Online appendix for: Search Frictions in International Goods Markets

Clémence Lenoir* Julien Martin† Isabelle Mejean‡

11 July 2022

OA.1. Online appendix: Theory

OA.1.1. Analytical details

Under the model's assumptions, the number of suppliers from j offering a price below p is drawn from a Poisson distribution of parameter $\lambda_{ij}\mu_{ij}(p) = \lambda_{ij}T_j\left(d_{ij}w_j\right)^{-\theta}p^{\theta} = \lambda_{ij}\upsilon_{ij}p^{\theta}$ where we define $\upsilon_{ij} \equiv T_j\left(d_{ij}w_j\right)^{-\theta}$ to alleviate notations. Likewise, the number of suppliers from any country offering a price below p is drawn from a Poisson distribution of parameter $\sum_j \lambda_{ij}\mu_{ij}(p) = \sum_j \lambda_{ij}T_j\left(d_{ij}w_j\right)^{-\theta}p^{\theta} = \kappa_i\Upsilon_i p^{\theta}$.

Preliminary results: The assumption of Poisson draws makes it possible to derive a number of useful properties for the distribution of prices offered to buyers in country i. The following theorem characterizes the joint distributions of prices offered to a particular buyer, when we note $P_i^{(n)}$ the n'th lowest price offer received by a buyer in i and $P_{ij}^{(n)}$ the n'th lowest price offer received by a buyer in i from a seller in j.[1]

Theorem 1. *The joint density of $P_i^{(n)}$ and $P_i^{(n+1)}$ is:*

$$g_{i,n,n+1}(p_n, p_{n+1}) = \frac{\theta^2}{(n-1)!}\left(\kappa_i\Upsilon_i\right)^{n+1}p_n^{\theta n-1}p_{n+1}^{\theta-1}\exp\left[-\kappa_i\Upsilon_i p_{n+1}^{\theta}\right]$$

*CREST-ENSAE. Address: CREST 5, Av Henry Le Chatelier, 91120 Palaiseau, France. Email address: clemence.lenoir@ensae-paristech.fr.

†UQAM, CREST, and CEPR. Address: Université du Québec à Montréal, C.P. 8888, succursale A, Montréal, QC H3C 3P8, Canada. Email address: martin.julien@uqam.ca.

‡Sciences Po and CEPR. Address: Sciences Po, Department of Economics, 28 rue des Saints Pères, 75006 Paris, France. Email address: isabelle.mejean@sciencespo.fr

1. Here, we closely follow the steps in Eaton and Kortum (2010), chapter 4.

for $0 < p_n \leq p_{n+1} < \infty$*. The marginal density of* $P_i^{(n)}$ *is:*

$$g_{i,n}(p) = \frac{\theta}{(n-1)!} (\kappa_i \Upsilon_i)^n p^{\theta n - 1} \exp\left[-\kappa_i \Upsilon_i p^\theta\right]$$

for $0 < p < \infty$*.*
Likewise, the joint density of $P_{ij}^{(n)}$ *and* $P_{ij}^{(n+1)}$ *is:*

$$g_{ij,n,n+1}(p_n, p_{n+1}) = \frac{\theta^2}{(n-1)!} (\lambda_{ij} \upsilon_{ij})^{n+1} p_n^{\theta n - 1} p_{n+1}^{\theta - 1} \exp\left[-\lambda_{ij} \upsilon_{ij} p_{n+1}^\theta\right]$$

for $0 < p_n \leq p_{n+1} < \infty$ *while the marginal density of* $P_{ij}^{(n)}$ *is:*

$$g_{ij,n}(p) = \frac{\theta}{(n-1)!} (\lambda_{ij} \upsilon_{ij})^n p^{\theta n - 1} \exp\left[-\lambda_{ij} \upsilon_{ij} p^\theta\right]$$

for $0 < p < \infty$*.*

Proof. Under the model's assumptions, the distribution of P_i given $P_i \leq \bar{p}$ is:

$$F(p|\bar{p}) = \begin{cases} \left(\frac{p}{\bar{p}}\right)^\theta & if \quad p \leq \bar{p} \\ 1 & if \quad p > \bar{p} \end{cases}$$

The probability that a price is less than p_n is $F(p_n|\bar{p})$ and the probability that a price is more than p_{n+1} is $(1 - F(p_{n+1}|\bar{p}))$. Hence, if the buyer has met with m sellers with price below $\bar{p}$, the probability that n are lower than p_n and $m - n$ are greater than p_{n+1} is:

$$Pr\left[P_i^{(n)} \leq p_n, P_i^{(n+1)} \geq p_{n+1} | m\right] = \binom{n}{m} F(p_n|\bar{p})^n (1 - F(p_{n+1}|\bar{p}))^{m-n}$$

Taking the negative of the cross-derivative of this expression with respect to p_n and p_{n+1} gives the joint density of $P_i^{(n)}$ and $P_i^{(n+1)}$, conditional on m:

$$g_{i,n,n+1}(p_n, p_{n+1}|\bar{p}, m) = \frac{m! F(p_n|\bar{p})^{n-1} (1 - F(p_{n+1}|\bar{p}))^{m-n-1} F'(p_n|\bar{p}) F'(p_{n+1}|\bar{p})}{(n-1)!(m-n-1)!}$$

for $p_{n+1} \leq p_n$ and $m \leq n+1$. For $m < n+1$, $g_{i,n,n+1}(p_n, p_{n+1}|\bar{p}, m) = 0$.

The number m of price quotes is drawn from a Poisson distribution with parameter $\kappa_i \Upsilon_i \bar{p}^\theta$. The expectation of the joint distribution unconditional on m is thus:

$$\begin{aligned} g_{i,n,n+1}(p_n, p_{n+1}|\bar{p}) &= \sum_{u=0}^{\infty} \frac{\exp\left[-\kappa_i \Upsilon_i \bar{p}^\theta\right] \left(\kappa_i \Upsilon_i \bar{p}^\theta\right)^u}{u!} g_{i,n,n+1}(p_n, p_{n+1}|\bar{p}, m) \\ &= \frac{F(p_n|\bar{p})^{n-1} \left(\kappa_i \Upsilon_i \bar{p}^\theta\right)^{n+1} \exp\left[-\kappa_i \Upsilon_i \bar{p}^\theta F(p_{n+1}|\bar{p})\right] F'(p_n|\bar{p}) F'(p_{n+1}|\bar{p})}{(n-1)!} \\ &= \frac{\theta^2}{(n-1)!} (\kappa_i \Upsilon_i)^{n+1} p_n^{\theta n - 1} p_{n+1}^{\theta - 1} \exp\left[-\kappa_i \Upsilon_i p_{n+1}^\theta\right] \end{aligned}$$

which is the expression in Theorem 1 for $\bar{p} \to \infty$. The marginal density comes immediately from:

$$g_{i,n}(p) = \int_p^\infty g_{i,n,n+1}(p, p_{n+1}) dp_{n+1}$$

$\square$

Theorem 1 thus characterizes the joint distribution of each pair of adjacent order statistics, in the overall subset of offers received by a buyer in i and in the subset of offers originating from j. These distributions solely depend on θ, $\kappa_i \Upsilon_i$ and $\lambda_{ij} \upsilon_{ij}$.

Another useful property of random variables described by the marginal distribution in Theorem 1 is summarized in Theorem 2:

THEOREM 2. *For each order n, the b'th moment ($b > -\theta n$) is:*

$$E\left[\left(P_i^{(n)}\right)^b\right] = (\kappa_i \Upsilon_i)^{-b/\theta} \frac{\Gamma\left[(\theta n + b)/\theta\right]}{(n-1)!}$$

where $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$ is the gamma function.

Likewise,

$$E\left[\left(P_{ij}^{(n)}\right)^b\right] = (\lambda_{ij} \upsilon_{ij})^{-b/\theta} \frac{\Gamma\left[(\theta n + b)/\theta\right]}{(n-1)!}$$

Proof. First consider $k = 1$:

$$\begin{aligned} E\left[\left(P_i^{(1)}\right)^b\right] &= \int_0^\infty p^b g_{i,1}(p) dp \\ &= \int_0^\infty (\kappa_i \Upsilon_i)\, \theta p^{\theta+b-1} \exp\left[-\kappa_i \Upsilon_i p^\theta\right] dp \end{aligned}$$

We now use a change of variable: $v = \kappa_i \Upsilon_i p^\theta$:

$$\begin{aligned} E\left[\left(P_i^{(1)}\right)^b\right] &= \int_0^\infty \left(\frac{v}{\kappa_i \Upsilon_i}\right)^{b/\theta} \exp\left[-v\right] dv \\ &= (\kappa_i \Upsilon_i)^{-b/\theta}\, \Gamma\left(\frac{\theta + b}{\theta}\right) \end{aligned}$$

which is defined for $\theta + b > 0$.

More generally:

$$E\left[\left(P_i^{(n)}\right)^b\right] = \int_0^\infty p^b g_{i,n}(p) dp$$

$\square$

Weibull distribution of prices: Based on Theorem 1, one can recover the distribution of the n'th lowest price $P_i^{(n)}$:

$$F_{i,n}(p) \equiv Pr[P_i^{(n)} \leq p] = 1 - \sum_{u=0}^{n-1} \frac{\left(\kappa_i \Upsilon_i p^\theta\right)^u}{u!} \exp\left[-\kappa_i \Upsilon_i p^\theta\right]$$

As is necessary for a cumulative distribution, $F_{i,n}(p)$ approaches 1 when p tends to infinity and $F'_{i,n}(p) = g_{i,n}(p)$.

In particular, the lowest price is distributed:

$$F_{i,1}(p) \equiv Pr[P_i^{(1)} \leq p] = 1 - \exp\left[-\kappa_i \Upsilon_i p^\theta\right]$$

which is the Weibull distribution used in the main text.

Likewise, the distribution of the lowest price received from j is:

$$F_{ij,1}(p) \equiv Pr[P_{ij}^{(1)} \leq p] = 1 - \exp\left[-\lambda_{ij} \upsilon_{ij} p^\theta\right]$$

Bilateral trade probabilities: Section 3.2 of the paper analyzes the share of buyers from i purchasing the product from country j. When the number of buyers is large enough, the share is also the expected value of $\mathbb{1}_{b_i j}^{(1)}$, a dummy variable that is equal to 1 if the lowest price received by buyer b_i originates from country j.

To derive this expected value, we first derive the probability that the lowest-cost seller from j that buyer b_i meets is the lowest cost supplier overall, conditional on a price p. By definition, this probability is equal to the probability that all lowest-cost sellers from a different country offer a price above p:

$$\prod_{j' \neq j} [1 - F_{ij',1}(p)] = exp\left[-p^\theta \sum_{j' \neq j} \lambda_{ij} \upsilon_{ij}\right]$$

Integrating over p gives the probability that the lowest-cost supplier met from j is the lowest-cost supplier met:

$$\begin{aligned} E\left[\mathbb{1}_{b_i j}^{(1)}\right] &= \int_0^\infty exp\left[-p^\theta \sum_{j' \neq j} \lambda_{ij} \upsilon_{ij}\right] dF_{ij,1}(p) \\ &= \frac{\lambda_{ij}\upsilon_{ij}}{\kappa_i \Upsilon_i} \left[1 - F_{i,1}(p)\right]_0^\infty \\ &= \frac{\lambda_{ij}\upsilon_{ij}}{\kappa_i \Upsilon_i} \end{aligned}$$

Bilateral trade shares: Under iso-elastic preferences, the nominal demand expressed by a buyer b_i is a function of the lowest price received, at the power $1 - \sigma$:

$$p_{b_i} c_{b_i} = \left(P_{b_i}^{(1)}\right)^{1-\sigma} \bar{X}_i$$

The expected value of bilateral imports from j to i can thus be written as the expected value of individual purchases, across buyers in i that end up purchasing the good from j, i.e. conditional on the lowest-cost supplier from j that the buyer has met offering a price below the lowest-cost supplier of any other country:

$$\begin{aligned}\mathbb{E}\left[p_{b_i}c_{b_i}|\mathbb{1}^{(1)}_{b_ij}=1\right] &= \bar{X}_i\int_0^\infty p^{1-\sigma}\exp\left[-p^\theta\sum_{j'\neq j}\lambda_{ij}\upsilon_{ij}\right]dF_{ij,1}(p)\\ &= \bar{X}_i\theta\lambda_{ij}\upsilon_{ij}\int_0^\infty p^{\theta-\sigma}\exp\left[-p^\theta\kappa_i\Upsilon_i\right]dp\end{aligned}$$

To derive bilateral trade shares, this expression must be compared with the expected value of individual purchases, irrespective of the source country:

$$\begin{aligned}\mathbb{E}\left[p_{b_i}c_{b_i}\right] &= \bar{X}_i\int_0^\infty p^{1-\sigma}dF_{i,1}(p)\\ &= \bar{X}_i\theta\kappa_i\Upsilon_i\int_0^\infty p^{\theta-\sigma}\exp\left[-p^\theta\kappa_i\Upsilon_i\right]dp\\ &= \bar{X}_i\left(\kappa_i\Upsilon_i\right)^{\frac{\sigma-1}{\theta}}\Gamma\left(\frac{\theta+\sigma-1}{\theta}\right)\end{aligned}$$

Taking the ratio of the two terms and simplifying gives:

$$\begin{aligned}\pi_{ij} &= \frac{\mathbb{E}\left[p_{b_i}c_{b_i}|\mathbb{1}^{(1)}_{b_ij}=1\right]}{\mathbb{E}\left[p_{b_i}c_{b_i}\right]}\\ &= \frac{\lambda_{ij}\upsilon_{ij}}{\kappa_i\Upsilon_i}\end{aligned}$$

As in Eaton and Kortum (2002), trade shares are fully summarized by the probability that any supplier from j ends up serving market i. The reason is that, conditional on the identity of the seller, the distribution of prices offered to buyers in i is the same whatever the origin of the seller. In this context, trade shares only depend on the likelihood that a seller from j is the lowest-cost supplier met by a buyer from i. In our model, the probability depends on country j's comparative advantage in market i (υ_{ij}/Υ_i) and the relative size of frictions (λ_{ij}/κ_i).

As discussed in the text, the semi-elasticity of this trade share with respect to the bilateral search parameter is unambiguously positive:

$$\begin{aligned}\frac{d\ln\pi_{ij}}{d\lambda_{ij}} &= \frac{1}{\lambda_{ij}}-\frac{1}{\kappa_i}\frac{d\kappa_i}{d\lambda_{ij}}\\ &= \frac{1-\pi_{ij}}{\lambda_{ij}}>0\end{aligned}$$

OA.1.2. Proof of proposition 2

The sensitivity of export probabilities to search frictions can be assessed through the following derivative:

$$\begin{aligned}\frac{\partial \ln \rho_{ij}(z)}{\partial \lambda_{ij}} &= \underbrace{\frac{\partial \ln \lambda_{ij}}{\partial \lambda_{ij}}}_{\text{Visibility channel}} + \underbrace{\frac{\partial \ln e^{-(w_j d_{ij})^{\theta} z^{-\theta} \kappa_i \Upsilon_i}}{\partial \lambda_{ij}}}_{\text{Competition channel}} \\ &= \frac{1}{\lambda_{ij}} - (d_{ij} w_j)^{\theta} z^{-\theta} \Upsilon_i \frac{d\kappa_i}{d\lambda_{ij}} \\ &= \frac{1}{\lambda_{ij}} - T_j \underline{z}^{-\theta} \left(\frac{z}{\underline{z}}\right)^{-\theta}.\end{aligned}$$

Depending on the current level of frictions (λ_{ij}), the expected number of firms in country j ($T_j \underline{z}^{-\theta}$) and the position of the firm in the productivity distribution ($\left(\frac{z}{\underline{z}}\right)^{-\theta}$), the derivative can be positive or negative. It is more positive for high values of z. At the limit:

$$\lim_{z \to +\infty} \frac{\partial \ln \rho_{ij}(z)}{\partial \lambda_{ij}} = \frac{1}{\lambda_{ij}}.$$

Instead, low-productivity sellers' export probability is less sensitive to frictions and can even be negatively affected by a decrease in frictions. Namely, if the level of frictions is such that $\lambda_{ij} > 1/T_j \underline{z}^{-\theta}$, that is, if frictions are not too strong so that buyers in expectation meet with at least one seller from j, a strictly positive mass of firms exists whose export probability decreases when search frictions are reduced: $\partial \ln \rho_{ij}(\underline{z})/\partial \lambda_{ij} < 0$, where $\rho_{ij}(\underline{z})$ denotes the export probability of the least productive firm.

This non-monotonicity is to be compared with the sensitivity of export probabilities to iceberg trade costs, which is instead unambiguously negative, less so for more productive sellers:

$$\begin{aligned}\frac{\partial \ln \rho_{ij}(z)}{\partial d_{ij}} &= -(c_j d_{ij})^{\theta} z^{-\theta} \Upsilon_i \kappa_i \left[\frac{\theta}{d_{ij}} + \frac{\partial \ln \Upsilon_i}{\partial d_{ij}} + \frac{\partial \ln \kappa_i}{\partial d_{ij}}\right] \\ &= -\frac{\theta}{d_{ij}} (c_j d_{ij})^{\theta} z^{-\theta} \Upsilon_i \kappa_i (1 - \pi_{ij}) < 0.\end{aligned}$$

These contrasted results are the key reason search frictions and iceberg costs can be identified separately in firm-level export patterns in this model. Larger iceberg trade costs decrease the probability of serving any buyer in the destination, less so for more productive sellers. By contrast, more search frictions are more costly for high-productivity firms, in relative terms.

OA.1.3. Plugging the model into a general equilibrium structure

The model developed in paper describes the matching equilibrium in a sector whose production costs are taken as exogenous as in partial equilibrium. We now show how this structure can be plugged into a general equilibrium framework. In this general equilibrium structure, the assumptions in the main text describe the matching of consumers and producers within a particular sector k. Following Atkeson and Burstein (2008) and Gaubert and Itskhoki (2018), the aggregate impact of the discretedness at product level is neglected by assuming the economy displays a continuum of products. A consumer in country i consumes a CES bundle of products:

$$c_{b_i} = \left[\int_0^1 \left(c_{b_i}^k\right)^{\frac{\sigma-1}{\sigma}} dk\right]^{\frac{\sigma}{\sigma-1}}$$

with σ the elasticity of substitution between product-level consumptions. Consumption at product level follows the assumptions in of our model. Each consumer b_i meets with a random number of potential suppliers for variety k, chooses the lowest-cost supplier met and pays the price:

$$p_{b_i}^k = \arg\min\left\{\frac{w_j d_{ij}^k}{z_{s_j}}; s_j \in \Omega_{b_i}^k\right\},$$

where $\Omega_{b_i}^k$ now denotes the set of producers of variety k met by buyer b_i and d_{ij}^k is the product-specific iceberg cost. The price of the input bundle is instead assumed homogenous across products and we will now interpret it as the wage rate: Firms produce out of labor with a constant returns to scale technology and labor is perfectly mobile across sectors.

To solve the model, it is assumed that individual consumers maximize aggregate consumption based on the expected price index, i.e. they neglect the aggregate impact of the randomness in the matching process:

$$\begin{cases} \max_{\{c_{b_i}^k\}_{k\in[0,1]}} & \left[\int_0^1 \left(c_{b_i}^k\right)^{\frac{\sigma-1}{\sigma}} dk\right]^{\frac{\sigma}{\sigma-1}} \\ s.t. & \mathbb{E}\left[\int_0^1 p_{b_i}^k c_{b_i}^k dk\right] \leq \frac{R_i}{B_i} \end{cases}$$

R_i is the country's aggregate income that we assume is shared equally across buyers. In equilibrium $R_i = w_i B_i + \Pi_i$ where Π_i denotes aggregate profits, assumed to be distributed lump-sum to all consumers.

Under these assumptions, the demand addressed to the lowest cost supplier met by consumer b_i writes:

$$p_{b_i}^k c_{b_i}^k = \left(\frac{p_{b_i}^k}{P_i}\right)^{1-\sigma} \frac{R_i}{B_i}$$

where $P_i = \mathbb{E}\left[\int_0^1 \left(p_{b_i}^k\right)^{1-\sigma} dk\right]^{\frac{1}{1-\sigma}}$ is the expected price index in country i.

This demand function is consistent with the assumption in the main text under the notation $\bar{X}_i \equiv \frac{R_i}{B_i} P_i^{\sigma-1}$. Aggregating these demand functions across buyers within a product implies:

$$
\begin{aligned}
X_i^k &= B_i \mathbb{E}\left[p_{b_i}^k c_{b_i}^k\right] = B_i \bar{X}_i \left(\kappa_i^k \Upsilon_i^k\right)^{\frac{\sigma-1}{\theta}} \Gamma\left(\frac{\theta+\sigma-1}{\theta}\right) \\
\pi_{ij}^k &= \frac{X_{ij}^k}{X_i^k} = \frac{T_j^k \left(d_{ij}^k w_j\right)^{-\theta^k}}{\Upsilon_i^k} \frac{\lambda_{ij}^k}{\kappa_i^k}
\end{aligned}
$$

where the notations are the same as in the main text except that we explicitly introduce the product dimension that was neglected to alleviate notations.

The model is closed using the trade balance condition. For each pair of countries i and j, we have:

$$
\int_0^1 \pi_{ij}^k dk = \int_0^1 \pi_{ji}^k dk
$$

These conditions for all pairs of countries define a system of equations that can be used to solve for equilibrium factor prices $\{w_j\}_{j=1\ldots N}$. Solving the model numerically requires estimates for all parameters, most notably the whole vector of search frictions λ_{ij}^k, $\forall k \in [0,1]$, $i = 1, \ldots N$ and $j = 1\ldots N$. In the main text, we propose a strategy to estimate λ_{ij}^k from firm-to-firm trade data. Because we have access to data for French exporters only, we are unable to recover the whole set of parameters necessary to solve the model in general equilibrium.

The general equilibrium extension discussed in this Section however gives intuition for how search frictions affect welfare in general equilibrium. The gravity structure at product-level makes it possible to compare our theoretical framework with other trade models displaying structural gravity, most notably the multi-sector extension of Eaton and Kortum (2002) in Caliendo and Parro (2015). As discussed in Costinot and Rodriguez-Clare (2014), a gravity structure, together with other common micro and macro restrictions, defines a general class of trade models which welfare predictions can be summarized using a simple formula that solely involves trade shares and measures of trade elasticities. As long as the search frictions we introduced in the model are technological constraints that the planner faces, so that the planner's solution coincides with the decentralized equilibrium, the results in Costinot and Rodriguez-Clare (2014) are likely to apply in our model as in the class of multi-sector models they discuss. By distorting the geography of trade, search frictions will thus affect the welfare benefits from trade.

OA.1.4. Alternative market structure assumption

Results in the main text rely on the assumption that firms price at their marginal cost. The randomness induced by matching frictions however increases the market power of the lowest-cost supplier met by a particular buyer. Even if the firm has announced a price at its marginal cost, it has an incentive to deviate ex-post and exploit its market power. As we now show, the model is flexible enough to handle more realistic price strategies.

Suppose that firms in each buyer's random choicest compete à la Bertrand. Under Bertrand competition, the lowest-cost producer ends up setting the price which is just sufficient to beat the second lowest-cost supplier, unless this price is above the monopoly price.[2] As in the baseline discussed in the main text, each buyer ends up purchasing the product from the lowest-cost supplier she has met. The price that she pays is however equal to the marginal cost of the second lowest-cost supplier.

In this setting, bilateral trade probabilities are the same as in the baseline case. The value of trade, conditional on a match, is however a function of the distribution of the second lowest-cost supplier, through the lowest-cost supplier's pricing function. As demonstrated in Bernard et al. (2003), endogenous mark-ups do not distort the geography of trade, and thus trade shares are still equal to bilateral trade probabilities. The reason is that the distribution of markups is the same whatever the origin of the firm setting this markup: within a destination, no source sells at systematically higher markups. This result is summarized in Theorem 3.

Theorem 3. *Under Bertrand competition, the distribution of markups set to buyers in country i writes:*

$$H_i(m) = P[M_i \leq m] = 1 - m^{-\theta}$$

with M_i the ratio of the price set by the lowest cost supplier over its marginal cost, $P_i^{(2)}/P_i^{(1)}$ if we keep the notations used in Appendix OA.1.1.

2. In the set-up under study, the monopoly markup is $\frac{\sigma}{\sigma-1}$ for $\sigma > 1$.

Proof. The distribution of the second to the first lowest cost $P_i^{(2)}/P_i^{(1)}$, conditional on $P_i^{(2)} = p_2$ is:

$$
\begin{aligned}
\mathbb{P}\left[\frac{P_i^{(2)}}{P_i^{(1)}} \leq m | P_i^{(2)} = p_2\right] &= \mathbb{P}\left[P_i^{(1)} \geq \frac{p_2}{m} | P_i^{(2)} = p_2\right] \\
&= 1 - \mathbb{P}\left[P_i^{(1)} \leq \frac{p_2}{m} | P_i^{(2)} = p_2\right] \\
&= 1 - \int_0^{\frac{p_2}{m}} \frac{g_{i,1,2}(p, p_2)}{g_{i,2}(p_2)} dp \\
&= 1 - \left(\frac{\frac{p_2}{m}}{p_2}\right)^{\theta} \\
&= 1 - (m)^{-\theta}
\end{aligned}
$$

The unconditional distribution $H_i(m)$ comes immediately. □

Whereas the baseline model sticks to the assumption of marginal cost pricing, the result in Theorem 3 shows that the main conclusions would be left unchanged if we instead assumed Bertrand competition among the random choiceset of firms met by a particular buyer. The pricing assumption instead has consequences for the dynamics of trade adjustment to shocks, that is studied into more details in Fontaine et al. (2021).

OA.1.5. Increasing meeting probabilities

One may wonder whether imposing the same meeting probability to all firms, whatever their productivity, is a key driver of the result. An alternative approach would assume the meeting probability to be increasing in the firm's productivity. Such increasing relationship would for instance emerge under endogenous search effort. In a reduced-form set-up, this assumption would imply that the meeting probability for a firm of productivity z in country j that seeks to serve market i can be summarized by

$$\lambda_{ij}(z) = f(\lambda_{ij}, z)$$

with $df(\lambda_{ij}, z)/d\lambda_{ij} > 0$ and $df(\lambda_{ij}, z)/dz > 0$, i.e. high-productivity firms meet more buyers on average but more structural frictions reduce meeting probabilities at each point of the productivity distribution.

Under such assumption, the probability for a firm with productivity z_{s_j} to serve a buyer in i is still equal to the probability of a match times the probability of being the lowest cost supplier, conditional on this match. However, the cross-derivative of $\rho_{ij}(z_{s_j})$ with respect to the (exogenous) search parameter and the firm's productivity now takes a more complicated form:

$$\frac{d^2\rho_{ij}(z_{s_j})}{d\lambda_{ij}dz_{s_j}} = \left[\frac{\rho_{ij}(z_{s_j})}{\lambda_{ij}}\frac{d^2 f(\lambda_{ij}, z_{s_j})}{d\lambda_{ij}dz_{s_j}} + \frac{\rho_{ij}(z_{s_j})}{\mathbb{P}()}\frac{d^2\mathbb{P}\left(\min_{s'_k\in\Omega_{b_i}}\left\{\frac{w_k d_{ik}}{z_{s'_k}}\right\} = s_j\right)}{d\lambda_{ij}dz_{s_j}}\right].$$

As in the benchmark case, the second term is likely to be negative and increasing in z_{s_j}. The second derivative of the probability of serving the buyer conditional on a match with respect to λ_{ij} and z_{s_j} is expected larger than in the baseline, however. The reason is that a reduction in frictions implies the typical buyer in i meets with more sellers and the additional sellers met are more productive, on average. From this point of view, the competitive channel is even more distortive in this case. However, a reduction in frictions also affects the relative meeting probabilities at different points of the distribution; that is, $d^2 f(\lambda_{ij}, z_{s_j})/d\lambda_{ij}dz_{s_j}$ might no longer be zero. From this, it comes that the distortive impact of frictions is likely to show up in this model as well, whenever the cross derivative of the meeting probability with respect to λ_{ij} and z_{s_j} is not too negative.

Figure OA.1.1 illustrates the impact of varying the meeting probability under a specific parametric assumption, which can be compared with Figure 2 of the paper. Namely, we simulate the model assuming:

$$\lambda_{ij}(z_{s_j}) = 2\lambda_{ij}\left[1-\left(\frac{z_{s_j}}{\underline{z}}\right)^{-\theta}\right] \tag{1}$$

Under this parametric assumption, the expected meeting probability is equal to λ_{ij}, as in the baseline, but now varies between 0 and $2\lambda_{ij}$ along the productivity distribution. Assuming the meeting probability to be increasing in the seller's productivity mechanically increases the likelihood that a high-productivity seller will end up serving a foreign buyer. As a consequence, the probability of a match is larger at the top of the distribution in the extended model than in the benchmark case. Whereas the level probabilities are different in the baseline and extended models, the extended model still displays the log-supermodularity in λ_{ij} and z_{s_j}, which we exploit in the empirical Section. For this reason, we believe that our empirical strategy is not threatened by the strong parametric assumption imposed to recover analytical results.

OA.1.6. A model of heterogeneous consumers

In this Section, we discuss the predictions of another class of models which the trade literature often compares with Ricardian models à la Eaton and Kortum (2002) or the monopolistic competition structure in Melitz (2003), namely models displaying horizontal differentiation and heterogeneous consumers. As discussed in Anderson et al. (1992) and Head and Mayer (2014), discrete choice

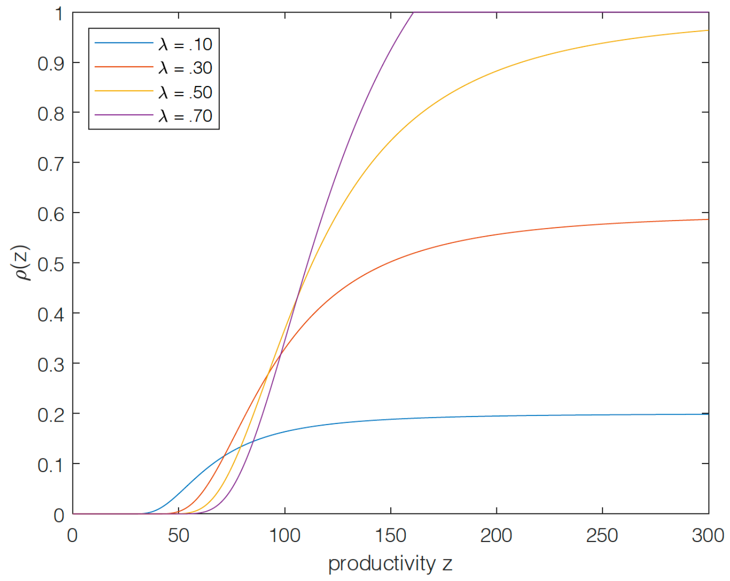


FIGURE OA.1.1. Probability of serving a buyer as a function of the seller's productivity under increasing meeting probabilities. This Figure illustrates how the probability of serving a buyer varies with the seller's productivity, for four different values of bilateral frictions.

models of demand for horizontally differentiated varieties can deliver a gravity structure under the adequate parametric assumptions. In this Section, we show how it is possible to interpret our predictions in the context of this class of models. We also show that the moment used to identify search frictions in the paper is likely to be orthogonal to the parameters of such a model.

Consider a model in which buyers display heterogeneous preferences with regard to the varieties produced by the (discrete) set of wordwide producers. More specifically, let us denote $\psi_{b_i s_j}$ a random variable characterizing the preference of buyer b_i with respect to variety s_j. Following the literature, we will assume that the preference parameters are independently drawn into a Pareto distribution of shape θ:

$$\mathbb{P}\left(\psi_{b_i s_j} \geq \psi\right) = \begin{cases} \left(\frac{\psi}{\underline{\psi}}\right)^{-\theta} & if \quad \psi \geq \underline{\psi} \\ 1 & if \quad \psi < \underline{\psi} \end{cases}$$

Let us further assume that consumers' preferences are systematically biased towards varieties produced in some countries. Namely, the number of varieties from j that deliver a preference parameter above ψ in country i is assumed to follow a Poisson distribution of parameter $T_j \lambda_{ij} \psi^{-\theta}$. T_j can be interpreted as the mean quality of varieties produced in j whereas λ_{ij} measures a dyadic preference bias.

Finally, suppose as in the paper's baseline model that the cost of serving market i from j is equal to $w_j d_{ij}$ and that the revenue of a consumer in country i is equal to $x_i \equiv R_i/B_i$. Under these assumptions, the indirect utility recovered from the consumption of variety s_j writes:[3]

$$V_{b_i s_j} = \frac{x_i}{w_j d_{ij}} \psi_{b_i s_j}$$

Using the properties of the Poisson distribution, it is straightforward to show that the number of varieties from j (resp. from any country) delivering an indirect utility above v is distributed Poisson of parameter $v^{-\theta} c_i^{\theta} T_j \lambda_{ij} (w_j d_{ij})^{-\theta}$ (resp. $v^{-\theta} c_i^{\theta} \sum_j T_j \lambda_{ij} (w_j d_{ij})^{-\theta}$).

In this model, the probability that a buyer b_i chooses a variety produced in country j, conditional on the variety delivering indirect utility above v, writes:

$$\mathbb{P}\left(V_{b_i}^{(1)} \; originates \; from \; j | V_{b_i}^{(1)} \geq v\right) = \frac{v^{-\theta} c_i^{\theta} \left(w_j d_{ij}\right)^{-\theta} T_j \lambda_{ij}}{v^{-\theta} c_i^{\theta} \sum_j \left(w_j d_{ij}\right)^{-\theta} T_j \lambda_{ij}}$$

As the probability is independent from v and homogenous across buyers, it is also equal to the probability that any buyer from i purchases the variety from j, which is also the trade share in this model in which the value of trade is homogenous across buyers in expectation:

$$\pi_{ij} = \frac{\left(w_j d_{ij}\right)^{-\theta} T_j \lambda_{ij}}{\sum_j \left(w_j d_{ij}\right)^{-\theta} T_j \lambda_{ij}} = \frac{T_j \left(w_j d_{ij}\right)^{-\theta}}{\Upsilon_i} \frac{\lambda_{ij}}{\kappa_i}$$

where $\Upsilon_i \equiv \sum_j \left(w_j d_{ij}\right)^{-\theta} T_j$ and $\kappa_i = \frac{\sum_j (w_j d_{ij})^{-\theta} T_j \lambda_{ij}}{\sum_j (w_j d_{ij})^{-\theta} T_j}$.

Based on bilateral trade shares, it is thus not possible to discriminate our model, that combines Ricardian comparative advantages with search frictions, and a model that displays (biased) heterogeneous preferences. As discussed in the main text though, this prediction of the model is not sufficient to identify search frictions anyway. Indeed, the geography of trade involves two dyadic components, d_{ij} and λ_{ij}, which cannot be separated in the data based on predicted and observed trade shares. What the discussion in this Section adds is that, even if we were able to control for iceberg costs, using this prediction of the model would not be desirable because the same structural equation arises from a completely different model in which λ_{ij} interprets as a preference parameter instead of a search friction.

3. As in the paper's model, it is implicitly assumed that the seller s_j serves buyer b_i at marginal cost. Assuming instead that the seller exploits her competitive advantage to price at the second lowest preference-adjusted marginal cost would complicate the analysis although results regarding the selection of firms into exporting would be left unaffected.

Our identification strategy does not exclusively relies on the gravity structure of the model though. Instead, we exploit the model's prediction regarding *individual* trade patterns. Namely, the moment used for identification is based on the expected number of firms from a given country that serve exactly M buyers in i:

$$h_{ij}(M) = \int C_{B_i}^{M} \rho_{ij}(s_j)^M (1 - \rho_{ij}(s_j))^{B_i - M} f(s_j) ds_j$$

where $f(s_j)$ is the pdf of the distribution of firms and $\rho_{ij}(s_j)$ is the probability that seller s_j serves any buyer in country i. Our empirical strategy uses the heterogeneity across sellers in their ability to reach foreign consumers, that is log-supermodular in firms' productivity and the level of search frictions.

In a model in which the only source of heterogeneity is consumers' preferences with respect to differentiated varieties, such heterogeneity does not exist and the moment used in the structural estimation becomes:

$$h_{ij}(M) = C_{B_i}^{M} \rho_{ij}^{M} (1 - \rho_{ij})^{B_i - M}$$

where

$$\rho_{ij} = \frac{T_j \left(w_j d_{ij}\right)^{-\theta}}{\Upsilon_i} \frac{\lambda_{ij}}{\kappa_i}$$

is homogenous across sellers and captures the likelihood that any seller from i is the prefered variety of any buyer from j.

OA.1.7. A model of buyer acquisition under monopolistic competition

Whereas our model is Ricardian in nature, an alternative interpretation of the buyer margin can be done in the context of an imperfect competition model *à la* Melitz (2003), as notably done by Bernard et al. (2018); Carballo et al. (2018). In this Section, we develop such a model using a structure and notations comparable to those used in our model to ease the comparison. The model introduces market penetration costs *à la* Arkolakis (2010) in the discrete version of the Melitz model proposed by Eaton et al. (2012). As in the paper's model, we abstract from any general equilibrium effects.

We start with the supply side structure used in our model, that features a discrete and random number of producers that are heterogeneous in their productivity. Remember that under our assumptions, borrowed from Eaton et al. (2012), the number of sellers from j that display a productivity above z is the realization of a Poisson variable with parameter $T_j z^{-\theta}$. Given exogenous input costs w_j and iceberg costs d_{ij} the number of firms serving market i at a cost below p is itself a Poisson variable of parameter $\mu_{ij}(p) = T_j \left(d_{ij} w_j / p\right)^{-\theta}$.

In the Ricardian framework, worldwide firms compete to serve market i with the same perfectly substituTable variety, which triggers prices towards marginal costs. In the monopolistic competition variant, we instead follow Eaton et al.

(2012), and assume that each seller offers a differentiated variety and faces a demand which is isoelastic. Equilibrium prices are then a constant mark-up over marginal costs:

$$p_{ij}(z_{s_j}) = \frac{\sigma}{\sigma - 1} \frac{d_{ij} w_j}{z_{s_j}}$$

$p_{ij}(z_{s_j})$ is the price set by s_j in country i, which is uniform across buyers within a destination if the residual demand elasticity is itself homogeneous. We assume this is the case and denote $\sigma > 1$ this elasticity.

In Eaton et al. (2012), sellers face a representative consumer in each market i and decide whether to serve the market or not, depending on the size of some fixed export cost F_{ij}. To introduce the buyer margin, we instead assume that i) sellers can serve a discrete number B_i of homogeneous buyers in the destination, that all display an iso-elastic demand function as in our model, and ii) the fixed cost of exporting is increasing in the number of buyers served:

$$F_{ij}\left(B_{ij}(z_{s_j})\right) = F_{ij} \times \frac{1 - \left(1 - \frac{B_{ij}(z_{s_j})}{B_i}\right)^{1-1/\lambda}}{1 - 1/\lambda}$$

where F_{ij} is a positive parameter, $B_{ij}(z_{s_j})/B_i$ measures the share of the market that seller s_j chooses to serve and $\lambda > 0$ measures the increasing cost of reaching a larger fraction of potential buyers.

Solving for the seller's optimal number of buyers served implies:

$$\frac{B_{ij}(z_{s_j})}{B_i} = Max\left\{0; 1 - \left(\frac{p_{ij}(z_{s_j})^{1-\sigma}}{\sigma} \frac{B_i \bar{X}_i}{F_{ij}}\right)^{-\lambda}\right\}$$

From this, it comes:

$$\begin{aligned}
\frac{\partial \ln B_{ij}(z_{s_j})}{\partial z_{s_j}} &= \lambda \left[1 - \frac{B_{ij}(z_{s_j})}{B_i}\right] \frac{\sigma - 1}{z_{s_j}} > 0 \\
\frac{\partial \ln B_{ij}(z_{s_j})}{\partial d_{ij}} &= \lambda \left[1 - \frac{B_{ij}(z_{s_j})}{B_i}\right] \frac{1 - \sigma}{d_{ij}} < 0 \\
\frac{\partial \ln B_{ij}(z_{s_j})}{\partial F_{ij}} &= \lambda \left[1 - \frac{B_{ij}(z_{s_j})}{B_i}\right] \frac{-1}{F_{ij}} < 0
\end{aligned}$$

In this model, the buyer margin is decreasing in both iceberg and fixed export costs, especially at the bottom of the productivity distribution. This feature of the model is illustrated in Figure OA.1.2, which reproduces Figure 2 of the paper in the context of the alternative model just discussed. As in the baseline model, the probability of serving a buyer is increasing in the seller's productivity. Reducing trade frictions, whether the fixed or the variable component of trade costs, however increases the export probability at every point of the productivity distribution. This is in contrast with our model which displays a non-linear impact of moving from a high to a low level of frictions.

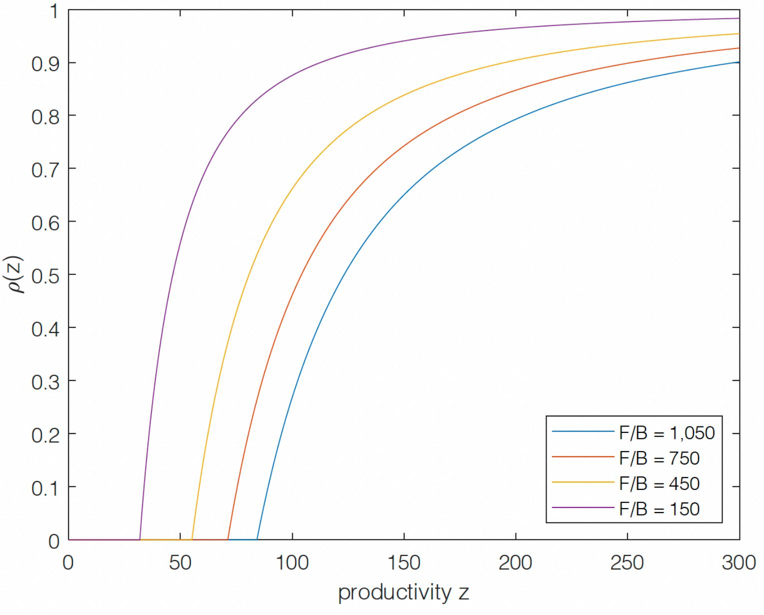


FIGURE OA.1.2. Probability of serving a buyer as a function of the seller's productivity in the model of buyer acquisition. This Figure illustrates how the probability of serving a buyer varies with the seller's productivity, for four different values of the fixed cost per buyer of serving market i.

In this model, the number of sellers choosing to serve exactly $M \leq B_i$ buyers is equal to the number of sellers which productivity satisfies the following conditions:

$$\begin{aligned} & \underline{z}(M) \leq z < \underline{z}(M+1) \\ \textit{where} \quad & \underline{z}(M) \equiv \left(\frac{B_i - M}{B_i}\right)^{\frac{-1}{\lambda(\sigma-1)}} A_{ij} \\ \textit{and} \quad & A_{ij} \equiv \frac{\sigma}{\sigma-1} w_j d_{ij} \left[\sigma \frac{F_{ij}}{B_i \bar{X}_i}\right]^{\frac{1}{\sigma-1}} \end{aligned}$$

The moment used to identify search frictions would thus capture the variance of the following ratios:

$$\begin{aligned} \frac{h_{ij}(M)}{h_{ij}(1)} &= \frac{\underline{z}(M)^{-\theta} - \underline{z}(M+1)^{-\theta}}{\underline{z}(1)^{-\theta} - \underline{z}(2)^{-\theta}} \\ &= \frac{(B_i - M)^{\frac{\theta}{\lambda(\sigma-1)}} - (B_i - M - 1)^{\frac{\theta}{\lambda(\sigma-1)}}}{(B_i - 1)^{\frac{\theta}{\lambda(\sigma-1)}} - (B_i - 2)^{\frac{\theta}{\lambda(\sigma-1)}}} \end{aligned}$$

OA.1.8. Two-sided heterogeneity

The model derived in the previous Section can further be extended to handle two-sided heterogeneity, as in Bernard et al. (2018). Suppose that the supply-side of the previous model is left unchanged but buyers in each destination are now heterogeneous in terms of their average demand:

$$c_{b_i}(p_{b_i}, \bar{X}_{b_i}) = p_{b_i}^{-\sigma} \bar{X}_{b_i}$$

In Bernard et al. (2018), the heterogeneity comes from buyers combining inputs into a final good sold to final consumers. In their setting, the heterogeneity in demand ultimately comes from a random productivity component that affects the demand addressed to buyers, and in turn their network of suppliers. For the purpose of the appendix, it will be sufficient to assume that buyers are born with a random demand level, drawn from a Pareto distribution with shape Γ. With a discrete number of such buyers, the number of buyers that draw a $\bar{X}_{b_i}$ above X is distributed Poisson of parameter $B_i\,(X/X_L)^{-\Gamma}$.

Following Bernard et al. (2018), sellers are assumed to decide whether to serve a buyer or not, given a fixed cost per buyer f_{ij}. Under these assumptions, a seller with productivity z chooses to serve all buyers which demand is sufficiently high to cover the fixed cost per buyer. At the margin:

$$\begin{aligned} & \pi_{ij}(z, \underline{\bar{X}}(z)) = 0 \\ \Leftrightarrow \quad & \underline{\bar{X}}(z) = \underbrace{f_{ij}\sigma\left(\frac{\sigma}{\sigma-1}d_{ij}w_j\right)^{\sigma-1}}_{F_{ij}} z^{1-\sigma} \end{aligned}$$

As discussed in Bernard et al. (2018), the model thus displays negative assortative matching: High productivity sellers can afford serving relatively small buyers. As a consequence, high-productivity sellers are also those that serve more buyers, consistent with the data. The assortative matching also implies that the relative share of sellers at different points of the distribution of outdegrees reflects the shape of the Pareto distributions that parametrize the heterogeneity of sellers and buyers.

In this setting, the unconditional probability that a particular buyer is served by a seller of productivity z is the probability that this buyer's demand parameter is above the seller's threshold:

$$\begin{aligned} \rho_{ij}(z) &= Pr\left[\bar{X}_{b_i} \geq \underline{\bar{X}}(z)\right] \\ &= Pr\left[\bar{X}_{b_i} \geq F_{ij} z^{1-\sigma}\right] \\ &= \left(\frac{F_{ij} z^{1-\sigma}}{X_L}\right)^{-\Gamma} \end{aligned}$$

The probability is depicted in Figure OA.1.3, which reproduces Figure 2 of the paper in the context of the model just discussed. Again, the probability of serving a buyer is increasing in the seller's productivity, consistent with the data. Here as well, and contrary to our model, reducing trade frictions, whether the fixed or the variable component of trade costs, increases the export probability whatever the firm's productivity.

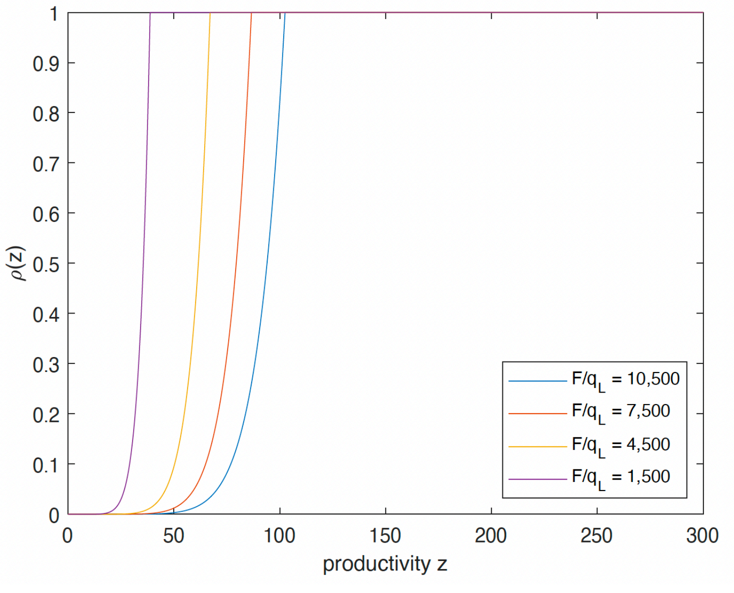


FIGURE OA.1.3. Probability of serving a buyer as a function of the seller's productivity under two-sided heterogeneity. This Figure illustrates how the probability of serving a buyer varies with the seller's productivity, for four different values of F_{ij}/q_L.

Finally, the number of sellers choosing to serve exactly $M \leq B_i$ buyers is equal to the number of sellers which productivity satisfies the following conditions:

$$\underline{z}(M) \leq z < \underline{z}(M+1)$$

$$\textit{where} \quad \underline{z}(M) \equiv \left(\frac{M}{B_i}\right)^{\frac{1}{\Gamma(\sigma-1)}} A_{ij}$$

$$\textit{and} \quad A_{ij} \equiv \frac{\sigma}{\sigma-1} w_j d_{ij} \left[\sigma \frac{f_{ij}}{X_L}\right]^{\frac{1}{\sigma-1}}$$

The moment used to identify search frictions would thus capture the variance of the following ratios:

$$
\begin{aligned}
\frac{h_{ij}(M)}{h_{ij}(1)} &= \frac{\underline{z}(M)^{-\theta} - \underline{z}(M+1)^{-\theta}}{\underline{z}(1)^{-\theta} - \underline{z}(2)^{-\theta}} \\
&= \frac{M^{\frac{-\theta}{\Gamma(\sigma-1)}} - (M+1)^{\frac{-\theta}{\Gamma(\sigma-1)}}}{1 - 2^{\frac{-\theta}{\Gamma(\sigma-1)}}}
\end{aligned}
$$

OA.2. Online appendix: Details on the empirical strategy

OA.2.1. Expected number of firms serving M buyers

Integrating the probability of having exactly M buyers along the distribution of productivities gives the expected number of firms from j with exactly M buyers in i:

$$
h_{ij}(M) = \int_{z_{min}}^{+\infty} C_{B_i}^{M} \rho_{ij}(z)^M (1 - \rho_{ij}(z))^{B_i - M} d\mu_j^Z(z).
$$

Using the following change of variable,

$$
\rho_{ij}(z) = \lambda_{ij} e^{-\frac{\lambda_{ij}}{\pi_{ij}} T_j z^{-\theta}},
$$

one can show that

$$
h_{ij}(M) = \frac{\pi_{ij}}{\lambda_{ij}} C_{B_i}^{M} \int_{\rho_{ij}(\underline{z})}^{\lambda_{ij}} \rho_{ij}(z)^{M-1} (1 - \rho_{ij}(z))^{B_i - M} d\rho_{ij}(z),
$$

where $\rho_{ij}(\underline{z})$ is the probability of the least productive firm in j serving a buyer in i.

If we assume $M > 0$, we can recognize a function of the family of the Beta function:

$$
h_{ij}(M) = \frac{\pi_{ij}}{\lambda_{ij}} C_{B_i}^{M} \left(B(\lambda_{ij}, M, B_i - M + 1) - B(\rho_{ij}(\underline{z}), M, B_i - M + 1) \right),
$$

with $B(\lambda_{ij}, M, B_i - M + 1) = \int_0^{\lambda_{ij}} \rho_{ij}(z)^{M-1} (1 - \rho_{ij}(z))^{B_i - M} d\rho_{ij}(z)$ being the incomplete beta function.

Using properties of the Beta function, notice that

$$
\begin{aligned}
B(M, B_i - M + 1) &= \frac{\Gamma(M)\Gamma(B_i - M + 1)}{\Gamma(M + B_i - M + 1)} = \frac{\Gamma(M)\Gamma(B_i - M + 1)}{\Gamma(B_i + 1)} \\
&= \frac{(M-1)!(B_i - M)!}{B_i!} = \frac{1}{M}\frac{(M)!(B_i - M)!}{B_i!} \\
&= \frac{1}{M}\frac{1}{C_{B_i}^M}.
\end{aligned}
$$

Then, the regularized incomplete beta function is

$$
I_{\lambda_{ij}}(M, B_i - M + 1) = \frac{B(\lambda_{ij}, M, B_i - M + 1)}{B(M, B_i - M + 1)} = B(\lambda_{ij}, M, B_i - M + 1) C_{B_i}^M M.
$$

Now, we can rewrite the expression for the mass of suppliers from j with M buyers in i with the help of the regularized incomplete beta function:

$$
h_{ij}(M) = \frac{\pi_{ij}}{\lambda_{ij}}\frac{1}{M}\left(I_{\lambda_{ij}}(M, B_i - M + 1) - I_{\rho_{ij}(\underline{z})}(M, B_i - M + 1)\right).
$$

Finally, note that if $\rho_{ij}(\underline{z})$ goes to 0, $I_{\rho_{ij}(\underline{z})}(M, B_i - M + 1)$ goes to 0 as well:

$$
\lim_{\rho_{ij}(\underline{z})\to 0} I_{\rho_{ij}(z_{min})}(M, B_i - M + 1) = \lim_{\rho_{ij}(\underline{z})\to 0}\int_0^{\rho_{ij}(\underline{z})} \rho_{ij}(z)^{M-1}(1 - \rho_{ij}(z))^{B_i - M} d\rho_{ij}(z) = 0.
$$

Using this property, one gets:

$$
h_{ij}(M) = \frac{\pi_{ij}}{\lambda_{ij}}\frac{1}{M} I_{\lambda_{ij}}(M, B_i - M + 1).
$$

OA.2.2. Choice of the empirical moment

Once normalized by the expected number of firms in the market $(T^k \underline{z}^{-\theta})$ to recover a convergent moment, the equation of $h_{ij}(M)$ can be used to estimate search frictions. Empirically, this moment however varies with distance, which potentially reflects the impact of other physical trade barriers on a firm's customer base in a destination. This sensitivity is illustrated in Table OA.2.1, which shows the correlation between various transformations of the empirical moment and distance from France, used as a proxy for iceberg trade costs.[4] The correlation between the number of firms with exactly M buyers in a destination and distance to the destination is negative and strongly significant. This finding is consistent with evidence that French sellers tend to serve fewer partners, if

4. For practical reasons detailed in the text, we restrict our attention to four values for $h_i^k(M)$, corresponding to the bottom of the distribution of sellers' degrees.

any at all, in more distant countries. This result should be expected from the model, as the π_i^k component entering $h_{ij}(M)$ is negatively correlated with iceberg trade costs d_i^k, which are likely to be increasing in distance. In principle, the correlation can be controlled for using readily available data for those trade shares.

Another option is to normalize the expected number of firms with M buyers with the destination-specific proportion of sellers with one buyer, i.e. compute the theoretical moment $h_i^k(M)/h_i^k(1)$ and compare it with its empirical counterpart. In theory, this convergent moment is useful to identify search frictions as it varies monotonically with λ_i^k (see Figure OA.3.6). Moreover, several ratios can be combined to identify precisely search frictions along a wide range of possible values. In the data, the corresponding empirical moments are still correlated with distance, which the model does not explain (see the second panel of Table OA.2.1). In principle, the normalization should neutralize the impact of trade shares, and thus of iceberg trade costs. A correlation between search frictions and distance may explain this result. However, iceberg trade costs may also affect the ratios through other channels, which the model does not encompass but the data reveal. To prevent such correlation from polluting our estimates of search frictions, we use an alternative moment that is not correlated with distance to France and is thus more likely to help us extract from the data information on pure search frictions.

Table OA.2.1. Correlation between various empirical moments and distance from France.

	log Distance	Std Dev.	Adjusted R-squared
Dependent Variable			
# sellers with:			
1 buyer	-15.92***	((1.51)	.698
2 buyers	-5.89***	(.575)	.535
3 buyers	-3.24***	(.374)	.417
4 buyers	-2.00***	(.261)	.334
# sellers (in relative terms with respect to the sellers with 1 buyer) with:			
2 buyers	.021**	(.009)	.339
3-4 buyers	-.027***	(.008)	.372
5+ buyers	-.121***	(.021)	.408
Variance of the relative shares of sellers:			
across M	.002	(.011)	.210
	Common language	Std Dev.	Adjusted R-squared
Dependent Variable			
# sellers with:			
1 buyer	45.37***	(5.258)	.682
2 buyers	17.05***	(2.083)	.518
3 buyers	9.85***	(1.526)	.417
4 buyers	6.11***	(1.045)	.334
# sellers (in relative terms with respect to the sellers with 1 buyer) with:			
2 buyers	-.08**	(.033)	.339
3-4 buyers	.07***	(.036)	.372
5+ buyers	.20***	(.033)	.401
Variance of the relative shares of sellers:			
across M	-.062**	(.028)	.210

Notes: Robust standard errors, clustered at the country level, are in parentheses, with ***, **, and * respectively denoting significance at the 1%, 5% and 10% levels. The last regression uses as right-hand-side variables the (log of) distance from France and the probability of citizens in France and the destination speaking the same language.

The moment chosen exploits information on the *dispersion* in the number of buyers served by sellers serving the same destination with the same product. Namely, the theoretical moment is defined as the variance in the $h_{ij}(M)/h_{ij}(1)$ ratios:

$$Var_i^k\left(\lambda_i^k\right) = \frac{1}{B_i^k - 1} \sum_{M=2}^{B_i^k} \left(\frac{h_i^k(M)}{h_i^k(1)} - \frac{1}{B_i^k - 1} \sum_{M=2}^{B_i^k} \frac{h_i^k(M)}{h_i^k(1)} \right)^2 . \tag{2}$$

This moment is also correlated positively with λ_i^k. As shown in the third panel of Table OA.2.1, the empirical counterpart of this moment is not correlated with distance. On the other hand, there is a significant correlation with the probability of citizens speaking the same language that can be interpreted as a proxy for frictions.

OA.2.3. Distribution of the Auxiliary Parameter

We work with the following convergent moments as auxiliary parameters:

$$\theta_{ij}(\lambda_{ij}, M) = \frac{h_{ij}(M)}{\sum\limits_{M=0}^{B_i} h_{ij}(M)} = \frac{1}{M} \frac{I_{\lambda_{ij}}(M, B_i - M + 1)}{\int_0^{\lambda_{ij}} \frac{(1-\rho_{ij}(z))^{B_i}}{\rho_{ij}(z)} d\rho_{ij}(z) + \sum\limits_{M=1}^{B_i} \frac{1}{M} I_{\lambda_{ij}}(M, B_i - M + 1)}, \tag{3}$$

that is, the proportion of firms from j having exactly M buyers in destination i. We first show the empirical counterparts of these auxiliary parameters are normally distributed. Then, we apply the delta method to work with the moment we chose to identify λ_{ij}. Finally, we discuss the asymptotic distribution of our estimator of λ_{ij}.

In line with our theoretical framework, we note $\left[\mathbb{1}\{B_{ij}(z_{s_j}) = M\}\right]_{s_j \in S_j}$, the vector of dummy variables that equal 1 whenever a firm in the sample has exactly M buyers in country i. The vector is of size S_j, the number of observations in the sample under consideration. The dummies are independent and identically distributed random variables of mean $\theta_{ij}(\lambda_{ij}, M)$ and of variance $\sigma_{ij}^2(M)$. This is true for all $M \in [0, B_i]$.[5]

The central limit theorem implies

$$\sqrt{S_j}\left(\hat{\theta}_{ij} - \theta_{ij}(\lambda_{ij})\right) \underset{S_j \to +\infty}{\xrightarrow{\mathcal{D}}} \mathcal{N}_B(0, \Sigma_{ij}), \tag{4}$$

5. Independence comes from the fact that sellers are independent from each other. Note this assumption could be relaxed because we could eventually use a version of the central limit theoreim based on weak dependence conditions. They are identically distributed ex ante as sellers draw their productivity in the same distribution and face the same degree of search frictions.

where

$$\hat{\theta}_{ij} = \begin{pmatrix} \frac{\sum_{s_j=1}^{S_j} \mathbb{1}\{B_{ij}(z_{s_j})=1\}}{S_j} \\ \frac{\sum_{s_j=1}^{S_j} \mathbb{1}\{B_{ij}(z_{s_j})=2\}}{S_j} \\ \dots \\ \frac{\sum_{s_j=1}^{S_j} \mathbb{1}\{B_{ij}(z_{s_j})=B_i\}}{S_j} \end{pmatrix} \quad \text{and} \quad \theta_{ij}(\lambda_{ij}) = \begin{pmatrix} \frac{h_{ij}(1)}{\sum_{M=0}^{B_j} h_{ij}(M)} \\ \frac{h_{ij}(2)}{\sum_{M=0}^{B_i} h_{ij}(M)} \\ \dots \\ \frac{h_{ij}(B_i)}{\sum_{M=0}^{B_i} h_{ij}(M)} \end{pmatrix}$$

respectively denote the vector of empirical and auxiliary parameters and Σ_{ij} is the variance-covariance matrix of the B_i random variables $\mathbb{1}\{B_{ij}(z_{s_j}) = M\}$, for $M \in \{1..., B_i\}$.

We then consider the function

$$\begin{array}{cccc} g & : & \mathbb{R}^{B_i} & \mapsto & \mathbb{R} \\ & & \begin{pmatrix} \theta_{ij}(\lambda_{ij},1) \\ \theta_{ij}(\lambda_{ij},2) \\ \dots \\ \theta_{ij}(\lambda_{ij},B_i) \end{pmatrix} & \to & Var\left(m_1 = \frac{\theta_{ij}(\lambda_{ij},2)}{\theta_{ij}(\lambda_{ij},1)}, m_2 = \frac{\sum_{M=3}^{6} \theta_{ij}(\lambda_{ij},M)}{\theta_{ij}(\lambda_{ij},1)}, m_3 = \frac{\sum_{M=7}^{B_i} \theta_{ij}(\lambda_{ij},M)}{\theta_{ij}(\lambda_{ij},1)} \right) \end{array}$$

where $Var(.)$ is the variance operator. g is derivable and verifies the property $\nabla g(\theta_{ij}(\lambda_{ij})) \neq 0$. Using the delta method, one can show an estimate of λ_{ij} based on $g(.)$ is asymptotically normal:

$$\sqrt{S_j}[g(\hat{\theta}_{ij}) - g(\theta_{ij}(\lambda_{ij}))] \underset{S_j \to +\infty}{\xrightarrow{\mathcal{D}}} \mathcal{N}\left(\begin{pmatrix} 0 \end{pmatrix}, \Omega(\theta_{ij}(\lambda_{ij})) = \nabla' g(\theta_{ij}(\lambda_{ij}))\Sigma_{ij}\nabla g(\theta_{ij}(\lambda_{ij}))\right) \tag{5}$$

where $\nabla g(\theta_{ij}(\lambda_{ij}))$ is of dimension $[B_i, 1]$ and is defined as

$$\begin{pmatrix} \frac{\partial g}{\partial \theta_{ij}(\lambda_{ij},1)} = -\frac{2}{3}\sum_{p=1}^{3} \frac{(m_p - \bar{m})m_p}{\theta_{ij}(\lambda_{ij},1)} \\ \frac{\partial g}{\partial \theta_{ij}(\lambda_{ij},2)} = \frac{2}{3}\frac{m_1 - \bar{m}}{\theta_{ij}(\lambda_{ij},1)} \\ \frac{\partial g}{\partial \theta_{ij}(\lambda_{ij},3)} = \frac{2}{3}\frac{m_2 - \bar{m}}{\theta_{ij}(\lambda_{ij},1)} \\ \dots \\ \frac{\partial g}{\partial \theta_{ij}(\lambda_{ij},6)} = \frac{2}{3}\frac{m_2 - \bar{m}}{\theta_{ij}(\lambda_{ij},1)} \\ \frac{\partial g}{\partial \theta_{ij}(\lambda_{ij},7)} = \frac{2}{3}\frac{m_3 - \bar{m}}{\theta_{ij}(\lambda_{ij},1)} \\ \dots \\ \frac{\partial g}{\partial \theta_{ij}(\lambda_{ij},B_i)} = \frac{2}{3}\frac{m_3 - \bar{m}}{\theta_{ij}(\lambda_{ij},1)} \end{pmatrix}$$

with $\bar{m} = \frac{1}{3}\sum_{p=1}^{3} m_p$.

In practice, our estimation is implemented in two steps. First, we use an estimation of the $\Omega(\widehat{\theta}_{ij})$ weight matrix using our observations $\nabla g(\hat{\theta}_{ij})$ and $\widehat{\Sigma_{ij}}$. Second, with the $\hat{\lambda}_{ij}$ estimated in the first step, we re-run our estimation with $\Omega(\theta(\hat{\lambda}_{ij}))$.

As proved in Gouriéroux et al. (1985), the variance of the GMM estimator of λ_{ij} is

$$\Sigma_{\lambda_{ij}} = \left[\frac{\partial g(\theta_{ij}(\lambda_{ij}))}{\partial \lambda_{ij}} \Omega(\theta_{ij}(\lambda_{ij}))^{-1} \frac{\partial g(\theta_{ij}(\lambda_{ij}))}{\partial \lambda_{ij}}\right]^{-1}$$

with

$$\begin{aligned}
\frac{\partial g(\theta_{ij}(\lambda_{ij}))}{\partial \lambda_{ij}} \quad = \quad & \tfrac{2}{3}(m_1 - \bar{m})\frac{\partial \theta_{ij}(\lambda_{ij}, 2)/\theta_{ij}(\lambda_{ij}, 1)}{\partial \lambda_{ij}} \\
& + \tfrac{2}{3}(m_2 - \bar{m})\textstyle\sum_{M=3}^{6} \dfrac{\partial \theta_{ij}(\lambda_{ij}, M)/\theta_{ij}(\lambda_{ij}, 1)}{\partial \lambda_{ij}} \\
& + \tfrac{2}{3}(m_3 - \bar{m})\textstyle\sum_{M=7}^{B_i} \dfrac{\partial \theta_{ij}(\lambda_{ij}, M)/\theta_{ij}(\lambda_{ij}, 1)}{\partial \lambda_{ij}}
\end{aligned} \quad .$$

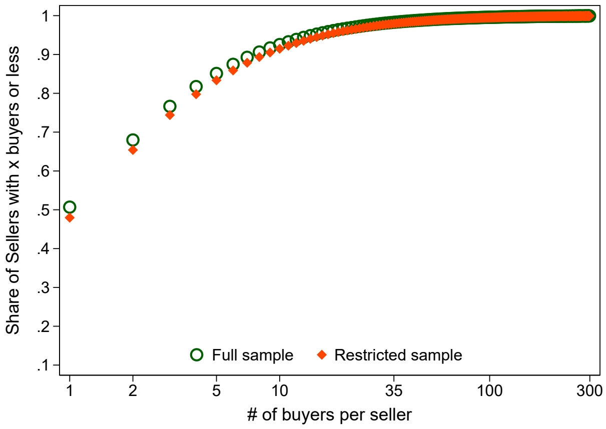


FIGURE OA.3.4. Number of buyers per seller, full and restricted sample. This Figure compares the number of buyers per seller, in the whole dataset and in the estimation dataset, restricted to the 70% of exporters that declare the product category of their exports ("Restricted sample").

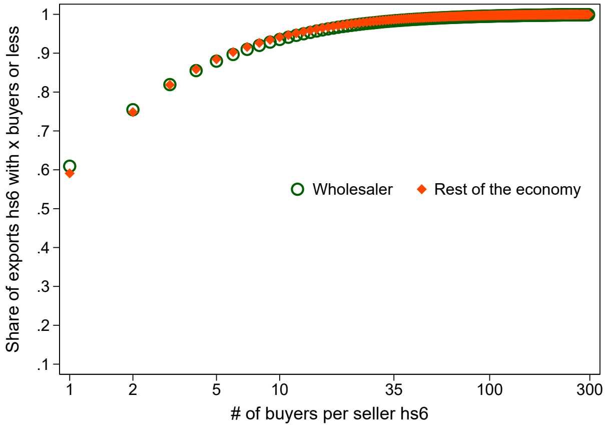


FIGURE OA.3.5. Number of buyers per seller, Wholesalers versus the rest of the economy. This Figure compares the number of buyers per seller, in the wholesaler sector and in the rest of the economy.

OA.3. Online appendix: Additional results

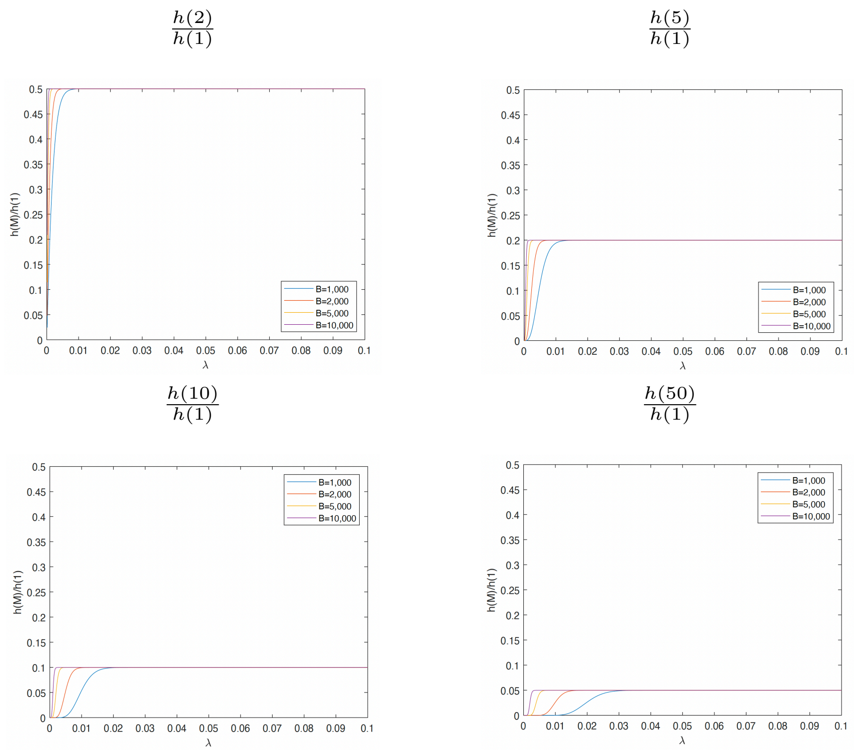


FIGURE OA.3.6. Identification power of the theoretical moments. This Figure shows the theoretical relationship between the underlying value of search frictions (λ, x-axis) and the share of firms with M buyers in the destination, in relative terms with respect to the expected number of firms with one buyer ($h(M)/h(1)$, y-axis). The relationship is derived conditional on the underlying number of buyers (B) and for various values of M, using the formula in equation (7) of the paper.

References

Anderson, S. P., de Palma, A., and Thisse, J.-F. (1992). *Discrete Choice Theory of Product Differentiation*. MIT Press.

Arkolakis, C. (2010). Market Penetration Costs and the New Consumers Margin in International Trade. *Journal of Political Economy*, 118(6):1151 – 1199.

Atkeson, A. and Burstein, A. (2008). Pricing-to-Market, Trade Costs, and International Relative Prices. *American Economic Review*, 98(5):1998–2031.

Bernard, A. B., Eaton, J., Jensen, J. B., and Kortum, S. (2003). Plants and productivity in international trade. *American Economic Review*, 93(4):1268–1290.

Bernard, A. B., Moxnes, A., and Ulltveit-Moe, K. H. (2018). Two-sided heterogeneity and trade. *Review of Economics and Statistics*, 100(3):424–439.

Table OA.3.2. French sellers and EU buyers, 2007

	Number of			Number of		
	Exporters	Importers	Pairs	Exporter-HS6	Importer-HS6	Triplets
	(1)	(2)	(3)	(4)	(5)	(6)
Overall	44,280	572,553	1,260,255	184,700	2,396,511	2,887,919
Austria	8,205	14,024	28,122	21,440	53,156	61,824
Belgium	29,485	71,173	213,952	97,520	379,141	482,851
Bulgaria	2,294	2,288	3,659	5,785	6,923	7,672
Cyprus	2,362	1,628	3,737	7,297	8,391	10,093
Czech Republic	6,848	6,118	13,198	16,572	21,516	25,221
Denmark	8,358	8,850	20,872	21,189	37,706	46,891
Estonia	1,803	1,235	2,495	5,247	5,525	6,411
Finland	5,258	5,173	11,599	13,758	22,235	26,391
Germany	24,651	117,942	236,574	73,889	392,438	464,008
Greece	7,793	11,261	25,415	26,120	55,664	68,628
Hungary	5,376	4,439	9556	12,957	16,347	18,719
Ireland	6,355	6,670	16,270	17,975	38,239	49,422
Italy	20,127	95,914	183,376	63,617	377,298	44,0634
Latvia	2,063	1,355	2,948	5,924	6,089	7,467
Lithuania	2,914	1,854	4,699	7,247	7,321	9906
Luxembourg	10,730	7,646	28,554	31,378	54,913	70,212
Malta	1,782	930	2,553	4,730	4,738	5,806
Netherlands	16,443	33,650	69,871	43624	131,986	158,602
Poland	9,734	12,860	30,232	24,736	43,577	52,732
Portugal	11,648	19,678	42,926	35,110	95,555	113,705
Romania	5,037	4,855	9,503	12,547	16,491	18,475
Slovakia	3,272	2,306	5,003	7,358	8,095	9,418
Slovenia	2,841	2,227	4,387	7,560	8,687	9,834
Spain	21,637	77,603	159,674	70,582	361,334	421,691
Sweden	7,684	10,220	20,426	20,274	39,616	45,815
UK	18,898	50,654	110,654	55,404	203,530	255,491

Notes: This Table gives the number of exporters, importers, exporter-importer pairs, exporter-HS6 product pairs, importer-HS6 product pairs, and importer-exporter-HS6 products triplets involved in a given bilateral trade flow. The data are for 2007 and are restricted to transactions with recorded CN8-products.

Caliendo, L. and Parro, F. (2015). Estimates of the Trade and Welfare Effects of NAFTA. *Review of Economic Studies*, 82(1):1–44.

Carballo, J., Ottaviano, G., and Volpe Martincus, C. (2018). The Buyer Margins of Firms' Exports. *Journal of International Economics*, 112(1):33–49.

Costinot, A. and Rodriguez-Clare, A. (2014). Trade Theory with Numbers: Quantifying the Consequences of Globalization. In Gopinath, G., Helpman, ., and Rogoff, K., editors, *Handbook of International Economics*, volume 4 of *Handbook of International Economics*, chapter 0, pages 197–261. Elsevier.

Eaton, J. and Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779.

Eaton, J. and Kortum, S. (2010). *Technology in the Global Economy: A Framework for Quantitative Analysis*.

Eaton, J., Kortum, S. S., and Sotelo, S. (2012). *International trade: Linking micro and macro*, volume II of *Advances in Economics and Econometrics: Tenth World Congress*. Applied economics edition.

Fontaine, F., Martin, J., and Mejean, I. (2021). Frictions and adjustments in firm-to-firm trade.

Table OA.3.3. Number of buyers per seller across destination countries

	Mean (1)	Median (2)	p75 (3)	Sh. with 1 buyer (4)
Austria	2.9	1	2	63%
Belgium	5.0	2	4	50%
Bulgaria	1.3	1	1	81%
Cyprus	1.4	1	1	80%
Czech Republic	1.5	1	1	76%
Denmark	2.2	1	2	66%
Estonia	1.2	1	1	85%
Finland	1.9	1	2	70%
Germany	6.3	2	4	50%
Greece	2.6	1	2	64%
Hungary	1.4	1	1	78%
Ireland	2.7	1	2	63%
Italy	6.9	1	3	53%
Latvia	1.3	1	1	84%
Lithuania	1.4	1	1	79%
Luxembourg	2.2	1	2	66%
Malta	1.2	1	1	86%
Netherlands	3.6	1	2	59%
Poland	2.1	1	2	67%
Portugal	3.2	1	2	62%
Romania	1.5	1	1	77%
Slovenia	1.3	1	1	82%
Slovakia	1.3	1	1	84%
Spain	6.0	1	3	54%
Sweden	2.3	1	2	67%
United Kingdom	4.6	1	3	53%
Across countries	15.6	3	10	32%

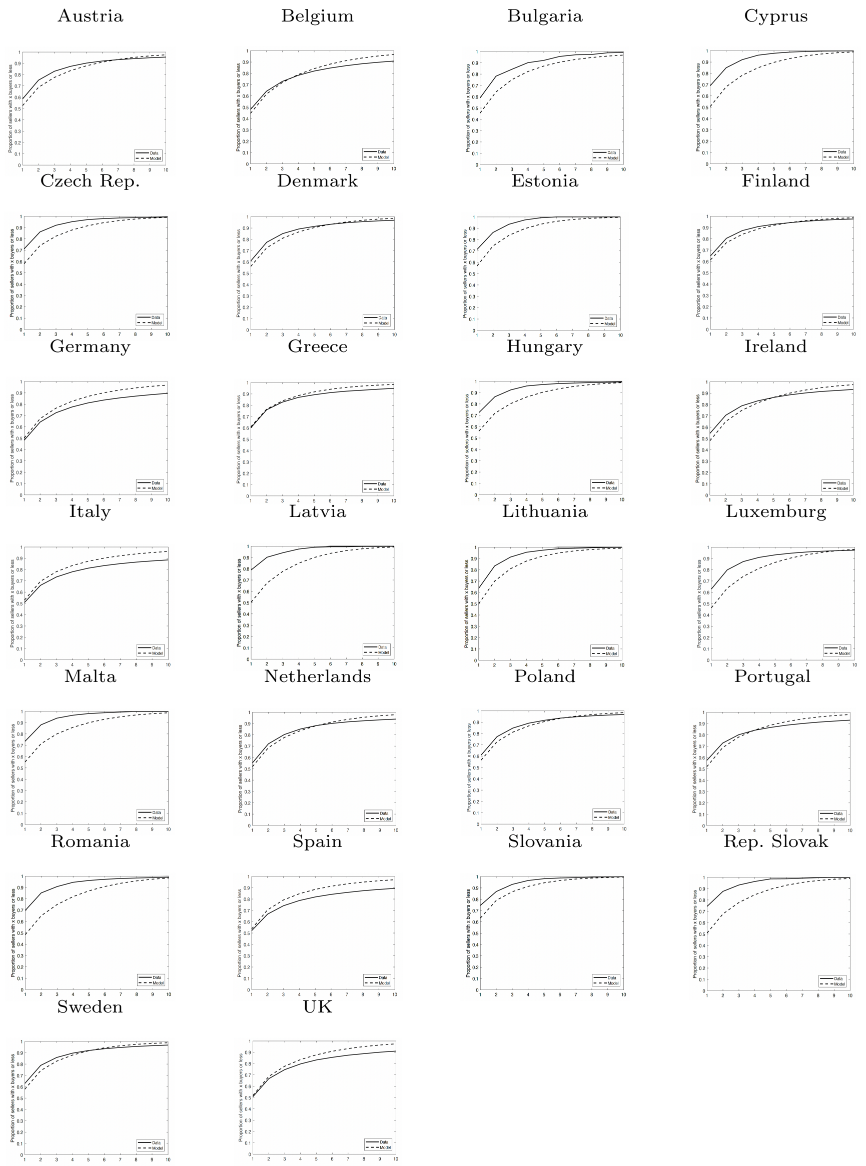
Notes: Columns (1)-(3) respectively report the mean, median, and third quartile number of buyers per seller in each destination. Column (4) gives the share of sellers having a unique buyer. A seller is defined as an exporter-HS6 product pair. The data are for 2007 and are restricted to transactions with recorded CN8-products.

Gaubert, C. and Itskhoki, O. (2018). Granular Comparative Advantage. NBER Working Papers 24807, National Bureau of Economic Research, Inc.

Gouriéroux, C., Monfort, A., and Trognon, A. (1985). Moindres carrés asymptotiques. *Annales de l'INSEE*, 58:91–122.

Head, K. and Mayer, T. (2014). *Gravity Equations: Workhorse,Toolkit, and Cookbook*, volume 4 of *Handbook of International Economics*, chapter 0, pages 131–195. Elsevier.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.



Notes: Observed and predicted CDF of sellers' numbers of buyers, by country. Predicted CDF are obtained using the model's definition of $h_i^k(M)$, at the country and product level, before aggregating across products using information on the relative number of producers of each good in France.

FIGURE OA.3.7. Model fit: Distribution of sellers' degrees