From Micro to Macro: Demand, Supply, and Heterogeneity in the Trade Elasticity

Maria Bas, Thierry Mayer, Mathias Thoenig

From Micro to Macro: Demand, Supply, and Heterogeneity in the Trade Elasticity*

Maria Bas[†] Thierry Mayer[‡] Mathias Thoenig [§]

February 19, 2016

Abstract

Models of heterogeneous firms with selection into export market participation generically exhibit aggregate trade elasticities that vary across country-pairs. Only when heterogeneity is assumed Pareto-distributed do all elasticities collapse into an unique elasticity, estimable with a gravity equation. This paper provides a theory-based method for quantifying country-pair specific elasticities when moving away from Pareto, i.e. when gravity does not hold. Combining two firm-level customs datasets for which we observe French and Chinese individual sales on the same destination market over the 2000-2006 period, we are able to estimate all the components of the dyadic elasticity: i) the demand-side parameter that governs the intensive margin and ii) the supply side parameters that drive the extensive margin. These components are then assembled under theoretical guidance to calculate bilateral aggregate elasticities over the whole set of destinations, and their decomposition into different margins. Our predictions fit well with econometric estimates, supporting our view that micro-data is a key element in the quantification of non-constant macro trade elasticities.

Keywords: trade elasticity, firm-level data, heterogeneity, gravity, Pareto, log-normal.

JEL Classification: F1

*This research has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) Grant Agreement No. 313522. We thank David Atkin, Dave Donaldson, Swati Dhingra, Ben Faber, Pablo Fagelbaum, Jean Imbs, Oleg Itskhoki, Jan de Loecker, Peter Morrow, Steve Redding, Andres Rodriguez Clare, Esteban Rossi-Hansberg, Katheryn Russ, Nico Voigtlander, and Yoto Yotov for useful comments on an early version, and participants at seminars in UC Berkeley, UCLA, Banque de France, WTO, CEPII, ISGEP in Stockholm, University of Nottingham, Banca d'Italia, Princeton University and CEMFI.

[†]University of Paris 1.

[‡]Corresponding author: Sciences Po, Banque de France, CEPII and CEPR. Email: thierry.mayer@sciencespo.fr. Postal address: 28, rue des Saints-Peres, 75007 Paris, France.

[§]Faculty of Business and Economics, University of Lausanne and CEPR.

1 Introduction

The response of trade flows to a change in trade costs, the aggregate trade elasticity, is a central element in any evaluation of the welfare impacts of trade liberalization. Arkolakis et al. (2012) recently showed that this parameter, denoted ε for the rest of the paper, is actually one of the (only) two sufficient statistics needed to calculate Gains From Trade (GFT) under a surprisingly large set of alternative modeling assumptions—the ones most commonly used by recent research in the field. Measuring those elasticities has therefore been the topic of a long-standing literature in international economics. The most common usage (and the one recommended by Arkolakis et al., 2012) is to estimate this elasticity in a *macro-level* bilateral trade equation referred to as structural gravity in the literature following the initial impulse by Anderson and van Wincoop (2003). In order for this estimate of ε to be relevant for a particular experiment of trade liberalization, it is crucial for this bilateral trade equation to be correctly specified as a structural gravity model with, in particular, a *unique* elasticity to be estimated across dyads.

Our starting point is that the model of heterogeneous firms with selection into export market participation (Melitz, 2003) will in general exhibit a *dyad-specific* elasticity, i.e. an ε_{ni}, which applies to each country pair. Only when heterogeneity is assumed Pareto-distributed[1] do all ε_{ni} collapse to a single ε. Under any other (commonly-used) distributional assumption, obtaining an estimate of the aggregate trade elasticity from a macro-level bilateral trade equation becomes problematic: first because a whole set of ε_{ni} has to be estimated, and second because structural gravity does not hold anymore. We argue that in this case quantifying trade elasticities at the aggregate level makes it necessary to use micro-level information. To this purpose, we combine sales of French and Chinese exporters on many destination-product combinations for which we also observe the applied tariff. We propose a theory-based method using this firm-level export data for estimating all the components of the dyad-specific trade elasticity: i) the demand-side parameter that governs the intensive margin and ii) the supply side parameters that drive the extensive margin. These components are then assembled under theoretical guidance to calculate the dyadic aggregate elasticities over the whole set of destination-products.

Taking into account cross-dyadic heterogeneity in trade elasticities is crucial for quantifying the expected impact of various trade policy experiments.[2] Consider the example of the current negotiations over a transatlantic trade agreement between the USA and the EU (TTIP). Under the simplifying assumption of a unique elasticity, whether the trade liberalization takes place with a proximate vs distant, large vs small economy, is irrelevant in terms of trade-promoting effect or welfare gains calculations. By contrast, our results suggest that the relevant ε_{ni} should be smaller (in absolute value) than if the United States were considering a comparable agreement with countries where the expected volume of trade is smaller. Regarding welfare, Head et al. (2014) and Melitz and Redding (2015) have shown theoretically that the GFT can be quite substantially mis-estimated if one assumes a constant trade elasticity when the "true" elasticity is variable (the margin of error can exceed 100 percent in both papers). The expected changes

[1] Unless otherwise specified, Pareto is understood here as the *un-truncated version* used by most of the literature. See Helpman et al. (2008) and Melitz and Redding (2015) for results with the truncated version, where the trade elasticity recovers a bilateral dimension.

[2] Imbs and Méjean (2015) and Ossa (2015) recently argued that another source of heterogeneity, the cross-sectoral one, raises important aggregation issues that matter for aggregate outcomes of trade liberalization. We abstract from this issue (which would reinforce the importance of heterogeneity for aggregate outcomes) in our paper, and mostly omit cross-sectoral variation in ε, apart from section 5.4 where we use industry-level estimates to show that both demand and supply side determinants enter aggregate elasticities.

in trade patterns and welfare effects of agreements such as TTIP will therefore be different compared to the unique elasticity case. One of the main objectives of our paper is to quantify how wrong can one be when making predictions based on a constant trade elasticity assumption.

Our approach maintains the traditional CES (σ) demand system combined with monopolistic competition. It features several steps that are structured around the following decomposition of aggregate trade elasticity into the sum of the intensive margin and the (weighted) extensive margin:

$$\varepsilon_{ni} = \underbrace{1-\sigma}_{\text{intensive margin}} + \underbrace{\frac{1}{\bar{x}_{ni}/x_{ni}^{\mathtt{MIN}}}}_{\text{min-to-mean}} \times \underbrace{\frac{d\ln N_{ni}}{d\ln \tau_{ni}}}_{\text{extensive margin}}, \tag{1}$$

The weight is the *mean-to-min ratio*, our observable measuring the dyadic dispersion of firm-level performance, that is defined as the ratio of average to minimum sales across markets. Intuitively, the weight of the extensive margin should be decreasing in easy markets where the increasing presence of weaker firms augments productivity dispersion. When assuming Pareto with shape parameter θ, the last part of the elasticity reduces to $\sigma - 1 - \theta$, and the overall elasticity becomes constant and reflects only the supply side homogeneity in the distribution of productivity: $\varepsilon_{ni}^P = \varepsilon^P = -\theta$ (Chaney, 2008). Without the Pareto assumption, one needs to calculate the two parts of the aggregate elasticity (1). We do so in two steps.

Our first step aims to estimate the demand side parameter σ using firm-level exports. Since protection is imposed on all firms from a given origin, higher demand and lower protection are not separately identifiable when using only one country of exports. With CES, firms are all confronted to the same aggregate demand conditions. Thus, considering a second country of origin enables to isolate the effects of trade policy, if the latter is discriminatory. We therefore combine shipments by French and Chinese exporters to destinations that confront those firms with different levels of tariffs. Our setup yields a *firm-level* gravity equation which raises serious estimation challenges. The main issue is the combination of a selection bias (inherent in any firm-level estimation of the Melitz (2003) model) with a very large set of fixed effects to be included in the regression. We use adapted versions of three estimators that have been proposed in the literature to deal with different aspects of the problem. Those three methods are evaluated with Monte Carlo simulations of our theoretical setup, before being implemented on our data. Our preferred estimates of the intensive margin trade elasticity imply an average value of σ around 5.

Our second and main step applies equation (1) and combines the estimate of the intensive margin ($\hat{\sigma}$) with the central supply side parameter—reflecting dispersion in the distribution of productivity—to obtain predicted aggregate elasticities of total export, number of exporters and average exports to each destination. It is important of course to implement this second step on the same dataset as the first step, to have consistent predictions on aggregate elasticity predictions. Those dyadic predictions (one elasticity for each exporter-importer combination) require knowledge of the bilateral export productivity cutoff under which firms find exports to be unprofitable. We also make use of the mean-to-min ratio to reveal those cutoffs. A key element of our procedure is the calibration of the productivity distribution. As an alternative to Pareto we consider the log-normal distribution that fits the micro-data on firm-level sales very well (the next section provides empirical evidence).[3]

[3]Head et al. (2014) provide evidence and references for several micro-level datasets that individual sales are much better

A side result of our paper is to discriminate between Pareto and log-normal as potential distributions for the underlying firm-level heterogeneity, suggesting that log-normal does a better job at matching the non-unique response of exports to changes in trade costs. Two pieces of evidence in that direction are provided. The first provides direct evidence that aggregate elasticities are non-constant across dyads. The second is a strong correlation across industries between firm-level and aggregate elasticities–at odds with the prediction of a null correlation under Pareto. We also find that the heterogeneity in trade elasticities is quantitatively important: Although the cross-dyadic average of bilateral elasticities is quite well approximated by a standard gravity model constraining the estimated parameter to be constant, deviations from this average level can be large. We show that under log-normal the ε_{ni} are larger (in absolute value) for pairs with low volumes of trade. Hence the trade-promoting impact of liberalization is expected to be larger for this kind of trade partners. For Chinese exports, assuming a unique elasticity would yield to underestimate the trade impact of a tariff liberalization by about 25% for countries with initially very small trade flows (Somalia, Chad or Azerbaijan for instance). By contrast, the error would be to overestimate by around 20% the exports created when the United States or Japan reduce their trade costs.

The next section relates our paper to existing work in the literature. Section 3 describes our model and empirical strategy. Section 4 deals with the estimation issues of the firm-level gravity regressions and reports the estimates of the intensive margin elasticity. Section 5 computes predicted macro-level trade elasticities and compares them with estimates from the Chinese and French aggregate export data. It also provides two additional pieces of evidence in favor of non-constant trade elasticities. The final section concludes.

2 Related Literature

In the empirical literature estimating trade elasticities, different approaches and proxies for trade costs have been used, with an almost exclusive focus on aggregate country or industry-level data. The gravity approach to estimating those elasticities mostly uses tariff data to estimate bilateral responses to variation in applied tariff levels. Most of the time, identification is in the cross-section of country pairs, with origin and destination determinants being controlled through fixed effects (Baier and Bergstrand (2001), Head and Ries (2001), Caliendo and Parro (2015), Hummels (1999), Romalis (2007)). A related approach consists in using the fact that most foundations of gravity predict the same coefficient on trade costs and domestic cost shifters to estimate that elasticity from the effect on bilateral trade of exporter-specific changes in productivity, export prices or exchange rates (Costinot et al. (2012) is a recent example).[4] Costinot et al. (2012) also use industry-level data for OECD countries, and obtains a preferred elasticity of -6.53 using productivity based on producer prices of the exporter as the identifying variable. Our paper has consequences for how to interpret those numbers in terms of underlying structural parameters.

approximated by a log-normal distribution when the entire distribution is considered (without left-tail truncation). Freund and Pierola (2015) is a recent example showing very large deviations from the Pareto distribution if the data is not vastly truncated for all of the 32 countries used.

[4] Other methodologies (also used for aggregate elasticities) use identification via heteroskedasticity in bilateral flows, and have been developed by Feenstra (1994) and applied widely by Broda and Weinstein (2006) and Imbs and Méjean (2015). Yet another alternative is to proxy trade costs using retail price gaps and their impact on trade volumes, as proposed by Eaton and Kortum (2002) and extended by Simonovska and Waugh (2011).

With a homogeneous firms model of the Krugman (1980) type in mind, the estimated elasticity turns out to reveal a demand-side parameter only, $1 - \sigma$ (this is also the case with Armington differentiation and perfect competition as in Anderson and van Wincoop (2003)). When instead considering heterogeneous firms à la Melitz (2003), the literature has proposed that the macro-level trade elasticity is driven solely by a supply-side parameter describing the dispersion of the underlying heterogeneity distribution of firms. This result has been shown with several demand systems (CES by Chaney (2008), linear by Melitz and Ottaviano (2008), translog by Arkolakis et al. (2010) for instance), but relies critically on the maintained assumption of a Pareto distribution. The trade elasticity then provides an estimate of the dispersion parameter of the Pareto, θ.[5] We show here that both existing interpretations of the estimated elasticities are too extreme: When the Pareto assumption is relaxed, the aggregate trade elasticity is a mix of demand and supply parameters.

There is a small set of papers that estimate the intensive margin elasticity at the exporter level. Berman et al. (2012) presents estimates of the trade elasticity with respect to real exchange rate variations across countries and over time using firm-level data from France. Fitzgerald and Haller (2015) use firm-level data from Ireland, real exchange rate and weighted average firm-level applied tariffs as price shifters to estimate the trade elasticity. The results for the impact of real exchange rate on firms' export sales are of a similar magnitude, around 0.8 to 1. Applied tariffs vary at the product-destination-year level. Fitzgerald and Haller (2015) create a firm-level destination tariff as the weighted average over all hs6 products exported by a firm to a destination in a year using export sales as weights. Relying on this construction, they find a tariff elasticity ranging widely from -1.7 to -24 in their baseline table. The preferred estimate of Berthou and Fontagné (2015), who use the response of the largest French exporters in the United States to the levels of applied tariffs is -2.5. We depart from those papers by using an alternative methodology to identify the trade elasticity with respect to applied tariffs; i.e. the differential treatment of exporters from two distinct countries (France and China) in a set of product-destination markets. We also spend time to describe the estimation issues involved in firm-level gravity regressions and provide the first rigorous evaluation of the alternative estimators available with Monte Carlo simulations using the canonical Melitz (2003) model as a DGP.

Our paper also relates to several recent papers studying patterns and consequences of heterogeneity in trade elasticities. Berman et al. (2012) and Gopinath and Neiman (2014) find that in order to predict correctly the aggregate patterns of trade adjustments to price shocks, one has to take into account firm-level heterogeneity with use of micro data. In both papers, heterogeneity matters because firms have different individual responses in export and/or import behavior. In particular, both papers find that the firm-level elasticity depends negatively on the size of the firm (because of variable markups). Our paper also finds that measuring aggregate trade responses requires usage of firm-level data. It is however for a different reason: In our case, heterogeneity in aggregate trade elasticities simply originates in a departure from the common assumption that productive efficiency is Pareto-distributed. While we do recognize that trade elasticities might differ across firms because of variable markups, our paper shows that this is

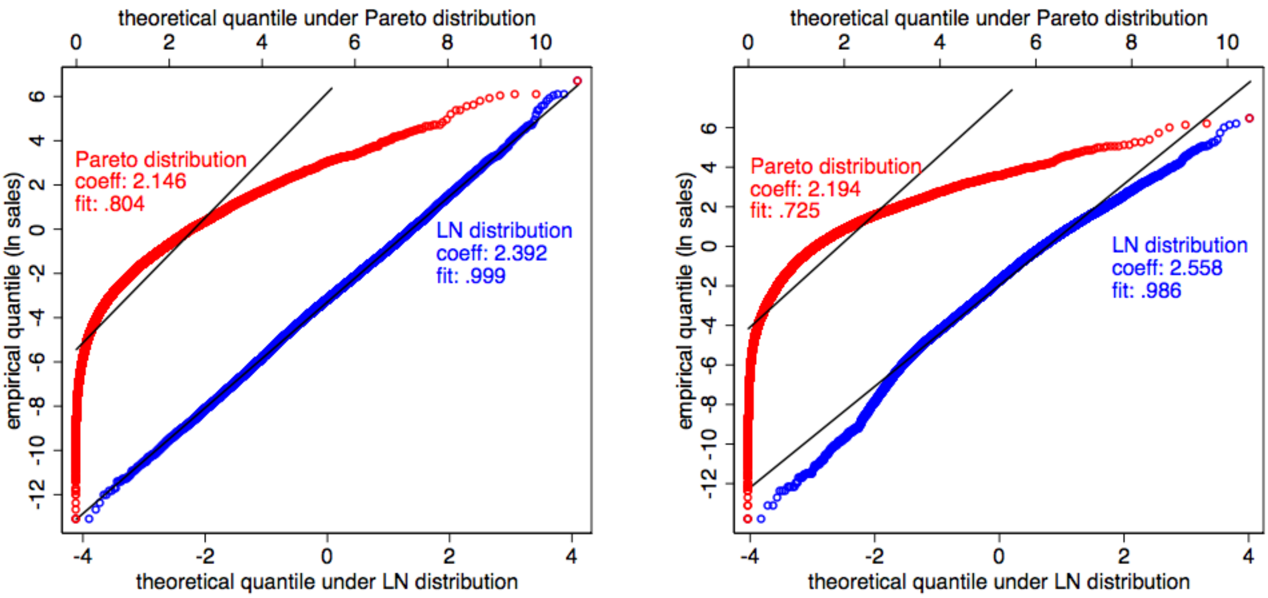
[5] This result of a constant trade elasticity reflecting the Pareto shape holds when maintaining the CES demand system but making other improvements to the model such as heterogeneous marketing and/or fixed export costs (Arkolakis, 2010; Eaton et al., 2011). In the Ricardian setup of Eaton and Kortum (2002), the trade elasticity is also a (constant) supply side parameter reflecting heterogeneity, but this heterogeneity takes place at the national level, and reflects the scope for comparative advantage.

not required to ensure that heterogeneity matters for the aggregate economy and investigates a different, complementary, channel.

We also contributes to the literature studying the importance of the distribution assumption of heterogeneity for trade patterns, trade elasticities and welfare. Head et al. (2014), Yang (2014), Melitz and Redding (2015) and Feenstra (2013) have recently argued that the simple gains from trade formula proposed by Arkolakis et al. (2012) relies crucially on the Pareto assumption, which mutes important channels of gains in the heterogenous firms case. Barba Navaretti et al. (2015) present gravity-based evidence that the exporting country fixed effects depends on characteristics of firms' distribution that go beyond the simple mean productivity, a feature incompatible with the usually specified Pareto heterogeneity. Fernandes et al. (2015) use customs data for numerous developing countries to show that a decomposition of total bilateral exports into intensive and extensive margins exhibits an important role for the latter, with patterns consistent with log-normally distributed heterogeneity and incompatible with (untruncated) Pareto. The alternatives to Pareto considered to date in welfare gains quantification exercises are i) the truncated Pareto by Helpman et al. (2008), Melitz and Redding (2015) and Feenstra (2013), and ii) the log-normal by Head et al. (2014), Fernandes et al. (2015) and Yang (2014). A key simplifying feature of Pareto is to yield a constant trade elasticity, which is not the case for alternative distributions. Helpman et al. (2008) and Novy (2013) have produced gravity-based evidence showing substantial variation in the trade cost elasticity across country pairs. Our contribution to that literature is to use the estimated demand and supply-side parameters to construct predicted bilateral elasticities for aggregate flows under the log-normal assumption, and compare their first moments to gravity-based estimates. It should be noted that there are other ways to generate bilateral trade elasticities. The most obvious is to depart from the simple CES demand system. Novy (2013) builds on Feenstra (2003), using the translog demand system with homogeneous firms to obtain variable trade elasticities. Atkeson and Burstein (2008) is another example maintaining CES demand, and generating heterogeneity in elasticities trough oligopoly. We choose here to keep the change with respect to the benchmark Melitz/Chaney framework to a minimal extent, keeping CES and monopolistic competition, while changing only the distributional assumption, comparing Pareto to log-normal.

Our interest in the log-normal distribution simply comes from data patterns. With CES demand and constant markups, Head et al. (2014) show that the distribution of sales in a given destination inherits the distribution of the firms' underlying performance variable (call it productivity). When the latter is distributed Pareto or log-normal, sales are also distributed Pareto and log-normal, the only substantial difference being a shift in the shape parameter of each of those distributions. While we don't observe productivity, we do observe the complete distribution of sales, and can inspect which distribution seems to fit the best. A very easy tool for that inspection is the Quantile-Quantile (QQ) regression, where the theoretical quantile under each alternative is regressed on the empirical quantile (log sales). QQ regressions are linear in both cases, and the slope reveals the shape parameter of the underlying distribution. Figure 1 reports those regressions for two sets of exporters (French and Chinese) used later in this paper. We focus on a major destination for each of those countries, Belgium and Japan respectively. The Pareto regression is in red and the log-normal one in blue. It is very clear in those plots that log-normal is a much better fit of the data for the overall sales distribution, making it a credible and natural alternative to Pareto.

Figure 1: Distribution of firm-level sales



(a) French firms in Belgium

(b) Chinese firms in Japan

Note: In the regressions reported, the dependent variable is the firm-level log of exports in 2000. The RHS is $\Phi^{-1}(\hat{F}_i)$ for log-normal and $\ln(1-\hat{F}_i)$ for Pareto, where $\hat{F}_i$ is the empirical CDF of log sales and Φ is the CDF of the standard normal. Under the usual CES (σ)/constant markup assumptions, coefficients have a direct interpretation in terms of structural parameters: $\frac{\sigma-1}{\theta}$ if productivity is Pareto with shape parameter θ, and $(\sigma-1)\nu$ for log-normal productivity with dispersion parameter ν.

3 Firm-level and aggregate-level trade elasticities: theory

We use the multi-country version of the Melitz (2003) theoretical framework. Country i hosts a set of heterogeneous firms facing a constant price elasticity (CES utility combined with iceberg costs) and contemplating exports to several destinations indexed by subscript n. In this setup, firm-level export value x depends upon the firm-specific unit input requirement (α), wages at home (w_i), and real expenditure in n, $A_n \equiv X_n P_n^{\sigma-1}$, with P_n the ideal CES price index relevant for sales in n. A_n is a measure of "attractiveness" of market n (expenditure discounted by the degree of competition on this market). There are trade costs associated with reaching market n, consisting of an observable iceberg-type part (τ_{ni}), and a shock that affects firms differently on each market, $b_{ni}(\alpha)$.[6] Monopolistic competition ensures a complete pass-through of trade costs into delivered prices, such that

$$x_{ni}(\alpha) = \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} [\alpha w_i \tau_{ni} b_{ni}(\alpha)]^{1-\sigma} A_n. \tag{2}$$

The *firm-level trade elasticity*, i.e. the individual reaction to a change in observable trade costs is $1-\sigma$. We describe the estimation procedure for this micro-level elasticity in section 4.

In order to obtain the aggregate trade elasticity, we start by summing, for each country pair, the sales equation (2) across all active firms:

$$X_{ni} = V_{ni} \times \left(\frac{\sigma}{\sigma-1}\right)^{1-\sigma} (w_i \tau_{ni})^{1-\sigma} A_n M_i^e, \tag{3}$$

where M_i^e is the mass of entrants and V_{ni} is a term which denotes a cost-performance index of exporters located in country i and selling in n. This index, introduced by Helpman et al. (2008), is characterized by

$$V_{ni} \equiv \int_0^{a_{ni}^*} a^{1-\sigma} g(a) da, \tag{4}$$

where $a \equiv \alpha \times b(\alpha)$ corresponds to the unitary labor requirement rescaled by the firm-destination shock. In equation (4), $g(.)$ denotes the PDF of the rescaled unitary labor requirement and a_{ni}^* is the rescaled labor requirement of the firm that just breaks even and therefore exports to market n. The solution for this cutoff firm is the cost satisfying the zero profit condition, i.e., $x_{ni}(a_{ni}^*) = \sigma w_i f_n$. Using (2), this cutoff is characterized by

$$a_{ni}^* = \frac{1}{\tau_{ni} f_n^{1/(\sigma-1)}} \left(\frac{1}{w_i}\right)^{\sigma/(\sigma-1)} \left(\frac{A_n}{\sigma}\right)^{1/(\sigma-1)}. \tag{5}$$

We are interested in the (partial) elasticity of aggregate trade value with-respect to variable trade costs, τ_{ni}. Partial means here holding constant origin-specific and destination-specific terms (income and price indices) as in Arkolakis et al. (2012) and Melitz and Redding (2015).[7] Using (3), we obtain the *aggregate*

[6] An example of such unobservable term would be the presence of workers from country n in firm α, that would increase the internal knowledge on how to reach consumers in n, and therefore reduce trade costs for that specific company in that particular market (b being a mnemonic for barrier to trade). Note that this type of random trade cost shock is isomorphic to assuming a firm-destination demand shock in this CES-monopolistic competition model.

[7] In practical terms, the use of importer and exporter fixed effects in gravity regressions (the main source of estimates of

trade elasticity:

$$\varepsilon_{ni} \equiv \frac{d \ln X_{ni}}{d \ln \tau_{ni}} = 1 - \sigma - \gamma_{ni}, \tag{6}$$

which uses the fact that $d \ln a^*_{ni} / d \ln \tau_{ni} = -1$. The γ_{ni} term, introduced by Arkolakis et al. (2012), describes how V_{ni} varies with an increase in the cutoff cost a^*_{ni}, that is an easier access of market n for firms in i:

$$\gamma_{ni} \equiv \frac{d \ln V_{ni}}{d \ln a^*_{ni}} = \frac{a^{*2-\sigma}_{ni} g(a^*_{ni})}{V_{ni}}. \tag{7}$$

Equations (6) and (7) show that the aggregate trade elasticity should, in general, not be constant across country pairs. They also make it clear that the aggregate elasticity is a combination of the firm-level response, the intensive margin $1 - \sigma$, and the contribution to total export changes due to entry and exit of firms into the export market, γ_{ni}.

In order to evaluate ε_{ni}, combining (7) with (4) reveals that we need to know the value of bilateral cutoffs a^*_{ni}. In order to obtain those, we define the following expression

$$\mathcal{H}(a^*_{ni}) \equiv \frac{1}{a^{*1-\sigma}_{ni}} \int_0^{a^*_{ni}} a^{1-\sigma} \frac{g(a)}{G(a^*_{ni})} da, \tag{8}$$

a monotonic, invertible function which has a straightforward economic interpretation in this model. It is the ratio of average over minimum performance (measured as $a^{*1-\sigma}$) of firms located in i and exporting to n. Using equations (2) and (3) reveals that this ratio also corresponds to the observed mean-to-min ratio of sales:

$$\frac{\bar{x}_{ni}}{x_{ni}(a^*_{ni})} = \mathcal{H}(a^*_{ni}). \tag{9}$$

In firm-level export datasets, the ratio of average to minimum trade flows for each destination country n is an observable. Calibrating $\mathcal{H}(.)$ (see Section 5.1) and using equation (9), one can reveal $\hat{a}^*_{ni}$, the estimated value of the export cutoff for i firms exporting to n as a function of the mean-to-min ratio of sales on each destination market:

$$\hat{a}^*_{ni} = \mathcal{H}^{-1} \left(\frac{\bar{x}_{ni}}{x^{\mathtt{MIN}}_{ni}} \right). \tag{10}$$

Equipped with the dyadic cutoff, we use equations (6) to (9) to obtain the aggregate trade elasticities

$$\varepsilon_{ni} = 1 - \hat{\sigma} - \frac{x^{\mathtt{MIN}}_{ni}}{\bar{x}_{n,i}} \times \frac{\hat{a}^*_{ni} g(\hat{a}^*_{ni})}{G(\hat{a}^*_{ni})}, \tag{11}$$

where $\hat{\sigma}$ is the estimate of the intensive margin (the demand-side parameter) obtained from the firm-level export equation. We also calculate two aggregate elasticities: the elasticity of the number of exporters N_{ni} (the so-called extensive margin) and the elasticity of average shipments $\bar{x}_{ni}$. The number of active exporters is closely related to the cutoff since $N_{ni} = M^e_i \times G(a^*_{ni})$, where M^e_i represents the mass of entrants (also absorbed by exporter fixed effects in gravity regressions). Differentiating and using (11)

the aggregate elasticity) holds w_i, M^e_i and A_n constant. While this is literally true under Pareto because w_i, M^e_i and A_n enter a^*_{ni} multiplicatively, deviating from Pareto adds a potentially complex interaction term through a non-linear in logs effect of monadic terms on the dyadic cutoff. We expect this effect to be of second order, an intuition confirmed by Monte Carlo simulations of the model under log-normal heterogeneity (feature in Appendix 3).

we can calculate the dyadic extensive margin of trade

$$\frac{d\ln N_{ni}}{d\ln \tau_{ni}} = -\frac{\hat{a}^*_{ni} g(\hat{a}^*_{ni})}{G(\hat{a}^*_{ni})}. \tag{12}$$

From the accounting identity $X_{ni} \equiv N_{ni} \times \bar{x}_{ni}$, we obtain the (partial) elasticity of average shipments to trade simply as the difference between the estimated aggregate elasticities, (11) and the estimated extensive margins, (12):

$$\frac{d\ln \bar{x}_{ni}}{d\ln \tau_{ni}} = \varepsilon_{ni} - \frac{d\ln N_{ni}}{d\ln \tau_{ni}} = 1 - \hat{\sigma} - \frac{\hat{a}^*_{ni} g(\hat{a}^*_{ni})}{G(\hat{a}^*_{ni})} \left(\frac{x^{\mathtt{MIN}}_{ni}}{\bar{x}_{n,i}} - 1 \right). \tag{13}$$

Combining (11) and (12), we can re-express aggregate elasticities as a function of the intensive and extensive margins and of the mean-to-min ratio:

$$\varepsilon_{ni} = \underbrace{1 - \hat{\sigma}}_{\text{intensive margin}} + \underbrace{\frac{1}{\bar{x}_{ni}/x^{\mathtt{MIN}}_{ni}}}_{\text{min-to-mean}} \times \underbrace{\frac{d\ln N_{ni}}{d\ln \tau_{ni}}}_{\text{extensive margin}}, \tag{14}$$

which is equation (1) presented in the introduction. This decomposition shows that the aggregate trade elasticity is the sum of the intensive margin and of the (weighted) extensive margin. The weight on the extensive margin depends only on the mean-to-min ratio, an observable measuring the dispersion of relative firm performance ($\mathcal{H}(a^*_{ni})$ in the model). Intuitively, the weight of the extensive margin should be decreasing when the market gets easier. Indeed easy markets have larger rates of entry, $G(a^*)$, and therefore increasing presence of weaker firms which augments dispersion measured as $\mathcal{H}(a^*_{ni})$. The marginal entrant in an easy market will therefore have less of an influence on aggregate exports, a smaller impact of the extensive margin. In the limit, the weight of the extensive margin becomes negligible and the whole of the aggregate elasticity is due to the intensive margin / demand parameter. In the (untruncated) Pareto case however this mechanism is not operational since $\mathcal{H}(a^*_{ni})$ and therefore the weight of the extensive margin is constant. In section 5, we implement our method with Pareto-distributed a as opposed to log-normal a, an alternative yielding non-constant dispersion of sales across destinations. Before estimating the aggregate trade elasticities, we however need to obtain an estimate of $\hat{\sigma}$, the parameter relevant in the firm-level trade elasticity.

4 Estimating the firm-level trade elasticity

4.1 Estimation issues

Three serious methodological challenges arise when estimating the individual response of export values to variation in tariffs while keeping a close link to theory.

The need for multiple origins: The first challenge is to separate the effect of trade costs from destination fixed effects. At this stage, it is useful to add a notation (p) indexing products for which we observe both the value exported by the firm, $x^p_{ni}(\alpha)$, and the bilateral tariff rate $t^p_{ni}(\alpha)$. Trade costs include both tariffs and other trade costs (distance D_{ni} for instance), and we assume the standard functional form such that $\tau^p_{ni} = (1 + t^p_{ni})D^\delta_{ni}$. From now on, we will use the term "market" to designate a product-destination combination. Taking logs of the demand equation (2), where $\epsilon^p_{ni}(\alpha) \equiv (b^p_{ni}(\alpha))^{1-\sigma}$ is our unobservable firm-market error term, a "firm-level gravity" equation is obtained:

$$\ln x^p_{ni}(\alpha) = (1-\sigma)\ln\left(\frac{\sigma}{\sigma-1}\right)+(1-\sigma)\ln(\alpha w_i)+(1-\sigma)\ln(1+t^p_{ni})+(1-\sigma)\delta\ln D_{ni}+\ln A^p_n+\ln\epsilon^p_{ni}(\alpha). \quad (15)$$

The objective is to estimate $1 - \sigma$ out of the impact of tariffs on firm-level sales. At this stage of the paper, we consider a unique σ, which can be interpreted as an average of elasticities that might vary across goods. We will come back to industry-specific elasticities in section 5.4. In the gravity literature, it has become common practice to capture A^p_n (a complex construction, that depends non-linearly upon σ) with market fixed effects. This is however not applicable if the dataset at hand covers only one origin country, since A^p_n and τ^p_n would then vary across the same dimension.[8] To remain theory-consistent, one therefore need to use at least two sets of exporters, based in countries that face different levels of tariffs applied by n. We do combine firm-level customs data for France and China ($i = [\mathrm{FR}, \mathrm{CN}]$), where the value of flows is available at the HS6 level, the same disaggregation level for which we measure bilateral tariffs t^p_{ni} using WITS. Proxies for D_{ni} include distance, contiguity, colonial linkage and common language all obtained from the CEPII gravity database. The data appendix A.1.1. gives more detail about each of those data sources.

The fixed effects curse: The second challenge relates to the number of fixed effects to be estimated. In addition to the market dimension (A^p_n), we need a set of fixed effects at the firm level to capture marginal costs (αw_i) (and more generally all other unobservable firm-level determinants of export performance, such as quality of products exported, managerial capabilities...). Since there are tens of thousands of exporters in each origin country and several hundred thousand destination-product combinations, the Least Square Dummy Variable–brute force–approach is not feasible. There are two alternative implementable solutions that we consider. The first is to estimate (15) directly using the high-dimensional procedure that was developed by labor economists to deal with the very large number of fixed effects implied by employer-employee data.[9]

$$\ln x^p_{ni}(\alpha) = \mathrm{FE}^\alpha_i + \mathrm{FE}^p_n + (1-\sigma)\ln(1+t^p_{ni}) + (1-\sigma)\delta\ln D_{ni} + \ln\epsilon^p_{ni}(\alpha) \quad (16)$$

We call this approach two-way fixed effects procedure, 2WFE, since we have two dimensions of unobserved heterogeneity to be controlled for (i.e. firm fixed effects FE_i and market fixed effects FE^p_n). The second solution is a ratio-type estimation inspired by Hallak (2006), Romalis (2007), Head et al. (2010), and Caliendo and Parro (2015) that removes observable and unobservable determinants for both firm-level and destination factors. This method uses four individual export flows to calculate ratios of ratios: an

[8] Most if not all papers estimating firm-level gravity rely on only one source of export flows, while still estimating the impact of exchange rate (Berman et al. (2012)), tariffs (Berthou and Fontagné (2015)) or both (Fitzgerald and Haller (2015)). The identification in those papers then comes from another dimension, usually time. However, this strategy requires to make the assumption that $X_n P_n^{\sigma-1}$ does not vary over time when τ_n does. This is inconsistent with a theory where trade costs enter the price index. Also the time dimension of variance in tariffs might be problematic since Fitzgerald and Haller (2015) note that the changes in tariffs over time are small relative to the cross-sectional variations.

[9] The fastest procedure to date available in Stata, `reghdfe`, has been developed by Sergio Correia building on Guimares and Portugal (2010)

approach referred to as TETRADS from now on. Consider a given French firm j and a Chinese firm ℓ exporting to both n and a reference country k. The Independence of Irrelevant Alternatives (IIA) property of the CES demand system allows to manipulate equation (2) to write the following tetrad:

$$\frac{x_n^p(\alpha_{j,\mathrm{FR}})/x_k^p(\alpha_{j,\mathrm{FR}})}{x_n^p(\alpha_{\ell,\mathrm{CN}})/x_k^p(\alpha_{\ell,\mathrm{CN}})} = \left(\frac{\tau_{n\mathrm{FR}}^p/\tau_{k\mathrm{FR}}^p}{\tau_{n\mathrm{CN}}^p/\tau_{k\mathrm{CN}}^p}\right)^{1-\sigma} \times \frac{\epsilon_n^p(\alpha_{j,\mathrm{FR}})/\epsilon_k^p(\alpha_{j,\mathrm{FR}})}{\epsilon_n^p(\alpha_{\ell,\mathrm{CN}})/\epsilon_k^p(\alpha_{\ell,\mathrm{CN}})}. \tag{17}$$

Denoting tetradic terms with a $\sim$ symbol, one can re-write equation (17) as an estimable equation

$$\ln \widetilde{x}_{\{j,n,k\}}^p = (1-\sigma)\ln\left(\widetilde{1+t_{\{n,k\}}^p}\right) + (1-\sigma)\delta \ln \widetilde{D}_{\{n,k\}} + \ln \widetilde{\epsilon}_{\{j,n,k\}}^p. \tag{18}$$

This approach simply involves a linear regression of log "tetraded" flows on log "tetraded" trade costs and does not require the estimation of any fixed effect. This method is therefore very simple computationally. It also lends itself easily to graphical analysis and will finally provide a natural test of non-constant aggregate elasticities in section 5.3.

Firm-level zeroes (selection bias): The third challenge is to account for the endogenous selection into different export markets across firms. Assuming that fixed export costs vary across markets and are paid using labor of the origin country, profits in this setup are given by $x_{ni}^p(\alpha)/\sigma - w_i f_n^p$. From equation (15), we see that a firm with a low cost (αw_i) can afford having a low draw on $\epsilon_{ni}^p(\alpha)$ and still export profitably to n. The same logic applies for large (high A_n^p), and easy to reach (low τ_{ni}^p) markets. Concerning our variable of interest, higher tariff observations will be associated with firms having drawn higher $\epsilon_{ni}^p(\alpha)$, thus biasing downwards our estimate of the trade elasticity. The solution to this selection bias is not trivial in our case where a large set of fixed effects is included. Although we are unaware of a "perfect" estimator, we propose three alternative methods, that we confront to Monte Carlo evidence of a simulated version of the model.

First, one can focus the regressions on firms that have such a large productivity that their idiosyncratic destination shock is of second order. Inspired by Mulligan and Rubinstein (2008), Paravisini et al. (2015) and Fitzgerald and Haller (2015) concentrate the analysis on large firms that serve almost all markets. This requires to decide on a variable likely to predict small levels of selection. Paravisini et al. (2015) use firm-level measures of total exports and credit, while Fitzgerald and Haller (2015) use a threshold of firm-level employment. We implement this approach with our data by restricting the sample to the largest exporter in each origin-product. Because this approach can accommodate our two-way fixed effects procedure (firm and destination) very easily, we call it 2WFE on top exporters. The second estimator relies upon the tetrads method, with a similar strategy of restricting attention to large exporting firms that are the least likely to be affected by the selection bias. When taking ratios of ratios of individual trade flows, we focus on the top exporters of each country,[10] and look at the evolution of their relative exports on different markets (compared to a reference country). We expect those two methods to give comparable results. The issue with both estimators is that they estimate the intensive margin on a reduced sub-sample of the largest firms. Those might have different trade elasticities, for reasons outside

[10] j and ℓ are chosen as the top exporters to k in value terms in equation (17).

of our model.[11] Our third estimator reinstates the full sample of exporters. Assuming a normally distributed $\ln \epsilon^p_{ni}(\alpha)$ in (15) yields a generalized structural tobit, that we will refer to as EK-tobit, since it was developed by Eaton and Kortum (2001). Crozet et al. (2012) apply EK-tobit to the heterogeneous exporter model by using the theoretical equation for minimum sales, $x^{p,\mathtt{MIN}}_{ni}(\alpha) = \sigma w_i f^p_n$, which therefore provides a natural estimate for the truncation point for each market. EK-tobit is the best estimator for our theoretical framework, with an important caveat: We must reduce the number of included fixed effects because it seems computational unfeasible to estimate a generalized tobit with the very large set of fixed effects our theory demands.[12]

We therefore have three possible estimators, 2WFE on top exporters, TETRADS on top exporters, and EK-Tobit. We now proceed to test for the performance of our three imperfect estimators meant to correct for the selection bias using Monte-Carlo simulations. The DGP uses equation (15) for the value $x_{ni}(\alpha)$ exported by 100000 firms divided into two origin countries and selling in 80 (to match the numbers we have in our sample). The fixed export costs f_n and the market size A_n are drawn from independent log-normal distributions calibrated to generate the same proportion of zero valued flows we find in the data (about 95.5%). A key aspect of the simulation lies in the choice of the respective importance of firm-level cost, α, and unobserved firm-destination shock $b_{ni}(\alpha)$ in firms' sales. For a given origin-destination, the firm-level sales distribution follows the distribution of $a = \alpha \times b(\alpha)$. Following the QQ regressions shown in Figure 1, we simulate a log-normal a. The key parameter of this distribution is its standard deviation, equal to the QQ regression coefficient divided by $\sigma - 1$ (see Head et al. (2014) for details). We set this according to the average of the two regression coefficients obtained for French and Chinese exporters in Figure 1. We calibrate the relative contribution of α and $b_{ni}(\alpha)$ fitting the correlation between the rank of a firm in total exports of the country and its rank in sales to each country n. Without the random term our model predicts a perfect correlation, while this correlation would approach zero if $b_{ni}(\alpha)$ is the only source of firm-level heterogeneity. Our data reveals that this correlation is around 66% in both the French and Chinese cases. Table 1 summarizes our Monte Carlo results based on 200 replications. Mean and standard deviation of the sampling distribution are reported for various statistics. When the exports are censored (setting unprofitable exports to 0), the correlation between τ_{ni} and the error term is about 14%, sufficient to create a massive bias in the estimated trade elasticity which falls to about half its true value (rows 4 and 5). EK-tobit with the appropriate set of fixed effects (row 6) recovers almost exactly the true coefficient. Perhaps more surprising, EK-tobit without any fixed effect also is very close to the true trade elasticity. This is due to the fact that the simulation assumes no correlation between τ_{ni} and either A_n or α. Since EK-tobit considers the full sample of potential flows, no selection bias can occur through that channel. However, there might be some correlation between τ_{ni} and A_n for instance in the true data, suggesting the need to introduce proxies of A_n in empirical implementations. Rows 8 and 9 report the results for the two other estimators. Both seems to be slightly biased, 10% for 2WFE on top exporters, 13% for TETRADS on top exporters, even though they don't differ significantly from the true

[11] Berman et al. (2012) show that several models featuring variable markups predict that large firms should face lower demand elasticities and therefore react less than small firms to a change in trade costs. Their finding that the response to exchange rate changes declines with productivity (confirmed by Chatterjee et al. (2013) for Brazilian exporters and Li et al. (2015) for Chinese exporters) suggests that the estimates in the present paper could be considered as a lower bound.

[12] Greene (2004) shows that the tobit model is much less subject to the incidental parameters problem than other non-linear models such as logit or probit. We also detect no sign of bias in our Monte Carlo simulations.

value of σ.

Table 1: Monte Carlo results: firm-level elasticities wrt to a change in trade costs

	mean	s.d
% of positive flows	0.048	0.008
correlation between global and local rank	0.660	0.015
correlation between $\ln \tau_{ni}$ and $\ln b_{ni}(\alpha)$	-0.137	0.041
σ 2WFE on full sample	5.000	0.004
σ 2WFE on censored sample	2.419	0.105
σ EK-tobit	4.997	0.014
σ EK-tobit (no FEs)	4.996	0.527
σ tetrads	4.336	0.959
σ 2WFE on top exporter	4.528	1.031
# obs full sample (and EK-tobit)	8000000	0
# obs censored sample	384401	60323.965
# obs tetrads	313.195	44.942
# obs 2WFE on top exporter	114.430	20.829

Note: True σ is set to 5. There are 200 replications, parameters on fixed costs of exports and size of the demand term have been calibrated so that the share of non-selected trade flows at the firm-destination level averages between 4 and 5 %. For each elasticity, the first column reports the average value, while the second reports standard deviations of elasticities across the 50 replications.

4.2 Empirical estimates of the intensive margin

We now turn to our observed sample. Table 2 reports the results based on our three estimators applied to the French and Chinese firm-level exports in 2000. The first two columns use 2WFE on top exporters, the next 2 use TETRADS on top exporters, while the last two use EK-Tobit.

While 2WFE on top is straightforward to implement, one needs to define reference k countries for Tetrads (equation 18). We choose those with two criteria in mind. First, these countries should be those that are the main trade partners of France and China in the year 2000, since we want to minimize the number of zero trade flows in the denominator of the tetrad. The second criteria relies on the variation in the tariffs effectively applied by the importing country to France and China. Hence, among the main trade partners, we retain those countries for which the average difference between the effectively applied ad valorem tariffs to France and China is greater. These two criteria lead us to select the following set of 8 reference countries: Australia, Canada, Germany, Italy, Japan, New Zealand, Poland and the UK.

When implementing EK-tobit, we need to fill in (with zero flows) the destinations that a firm found unprofitable to serve. The set of potential destinations for each product is given by all countries where at least one firm exported that good. When estimating equation (15) through EK-tobit, we proxy for $\ln A_n^p$ with destination n fixed effects, and for firm-level determinants α with the count of markets served by the firm. An origin country dummy for Chinese exporters account for all differences across the two groups, such as wages, w_i.

For each of the three methods of estimation, the first column includes tariffs and the usual set of gravity variables (distance, contiguity, colonial link and common language). The second column adds a term for Regional Trade Agreements (RTA). The idea is to control for potential non-tariff barriers to trade that could be correlated with ad valorem applied tariffs. This is a particularly demanding specification, since a lot of the variance in tariffs should come from the distinction between RTA members facing zero tariffs and non-member pairs facing positive ones.

The TETRAD and 2WFE methods on top exporters in columns (1) to (4) show quite similar patterns of results, as expected from the similarity in approach and Monte Carlo results. Distance has the usual negative coefficient, contiguity enters strongly positive, while colonial link and common language have a much more volatile and mostly insignificant effect at the firm level. RTAs enter with a very comparable and strong effect in both methods (approximately tripling trade flows), with the expected effect of reducing the impact of tariffs. Overall the three methods point to similar coefficients with a reasonable value of σ averaging 5.5 across the 6 columns. This turns out to be a central value within the small set of papers estimating response of firm-level flows to applied tariffs. Berthou and Fontagné (2015) obtain coefficients that would imply a preferred value of $\sigma = 3.5$ and report a larger response when restricting the sample to the largest exporters, as expected from our analysis of selection bias above. Fitzgerald and Haller (2015) report a very strong variance in firm-level response to tariffs, with implied value of σ strongly rising when restricting regressions to the largest firms. Their benchmark table imply σ ranging from 2.7 to around 25, the latter being relevant for the biggest firms in the most popular markets.

The working paper version of this paper concentrated on the tetrads aproach to estimation and provided many robustness checks of the micro trade elasticity. We briefly report a summary of those results here. There are essentially two sources of variance of tariffs in our setting: across products and across destinations. When focusing on the cross-destination dimension through the use of product fixed effects, the coefficients for the applied tariffs $(1 - \sigma)$ range from -6 to -3.2. Restricting the sample to destination countries which apply non-MFN tariffs to France and China (Australia, Canada, Japan, New Zealand and Poland), yields results of similar magnitude, ranging from -5.47 to -3.24. We also consider two additional cross-sectional samples, one after China entry into WTO (2001), the other for the final year for which we have Chinese customs data (2006). Here again the results are qualitatively robust, although the coefficients on tariffs are lower as expected since the difference of tariffs applied to France and China by destination countries is much reduced after 2001. With the caveat in mind that entry into WTO combined with patchy data over time for many countries makes panel estimation quite difficult, we consider it over the 2000-2006 period. The coefficients are more volatile, but somehow close to the findings from the baseline cross-section estimations in 2000 (they range from -5.26 to -1.80). This set of robustness estimates combined with the ones in Table 2 points to a central value of the demand side parameter of our model located around 5. We use $\hat{\sigma} = 5$ in the coming section in order to calculate aggregate trade elasticities relevant for the same sample of French and Chinese exporters.

5 Aggregate trade elasticities

5.1 Quantifying the theoretical aggregate trade elasticities

In this section, we provide a theory-consistent methodology for inferring, from firm-level data, the *aggregate* elasticities of trade with respect to trade costs. Those elasticities (the reactions of total trade, number of exporters and average exports to tariffs) are characterized by equations (12), (13) and (14). In each of those, the distribution of rescaled labor requirement, the CDF $G(a)$ enters prominently. Spec-

Table 2: Intensive margin elasticities in 2000: 3 methods

Dependent variable:	firm-level exports					
Estimator:	2WFE on top		Tetrad on top		EK-Tobit	
	(1)	(2)	(3)	(4)	(5)	(6)
ln (1 + Applied Tariff)	-5.15^{b}	-2.96	-5.74^{a}	-3.83^{a}	-5.45^{a}	-5.44^{a}
	(2.03)	(2.06)	(0.76)	(0.71)	(0.26)	(0.26)
ln Distance	-0.51^{a}	-0.17^{a}	-0.47^{a}	-0.15^{a}	-1.73^{a}	-1.66^{a}
	(0.04)	(0.06)	(0.03)	(0.04)	(0.04)	(0.06)
Common language	-0.36^{a}	-0.10	0.10	0.39^{a}	1.86^{a}	1.93^{a}
	(0.11)	(0.12)	(0.09)	(0.09)	(0.12)	(0.12)
Contiguity	1.00^{a}	0.90^{a}	0.58^{a}	0.52^{a}	1.43^{a}	1.41^{a}
	(0.09)	(0.09)	(0.08)	(0.07)	(0.11)	(0.11)
Colony	0.46	-0.14	0.27	-0.24	3.50^{a}	3.41^{a}
	(0.54)	(0.54)	(0.29)	(0.29)	(0.21)	(0.21)
RTA		1.15^{a}		1.06^{a}		0.23^{c}
		(0.18)		(0.12)		(0.13)
ln # of dest. by firm					1.69^{a}	1.69^{a}
					(0.02)	(0.02)
Chinese exporter					0.59^{a}	0.63^{a}
					(0.06)	(0.05)
sigma						
Constant					8.68^{a}	8.69^{a}
					(0.02)	(0.02)
Observations	14044		37396		49067922	
R^2	0.777	0.779	0.137	0.143	0.829/0.08	0.829/0.08

Notes: ^{a}, ^{b} and ^{c} denote statistical significance levels of one, five and ten percent respectively. Standard errors are clustered by destination×reference country for columns (3) and (4) (tetrads), and by HS6-origin-destination for columns (5) and (6). Those two columns also add fixed effects for each of the 82 destination countries in the sample and compute the R^2 as the squared correlation between the predicted and actual values of the dependent variable. The second R^2 does the same calculation on positive trade flows.

ifying $G(a)$ is also necessary to inverse the $\mathcal{H}(a^*)$ function, and reveal the bilateral cutoffs required to compute the bilateral trade elasticities.

Pareto-distributed rescaled productivity $\varphi \equiv 1/a$ translates into a power law CDF for a, with shape parameter θ. A log-normal distribution of a retains the log-normality of productivity (with location parameter μ and dispersion parameter ν) but with a change in the log-mean parameter from μ to $-\mu$. Under those two distributional assumptions the CDFs for a are therefore given by

$$G^{\mathrm{P}}(a) = \left(\frac{a}{\bar{a}}\right)^{\theta}, \qquad \text{and} \qquad G^{\mathrm{LN}}(a) = \Phi\left(\frac{\ln a + \mu}{\nu}\right), \tag{19}$$

where we use Φ to denote the CDF of the standard normal. Simple calculations using (19) in (8), and detailed in Appendix 2, show that the resulting formulas for $\mathcal{H}$ are

$$\mathcal{H}^{\mathrm{P}}(a^*_{ni}) = \frac{\theta}{\theta - \sigma + 1}, \qquad \text{and} \qquad \mathcal{H}^{\mathrm{LN}}(a^*_{ni}) = \frac{h[(\ln a^*_{ni} + \mu)/\nu]}{h[(\ln a^*_{ni} + \mu)/\nu + (\sigma - 1)\nu]}, \tag{20}$$

where $h(x) \equiv \phi(x)/\Phi(x)$, the ratio of the PDF to the CDF of the standard normal.

Calculating $G^{\mathrm{P}}(.)$, $G^{\mathrm{LN}}(.)$, $\mathcal{H}^{\mathrm{P}}(.)$ and $\mathcal{H}^{\mathrm{LN}}(.)$ requires knowledge of underlying key supply-side distribution parameters θ and ν.[13] For those, we rely on the estimates from the QQ regressions represented in figure 1 combined with our estimate of the CES ($\hat{\sigma} = 5$) from the preceding section. The log-normal case is simple, since it is a very good fit to the overall distribution, we simply have $\hat{\nu}_{\mathrm{FRA}} = 2.392/(\hat{\sigma}-1) = 0.797$ and $\hat{\nu}_{\mathrm{CHN}} = 2.558/(\hat{\sigma} - 1) = 0.853$. The Pareto case is more tricky since the implied values of $\hat{\theta}$ for the overall estimation of the QQ regression are incompatible with finite values of the price index. We therefore concentrate on the part of the distribution where the Pareto QQ relationship is approximately linear with a slope satisfying $\hat{\theta} > \hat{\sigma} - 1$, that is the extreme right tail (like all papers we know of that estimate Pareto shape parameters on sales data). Concentrating on the top 1% of sales, we obtain $\hat{\theta}_{\mathrm{FRA}} = (1/0.779)(\hat{\sigma} - 1) = 5.134$ and $\hat{\theta}_{\mathrm{CHN}} = (1/0.618)(\hat{\sigma} - 1) = 6.472$.

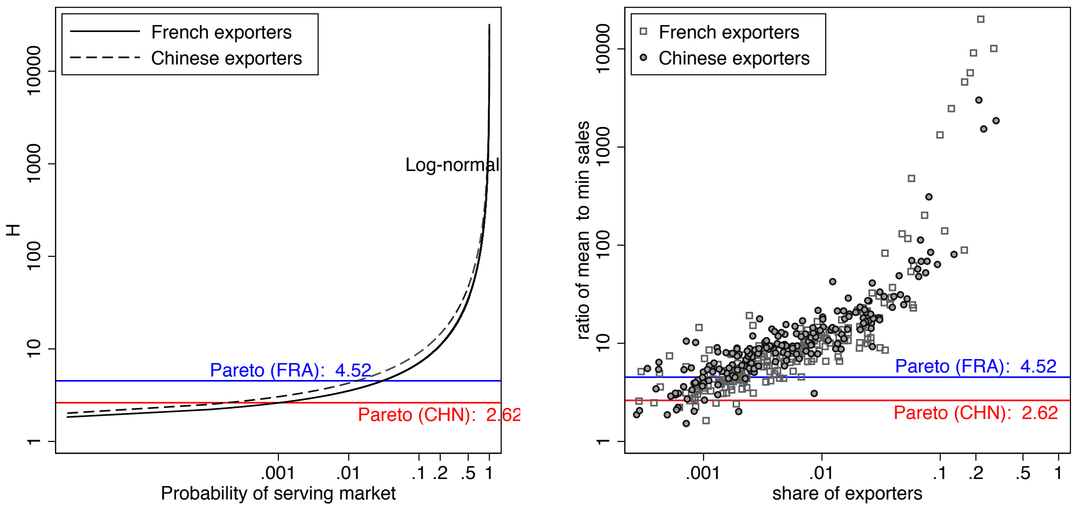
Panel (a) of Figure 2 depicts the theoretical relationship between the ratio of average to minimum sales, $\mathcal{H}(a^*_{ni})$, and the probability of serving the destination market , $G(a^*_{ni})$, spanning over values of the cutoff a^*_{ni}. Under Pareto heterogeneity, $\mathcal{H}$ is constant but this property of scale invariance is specific to the Pareto: Indeed it is increasing in G under log-normal. Panel (b) of figure 2 depicts the empirical counterpart of this relationship as observed for French and Chinese exporters in 2000 for all countries in the world. On the x-axis is the share of exporters serving each of those markets.[14] Immediately apparent is the non-constant nature of the mean-to-min ratio in the data, contradicting the Pareto prediction. This finding is very robust when considering alternatives to the minimum sales (which might be noisy if only because of statistical threshold effects) for the denominator of $\mathcal{H}$, that is different quantiles of the export distribution (results available upon request).[15]

[13] We show in Appendix 2 that the values taken by $\bar{a}$ and μ do not affect calculations of the trade elasticity.

[14] While this is not exactly the empirical counterpart of $G(a^*_{ni})$, the x-axis of panel (a), those two shares differ by a multiplicative constant, leaving the *shape* of the (logged) relationship unchanged.

[15] In a further effort to minimize noise in the calculation of the mean-to-min ratio, the figures are calculated for each of the 99 HS2 product categories and averaged. In the rest of the section, we will stick to this approach for the calculation of elasticities, done at the sector level before being averaged, which also simplifies exposition.

Figure 2: Theoretical and Empirical Mean-to-Min ratios



(a) theory (b) data

Figure 3 turns to the predicted trade elasticities under the two alternative distributions.[16] Functional forms (19) and (20) combined with (11), are used to deliver the two aggregate elasticities $\varepsilon^{\mathrm{P}}_{ni}$ and $\varepsilon^{\mathrm{LN}}_{ni}$:

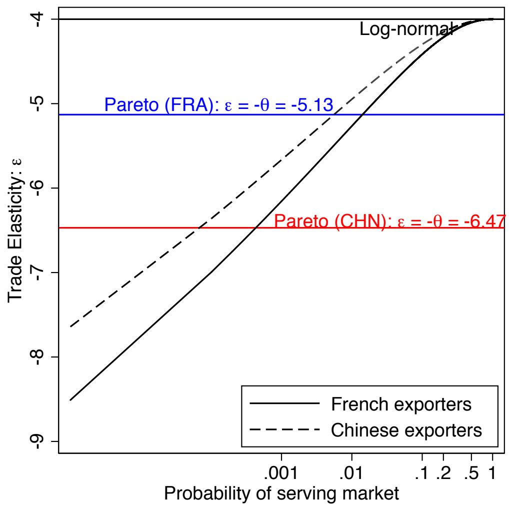
$$\varepsilon^{\mathrm{P}}_{ni} = -\theta, \qquad \text{and} \qquad \varepsilon^{\mathrm{LN}}_{ni} = 1 - \sigma - \frac{1}{\nu} h\left(\frac{\ln a^*_{ni} + \mu}{\nu} + (\sigma - 1)\nu\right). \tag{21}$$

Parallel to figure 2, panel (a) of figure 3 shows the theoretical relationship between those elasticities and $G(a^*_{ni})$, while panel (b) plots the same elasticities evaluated for each individual destination country against the empirical counterpart of $G()$. Again, the Pareto case has a constant prediction (one for each exporter), while log normal predicts a trade elasticity that is declining (in absolute value) with easiness of the market. Panel (b) confirms the large variance of trade elasticities according to the share of exporters that are active in each of the markets. It also shows that the response of aggregate flows to trade costs is reduced (in absolute value) when the market becomes easier. The intuition is that for very difficult markets, the individual reaction of incumbent firms is supplemented with entry of exporters selected among the most efficient firms. The latter effect becomes negligible for the easiest markets, yielding ε to approach the intensive margin. This mechanism becomes very clear when looking at the patterns of the extensive margin and average export elasticities in figure 4.

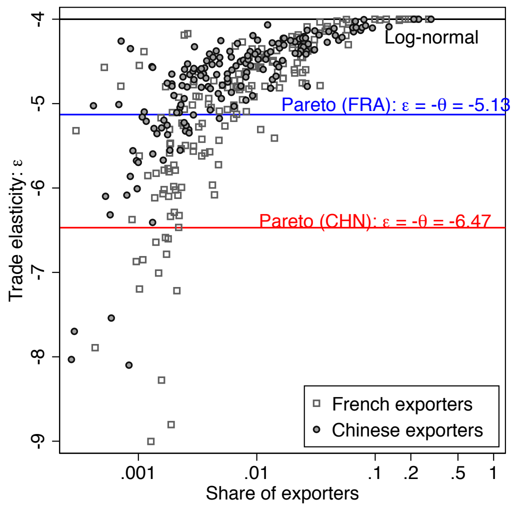
The predicted elasticity on the extensive margin is also rising with market toughness as shown in panel (a) of figure 4. The inverse relationship is true for average exports (panel b). When a market is very easy and most exporters make it there, the extensive margin goes to zero, and the response of average exports goes to the value of the intensive margin (the firm-level response), $1 - \sigma$, as shown in figure 4 when the share of exporters increase. While this should intuitively be true in general, Pareto does not allow for this change in elasticities across markets, since the response of average exports should be uniformly 0, while the total response is entirely due to the (constant) extensive margin. In Table

[16]See Appendix 2 for details.

Figure 3: Predicted trade elasticities: $\varepsilon_{n\text{FR}}$ and $\varepsilon_{n\text{CN}}$



(a) theory



(b) data

3, we compute the average value and standard deviation of bilateral trade elasticities calculated using log-normal, and presented in figures 3 and 4. The first column presents the statistics for the French exporters' sample, the second one is the Chinese exporters' case, and the last column averages those. The mean elasticities obtained vary slightly between France and China, but the dominant feature is that the total elasticity is in neither case confined to the extensive margin. In both cases, average exports are predicted to react strongly to trade costs, a pattern we will confirm on actual data in the next subsection.

Table 3: Predicted bilateral trade elasticities (LN distribution)

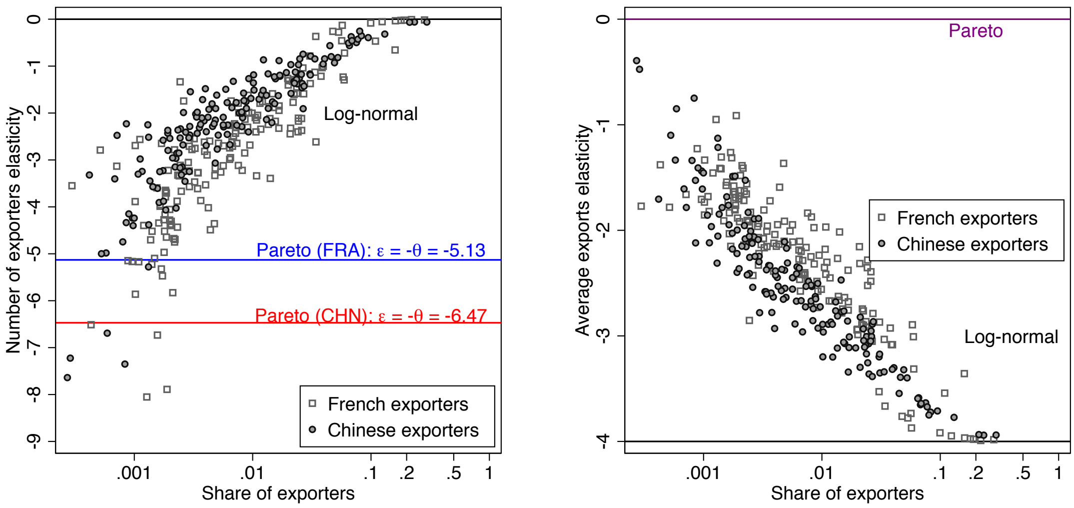
LHS	France	China	Average
Total flows	-5.14	-4.792	-4.966
	(1.069)	(.788)	(.742)
Number of exporters	-2.866	-2.274	-2.57
	(1.657)	(1.472)	(1.335)
Average flows	-2.274	-2.517	-2.396
	(.687)	(.731)	(.64)

Notes: This table presents the predicted elasticities (mean and s.d.) on total exports, the number of exporting firms, and average export flows. Required parameters are σ, the CES, and ν, the dispersion parameter of the log normal distribution.

Figure 5 groups our bilateral trade elasticities (ε_{ni}) into ten bins of export shares for both France and China in a way similar to empirical evidence by Novy (2013), which reports that the aggregate trade cost elasticity decreases with bilateral trade intensity.[17] The qualitative pattern is very similar here, with the

[17] Although Novy (2013) estimates variable distance elasticity, his section 3.4 assumes a constant trade costs to distance parameter to focus on the equivalent of our ε_{ni}.

Figure 4: Predicted elasticities: extensive and average exports

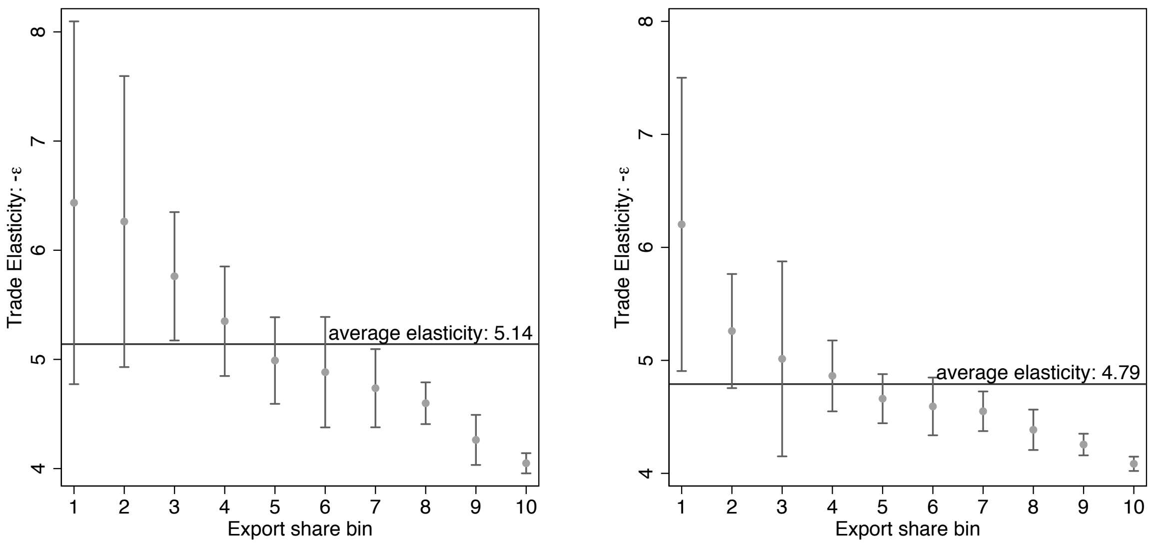


(a) extensive margin (b) average exports

bilateral elasticity decreasing in absolute value with the share of exports going to a destination. One can use this variance in ε_{ni} to quantify the error that a practitioner would make when assuming a constant response of exports to a trade liberalization episode. Taking China as an example, decreasing trade costs by one percent would raise flows by around 6.5 percent for countries like Somalia, Chad or Azerbaijan (first bin of Chinese exports) and slightly more than 4 percent for the USA and Japan (top bin). Since the estimate that would be obtained when imposing a unique elasticity would be close to the average elasticity (4.79), this would entail about 25 percent underestimate of the trade growth for initially low traders (1.7/6.5) and an overestimate of around 20 percent (0.8/4) for the top trade pairs.[18]

[18]We thank Steve Redding for suggesting this quantification.

Figure 5: Variance in trade elasticities: $\varepsilon_{n\mathrm{FR}}$ and $\varepsilon_{n\mathrm{CN}}$



(a) French exports (b) Chinese exports

5.2 Comparison with macro-based estimates of trade elasticities

We now can turn to empirical estimates of aggregate elasticities to be compared with our predictions. Those are obtained using aggregate versions of our estimating tetrad equations presented above, which is very comparable to the method most often used in the literature: a gravity equation with country fixed effects and a set of bilateral trade costs covariates, on which a constant trade elasticity is assumed.[19] Column (1) of Table 4 uses the same sample of product-markets as in our benchmark firm-level estimations and runs the regression on the tetrad of aggregate rather than individual exports. Column (2) uses the same covariates but on the count of exporters, and column (3) completes the estimation by looking at the effects on average flows.[20] An important finding is that the effect on average trade flow is estimated at -2.55, and is significant at the 1% level, contrary to the Pareto prediction (in which no variable trade cost should enter the equation for average flows).[21] This finding is robust to controlling for RTA (column 6) or constraining the sample to positive tariffs (column 9). The estimated median trade elasticity on total flows over all specifications at -4.79, is very close from the -5.03 found as the median estimate in

[19]Note that the gravity prediction on aggregate flows where origin, destination, and bilateral variables are multiplicatively separable and where there is a unique trade elasticity is only valid under Pareto. The heterogeneous elasticities generated by deviating from Pareto invalidate the usual gravity specification. Our intuition however is that the elasticity estimated using gravity/tetrads should be a reasonable approximation of the *average* bilateral elasticities. In order to verify this intuition, we run Monte Carlo simulations of the model with log-normal heterogeneity and find that indeed the average of micro-based heterogeneous elasticities is very close to the unique macro-based estimate in a gravity/tetrads equation on aggregate flows. Description of those simulations are in Appendix 3.

[20]Note that the three dependent variables are computed for each hs6 product-destination, and therefore that the average exports do not contain an extensive margin where number of products would vary across destinations.

[21]Fernandes et al. (2015) also show that the Melitz model combined with log-normal productivity can explain the reaction of average flows to distance. They refer to that response as the "intensive margin puzzle". We prefer to keep the terminology "intensive" for the firm-level response, and while we measure the trade elasticity directly through the impact of tariffs rather than distance, our results are totally in line with their main finding.

the literature by Head and Mayer (2014).

Table 4: Elasticites of total flows, count of exporters and average trade flows.

	Tot. (1)	# exp. (2)	Avg. (3)	Tot. (4)	# exp. (5)	Avg. (6)	Tot. (7)	# exp. (8)	Avg. (9)
ln (1+ Applied Tariff)	-6.84^a	-4.29^a	-2.55^a	-4.00^a	-1.60^b	-2.41^a	-4.79^a	-2.13^a	-2.66^a
	(0.82)	(0.66)	(0.54)	(0.73)	(0.63)	(0.50)	(0.84)	(0.50)	(0.54)
ln Distance	-0.85^a	-0.61^a	-0.24^a	-0.51^a	-0.28^a	-0.23^a	-0.85^a	-0.60^a	-0.25^a
	(0.04)	(0.03)	(0.02)	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)	(0.02)
Contiguity	0.62^a	0.30^a	0.32^a	0.53^a	0.21^a	0.32^a	0.64^a	0.35^a	0.29^a
	(0.12)	(0.09)	(0.06)	(0.11)	(0.07)	(0.05)	(0.12)	(0.09)	(0.06)
Colony	0.93^a	0.72^a	0.20^a	0.38^a	0.20^b	0.17^b	1.12^a	0.94^a	0.17^b
	(0.11)	(0.10)	(0.06)	(0.13)	(0.10)	(0.08)	(0.14)	(0.11)	(0.07)
Common language	0.09	0.16^c	-0.07	0.39^a	0.44^a	-0.05	-0.02	0.05	-0.07
	(0.09)	(0.09)	(0.06)	(0.09)	(0.08)	(0.06)	(0.10)	(0.08)	(0.06)
RTA				1.10^a	1.04^a	0.06			
				(0.11)	(0.06)	(0.07)			
Observations	99645	99645	99645	99645	99645	99645	41376	41376	41376
R^2	0.319	0.537	0.063	0.331	0.575	0.063	0.311	0.505	0.066
rmse	1.79	0.79	1.47	1.77	0.76	1.47	1.65	0.75	1.34

Notes: All estimations include fixed effects for each product-reference importer country combination. Standard errors are clustered at the destination-reference importer level. The dependent variable is the tetradic term of the logarithm of total exports at the hs6-destination-origin country level in columns (1), (4) and (7); of the number of exporting firms by hs6-destination and origin country in columns (2), (5) and (8) and of the average exports at the hs6-destination-origin country level in columns (3), (6) and (9). Applied tariff is the tetradic term of the logarithm of applied tariff plus one. Columns (7) to (9) present the estimations on the sample of positive tetraded tariffs and non-MFN tariffs. ^{a}, ^{b} and ^{c} denote statistical significance levels of one, five and ten percent respectively.

Under Pareto, the aggregate elasticity should reflect fully the one on the number of exporters, and there should be no impact of tariffs on average exports. This prediction of the Pareto distribution is therefore strongly contradicted by our results. As a first pass at assessing whether, the data support the log-normal predictions, we compare the (unique) macro-based elasticity obtained in Table 4, with the corresponding average of bilateral elasticities shown in Table 3 of the preceding sub-section. The numbers obtained are quite comparable when the effects of RTAs are taken into account (columns (4) to (6)) or with positive tetrad tariffs (columns (7) to (9)). Although this is not a definitive validation of the heterogenous firms model with log-normal distribution, our results clearly favor this distributional assumption over Pareto, and provides support for the empirical relevance of non-constant trade elasticities.

5.3 Direct evidence of non-constant trade elasticities

We can further use tetrads on aggregate trade flows in order to show *direct* empirical evidence of non-constant trade elasticities. Using aggregate bilateral flows from equation (3), and building tetrads with a procedure identical to the one used at the firm level, we obtain the (FR,CN, n, k)–tetrad of aggregate exports

$$\widetilde{X}_{\{n,k\}} \equiv \frac{X_{n\mathrm{FR}}/X_{k\mathrm{FR}}}{X_{n\mathrm{CN}}/X_{k\mathrm{CN}}} = \left(\frac{\tau_{n\mathrm{FR}}/\tau_{k\mathrm{FR}}}{\tau_{n\mathrm{CN}}/\tau_{k\mathrm{CN}}}\right)^{1-\sigma} \times \frac{V_{n\mathrm{FR}}/V_{k\mathrm{FR}}}{V_{n\mathrm{CN}}/V_{k\mathrm{CN}}} \tag{22}$$

Taking logs, differentiating with respect to tariffs and using the expression for the cutoff (5), we obtain

$$\begin{aligned} d\ln \widetilde{X}_{\{n,k\}} &= (1-\sigma-\gamma_{n\mathrm{FR}}) \times d\ln \tau_{n\mathrm{FR}} - (1-\sigma-\gamma_{k\mathrm{FR}}) \times d\ln \tau_{k\mathrm{FR}} \\ &\quad - (1-\sigma-\gamma_{n\mathrm{CN}}) \times d\ln \tau_{n\mathrm{CN}} + (1-\sigma-\gamma_{k\mathrm{CN}}) \times d\ln \tau_{k\mathrm{CN}}, \end{aligned} \tag{23}$$

where γ_{ni} is the elasticity of the cost performance index to a rise in the easiness of the market, defined in (7). For general distributions of heterogeneity, this elasticity is not constant across dyads as it depends on the dyad-specific cutoff a^*_{ni}. Hence, our interpretation of equation (23) is that the contribution to the (tetraded) total exports of a change in bilateral tariffs is larger for dyads that have a larger elasticity. Under Pareto, this elasticity is constant across dyads, $\gamma^{\mathrm{P}}_{ni} = 1-\sigma+\theta$. Combined with equation (23) this leads to

$$\widetilde{\varepsilon}^{\mathrm{P}}_{\{n,k\}} = \frac{d\ln \widetilde{X}_{\{n,k\}}}{d\ln \widetilde{\tau}_{\{n,k\}}} = -\theta, \tag{24}$$

where $\widetilde{\tau}_{\{n,k\}}$ is the vector of tetraded trade costs. This formula states that under Pareto, the elasticity of aggregate tetraded exports to tetraded tariffs is equal to the supply-side parameter θ. This transposes to the tetrad environment the well-known result of Chaney (2008) on gravity. Under non-Pareto heterogeneity, the four elasticities in (23) will remain different, a prediction we can put to a test. Results are shown in Table 5, where we pool observations for the years 2000 to 2006. Columns (1), (2) and (3) are the equivalent of the first three columns in Table 4, with the trade costs tetrads being split into its four components and the coefficients allowed to differ. The coefficients on tariffs to the destination country n show that the elasticity when considering France and China as an origin country differ significantly, consistent with the non-Pareto version of heterogeneity. Coefficients related to the reference importer k also differ significantly from each other, supporting further heterogeneity in the trade elasticities. A related approach is to confine identification on the destination country, neutralizing the change of reference country with a k fixed effect. Those results are shown in columns (4) to (6), where again most of the tariff elasticities differ across origin countries.[22]

[22]Table A.2 shows those same estimations for the two extreme years of our sample, 2000 and 2006, with significant evidence of non-constant elasticities in most cases.

Table 5: Non-constant trade elasticity

Dependent variable:	Tot.	# exp.	Avg.	Tot.	# exp.	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)
ln (1+Applied Tariff$_{n,\mathrm{FR}}$)	-4.25^a	-2.90^a	-1.34^a	-4.06^a	-2.76^a	-1.30^a
	(0.27)	(0.25)	(0.18)	(0.26)	(0.23)	(0.17)
ln (1+Applied Tariff$_{n,\mathrm{CN}}$)	3.43^a	1.87^a	1.56^a	3.30^a	1.76^a	1.53^a
	(0.27)	(0.25)	(0.17)	(0.26)	(0.23)	(0.17)
ln (1+Applied Tariff$_{k,\mathrm{FR}}$)	7.11^a	6.60^a	0.52^b			
	(0.36)	(0.20)	(0.24)			
ln (1+Applied Tariff$_{k,\mathrm{CN}}$)	-3.79^a	-2.14^a	-1.66^a			
	(0.40)	(0.26)	(0.23)			
Observations	1077652	1077652	1077652	1085643	1085643	1085643
R^2	0.346	0.587	0.080	0.349	0.593	0.081
rmse	2.41	1.01	2.05	2.41	1.01	2.05

Notes: All estimations include a product and year fixed effects and the four components (n, FR; n, CN; k, FR; and k, CN) of each gravity control (distance, common language, contiguity and colony). In all estimations standard errors are clustered at the destination-reference country and year level.

5.4 Micro and Aggregate elasticities at the industry level

As a last exercise, we provide evidence that both demand and supply determinants enter the aggregate elasticity by looking at industry-level estimates. For each good, we can estimate a firm-level and an aggregate elasticity to tariffs. Under the Pareto assumption, those two elasticities have no reason to be correlated, since the micro elasticity is a measure of (inverse) product differentiation, while the macro one is capturing homogeneity in firms' productive efficiency. Under alternative distributions like the log-normal, the aggregate elasticity includes both determinants and therefore should be correlated with the micro one (equation 1).

We run our micro and macro-level tetrad estimations over 2000-2006 for each 2-digit ISIC industry separately including destination-reference country and year fixed effects. Table 6 presents the results. Columns (1) and (2) show the coefficients for the micro-level elasticity while columns (3) and (4) report the estimates of the aggregate elasticity using the tetrad term of total exports by product-destination-reference country and year. Columns (2) and (4) restrict the sample to EU destinations. Each cell reports the coefficient on the applied tariffs tetrad by industry with associated degree of statistical significance.

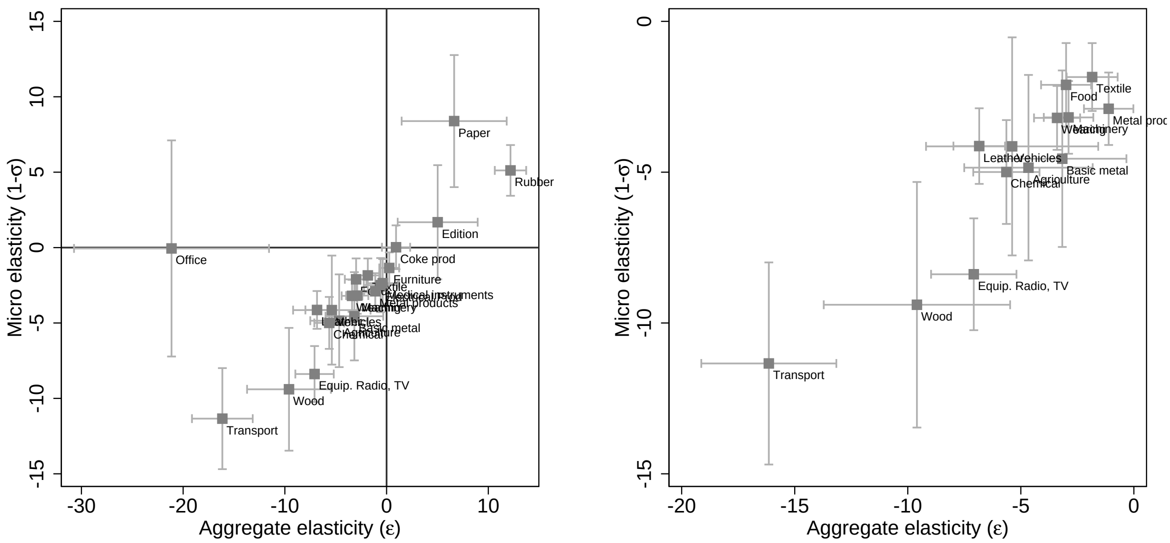
Estimates at the industry level yield coefficients of the intensive margin elasticity that average to -2.67 (column 1). The coefficients of the aggregate elasticity have a mean of -3.22 (column 3). More important for our main investigation, the intensive and aggregate elasticities are correlated (pairwise correlations are .68 for the full sample, and .74 for the EU one). Figure 6 shows graphical evidence of those correlations and exhibits overwhelming evidence in favor of the aggregate elasticity including demand side determinants.

Table 6: Micro and Aggregate elasticities by industry: 2000-2006

	Micro		Aggregate	
Dependent variable:	firm-level exports		total exports	
	(1)	(2)	(3)	(4)
	Full	EU	Full	EU
Agriculture	-4.85^{a}	-3.77	-4.66^{a}	-4.9^{c}
Food	-2.1^{a}	-3.09^{a}	-3^{a}	-3.57^{a}
Textile	-1.85^{a}	-5.74^{a}	-1.84^{a}	-2.97^{b}
Wearing	-3.2^{a}	-3.23^{a}	-3.4^{a}	-3.64^{a}
Leather	-4.13^{a}	-5.96^{a}	-6.84^{a}	-8.72^{a}
Wood	-9.4^{a}	-20.88^{a}	-9.59^{a}	-15.33^{a}
Paper	8.39^{a}	6.35	6.64^{b}	8.34
Edition	1.68	2.2	5.02^{b}	3.86
Coke prod	.02	1.39	.93	1.21
Chemical	-4.99^{a}	-7.29^{a}	-5.64^{a}	-7.66^{a}
Rubber	5.12^{a}	4.8^{a}	12.18^{a}	16.25^{a}
Basic metal	-4.55^{a}	-7.35^{a}	-3.17^{c}	-6.27^{c}
Metal products	-2.9^{a}	-6.33^{a}	-1.11^{c}	-3.32^{a}
Machinery	-3.18^{a}	-6.92^{a}	-2.89^{a}	-5.06^{a}
Office	-.06	-4.88	-21.14^{a}	-31.42^{b}
Electrical Prod	-2.49^{b}	-8.44^{a}	-.59	-1.34
Equip. Radio, TV	-8.38^{a}	-10.66^{a}	-7.08^{a}	-8.1^{a}
Medical instruments	-2.36^{a}	-3.9^{b}	-.38	-.29
Vehicles	-4.14^{c}	12.69^{c}	-5.38^{a}	15.52^{a}
Transport	-11.34^{a}	-18.28^{a}	-16.15^{a}	-17.37^{a}
Furniture	-1.35^{b}	-3.75^{a}	.26	0

Notes: All estimations are run by industry 2 digit. The cells report the coefficient on the applied tariffs tetrad by industry. All estimations include destination-reference country and year fixed effects. Standard errors are clustered by product-reference country and year. All estimations include a constant that is not reported. Applied tariff is the tetradic term of the logarithm of applied tariff plus one. ^{a}, ^{b} and ^{c} denote statistical significance levels of one, five and ten percent respectively.

Figure 6: Aggregate and intensive margin elasticities by industry 2 digit



(a) Full set

(b) Coeff sig. < 0

6 Conclusion

We argue in this paper that knowledge of the firm-level response to trade costs is key for understanding aggregate export reaction. In other words, we need micro-level data to uncover the macro-level impact of trade costs, a central element in any trade policy evaluation. This need for micro data is presumably true with the vast majority of possible heterogeneity distribution assumptions. There is one exception however where micro data is not needed: The (untruncated) Pareto distribution. The literature has been concentrating on that exception for reasons of tractability that are perfectly legitimate. In particular, it maintains the simple log-linear gravity equation with a constant trade elasticity to be estimated with macro data. However, the evidence presented in our paper points to systematic variation in bilateral aggregate trade elasticities that is both substantial and compatible with log-normal heterogeneity (in addition to be strongly preferred when looking at the micro-level distribution of export sales). We find in particular that the average values of bilateral trade elasticities obtained under a log-normal calibration are close to the empirical gravity estimates. By contrast, the Pareto-based calibration leads to predictions that seem invalidated by the data. Namely, the invariance of average shipments to ad-valorem tariff variations, the lack of correlation between firm-level and aggregate elasticities estimated industry by industry, and the constant aggregate trade elasticities. We are therefore tempted to call for a "micro approach" to estimating those elasticities as opposed to the "macro approach" that uses gravity specified so as to estimate a constant elasticity.

The micro- and macro- approaches differ substantially in several respects and both have positive and negative aspects. On the one hand, gravity is a more direct and parsimonious route for estimating aggregate elasticities: (i) parametric assumptions are reduced to a minimum while our micro-based procedure depends on the calibration of the productivity distribution; (ii) gravity is less demanding in terms of data and makes possible the use of easily accessible dataset of bilateral aggregate trade flows. On the down side, the log-linear specification of gravity is inconsistent with theory when dropping Pareto-heterogeneity. Also gravity provides, for each origin country, only a cross-destination average of elasticities while the micro-based approach provides the full cross-dyadic distribution of elasticities. Our Monte Carlo simulations show that gravity actually approximates this average of the true underlying bilateral elasticities quite decently. However, the gravity-predicted impact can be a bad approximation when trade liberalization occurs between countries that have either very low or very high levels of trade initially. When interested in policy experiments that involve this type of country pairs (distant and small countries or proximate and large ones for instance), one might strongly prefer the micro-approach.

References

Anderson, J. E. and van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, 93 (1):170–192.

Arkolakis, C. (2010). Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy*, 118(6):1151 – 1199.

Arkolakis, C., Costinot, A., and Rodríguez-Clare, A. (2010). Gains from trade under monopolistic competition: A simple example with translog expenditure functions and pareto distributions of firm-level productivity. manuscript.

Arkolakis, C., Costinot, A., and Rodriguez-Clare, A. (2012). New trade models, same old gains? *American Economic Review*, 102(1):94–130.

Atkeson, A. and Burstein, A. (2008). Pricing-to-market, trade costs, and international relative prices. *The American Economic Review*, 98(5):1998–2031.

Baier, S. and Bergstrand, J. H. (2001). The growth of world trade: tariffs, transport costs, and income similarity. *Journal of International Economics*, 53 (1):1–27.

Barba Navaretti, G., Bugamellli, M., Forlani, E., and Ottaviano, G. (2015). It takes (more than) a moment: Revisiting the link between firm productivity and aggregate exports. University of Milan Mimeo.

Berman, N., Martin, P., and Mayer, T. (2012). How do different exporters react to exchange rate changes? *The Quarterly Journal of Economics*, 127 (1):437–492.

Berthou, A. and Fontagné, L. (2015). Variable trade costs, composition effects, and the intensive margin of trade. Technical Report 543, Banque de France.

Broda, C. and Weinstein, D. (2006). Globalization and the gains from variety. *Quarterly Journal of Economics*, 121 (2):541–585.

Bury, K. (1999). *Statistical distributions in engineering*. Cambridge University Press.

Caliendo, L. and Parro, F. (2015). Estimates of the trade and welfare effects of nafta. *Review of Economic Studies*, 82(1):1–44.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review*, 98 (4):1707–1721.

Chatterjee, A., Dix-Carneiro, R., and Vichyanond, J. (2013). Multi-product firms and exchange rate fluctuations. *American Economic Journal: Economic Policy*, 5(2):77–110.

Costinot, A., Donaldson, D., and Komunjer, I. (2012). What goods do countries trade? a quantitative exploration of ricardo's ideas. *Review of Economic Studies*, 79(2):581–608.

Crozet, M., Head, K., and Mayer, T. (2012). Quality sorting and trade: Firm-level evidence for french wine. *Review of Economic Studies*, 79(2):609 – 644.

Eaton, J. and Kortum, S. (2001). Trade in capital goods. *European Economic Review*, 45(7):1195–1235.

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

Eaton, J., Kortum, S., and Kramarz, F. (2011). An anatomy of international trade: Evidence from French firms. *Econometrica*, 79(5):1453–1498.

Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *American Economic Review*, 84:157–177.

Feenstra, R. C. (2003). A homothetic utility function for monopolistic competition models. *Economics Letters*, 78(2):79–86.

Feenstra, R. C. (2013). Restoring the product variety and pro-competitive gains from trade with heterogeneous firms and bounded productivity. UC Davis Mimeo.

Fernandes, A., Klenow, P., Meleshchuk, S., Pierola, M. D., and Rodriguez-Clare, U. A. (2015). The intensive margin in trade: Moving beyond pareto. mimeo.

Fitzgerald, D. and Haller, S. (2015). Exporters and shocks: Dissecting the international elasticity puzzle. manuscript.

Freund, C. and Pierola, M. D. (2015). Export superstars. *Review of Economics and Statistics*.

Gopinath, G. and Neiman, B. (2014). Trade adjustment and productivity in large crises. *American Economic Review*, 104(3):793–831.

Greene, W. (2004). Fixed effects and bias due to the incidental parameters problem in the tobit model. *Econometric Reviews*, 23(2):125–147.

Guimares, P. and Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. *Stata Journal*, 10(4):628–649(22).

Hallak, J. C. (2006). Product quality and the direction of trade. *Journal of International Economics*, 68 (1):238–265.

Head, K. and Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook. In Helpman, E., Gopinath, G., and Rogoff, K., editors, *Handbook of International Economics*, volume 4. Elsevier.

Head, K., Mayer, T., and Ries, J. (2010). The erosion of colonial trade linkages after independence. *Journal of International Economics*, 81 (1):1–14.

Head, K., Mayer, T., and Thoenig, M. (2014). Welfare and trade without pareto. *American Economic Review*, 104(5):310–16.

Head, K. and Ries, J. (2001). Increasing returns versus national product differentiation as an explanation for the pattern of us-canada trade. *American Economic Review*, 91(4):858–876.

Helpman, E., Melitz, M., and Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics*, 123(2):441–487.

Hummels, D. (1999). Towards a geography of trade costs. Technical Report 17, GTAP Working Paper.

Imbs, J. and Méjean, I. (2015). Elasticity optimism. *American Economic Journal: Macroeconomics*, 7 (3):43–83.

Jawitz, J. W. (2004). Moments of truncated continuous univariate distributions. *Advances in Water Resources*, 27:269–281.

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *American Economic Review*, 70:950–959.

Li, H., Ma, H., and Xu, Y. (2015). How do exchange rate movements affect chinese exports?a firm-level investigation. *Journal of International Economics*, 97(1):148–161.

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

Melitz, M. and Ottaviano, G. (2008). Market size, trade, and productivity. *Review of Economic Studies*, 75(1):295–316.

Melitz, M. J. and Redding, S. J. (2015). New trade models, new welfare implications. *American Economic Review*.

Mulligan, C. B. and Rubinstein, Y. (2008). Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics*, pages 1061–1110.

Novy, D. (2013). International trade without CES: Estimating translog gravity. *Journal of International Economics*, 89(2):271–282.

Ossa, R. (2015). Why trade matters after all. *Journal of International Economics*.

Paravisini, D., Rappoport, V., Schnabl, P., and Wolfenzon, D. (2015). Dissecting the effect of credit supply on trade: Evidence from matched credit-export data. *The Review of Economic Studies*, 82(1):333–359.

Redding, S. J. (2011). Theories of heterogeneous firms and trade. *Annual Review of Economics*, 3:77–105.

Romalis, J. (2007). Nafta's and cusfta's impact on international trade. *Review of Economics and Statistics*, 89(3):416 – 435.

Simonovska, I. and Waugh, M. E. (2011). The elasticity of trade: Estimates and evidence. *Working Paper 16796, NBER.*

Yang, M.-J. (2014). Firm heterogeneity and trade in a distorted gravity world. University of Washington Mimeo.

Appendix 1 Empirical Appendix

A.1.1. Data

We combine French and Chinese firm-level datasets from the corresponding customs administrations which report export value by firm at the hs6 level for all destinations in 2000. The firm-level customs datasets are matched with data on tariffs effectively applied to each exporting country (China and France) at the same level of product disaggregation for each destination. Focusing on 2000 allows us to exploit variation in tariffs applied to each exporter country (France/China) at the product level by the importer countries since it precedes the entry of China into WTO at the end of 2001. We exploit the variation over time of trade and tariffs from 2000 to 2006 in a set of robustness checks contained in the working paper version.

Trade: The French trade data comes from the French Customs, which provide annual export data at the product level for French firms.[23] The customs data are available at the 8-digit product level Combined Nomenclature (CN) and specify the country of destination of exports. The free on board (f.o.b) value of exports is reported in euros and we converted those to US dollars using the real exchange rate from Penn World Tables for 2000. The Chinese transaction data comes from the Chinese Customs Trade Statistics (CCTS) database which is compiled by the General Administration of Customs of China. This database includes monthly firm-level exports at the 8-digit HS product-level (also reported f.o.b) in US dollars. The data is collapsed to yearly frequency. The database also records the country of destination of exports. In both cases, export values are aggregated at the firm-product(hs6)-destination level in order to match with applied tariffs information that are available at the origin-product(hs6)-destination level.

Tariffs: Tariffs come from the WITS (World Bank) database.[24] We rely on the ad valorem rate effectively applied at the hs6 level by each importer country to France and China. In our cross-section analysis performed for the year 2000 before the entry of China into the World Trade Organization (WTO), we exploit different sources of variation within hs6 products across importing countries on the tariff applied to France and China. The first variation naturally comes from the European Union (EU) importing countries that apply zero tariffs to trade with EU partners (like France) and a common external tariff to extra-EU countries (like China). The second source of variation in the year 2000 is that several non-EU countries applied the Most Favored Nation tariff (MFN) to France, while the effective tariff applied to Chinese products was different (since China was not yet a member of WTO).

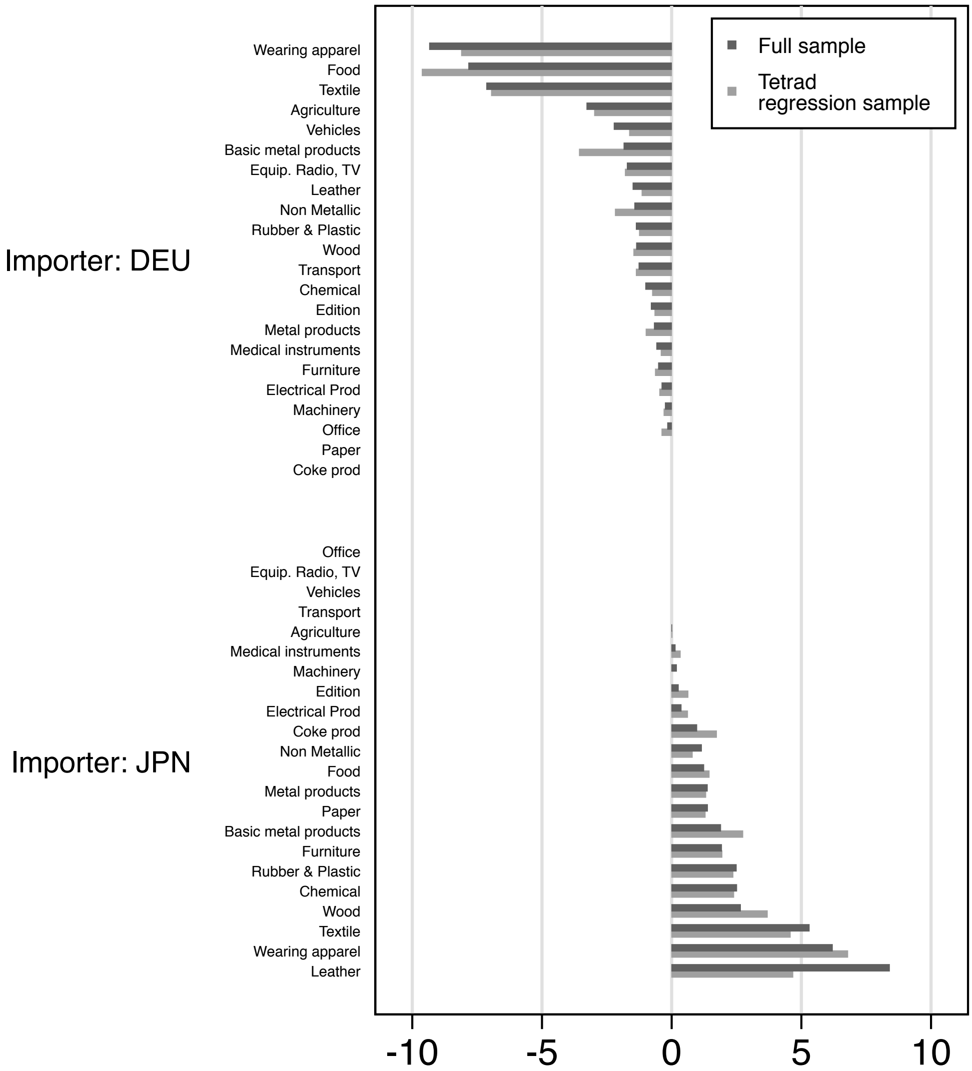
Gravity controls: In all estimations, we include additional trade barriers variables that determine bilateral trade costs, such as distance, common (official) language, colony and common border (contiguity). The data come from the CEPII distance database.[25] We use the population-weighted great circle distance between the set of largest cities in the two countries.

[23]This database is quite exhaustive. Although reporting of firms by trade values below 250,000 euros (within the EU) or 1,000 euros (rest of the world) is not mandatory, there are in practice many observations below these thresholds.

[24]Information on tariffs is available at `http://wits.worldbank.org/wits/`

[25]This dataset is available at `http://www.cepii.fr/anglaisgraph/bdd/distances.htm`

Figure A.1: Average percentage point difference between the applied tariff to France and China across industries by Germany and Japan (2000)



Source: Authors' calculation based on Tariff data from WITS (World Bank).

Table A.1: Average percentage point difference between the applied tariff to France and China across industries by Germany and Japan (2000)

Reference importer:	Germany		Japan	
	Full sample	Tetrad regression sample	Full sample	Tetrad regression sample
Agriculture	-3.27	-2.98	.01	.02
Food	-7.83	-9.63	1.24	1.45
Textile	-7.14	-6.95	5.31	4.58
Wearing apparel	-9.34	-8.11	6.2	6.79
Leather	-1.5	-1.15	8.4	4.68
Wood	-1.36	-1.47	2.66	3.69
Paper	0	0	1.39	1.3
Edition	-.79	-.65	.26	.64
Coke prod	0	0	.97	1.73
Chemical	-1.01	-.74	2.51	2.4
Rubber & Plastic	-1.37	-1.25	2.5	2.37
Non Metallic	-1.43	-2.18	1.16	.8
Basic metal products	-1.84	-3.56	1.89	2.75
Metal products	-.67	-.99	1.38	1.32
Machinery	-.25	-.3	.19	0
Office	-.16	-.38	0	0
Electrical Prod	-.38	-.47	.37	.62
Equip. Radio, TV	-1.72	-1.79	0	0
Medical instruments	-.58	-.41	.14	.34
Vehicles	-2.22	-1.63	0	0
Transport	-1.27	-1.37	0	0
Furniture	-.51	-.63	1.93	1.95

A.1.2. Non-constant trade elasticity

Table A.2: Non-constant trade elasticity

	2000			2006		
Dependent variable:	Tot.	# exp.	Avg.	Tot.	# exp.	Avg.
	(1)	(2)	(3)	(4)	(5)	(6)
Applied Tariff$_{n,\mathrm{FR}}$	-5.74^{a}	-3.41^{a}	-2.34^{a}	-4.31^{a}	-3.52^{a}	-0.79^{b}
	(1.02)	(0.69)	(0.72)	(0.53)	(0.57)	(0.37)
Applied Tariff$_{n,\mathrm{CN}}$	5.08^{a}	2.55^{a}	2.53^{a}	2.16^{a}	1.55^{a}	0.61
	(0.99)	(0.71)	(0.66)	(0.57)	(0.58)	(0.38)
Observations	99745	99745	99745	218036	218036	218036
R^2	0.357	0.590	0.093	0.339	0.594	0.072
rmse	2.42	1.00	2.02	2.34	0.98	2.01

Notes: All estimations include a product and reference country fixed effects and the four components (nFR, nCN, kFR, and kCN) of each gravity control (distance, common language, contiguity and colony). In all estimations standard errors are clustered at the destination-reference country.

Appendix 2 Theoretical derivations of V, γ and $\mathcal{H}$ under Pareto and log-Normal distributions

The central relationship (6) makes it clear that the heterogeneity of aggregate trade elasticity comes entirely from the term γ_{ni} that stems from endogenous selection of firms into export markets (see equation 7). In turn, γ_{ni} depends on the cost-performance index V_{ni} as defined by

$$V_{ni} \equiv \int_0^{a^*_{ni}} a^{1-\sigma} g(a) da.$$

We therefore need to understand how these γ and V terms behave under alternative distributional assumptions on productivity, i.e. Pareto and log-normal. One important advantage of Pareto, pointed out by Redding (2011) is that if φ is Pareto (θ) then φ^r is Pareto also. The shape parameter becomes θ/r. This advantage is shared by the log-normal. If φ is log-$\mathcal{N}(\mu, \sigma^2)$ then φ^r is log-$\mathcal{N}(r\mu, r^2\sigma^2)$. This follows from a more general reproductive property reported by (Bury, 1999, p. 156).

If productivity is Pareto then the rescaled unit input requirement a has PDF $g(a) = \theta a^{\theta-1}/\bar{a}^{\theta}$, which translates into

$$V^{\mathrm{P}}_{ni} = \frac{\theta a^{*\theta-\sigma+1}_{ni}}{\bar{a}^{\theta}(\theta-\sigma+1)}. \tag{A.1}$$

The elasticity of V^{P}_{ni} with respect to a^* is

$$\gamma^{\mathrm{P}}_{ni} = \theta - \sigma + 1 > 0. \tag{A.2}$$

Hence, Pareto makes all the γ_{ni} terms be the same, and therefore transforms an expression generally yielding heterogeneous trade elasticities into a one-parameter elasticity $\frac{d\ln X_{ni}}{d\ln \tau_{ni}} = \theta$, that is related to the supply side of the economy only.

When productive efficiency is distributed log-normally, things are very different.[26] For φ ∼log-$\mathcal{N}(\mu,\nu)$, the distribution of rescaled unit input requirements is a ∼log-$\mathcal{N}(-\mu,\nu)$. (Jawitz, 2004, Table 1) expresses m_r, the absolute rth truncated moment in terms of the error function (erf). We convert his expression to be in terms of the more familiar, $\Phi()$, the CDF of the standard normal, using the relationship $\text{erf}(x) = 2\Phi(x\sqrt{2}) - 1$. For $x \sim$ log-$\mathcal{N}(\mu,\nu)$ truncated between lower limit ℓ and upper limit u the Jawitz formula can be expressed as

$$m_r = \exp(r\mu + r^2\nu^2/2)\left[\Phi\left(\frac{\ln u - \mu - r\nu^2}{\nu}\right) - \Phi\left(\frac{\ln \ell - \mu - r\nu^2}{\nu}\right)\right]$$

We are considering the distribution of α which is the inverse of productivity so it has distribution log-$\mathcal{N}(-\mu,\nu)$ and it has a lower limit $\ell = 0$ and an upper limit $u = \alpha^*$. The limit of $\Phi(x)$ as $x \to -\infty$ is zero so the second term involving $\ln \ell$ disappears. Replacing μ with $-\mu$ and r with $1-\sigma$, we obtain:

$$V_{ni}^{\text{LN}} = \exp[(\sigma-1)\mu + (\sigma-1)^2\nu^2/2]\Phi[(\ln a_{ni}^* + \mu)/\nu + (\sigma-1)\nu], \tag{A.3}$$

Differentiating $\ln V_{ni}^{\text{LN}}$ with respect to $\ln a_{ni}^*$,

$$\gamma_{ni}^{\text{LN}} = \frac{1}{\nu}h\left(\frac{\ln a_{ni}^* + \mu}{\nu} + (\sigma-1)\nu\right), \tag{A.4}$$

where $h(x) \equiv \phi(x)/\Phi(x)$, the ratio of the PDF to the CDF of the standard normal. Thus γ_{ni} is no longer the constant $1-\sigma+\theta$ which obtains for productivity distibuted Pareto with shape parameter θ. Bilateral elasticities therefore write as

$$\varepsilon_{ni}^{\text{P}} = -\theta, \qquad \text{and} \qquad \varepsilon_{ni}^{\text{LN}} = 1-\sigma-\frac{1}{\nu}h\left(\frac{\ln a_{ni}^* + \mu}{\nu} + (\sigma-1)\nu\right). \tag{A.5}$$

The $\mathcal{H}$ function is a central element of our calibration procedure, as summarized by relationship (9), that reveals cutoffs and therefore aggregate bilateral elasticities. Comparing (4), (7) and (8) we see that $\mathcal{H}$ and γ are closely related

$$\gamma_{ni} \times \mathcal{H}(a_{ni}^*) = a_{ni}^* \frac{g(a_{ni}^*)}{G(a_{ni}^*)} \tag{A.6}$$

With Pareto, we make use of (A.2) to obtain

$$\mathcal{H}^{\text{P}}(a_{ni}^*) = \frac{\theta}{\theta-\sigma+1}, \tag{A.7}$$

[26] There are a number of useful properties of the normal distribution that we use here:

1. $\Phi(-x) = 1 - \Phi(x)$
2. $\phi(-x) = \phi(x)$
3. $\phi'(x) = -x\phi(x)$
4. $\Phi'(x) = \phi(x)$
5. $\partial(x\Phi(x) + \phi(x))/\partial x = \Phi(x) > 0$

With a log-normal productivity, equation (A.4) leads to

$$\mathcal{H}^{\mathrm{LN}}(a^*_{ni}) = \frac{h[(\ln a^*_{ni} + \mu)/\nu]}{h[(\ln a^*_{ni} + \mu)/\nu + (\sigma - 1)\nu]}, \tag{A.8}$$

Table A.3 summarizes all formulas for the variables used in this paper under both distributions.

An attractive feature of our quantification procedure relates to the small number of relevant parameters to be calibrated. Under Pareto, equations (A.2) and (A.7) show that only the shape parameter θ matters. Similarly, under a log-normal, only the calibration of the second-moment of the distribution, ν, is necessary for inverting the $\mathcal{H}$ function to reveal the cutoff and for quantifying the aggregate elasticity: This last point stems from the fact that shifting the first moment, μ, affects (A.4) and (A.8) in an identical way and so has no impact on the quantification.

Table A.3: Pareto vs Log-Normal: key variables.

Variable	Pareto	Log-Normal
PDF: $g(a)$	$\frac{\theta a^{\theta-1}}{\bar{a}^\theta}$	$\phi\left(\frac{\ln a+\mu}{\nu}\right)/a\nu$
CDF: $G(a)$	$\frac{a^\theta}{\bar{a}^\theta}$	$\Phi\left(\frac{\ln a+\mu}{\nu}\right)$
$V_{ni}(a^*) \equiv \int_0^{a^*} a^{1-\sigma} g(a) da$	$\frac{\theta a^{*\theta-\sigma+1}_{ni}}{\bar{a}^\theta(\theta-\sigma+1)}$	$\exp\left[(\sigma-1)\mu + \frac{(\sigma-1)^2\nu^2}{2}\right]\Phi\left[\frac{(\ln a^*_{ni}+\mu)}{\nu} + (\sigma-1)\nu\right]$
$\gamma_{ni} \equiv \frac{d\ln V_{ni}}{d\ln a^*_{ni}}$	$\theta - \sigma + 1$	$\frac{1}{\nu}h\left(\frac{\ln a^*_{ni}+\mu}{\nu} + (\sigma-1)\nu\right)$
$\frac{d\ln X_{ni}}{d\ln \tau_{ni}} = 1 - \sigma - \gamma_{ni} = \epsilon_{ni}$	$-\theta$	$1 - \sigma - \frac{1}{\nu}h\left(\frac{\ln a^*_{ni}+\mu}{\nu} + (\sigma-1)\nu\right)$
$\mathcal{H}(a^*) \equiv \frac{V(a^*)}{a^{*1-\sigma}G(\alpha^*)}$	$\frac{\theta}{\theta-\sigma+1}$	$\frac{h[(\ln a^*_{ni}+\mu)/\nu]}{h[(\ln a^*_{ni}+\mu)/\nu+(\sigma-1)\nu]}$

Note: The Pareto parameters of the unit input requirement distributions are θ and $\bar{\alpha}$. For the log normal distribution, when $\varphi \sim$log-$\mathcal{N}(\mu,\nu)$, the distribution of rescaled unit input requirements is $a \sim$log-$\mathcal{N}(-\mu,\nu)$. We define $h(x) \equiv \phi(x)/\Phi(x)$, a non-increasing function.

Appendix 3 Micro/Macro based estimations: Monte Carlo Evidence

In Section 5.2 we find that the macro-based estimate of the aggregate trade elasticity is quantitatively close to the cross-dyadic average of the micro-based estimates when heterogeneity is calibrated as being log-normal. We interpret this finding as an empirical support in favor of this distributional assumption. In this section we substantiate this last statement by embracing a more theoretical perspective. This is an important step in the argument because the theoretical relationship between the macro- and the micro-based estimates of the elasticities is unknown (except under Pareto where they are unambiguously equal). Hereafter we provide simulation-based evidence that the similarity between micro- and macro-based estimates is not accidental, even under log-normal heterogeneity.

We proceed with Monte Carlo (MC) simulations of our generic trade model with heterogeneous firms.

In the baseline simulations we generate fake bilateral trade for 10 countries and 1 million active firms per country. Our data generating process uses the firms' sales in equation (2). Firm-level heterogeneity in terms of rescaled labor requirement, $a \equiv \alpha \times b(\alpha)$ is assumed to be Pareto or Log-normally distributed with a set of parameters identical to the ones used in our empirical analysis (section 5.1). We also retain $\sigma = 5$ as the parameter for the intensive margin. Without loss of generality, in this partial equilibrium framework, we normalize the nominal wage, $w = 1$, and we draw A_n/f_{ni}, i.e. the dyadic ratio of destination n attractiveness over entry cost from a log-normal distribution. This distribution is calibrated such as to match an average dyadic share of exporting firms of 10 percent. Finally the applied-tariffs $\tau_{ni} = 1 + t_{ni}$ are drawn from an uniform distribution over the range $[1, 2]$.

In each MC draw, we first generate a matrix of firm-level trade flows that are non-zero when sales exceed the bilateral entry cost, i.e. $x_{ni}(a) > \sigma w f_{ni}$. In a first stage we infer from this fake trade dataset the micro-based estimates of the aggregate trade elasticities by applying the methodology of Section 5.1: We first retrieve min-to-mean ratios for all country-pairs and then compute the corresponding set of theoretical dyadic elasticities (equations 10 and 11). In a second stage, we turn to the macro-based estimates of the trade elasticity. To this purpose we collapse firm-level trade flows at the country-pair level to construct a matrix of bilateral aggregate trade. We then run gravity regressions (both using country fixed effects and tetrads) and retrieve the point estimate of applied tariffs. Hence, for each draw, we obtain one macro-based estimate of the trade elasticity that we compare to the cross-dyadic average of the micro-based elasticities. This procedure is replicated 50 times. Notice that it is computationally demanding as we have to manipulate very large trade matrices (1 million firms $\times$ 10 origin countries $\times$ 10 destination countries) .

Table A.4: Monte Carlo results: elasticities wrt to a change in trade costs

Distribution:	Log-Normal				Pareto			
# firms per country:	1K	10K	100K	1M	1K	10K	100K	1M
total exports (micro)	-4.69	-4.57	-4.56	-4.55	-5.74	-5.36	-5.23	-5.18
	(0.60)	(0.38)	(0.34)	(0.33)	(0.87)	(0.37)	(0.21)	(0.13)
total exports (macro/tetrads)	-4.80	-4.60	-4.59	-4.58	-5.55	-5.37	-5.21	-5.18
	(0.66)	(0.29)	(0.09)	(0.03)	(0.81)	(0.56)	(0.35)	(0.22)
total exports (macro/FE)	-4.65	-4.57	-4.55	-4.55	-5.59	-5.31	-5.22	-5.17
	(0.20)	(0.09)	(0.03)	(0.01)	(0.29)	(0.16)	(0.15)	(0.08)
nb exporters (macro/FE)	-3.20	-3.16	-3.16	-3.16	-5.19	-5.15	-5.14	-5.13
	(0.09)	(0.03)	(0.01)	(0.00)	(0.13)	(0.05)	(0.02)	(0.01)
avg. exports (macro/FE)	-1.45	-1.41	-1.40	-1.39	-0.40	-0.15	-0.08	-0.04
	(0.17)	(0.08)	(0.03)	(0.01)	(0.26)	(0.16)	(0.15)	(0.08)

Notes: 50 replications for each cell, parameters on fixed costs of exports and size of the demand term have been calibrated so the share of exporters averages to 10-11% in all simulations. For each elasticity, the first line reports the average value. Standard deviations are in parentheses. For the micro elasticity, the number in parentheses is the average of standard deviations of the elasticity in each draw (quantifying the degree of heterogeneity in bilateral elasticities). For the macro elasticities, we report the standard deviation of elasticities across the 50 replications.

The simulation results are displayed in Table A.4 for log-normal (col.1-col.4) and Pareto (col.5-col.8) and for different degrees of firm scarceness (from 1000 to 1 million firms per country). Each column reports averages and standard errors across replications.

Our baseline simulation under Pareto (col. 8) shows that the simulated economy with 1 million firms conforms to the theoretical prediction of a model with a continuum of firms. The micro-based estimates of the aggregate trade elasticity are relatively homogeneous across dyads (the second row reports the mean value of the standard deviation within each draw) and their average (first row) is close to the macro-based estimates of the elasticities retrieved from tetrad-like specification (third row) or standard gravity (fourth row). Finally the elasticity of the average export (last row) is not significantly different from zero, as expected from the theoretical prediction associated with Pareto heterogeneity and a continuum of firms. We conclude from this exercise that scarceness does not seem to play a central role in our fake sample of 1 million firms with 10 percent of exporters.

From the baseline simulation under log-normal (Column 4) we see that the macro-based estimate of the aggregate elasticity and the cross-dyadic average of the micro-based estimates are quantitatively very close - i.e. equality cannot be rejected. This constitutes the main result of our Monte Carlo approach. It confirms that the similarity between micro- and macro-based estimates in section 5.2 can be safely interpreted as supportive of the log-normal distribution. Notice that the magnitude of the simulation results on the three macro-based elasticities (total exports, count of exporters and average exports) is also close to what we obtain with the sample of French and Chinese firms. This is remarkable given that our Monte Carlo approach is minimal and shares only few features with the true data, i.e. the parameters of firm-level heterogeneity and the share of exporters.