Redding, and Schott (2007) and use $\epsilon = 3.8$ and $\gamma = 3.4$. This results in

estimates of $\theta_3 = 0.35$ and $\theta_4 = 0.08$.[12] This latter estimate implies that destination country GDP has very little effect on variable trade costs.[13]

The pattern of the estimated effects of distance are perhaps a bit surprising. The most obvious trade costs that are directly related to distance are transport costs. And indeed, our estimate of an elasticity of 0.35 for variable trade costs with respect to distance is very similar to the estimates of the effect of distance on various types of transport costs (rail, shipping, air) presented by Hummels (2001). However, the far larger effect of distance on fixed costs trade costs suggests a more complicated set of barriers that appear to increase with distance.

5.2. Decomposing the Distance Effect on Trade

Equations (20) and (21) can be combined to provide a full expression for total export sales from our model economy to country j as a function of distance and GDP, illustrating the separate roles these factors play through their effects on fixed and variable costs:

$$S_j = \left(\frac{\gamma\epsilon}{\gamma - \epsilon + 1}\right) w^{\frac{\epsilon(1-\gamma)-1}{\epsilon-1}} P_j^{\gamma} \left(\mu Y_j\right)^{\frac{\gamma}{\epsilon-1}} \left(d_j^{\theta_3} Y_j^{\theta_4}\right)^{-\gamma} \left(d_j^{\theta_1} Y_j^{\theta_2}\right)^{\frac{\epsilon-1-\gamma}{\epsilon-1}} \tag{23}$$

Combining this equation with our estimates of the trade costs functions and Bernard, Redding, and Schott's values for ϵ and γ, we can decompose the effect of distance on exports (reported in Table 1 to be -1.36) into the effect due to its impact on fixed trade costs and the effect due to its impact on variable trade costs. The total distance elasticity of -1.36 turns out to be composed of an effect of $-0.17 \left(= -\frac{\epsilon-1-\gamma}{\epsilon-1}\theta_1\right)$ due to fixed costs, while the remaining $-1.19\,(= -\gamma\theta_3)$ is due to variable costs. Thus, the impact of distance on total exports works primarily through the channel of variable costs, even though these are less sensitive to distance than are fixed costs. This result is obtained because fixed costs only influence exports by adding or removing marginal products with lower sales, while variable trade costs influence both the entry decision and subsequent sales for all firms.

This last calculation sheds some interesting light on the effect of distance on exports. In this model, it is the presence of fixed costs that generates the extensive margin: Without these costs, the model would predict that all firms would export to all markets. And the BJRS evidence shows that the extensive margin is what determines the negative effect of distance on exports (see the

[12] Experimentation with other values of ϵ and γ with the ranges reported in the literature, while maintaining the necessary assumption of $\gamma > \epsilon - 1$, gave relatively similar values.

[13] We should note that this calculation could have been done somewhat differently. Chaney (2008) applies this model to determine the price level in all countries so that P_j depends on Y_j. Once this adjustment is made, the overall elasticity of N_j with respect to Y_j becomes one, as in the traditional gravity equation. Applying our calculation to this version of the equation, the elasticity of variable trade costs with respect to destination country GDP becomes $\theta_4 = 0.02$, further confirming our conclusion of a weak relationship.

elasticity of -2.2 in Table 1). Thus, one might expect that the effect of distance on exports largely reflects its effect on fixed trade costs. However, it turns out that the distance elasticity is mainly due to the effect of distance on variable trade costs.

One way to understand this result is through the formula for the cut-off value of productivity for exporting, equation (7). While it is the existence of fixed costs that implies the existence of a productivity cutoff, the subsequent importance of this margin depends more on variable trade costs than on fixed costs. Equation (7) shows that the elasticity of the cutoff with respect to variable costs is one, while the elasticity with respect to fixed costs is $\frac{1}{\epsilon-1}$, which our estimates suggest is about one-third.

6. Conclusions

That distance inhibits trade is one of the most robust findings in empirical international economics. The recent finding from detailed data on US firms of Bernard, Jensen, Redding and Schott (2007) that export sales per product tend to increase with distance is therefore quite surprising. This finding also runs counter to the predictions of popular models of the distance-trade relationship such the model of Anderson and van Wincoop (2003) which features homogenous firms within each country and tarriff-like trade costs: Such models would predict that sales of all products should decline with distance. In this paper, we have shown that this apparently counterintuitive finding is consistent with models of international trade that assume firm heterogeneity in productivity and fixed costs, such as the models of Melitz (2003) and Chaney (2008).

Heterogeneity in productivity helps to explain the finding because those firms engaged in exporting to distant locations are predicted to be more productive than those that only export close to home. However, the presence of fixed trade costs is also crucial. For instance, other models featuring heterogeneous productivity, such as Eaton and Kortum (2002) can generate intensive and extensive margins in trade without fixed costs by invoking different assumptions about preferences—that all economies can produce the same set of goods and trade only takes place when one country can be the cheapest supplier of a product to another country. However, the Eaton-Kortum model does not predict that sales per product should increase with distance. Thus, the BJRS results appear to favor models that incorporate fixed trade costs that increase with distance, so that firms that export to distant markets need to sell enough to cover these costs.

We have also shown how evidence on numbers of products exported and average sales per product can be combined with the Melitz-Chaney model to estimate the effects of distance on fixed and variable trade costs. We find that fixed trade costs appear to rise more with distance than variable costs. However, while fixed costs are necessary for this model to generate the extensive

margin of trade and it is through this margin that the distance affects trade negatively, we estimate that the elasticity of trade with respect to distance is largely due to its effect on variable costs.

References

[1] Anderson, James and Eric van Wincoop (2003). "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170-92.

[2] Anderson, James and Eric van Wincoop (2004). "Trade Costs," *Journal of Economic Literature*, 42, 691-751.

[3] Axtell, Robert (2001). "Zipf Distribution of US Firm Sizes," *Science*, 7 September 2001, pp.1818-1820.

[4] Baldwin, Richard and James Harrigan (2007). Zeros, Quality and Space: Trade Theory and Trade Evidence, NBER Working Paper 13214.

[5] Bernard, Andrew, Stephen Redding, and Peter Schott (2006). "Multiproduct Firms and Trade Liberalization," mimeo, Dartmouth.

[6] Bernard, Andrew, J. Bradford Jensen, Stephen Redding, and Peter Schott (2007). "Firms in International Trade," forthcoming, *Journal of Economic Perspectives*.

[7] Chaney, Thomas (2008). "Distorted Gravity: Heterogeneous Firms, Market Structure and the Geography of International Trade," *American Economic Review*, 98, 1707-1721.

[8] Eaton, Jonathan and Samuel Kortum (2002). "Technology, Geography and Trade" *Econometrica* 70(5) pp. 1741-1779.

[9] Eaton, Jonathan, Samuel Kortum and Francis Kramarz (2004). "Dissecting Trade: Firms, Industries and Export Destinations," *American Economic Review* Vol. 94, No.2, pp. 150-152.

[10] Eckel, Carsten and J. Peter Neary (2008). "Multi-Product Firms and Flexible Manufacturing in the Global Economy," mimeo, Oxford University.

[11] Gabaix, Xavier (1999). "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, Vol.114, No.3, pp.739-67.

[12] Helpman, Elhanan, Marc Melitz, and Stephen Yeaple (2004). "Export versus FDI with Heterogenous Firms," *American Economic Review*, Vol. 94, 300-316.

[13] Helpman, Elhanan, Marc Melitz and Yona Rubinstein (2008). "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123, 441-487.

[14] Hummels, David (2001). "Towards a Geography of Trade Costs," mimeo, Purdue.

[15] Hummels, David (2007). "Transportation Costs and International Trade Over Time," forthcoming, *Journal of Economic Perspectives*.

[16] Hummels, David and Alexandre Skiba (2004). "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture," *Journal of Political Economy*, 112, 1384-1402.

[17] Hummels, David and and Peter Klenow (2005). "The Variety and Quality of a Nation's Exports," *American Economic Review*, 95, 704-723.

[18] Kortum, Samuel (1997). "Research, Patenting and Technological Change," *Econometrica* Vol.65, No.6, pp. 1389-1419.

[19] Lawless, Martina (2007). Firm Export Dynamics and the Geography of Trade, MPRA Paper 10008, University Library of Munich, Germany.

[20] Lawless, Martina (2008). Deconstructing Gravity: Trade Costs and Extensive and Intensive Margins," MPRA Paper 10230, University Library of Munich, Germany.

[21] Melitz, Marc (2003). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica* Vol. 71, No.6, pp. 1695-1725.

[22] Nocke, Volker and Stephen Yeaple (2008). Globalization and the Size Distribution of Multiproduct Firms, mimeo, Oxford University.

A Labour Market Equilibrium

This appendix demonstrates how labour market equilibrium implies the wage rate described in equation (17) in the text.

The sales generated in market j by a firm with productivity a require the following number of units of labour:

$$L_j(a) = \frac{\tau_j}{a} q_j(a) + F_j \tag{24}$$

This firm's sales in market j are

$$s_j(a) = p_j(a) q_j(a) = \frac{\epsilon}{\epsilon - 1} \frac{\tau_j w}{a} q_j(a) \tag{25}$$

Combining these two expressions, one can re-write the number of units of labour used as

$$L_j(a) = \left(\frac{\epsilon - 1}{\epsilon}\right)\frac{s_j(a)}{w} + F_j \tag{26}$$

Recalling that labour supply is normalized to one, labour market equilbrium can be described by adding up this labour requirement over all firms equating the sum to one:

$$\sum_{j=1}^{N}\left[\left(\frac{\epsilon - 1}{\epsilon}\right)\frac{S_j}{w} + F_j N_j\right] = 1 \tag{27}$$

But we know from equation (14) that

$$F_j N_j = \frac{\gamma - \epsilon + 1}{\gamma\epsilon}\frac{S_j}{w} \tag{28}$$

So the formula describing labour market equilibrium can be re-written as

$$1 = \sum_{j=1}^{N}\left[\left(\frac{\epsilon - 1}{\epsilon} + \frac{\gamma - \epsilon + 1}{\gamma\epsilon}\right)\frac{S_j}{w}\right] \tag{29}$$

$$1 = \sum_{j=1}^{N}\left(\frac{\gamma\epsilon - \epsilon + 1}{\gamma\epsilon}\right)\frac{S_j}{w} \tag{30}$$

$$1 = \sum_{j=1}^{N}\left(\frac{\gamma\epsilon - \epsilon + 1}{\gamma\epsilon}\right)\left(\frac{\gamma\epsilon}{\gamma - \epsilon + 1}\right)\left(\frac{P_j}{\tau_j}\right)^{\gamma}(\mu Y_j)^{\frac{\gamma}{\epsilon-1}}(F_j)^{\frac{\epsilon-1-\gamma}{\epsilon-1}} w^{\frac{-\gamma\epsilon}{\epsilon-1}} \tag{31}$$

$$\tag{32}$$

From this, we obtain that labour market equilibrium requires a wage of

$$w = \left(\sum_{j=1}^{N}\left(\frac{\gamma\epsilon - \epsilon + 1}{\gamma - \epsilon + 1}\right)\left(\frac{P_j}{\tau_j}\right)^{\gamma}(\mu Y_j)^{\frac{\gamma}{\epsilon-1}} F_j^{\frac{\epsilon-1-\gamma}{\epsilon-1}}\right)^{\frac{\epsilon-1}{\gamma\epsilon}} \tag{33}$$

which is the required result.