Structural gravity equations with intensive and extensive margins

Matthieu Crozet *University of Reims*
Pamina Koenig *University of Paris West*

Abstract. Recent trade models with heterogeneous firms have changed the interpretation of gravity equations. Chaney (2008) shows that the effect of distance on the number of exporters and average exports depends on key parameters characterizing the elements of market structure. We use firm-level export data to estimate the structural parameters of Chaney's model. Controlling for the fixed costs of exporting, our estimated parameters match, for 28 out of 34 industries, the model's theoretical predictions. Our industry parameters allow us to evaluate separately the effect of transport costs and tariffs on trade, without having to resort to detailed data on trade frictions. JEL classification: F12

Equations structurelles de gravité avec marges intensives et extensives. Les modèles de commerce international avec firmes hétérogènes ont modifié l'interprétation des équations de gravité (Chaney 2008). L'effet de la distance sur le nombre d'exportateurs et la valeur moyenne exportée dépend d'un ensemble de paramètres caractérisant l'organisation du marché. Cet article mobilise des données individuelles sur les exportateurs français pour conduire une estimation structurelle de ces paramètres. Il confirme la pertinence du cadre théorique; pour 28 secteurs sur 34, les paramètres estimés sont conformes aux attentes théoriques. Il montre aussi que les coûts de transport ont un impact très différent de celui des droits de douanes, et que l'impact de l'ouverture commerciale dépend fortement de l'organisation des marchés.

Crozet is also affiliated with CEPII; Koenig is also affiliated with Paris School of Economics.
We are grateful to Thomas Chaney, Andrew Clark, Vincent Rebeyrol, Sandra Poncet, and three anonymous referees for helpful advice. We also thank participants at ETSG 2008, and seminars at CEPII and the University of Grenada. The second author thanks the Centre de Recherche en Economie et en STatistique (CREST). We also acknowledge financial support from the ACI - *Dynamiques de concentration des activités économiques dans l'espace mondial*. Email: matthieu.crozet@cepii.fr; pkoenig@u-paris10.fr.

Canadian Journal of Economics / Revue canadienne d'Economique, Vol. 43, No. 1
February / février 2010. Printed in Canada / Imprimé au Canada

0008-4085 / 10 / 41–62 /

1. Introduction

The question of how trade barriers affect trade is central to understanding the structure of world trade and the production of economically important information such as the trade impact of regional trade agreements, non-tariff barriers and other trade frictions. Since its first application to international trade by Tinbergen (1962), the gravity equation has been a leading tool for the estimation of bilateral trade flows. The basic form relates trade to the economic size of the trading countries and the geographic distance between them. Over time the specification has been improved both theoretically and empirically.

Current work on trade barriers using gravity equations faces two key challenges. First, while it is easy to measure the distance elasticity of trade, it is much harder to evaluate the trade-cost elasticity of trade, owing to the lack of direct measures and detailed data on trade frictions. Second, the number of parameters relating trade flows to trade barriers, and their interpretation, has changed in recent trade models with heterogeneous firms (Melitz 2003; Chaney 2008). In the latter, only a subset of heterogeneous firms will export at a given level of trade costs. As trade costs fall, two mechanisms are at work: incumbent exporters increase their volume of sales (the intensive margin), and new firms enter the export market (the extensive margin). The key parameters governing the intensive margin are the same as those in trade models with homogeneous firms: the elasticity of substitution between goods and the elasticity of trade costs with respect to distance. However, the extensive margin depends not only on these two parameters, but also on the distribution of productivity across firms. At the aggregate level, the effect of trade costs on trade is the sum of the trade-cost effect at the intensive and extensive margins. Chaney (2008) and Melitz and Ottaviano (2008) show that the elasticity of aggregate trade flows with respect to distance therefore can be interpreted not as the elasticity of substitution between goods, but rather as the degree of firm heterogeneity.

In this paper, we use French firm-level export information to estimate the structural parameters of Chaney's trade model. This produces estimates of the three structural parameters behind trade flows, avoiding the trade-barrier measurement problem. We consider three equations: a gravity equation at the firm level (providing a combination of the demand elasticity and the distance elasticity); an export-selection equation (yielding a combination of the three unknown parameters); and a rank-size distribution of productivity across firms (providing a combination of the demand elasticity and heterogeneity parameters). The estimated parameters confirm the importance of the intensive and extensive margins of international trade and are consistent, for 28 out of 34 industries, with the theoretical trade model with heterogeneous firms of Chaney (2008) and Melitz and Ottaviano (2008). We use the estimated parameters to compute the contributions of transport costs and tariffs to trade and to disentangle the roles of the intensive and extensive margins.

The paper is structured as follows. Section 2 sketches the theoretical model, and section 3 explains the empirical strategy. Section 4 presents the data and the regression estimates regarding the intensive and extensive margins. We highlight the importance of distance in explaining trade, and the predominant role of the extensive margin. Section 5 presents the structural estimates of the three parameters: the elasticity of substitution, the distance elasticity of trade costs, and the degree of firm heterogeneity; this section also illustrates the effects of distance and trade barriers on trade. Section 6 concludes.

2. The model

We briefly present a simple model of international trade with heterogeneous firms, bringing out the main features of Chaney (2008). We highlight the expressions for the trade-cost elasticity of the extensive and the intensive margins.

2.1. Production and consumption

We consider a Home country facing R foreign markets. The Home country produces H differentiated goods and a homogeneous numéraire good. In the H manufacturing industries, firms engage in Dixit-Stiglitz monopolistic competition. All consumers have the same CES utility function:

$$U_j = q_{0j}^{\mu_0} \Pi_{k=1\ldots H} \left(\int_{i \in \Omega_{kj}} q_{kij}^{\sigma_k - 1} \right)^{\mu_k \frac{\sigma_k}{\sigma_k - 1}}, \tag{1}$$

where q_{kij} is the quantity of the variety i of good k demanded by a representative consumer in country j; Ω_{kj} is the set of all varieties of good k available in country j; σ_k is the elasticity of substitution between the varieties of good k; q_{0j} is the consumption of the numéraire good; and μ_0 and μ_k are positive parameters such that $(\mu_0 + \sum_k \mu_k = 1)$. As the empirical analysis considers each industry separately, we drop the subscript k for notational convenience in this section.

There are N firms in the Home country. To produce and sell on a foreign market, each firm incurs a firm-specific marginal cost and a destination-country-specific fixed cost. As our data cover French firm exports, all firms are located in the same exporting country: in the following the subscript i denotes a French firm and the subscript j a foreign country. The export fixed cost is considered to be identical for all firms exporting to the same destination country. For a firm i with marginal cost a_i, the total cost of supplying consumers in country j with $q(a_i)$ units of good is $TC_{ij}(a_i) = q(a_i)a_i + C_j$. We assume 'iceberg' transportation costs: $\tau_j > 1$ units of good have to be shipped in order to ensure that one unit arrives in country j.

As is usual in the Dixit-Stiglitz monopolistic competition framework, the profit-maximizing price is a constant mark-up over marginal cost. Hence, the delivered price on market j of a good produced by a firm with marginal cost a_i is

$$p_{ij}(a_i) = \frac{\sigma}{\sigma - 1} a_i \tau_j. \tag{2}$$

Let E_j denote the total expenditure in country j on the relevant industry, and P_j the price index in country j. We can then show from (1) and (2) that the demand emanating from country j for a given variety i is

$$m_{ij}(a_i) = p_{ij}(a_i) q_{ij}(a_i) = \left(\frac{p_{ij}(a_i)}{P_j} \right)^{1-\sigma} E_j. \tag{3}$$

2.2. Trade costs and the intensive and extensive margins of trade

The marginal cost a is assumed to have a Pareto distribution, bounded between 0 and 1, with a scaling parameter $\gamma \geq 1$.[1] Hence, marginal cost is distributed as $P(\tilde{a} < a) = F(a) = a^{\gamma}$ and $dF(a) = f(a) = \gamma a^{\gamma - 1}$. The parameter γ is an inverted measure of firm heterogeneity. A value of γ close to one implies an almost uniform distribution of marginal costs between 0 and 1; as γ goes to infinity, the distribution becomes more concentrated.

For marginal cost a_i, the profit from market j is $\pi_{ij}(a_i) = m_{ij}(a_i) - TC_{ij}(a_i)$. Using profit-maximizing prices (equation (2)), we obtain

$$\pi_{ij}(a_i) = m_{ij}(a_i) \frac{1}{\sigma} - C_j = \left(\frac{\sigma}{\sigma - 1} \frac{a_i \tau_j}{P_j} \right)^{1-\sigma} E_j - C_j. \tag{4}$$

Individual profit drives the decision to export to country j. This increases with destination market size (E_j), and falls with impediments to trade (τ_j and C_j). As is standard in monopolistic-competition models, the importing-country price index (P_j) enters positively in the expressions for both trade flows and export profits. This price index captures the influence of the greater competition that occurs in more central markets.[2]

We denote by $\bar{a}_j$ the marginal-cost level that ensures that the revenue from sales in country j just equals the total cost of exporting. From (4), this threshold value is

$$\bar{a}_j = \lambda_j \left(\frac{1}{C_j} \right)^{1/(\sigma-1)} \frac{1}{\tau_j}, \tag{5}$$

1 We assume that $\gamma > \sigma - 1$.

2 If N_{hj} is the number of firms selling their varieties on market j, then the price index is $P_j = (\int_1^{N_{hj}} (p_{ij}(a_i))^{1-\sigma} di)^{1/1-\sigma}$.

with

$$\lambda_j = \frac{\sigma - 1}{\sigma}(E_j)^{1/(\sigma-1)} P_j.$$

All firms with marginal cost less than or equal to $\bar{a}_j$ will export to j. The total number of exporting firms is thus

$$N_j = \int_0^{\bar{a}_j} Nf(a)\,da = \left[N\frac{\gamma}{\gamma - 1}\lambda_j^{\gamma}\right]\left(\frac{1}{C_j}\right)^{\gamma/(\sigma-1)} \tau_j^{-\gamma}. \tag{6}$$

Hence the value of bilateral trade from country H to market j is given by

$$\begin{aligned} M_j &= \int_0^{\bar{a}_j} Nm_{ij}(a_i) f(a)\,da \\ &= \Theta \frac{E_j}{P_j^{1-\sigma}} N\left(C_j\right)^{-\frac{[\gamma-(\sigma-1)]}{\sigma-1}} \left(\tau_j\right)^{-\gamma}, \end{aligned} \tag{7}$$

where

$$\Theta = \left(\frac{\sigma}{\sigma - 1}\right)^{1-\sigma} \left(\frac{\gamma}{\gamma - (\sigma - 1)}\right) \lambda_j^{\gamma-(\sigma-1)}.$$

This export equation is very similar to the traditional gravity equation in the Dixit-Stiglitz-Krugman (DSK) framework. Bilateral trade flows increase with the demand in destination country (E_j) and supply capacity in the exporting country (N). Trade is also a decreasing function of trade costs, τ_j. There are, nonetheless, two main differences from the standard DSK gravity equation. First, the fixed cost of entering the foreign market appears logically as an additional factor. Second, the trade-cost elasticity of trade differs significantly from that under homogeneous firms. It is straightforward from (7) that

$$\frac{\partial M_j}{\partial \tau_j}\frac{\tau_j}{M_j} = -\gamma.$$

The trade-cost elasticity of trade does not depend on the price elasticity here, whereas it is $(1 - \sigma)$ in the DSK model.[3] This is one of the most striking results of the model and requires us to reconsider the plentiful empirical and theoretical literature relating σ to central features of international trade, such as the impact of trade costs and border effects.

3 Our model here makes one simplifying assumption. Following Chaney (2008), we imagine that τ_j does not affect E_j and P_j; that is, we implicitly assume that the exporting country is 'small,' with negligible influence on the world economy.

To understand why firm heterogeneity affects this elasticity of trade flows, we consider trade margins. Trade costs affect trade via both the intensive and extensive margins. Lower τ_j increases both the number of firms exporting to country j (equation (6)) and the volume each firm exports (equation (3)). These are the extensive and intensive margins of trade, respectively. Our definition of trade margins differs from that used in most empirical work (see Hillberry and Hummels 2008; Mayer and Ottaviano 2007), where aggregate trade is decomposed into the number of exporters and the average shipment per exporter. This implicitly attributes the average export value to each individual shipment. Here, as in Chaney (2008), the extensive margin is defined as the value shipped by the marginal exporter, which is less than the average shipment.

Letting $\varepsilon_{\tau_j}^{M_j}$ denote the trade-cost elasticity of total trade, and $\varepsilon_{\tau_j}^{INT_j}$ and $\varepsilon_{\tau_j}^{EXT_j}$ be the trade-cost elasticities of the intensive and extensive margins, we have

$$\varepsilon_{\tau_j}^{M_j} = \varepsilon_{\tau_j}^{INT_j} + \varepsilon_{\tau_j}^{EXT_j} = -\gamma. \tag{8}$$

For the trade-cost elasticity of the extensive margin ($\varepsilon_{\tau_j}^{EXT_j}$), we cannot consider only equation (6). Economic integration will increase the number of exporting firms, as suggested by equation (6), but as τ_j falls further, the firms entering the export market will be less efficient and export less. Thus, the impact of a marginal reduction in trade costs on the extensive margin is given by the increase in the number of exporting firms multiplied by the quantity exported by the threshold firm (i.e., that with marginal cost $\bar{a}_j$):[4]

$$\varepsilon_{\tau_j}^{EXT_j} = \left[Nm_{ij}(\bar{a}_j) f(\bar{a}_j) \frac{\partial \bar{a}_j}{\partial \tau_j} \right] \frac{\tau_j}{M_j}.$$

Using (3), (5), and (6), we obtain, after some manipulation,

$$\varepsilon_{\tau_j}^{EXT_j} = -\left[\gamma - (\sigma - 1)\right]. \tag{9}$$

Finally, we can use equation (8) to obtain the trade-cost elasticity of the intensive margin:

$$\varepsilon_{\tau_j}^{INT_j} = -(\sigma - 1). \tag{10}$$

This elasticity is identical to the trade-cost elasticity of total trade in Krugman-type (1980) models of trade with homogeneous firms.

Chaney's model emphasizes the role of firm heterogeneity in international trade and shows that economic integration has different effects across industries. First, lower trade costs have greater effects on trade in less-heterogeneous industries (i.e., those in which γ is larger). Second, the decomposition of the effect of

4 Chaney (2008) provides a more explicit decomposition of total trade into intensive and extensive margins.

trade integration differs according to the degree of good differentiation. In industries with highly differentiated products (i.e., where σ is relatively low), trade integration allows the entry of a large number of firms, each of which has a relatively small market share: trade grows mainly via the extensive margin. On the contrary, in industries with homogeneous goods, lower trade costs work through the intensive margin, as less efficient firms will have more difficulty in entering export markets. Hence, only a small number of firms become new exporters.

3. Empirical strategy

We use trade data for a large number of exporting firms to estimate the parameters determining the influence of distance on both total trade and each trade margin. We assume a very simple trade-cost function: $\tau_j = \theta D_j^{\delta}$, where θ is a positive constant, D_j is the distance between the Home country and j, and δ is a strictly positive coefficient (see, e.g., Hummels 2001b; Anderson and Van Wincoop 2004). The distance elasticities of trade margins are

$$\begin{aligned} \varepsilon_{d_j}^{INT_j} &= -\delta(\sigma - 1) \\ \varepsilon_{d_j}^{EXT_j} &= -\delta\left[\gamma - (\sigma - 1)\right]. \end{aligned} \tag{11}$$

We appeal to a three-step method to estimate σ, δ, and γ. First, we estimate the probability that a firm exports, from which we have $\delta\gamma$. Second, we derive $-\delta(\sigma - 1)$ from estimated gravity equations on individual exports. Finally, to identify all three parameters, we estimate the Pareto distribution, that is, the relationship between individual productivity and production, to obtain an estimate of $-[\gamma - (\sigma - 1)]$.

The first step obtains $-\delta\gamma$ from the estimated effect of distance to foreign countries on the firm export probability. Equation (5) shows the maximum marginal cost at which firms export. When we use the definition of the Pareto distribution and reintroduce the k subscript, the probability that firm i with marginal cost a_i exports to country j at year t is

$$\text{Prob}[Exp_{kjt}(a_i)] = P(a_i < \bar{a}_j) = \left[\lambda_j \left(\frac{1}{C_j}\right)^{1/(\sigma-1)} \frac{1}{\tau_j}\right]^{\gamma}.$$

We estimate this equation taking the log of trade costs and introducing firm fixed effects and importing country and time dummies.[5] For the country-year fixed effects not to eliminate the distance variable, the latter needs to be firm specific, so that there is variation within France.

5 Note that this expression does not depend on the firm characteristics, a_i. It is likely, however, that trade costs τ_j contain some unobserved idiosyncratic individual characteristics (e.g., firm networks or former experience on export markets). These will be captured by the firm fixed effects.

For the first step of our estimation strategy, we estimate the following logit equation for each industry k:

$$\text{Prob}[Exp_{kjt}(a_i)] = -\delta\gamma \ln D_{ij} + e_i + e_{jt} + \nu_{ikjt}, \tag{12}$$

where D_{ij} is the distance between firm i and country j; e_i is a firm fixed effect; and e_{jt} is an import country-year fixed effect that controls for foreign market size, prices, and export fixed cost.

The second step consists in estimating the determinants of the individual export value from equation (3). Log-linearizing equation (3) produces the following estimable gravity equation for individual firms:

$$\ln[m_{kjt}(a_i)] = -\delta(\sigma - 1)\ln(D_{ij}) + e_i + e_{jt}, \tag{13}$$

where $m_{kjt}(a)$ is the value shipped by firm i to market j in year t. From the theoretical framework, the distance coefficient in this equation is the distance elasticity of the intensive margin, $-\delta(\sigma - 1)$.

We compute the three trade-elasticity parameters, δ, σ, and γ, in step 3, which consists in estimating the Pareto distribution. It can be shown from equations (2) and (3) that, for each firm with productivity $1/a_i$, the cumulative production of all firms with a higher productivity is $X = \lambda(1/a_i)^{-[\gamma-(\sigma-1)]}$. We estimate the coefficient $-[\gamma - (\sigma - 1)]$ using the same set of French firms as in steps 1 and 2. We generate a proxy for firm productivity[6] and, by year and industry, sort firms from the most to the least productive. For each firm, we then compute the cumulative production of all firms of lower rank. Regressing log cumulative production on the log of individual TFP[7] produces an estimate of $-[\gamma - (\sigma - 1)]$ for each industry. We combine the three estimates $-\delta\gamma$, $-\delta(\sigma - 1)$, and $-[\gamma - (\sigma - 1)]$ to calculate γ, σ, and δ. The following section describes the data required to carry out these estimations.

4. The trade data

We now describe the construction of the export database, highlighting the intra-national variability of distance. We then use our data to compute trade margins and show that distance affects both export flows and the decision to export.

4.1. Individual export data and intra-national distance

To estimate equations (12) and (13) we require firm-level export data, with controls for both the country-specific fixed cost and price index, so that distance only picks up the variable trade cost. Individual export data from France are collected

6 The proxy for individual TFP is based on the Olley-Pakes procedure.
7 All regressions include year fixed effects.

by French Customs and are available from INSEE. This database contains the value of exports by firm and destination country for each firm located in the French metropolitan area, from 1986 to 1992.

The structural estimation procedure emphasizes the impact of variable trade costs on trade. We assume that the fixed cost is specific to each importing country and introduce country fixed effects. Distance is identified from the location of the exporting firm within France, and thus captures only movements in variable trade costs. For this variable cost to play a significant role in export decisions, we consider only adjacent countries: Belgium-Luxembourg, Germany, Switzerland, Italy, and Spain.

The distance between each exporting firm and each export market is computed as follows. We assign to each importing country an exit-city located at the border and compute distance as the sum of the intranational and the international distances. The international distance is the distance between the exit-city and the destination country. This is computed as a population-weighted sum of the distances between the exit-city and the regions of the destination country. The intra-national distance is that between the firm and the exit-city and is calculated using very detailed data from INSEE on firm location. Each firm has one or more establishments or plants, which can be production plants or headquarters. Ideally, intra-national distance should proxy trade costs between the exporting plant and the French border. However, while we know where all plants are located, we do not know from which plant exports originate. This is not a problem for smaller firms, all of whose establishments are located in the same region (in which case we use the headquarters address), but can pose problems for larger firms, which may have a number of production plants in different parts of the country. Continental France is divided into 21 administrative *régions* (these *régions* have an average size of 25,500 km^2).[8] In the following, we restrict the sample to firms, all of whose plants are all located in one *région* only. This minimizes measurement error with respect to intranational distance while not restricting the analysis to single-plant firms.

Figures 1 and 2 highlight that there is substantial variation in intra-national distance, showing, for each *employment area*,[9] the share of exporters in the population of manufacturing firms, and the mean value of their shipments. Darker shading denotes greater mean individual trade flows (figure 1) and a higher share of exporting firms (figure 2). The negative effect of distance on trade stands out, as most of the darker regions are located close to the relevant border: the Pyrenees (South West) for Spain; Rhône-Alpes, Provence-Côte d'Azur, and Franche-Comté (South East) for Italy and Switzerland; Alsace, Lorraine, and Champagne-Ardennes (East) for Germany; and Nord-Pas de Calais, Picardie, and Ardennes (North-East) for Belgium. Distance clearly influences firm exports, even within the French territory.

8 These *régions* are the NUTS2 level in the Eurostat nomenclature.

9 *Employment areas* (of which there are 348) are an additional level used by INSEE and are defined by workers' commuting patterns.

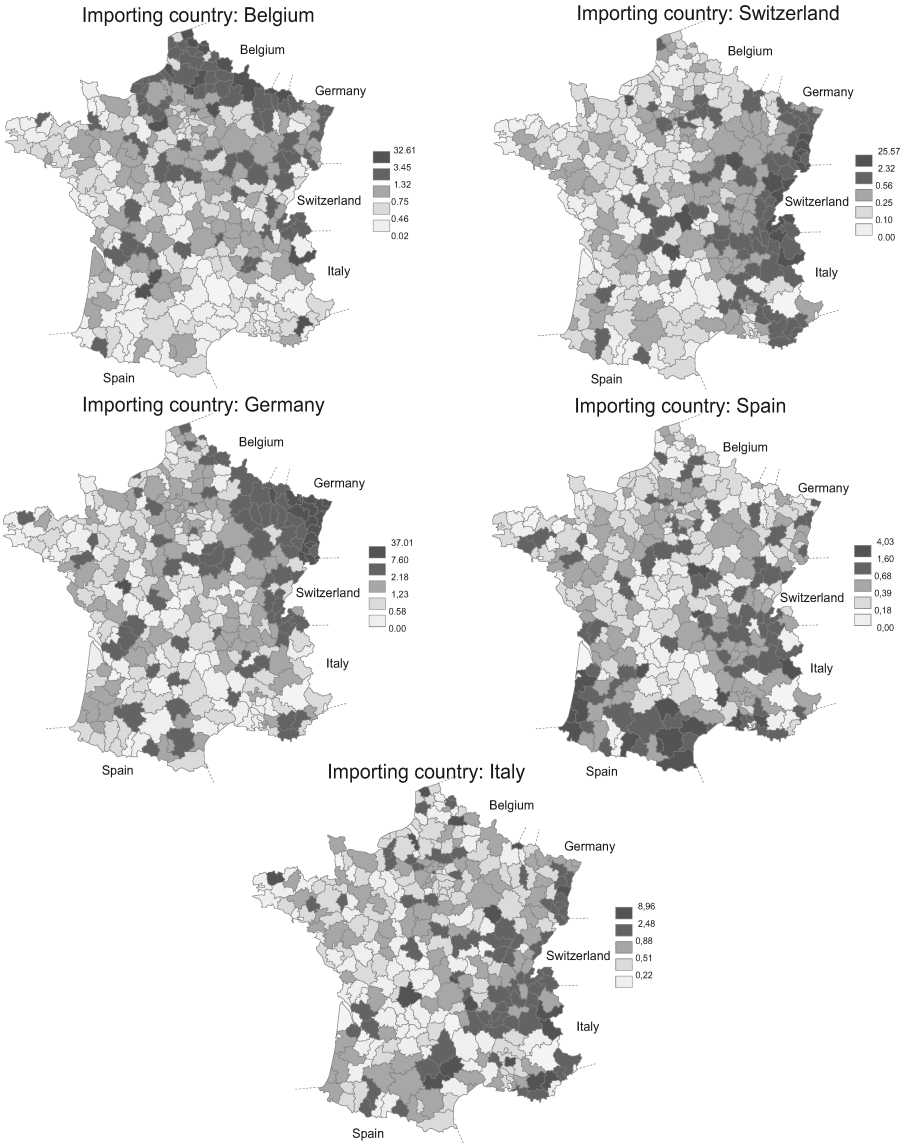


FIGURE 1 Mean value of individual-firm exports (single-region firms, 1992)

Finally, to add further information on firm characteristics, we merge firm-level exports with the Annual French Business Surveys (Enquêtes Annuelles d'Entreprises, EAE), also available from INSEE. Individual firms are identified via a 9-digit identifier, called the Siren. The Business Surveys provide information on the industrial sector, total employment, and the address of the firm. This

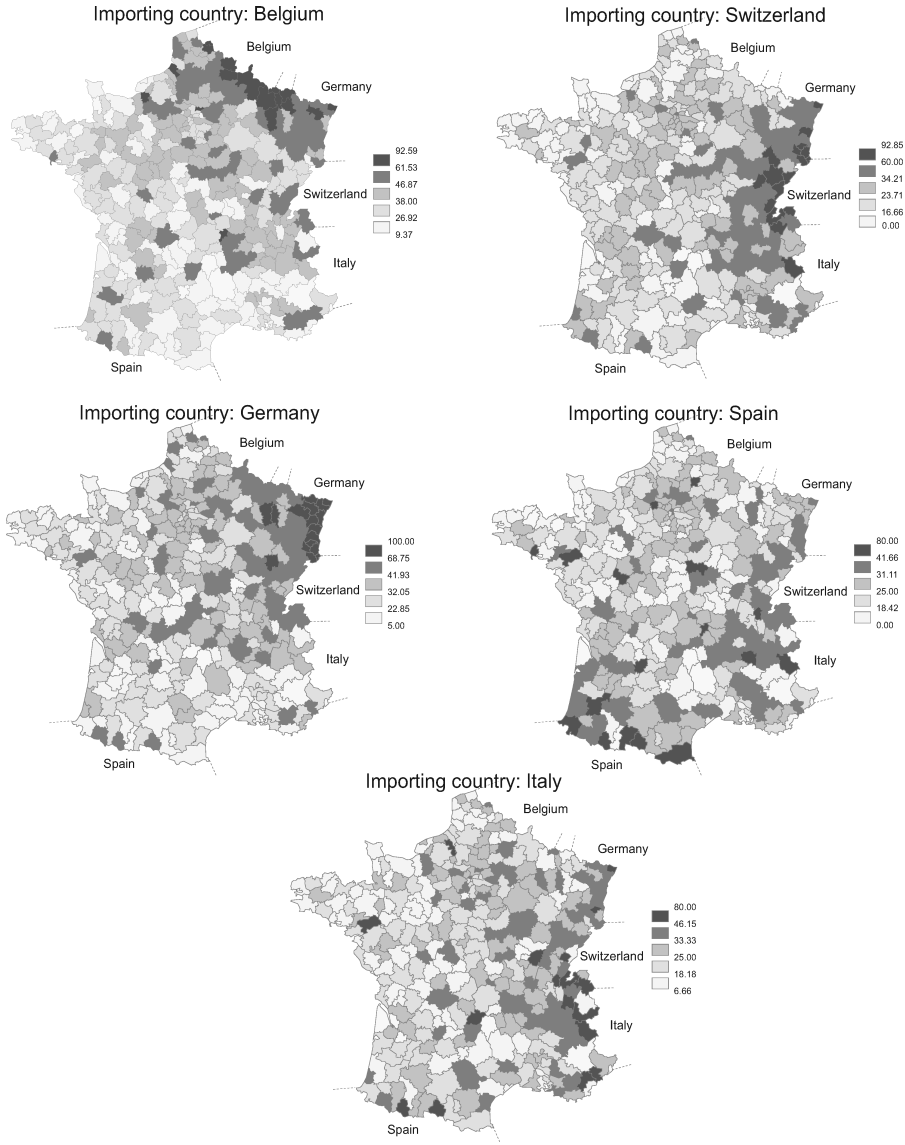


FIGURE 2 Percentage of firms that export (single-region firms, 1992)

merging does raise an additional selection issue, however, as firm-level information is available only for firms with more than 20 employees. The address of the firm being central for distance, we thus restrict the trade data to firms with more than 20 employees.

TABLE 1
Summary statistics

Year	No. firms	% of exporters	% of single-region firms	% of exporters among single-region firms
1986	22553	68.9	81.1	64.9
1987	22859	69.3	81.2	65.4
1988	23604	69.5	81.1	65.6
1989	23066	69.3	79.5	66.1
1990	23089	68.4	83.4	65.1
1991	24080	67.8	83.3	64.5
1992	23494	68.6	83.2	65.5

We tend to retain medium-size firms. Larger firms are more likely to be multi-region, and smaller firms do not appear in the EAE. Table 1 presents some descriptive statistics. In each year, over 83% of firms are single region. The share of exporters among those firms is lower than in the whole sample, but the difference remains small (about 65%, against 68% for the whole sample). Single-region firms account for about 78% of total exports. The potential estimation bias from our data restrictions should be recognized. We cannot explicitly evaluate the size of this bias, as we do not know the within-France location of smaller firms or from which region exports originate for multi-region firms. However, the resulting sample of firms fits the theoretical models well. The assumption that firm size is Pareto distributed plays an important role in our framework, and our restricted sample performs slightly better in this respect than does the overall sample. The distribution of manufacturing firms in developed countries appears to correspond better to a Pareto law in the case of medium-size firms (see, e.g., Axtell 2001; Cabral and Mata 2003). Further, the Pareto distribution is a power law, predicting a linear relationship between the log of the rank and the log of firm size. Rank size regressions on our sample of medium-size firms produce satisfactory results: the R-squared statistics by industry are relatively high, ranging from 0.47 to 0.92, with a mean value of 0.706. The same regressions carried out on all firms with over 20 employees yield less satisfactory results, with a smaller R-squared in all but two industries.[10]

4.2. A first look at trade margins

Before turning to the structural estimation, we relate the patterns in our data to existing results in the literature. The empirical trade literature provides simple decompositions of international trade flows and documents the existence of intensive and extensive trade margins. Eaton, Kortum, and Kramarz (2004) analyze how the French market share abroad affects the nature of bilateral trade.

10 These regression results are not reported here, but are available upon request.

Hillberry and Hummels (2008) provide a decomposition of the distance effect on intra-national US shipments, and Mayer and Ottaviano (2007) consider French and Belgian export flows. Bernard et al. (2007) use US export data at the firm level to estimate the impact of distance on the intensive and extensive margins (decomposed into the number of firms and the number of products).[11]

We estimate two different distance-trade elasticities: those regarding the number of exporters and the average individual-firm volume of trade. In this section, we do not restrict the database to neighbouring countries only: estimations are carried out on aggregate French exports to 159 countries, by industry, over all exporters and for single-region exporters only. We differ from Eaton, Kortum, and Kramarz (2004), as they focus on the trade effect of a change in French market share abroad, and not on distance. We are interested in the *distance* elasticity of trade flows (and in the decomposition of this elasticity); hence, we follow the methodology proposed by Hillberry and Hummels (2008) in their decomposition of intranational US shipments. Thus we decompose, for each industry k, the aggregate volume of trade from France to a given country j into the number of shipments (N_{kjt}) and the average value per shipment ($\bar{m}_{kjt}$), as follows:

$$M_{kjt} = N_{kjt}\bar{m}_{kjt}.$$

Taking logs, we have

$$\ln M_{kjt} = \ln N_{kjt} + \ln \bar{m}_{kjt}. \tag{14}$$

We analyze how each component varies with distance, and regress each term in equation (14) on distance, controlling for the size of importing countries via their current GDP.

We introduce three dummy variables reflecting cultural proximity between France and the importing country: French being spoken by at least 9% of the population ($French_j$), contiguity, and the destination country being a former French colony ($Colony_j$).[12] These will capture part of the fixed cost of exporting, C_j, which helps determine trade (equation (7)). We include a full set of industry and year dummies, and estimate the following equation:

$$\begin{aligned}\ln(Margin_{kjt}) = {} & \alpha_1 \ln D_j + \alpha_2 \ln GDP_{kjt} \\ & + French_j + Contig_j + Colony_j + e_k + e_t + v_{kjt},\end{aligned} \tag{15}$$

where e_k and e_t are industry and year fixed effects; v_{kjt} is an error term; and $\ln(Margin_{kjt})$ is, in turn, the log of average value per shipment and the log of the number of shipments. As OLS is linear, the coefficient on total trade will be equal

11 Helpman, Melitz, and Rubinstein (2008) estimate a structural trade-margins model. However, their trade-margin elasticities are based on aggregated trade flows for 158 countries rather than firm-level exports.

12 These three variables were obtained from the CEPII (*www.cepii.fr*).

TABLE 2
Decomposition of French aggregate industrial exports (34 industries, 159 countries, 1986 to 1992)

	All firms > 20 employees		Single-region firms > 20 employees	
	(1) Average shipment $\ln(M_{kjt}/N_{kjt})$	(2) Number of shipments $\ln(N_{kjt})$	(3) Average shipment $\ln(M_{kjt}/N_{kjt})$	(4) Number of shipments $\ln(N_{kjt})$
$\ln(GDP_{kj})$	0.461^a	0.417^a	0.421^a	0.417^a
	(0.007)	(0.007)	(0.007)	(0.008)
$\ln(Dist_j)$	-0.325^a	-0.446^a	-0.363^a	-0.475^a
	(0.013)	(0.009)	(0.012)	(0.009)
$Contig_j$	-0.064^c	-0.007	0.002	0.190^a
	(0.035)	(0.032)	(0.038)	(0.036)
$Colony_j$	0.100^a	0.466^a	0.141^a	0.442^a
	(0.032)	(0.025)	(0.035)	(0.027)
$French_j$	0.213^a	0.991^a	0.188^a	1.015^a
	(0.029)	(0.028)	(0.032)	(0.028)
N	23553	23553	23553	23553
R^2	0.480	0.591	0.396	0.569

NOTES: These are OLS estimates with year and industry dummies. Robust standard errors in parentheses with superscripts a, b, and c denoting significance at the 1%, 5%, and 10% level, respectively.

to the sum of the coefficients on the two margins. This shows us how much of the aggregate distance effect comes from average shipments per firm and how much from the number of shipments.

Table 2 shows the results over all firms (columns 1 and 2) and for single-region firms with more than 20 employees only (columns 3 and 4). The coefficients in the two sets of regressions are very similar, except for those on contiguity. Our sample of single-region firms thus exhibits the same trade patterns as the full set of firms regarding the decomposition of trade margins. Being close to a border country matters more for single-region than for multi-region firms, as reflected in the significant coefficient in column (4) compared with the insignificant estimate in column (2). Second, the estimated coefficients of the gravity equation are of the expected sign. GDP has a significant positive effect on both the volume exported by firms and the number of exporters. Distance always attracts a negative estimate. A common colonial history and sharing the same language increase both the intensive and the extensive margins. Third, the distance decomposition shows a somewhat greater effect on the extensive margin, both for all firms and for our sample. About 57% of the distance effect on trade works through the extensive margin (i.e., $(0.475/(0.363 + 0.475)) * 100 \simeq 56.7$); the remaining 43% reflects larger average shipments per firm.

Previous work (Eaton, Kortum, and Kramarz 2004; Bernard et al. 2007) finds qualitatively similar results, with the extensive margin being more important than

the intensive margin. However, it is difficult to compare our results exactly, as the methodologies differ. The most comparable results are found in Hillberry and Hummels (2008), who analyze trade flows within the United States, and Mayer and Ottaviano (2007), who study French and Belgian individual export flows. These show, respectively, that 96% and 75% of the distance effect on trade comes from the extensive margin. Our results are thus consistent with theirs, although we find a somewhat lower figure (since we do not consider firms with less than 20 employees, we lose part of the extensive margin). We also lose part of the intensive margin, but probably less than for the extensive margin, as small firms export less than the average export volume. Structural estimation of the three gravity parameters should produce a figure for the extensive margin closer to that estimated in Mayer and Ottaviano (2007). This is what we find in the following section, with a figure for the share of the extensive trade margin greater than 57%.

The elasticity decompositions in table 2 indicate that both trade margins exist. Further, the similar results for all firms and single-region firms is reassuring with respect to potential estimation bias. We now turn to the estimation of the structural parameters of the model, controlling for the fixed cost of exporting and using a definition of trade margins that conforms better to the theoretical model.

5. Structural gravity parameters

The firm-level estimation here includes destination-country fixed effects. We first discuss the estimated industry gravity parameters and then use the estimates to illustrate the sectoral elasticities of trade flows to trade barriers and distance.

5.1. Results by industry

We estimate the influence of distance on the firm-level export probability and export levels separately for each industry. All regressions include the contiguity dummy (for firms in regions sharing a common border with importing countries), country-year fixed effects, and firm fixed effects. This produces sectoral estimates for $\delta\gamma$ and $\delta(\sigma - 1)$, respectively. The estimated coefficients are shown in columns (1) and (2) of table 3.[13] Distance has a significant negative impact on export probability for all industries and a significant negative impact on individual export volume for all but six industries (pharmaceuticals, office equipment, electrical equipment, aeronautical building, precision instruments, and the garment industry). The average coefficient for export probability by industry ($-\delta\gamma$) is -1.41, which lies in the typical range of values for distance elasticities using

13 Column 1 in table 3 shows the marginal coefficients of the individual export probability.

TABLE 3
The structural parameters of the gravity equation (firm-level estimations)

Industry	Code	(1) $P[Export > 0]$ $-\delta\gamma$	(2) Export value $-\delta(\sigma - 1)$	(3) Pareto# $-[\gamma - (\sigma - 1)]$	(4) γ	(5) σ	(6) δ
Iron and steel	10	-5.51^a	-1.71^a	-1.36	1.98	1.62	2.78
Steel processing	11	-1.50^a	-0.99^a	-1.74	5.10	4.36	0.29
Metallurgy	13	-2.14^a	-0.73^a	-1.85	2.82	1.97	0.76
Minerals	14	-2.98^a	-0.91^a	-2.86	4.11	2.25	0.72
Ceramic and building mat.	15	-2.63^a	-0.76^a	-1.97	2.76	1.79	0.95
Glass	16	-2.33^a	-0.58^a	-2.13	2.84	1.70	0.82
Chemicals	17	-1.81^a	-0.76^a	-1.09	1.89	1.80	0.95
Speciality chemicals	18	-0.97^a	-0.34^a	-1.39	2.13	1.74	0.46
Pharmaceuticals	19	-1.19^a	-0.14	-1.40			
Foundry	20	-1.72^a	-0.85^a	-2.37	4.68	3.31	0.37
Metal work	21	-1.19^a	-0.36^a	-2.43	3.48	2.05	0.34
Agricultural machines	22	-2.06^a	-0.57^a	-2.39	3.31	1.92	0.62
Machine tools	23	-1.29^a	-0.48^a	-2.47	3.92	2.45	0.33
Industrial equipment	24	-1.25^a	-0.48^a	-1.97	3.21	2.24	0.39
Mining/civil egnring eqpmt	25	-1.37^a	-0.46^a	-1.90	2.86	1.96	0.48
Office equipment	27	-0.52^a	-1.02	-1.57			
Electrical equipment	28	-0.80^a	-0.14	-2.34			
Electronical equipment	29	-0.77^a	-0.24^a	-1.63	2.34	1.71	0.33
Domestic equipment	30	-0.94^a	-0.14^a	-2.13	2.51	1.37	0.38
Transport equipment	31	-1.40^a	-0.55^a	-2.23	3.69	2.46	0.38
Ship building	32	-3.69^a	-2.67^a	-1.52	5.53	5.01	0.67
Aeronautical building	33	-0.78^a	-0.13	-3.27			
Precision instruments	34	-1.07^a	0.08	-1.63			
Textile	44	-1.17^a	-0.30^a	-1.37	1.84	1.47	0.64
Leather products	45	-1.24^a	-0.44^a	-1.63	2.53	1.90	0.49
Shoe industry	46	-0.42^a	-0.29^a	-2.30	7.31	6.01	0.06
Garment industry	47	-0.33^a	0.13	-1.04			
Mechanical woodwork	48	-2.14^a	-0.20^a	-1.50	1.65	1.15	1.29
Furniture	49	-1.43^a	-0.37^a	-2.25	3.04	1.79	0.47
Paper & Cardboard	50	-1.45^a	-0.76^a	-1.76	3.71	2.95	0.39
Printing and editing	51	-1.40^a	-0.70^a	-1.24	2.46	2.22	0.57
Rubber	52	-1.26^a	-0.80^a	-2.52	6.93	5.41	0.18
Plastic processing	53	-1.24^a	-0.51^a	-1.60	2.70	2.11	0.46
Miscellaneous	54	-0.91^a	-0.33^a	-1.22	1.92	1.70	0.47
Trade-weighted mean		-1.41	-0.53	-1.86	3.09	2.25	0.58

NOTES: Superscripts a, b, and c denote significance at the 1%, 5%, and 10% level respectively. #: All coefficients in this column are significant at the 1% level. Estimations include the contiguity variable.

aggregate trade flows.[14] This similarity conforms to theory, which predicts the same distance coefficient for the firm-level export probability (equation (7)) and the aggregate volume of exports (equation (12)).

We use the preceding estimates and the results of the Pareto estimation (our step 3) to compute the three parameters of the trade elasticities, δ, σ, and γ. We obtain consistent coefficients for the large majority of industries: 28 out of 34 in table 3. These industries account for 78.3% of total manufacturing exports by French firms with more than 20 employees, and 79.2% of the number of these firms. For all industries with significant coefficients, the estimates are consistent with theory in both sign and size: the values of σ are strictly greater than 1, and those of γ are greater than $\sigma - 1$.

The values of σ reported in table 3 range between 1.15 and 6.01, with an average figure of 2.25. These values are smaller than those in the recent literature. Broda and Weinstein (2006) report values between 4 and 6.8 when estimating the parameters on three-digit data. Eaton and Kortum's (2002) results lie around the average value of 8.3, and Erkel-Rousse and Mirza (2002) obtain a mean value of 3.7. Hummels's (2001a) average value is 5.6, and that in Head and Ries (2001) is 7.9.[15] Nevertheless, the comparison of our sectoral estimates of σ with those in Broda and Weinstein (2006) is satisfactory. The latter provide import-demand elasticities for SITC Rev.2 3-digit products; we aggregate these data to match our sectoral classification, taking the median value of their estimates. Panel (a) of figure 3 plots our values for σ against those obtained by Broda and Weinstein. Although there is a positive relationship between the two series, the simple cross-industry correlation is not significant. However, the Spearman rank correlation coefficient of 0.53 is significant at the 1% level.

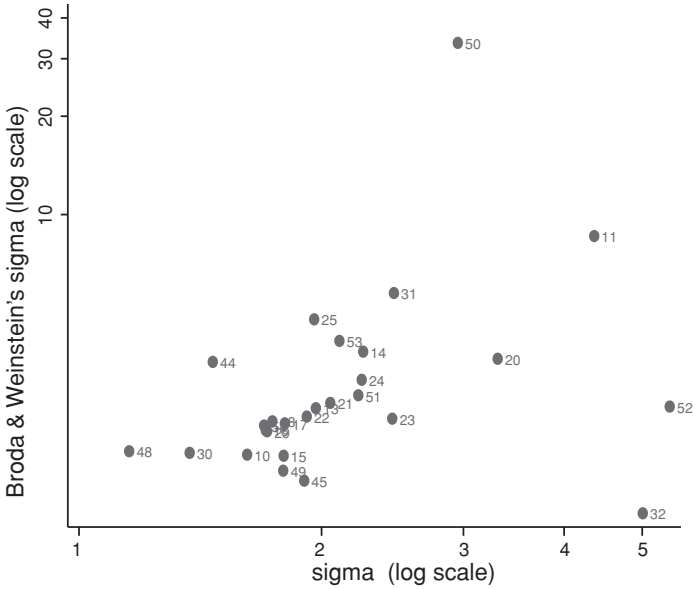
Our average value for δ is 0.58. This is close to the estimates in the existing literature based on international transport costs. The mean value of the distance elasticity of trade costs in Radelet and Sachs (1998) is 0.13. Glaeser and Kohlhase's (2003) estimate of the same parameter is 0.3, and Hummels's (2007) average estimate is 0.2. Our figure is larger, which is likely due to the fact that we only consider continental shipments. Road transport decays more strongly with distance: for instance, Combes and Lafourcade (2005) obtain an elasticity of 0.8 using road transport costs within France. For a closer comparison, we use freight rates for bilateral trade by road. These data do not exist for Europe, so we use sectoral freight rates between the U.S. and Canada, constructed from two sources: the NBER US import database compiled by Feenstra, Romalis, and Schott (2002), and those in Hummels (2007).[16] Panel (b) in figure 3

14 The meta-analysis of Disdier and Head (2008) surveys 1,466 gravity estimations and obtains a mean coefficient value of −0.91, with 90% of estimates lying between 0.28 and 1.55.

15 Note that Hummels (2001a) and Head and Ries (2001) estimate the impact of trade barriers on bilateral trade flows, which, according to the model, should be interpreted as a measure of γ for each industry.

16 We thank Daniel Mirza and Emmanuel Milet for computing and providing us with the freight data.

(a)



(b)

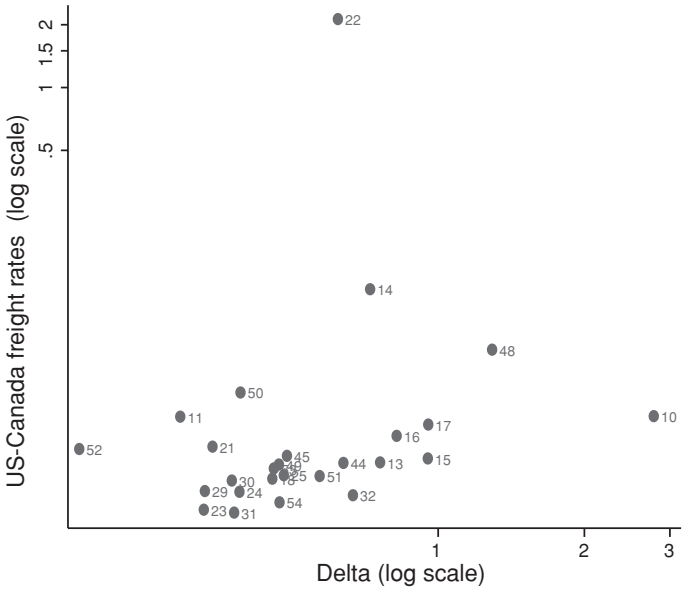


FIGURE 3 Comparison of our results for σ and δ with those of Broda and Weinstein (2003)

displays our values for δ together with the US-Canada freight rates. There is again a positive relationship, although the Pearson correlation between our δs and the North-American freight rates is not significant. However, the Spearman rank correlation of 0.41 is significant at the 5% level.

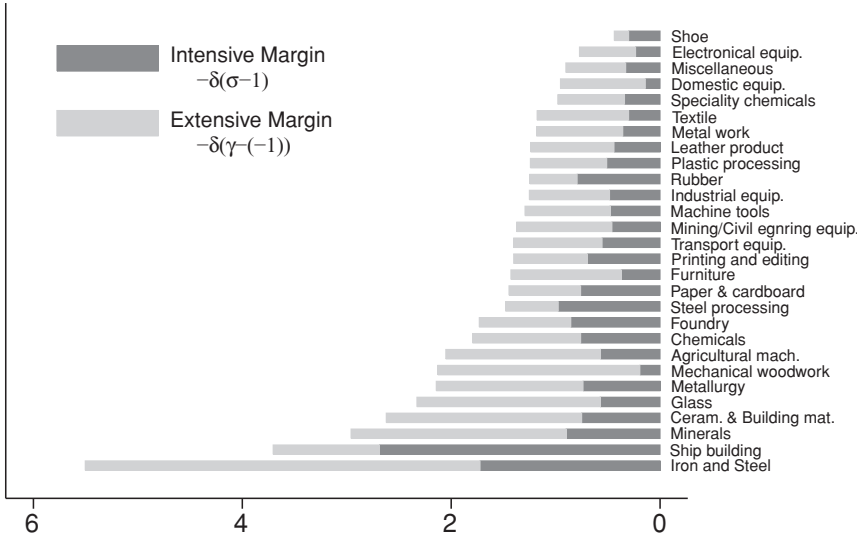
5.2. The impact of lower trade barriers

A simple way of illustrating these results is to calculate the impact of shorter distances, or lower trade barriers, on French trade. The results by industry are presented in figure 4. Shorter distances increase the extensive margin by $\delta[\gamma - (\sigma - 1)]$ and the intensive margin by $\delta(\sigma - 1)$. Panel (a) in figure 4 shows the decomposition of the aggregate trade elasticity to distance across industries. The width of the bar represents the sum of the distance elasticities of the extensive margin (the grey part of the bar) and the intensive margin (the black part of the bar). The effect of freight costs stands out clearly. The effect of distance is strongest for iron and steel, ship building, minerals, ceramic and building materials, and glass. These industries produce relatively heavy goods (high δ). On the other hand, the effect of distance on aggregate trade is the smallest for speciality chemicals, domestic equipment, electronical equipment, and the shoe industry.

Our results allow us to disentangle the trade effects of a change in distance from those reflecting trade costs such as tariffs. Lowering tariffs increases the extensive margin by $[\gamma - (\sigma - 1)]$ and the intensive margin by $(\sigma - 1)$. Panel (b) in figure 4 displays the decomposition into the number of exporters and individual export volume of the elasticity of trade margins to the variable trade cost τ, computed for all industries for which we obtained consistent estimates. The industry ranking is very different from that in panel (a). The industries with the greatest tariff effect are the shoe industry, rubber, ship building, steel processing and foundry. Table 3 shows that these industries are characterized by above-average γs, so that the overall effect of trade barriers on trade flows depends mostly on γ. On the other hand, the least sensitive industries are mechanical woodwork, textiles, chemicals, iron and steel, and speciality chemicals. The considerable differences between panels (a) and (b) of figure 4 demonstrate that inferring the consequences of trade liberalization from distance coefficients is potentially very hazardous.

Finally, we use our estimated parameters δ, σ, and γ to reconstruct the share of the extensive margin in the overall effect of distance or trade barriers on trade $(1 - [(\sigma - 1)/\gamma])$, which is the key figure in the literature. The trade-weighted average share across all industries for which coefficients are consistent is 61.96%. This number is higher than that obtained via the decomposition in section 4 (57%). Hence, despite our sample restrictions, we obtain a number which is closer to that in Mayer and Ottaviano (2007). Note that our definition of the extensive margin, which accounts only for the contribution of marginal exporters to total trade, will produce a smaller figure than that from the trade decomposition in Hillberry and Hummels (2008) and Mayer and Ottaviano (2007). Their decomposition implicitly assumes that all firms perform identically on export markets. We do

(a) Impact of distance on trade margins



(b) Impact of a tariff on trade margins

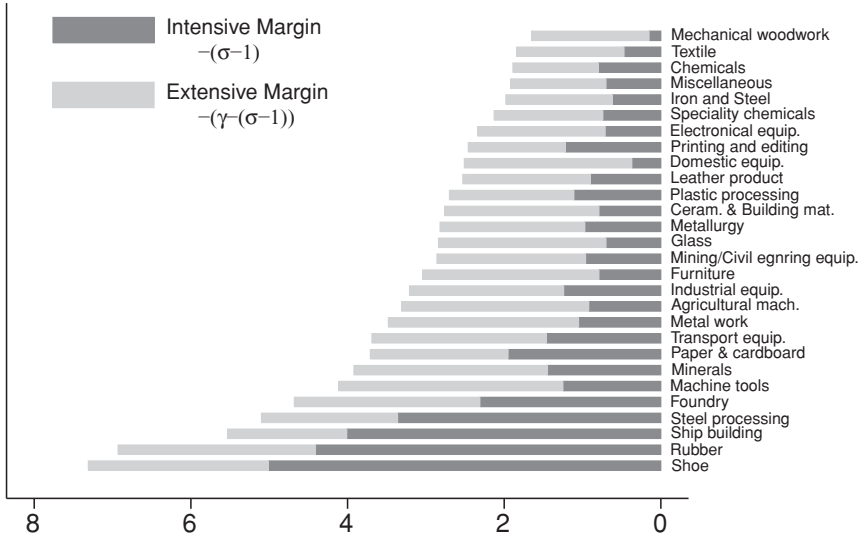


FIGURE 4 The estimated impact of trade barriers and distance on trade margins, by industry

indeed find a smaller share of extensive margin than do Mayer and Ottaviano (75%) and Hillberry and Hummels (96%).

There is considerable variation in the share of the extensive margin in the overall distance-trade elasticity in figure 4, ranging from 27.5% for ship building to 90.7% for mechanical woodwork. Rubber and textiles are good examples of the importance of trade margins. Their overall effect of distance on trade is similar (10% shorter distances increase trade by around 12%). However, for textiles this increase comes mostly from entry by new exporters (the extensive margin accounts for 75% of the overall increase), while this figure is only 36% for rubber. Industries thus have widely varying responses to trade barriers, beyond the overall consequences to trade. Trade policies will thus affect market structure differently by industry. Firm heterogeneity is therefore central to the analysis of trade policy.

6. Conclusion

The empirical literature on the effects of trade integration on trade has been considerably affected by new trade models with heterogeneous firms. This paper uses firm export-behaviour equations to estimate the three key parameters underlying trade in gravity models with heterogeneous firms (Chaney 2008): the elasticity of substitution (σ), the distance elasticity of trade costs (δ), and the degree of firm heterogeneity (γ).

We consider French exports to border countries, controlling for destination-country fixed costs to isolate the effect of variable trade costs. We estimate three equations: a gravity equation at the firm level, an export-selection equation, and a rank-size distribution of productivity across firms. These yield a combination of δ, σ, and γ for each sector. For a large majority of sectors, the estimated gravity parameters are consistent in sign and size with theory: the values of σ are strictly greater than 1, and the γs are greater than $\sigma - 1$. These industries account for 78.3% of total French manufacturing exports by firms with more than 20 employees.

This paper contributes to the existing literature in two ways. First, our structural estimates confirm the theoretical predictions from Chaney (2008) for a large majority of industries. Second, we use our estimates to describe the widely varying impacts of trade barriers and distance on trade across industries. We show that a comparable decrease in transport costs or in tariffs would have a very different impact on trade, depending on the elements of industrial structure.

References

Anderson, James. E., and Eric Van Wincoop (2004) 'Trade costs,' *Journal of Economic Literature* 42, 691–751

Axtell, R.L. (2001) 'Zipf distribution of U.S. firm sizes,' *Science* 293, 1818–20

Bernard, Andrew B., J. Bradford Jensen, Steve Redding, and Peter K. Schott (2007) 'Firms in international trade,' *Journal of Economic Perspectives* 21, 105–30
Broda, Christian, and David. E. Weinstein (2006) 'Globalization and the gains from variety,' *Quarterly Journal of Economics* 121, 541–85
Cabral, Luis M.B., and Mata José (2003) 'On the evolution of the firm size distribution: facts and theory,' *American Economic Review* 93, 1075–90
Chaney, Thomas (2008) 'Distorted gravity: the intensive and extensive margins of international trade,' *American Economic Review* 98, 1707–21
Combes, Pierre-Philippe, and Miren Lafourcade (2005) 'Transport costs: measures, determinants and regional policy implications for France,' *Journal of Economic Geography* 5, 319–49
Disdier, Anne-Celia, and Keith Head (2008) 'The puzzling persistence of the distance effect on bilateral trade,' *Review of Economics and Statistics* 90, 37–41
Eaton, Jonathan, and Sam S. Kortum (2002) 'Technology, geography, and trade,' *Econometrica* 70, 1741–79
Eaton, Jonathan, Sam S. Kortum, and F. Kramarz (2004) 'Dissecting trade: firms, industries, and export destinations,' *American Economic Review, Papers and Proceedings* 93, 150–4
Erkel-Rousse, Hélène, and Daniel Mirza (2002) 'Import price-elasticities: reconsidering the evidence,' *Canadian Journal of Economics* 35, 282–306
Feenstra, Robert C., John Romalis, and Peter K. Schott (2002) 'U.S. imports, exports, and tariff data, 1989–2001,' NBER Working Paper No. 9387
Glaeser, Edward L., and Janet E. Kohlhase (2003) 'Cities, regions and the decline in transport costs,' Discussion Paper 2014, Harvard Institute of Economic Research
Head, Keith, and John Ries (2001) 'Increasing returns versus national product differentiation as an explanation for the pattern of US-Canada trade,' *American Economic Review* 91, 858–76
Helpman, Elhanan, Marc Melitz, and Yona Rubinstein (2008) 'Estimating trade flows: trading partners and trading volumes,' *Quarterly Journal of Economics* 123, 441–87
Hillberry, Russel, and David Hummels (2008) 'Trade responses to geographic frictions: a decomposition using micro-data,' *European Economic Review* 52, 527–50
Hummels, David (2001a) 'Toward a geography of trade costs,' mimeo
— (2001b) 'Time as a trade barrier,' mimeo
— (2007) 'Transportation costs and international trade in the second era of globalization,' *Journal of Economic Perspectives* 21, 131–54
Krugman, Paul R. (1980) 'Scale economies, product differentiation, and the pattern of trade,' *American Economic Review* 70, 950–9
Mayer, Thierry, and Gianmarco I.P. Ottaviano (2007) 'The happy few: new facts on the internationalisation of European firms,' Bruegel-CEPR EFIM 2007 Report, Bruegel Blueprint Series
Melitz, Marc (2003) 'The impact of trade on intra-industry reallocations and aggregate industry productivity,' *Econometrica* 71, 1695–725
Melitz, Marc, and Gianmarco I.P. Ottaviano (2008) 'Market size, trade and productivity,' *Review of Economic Studies* 75, 295–316
Radelet, Steven, and Jeffrey Sachs (1998) 'Shipping costs, manufactured exports and economic growth,' Center for International Development, mimeo
Tinbergen, Jan (1962) *Shaping the World Economy* (New York: Twentieth Century Fund)