A Tractable Model of Trade with Flexible Cost Structure *

Konstantin Kucheryavyy† Ruo Shangguan‡

U of Tokyo Waseda University

June 7, 2021

Abstract

We introduce variable marginal costs of exporting in a heterogeneous firms trade model *à la* Melitz (2003). In our setup, the marginal costs of exporting depend on the quantity shipped in addition to the standard iceberg trade costs. Under the Pareto distribution of firms' productivities, our model implies a tractable gravity equation and an expression for welfare gains from trade, for which the Arkolakis *et al.* (2012) formula for gains from trade is a special case. This costs structure can be micro-found through a firm's inventory management problem, and the key parameter can be estimated using the frequency of shipment of exporters. Under the log-normal distribution of firms' productivities, we calibrate all trade costs using Chinese transaction-level data. The ad-valorem equivalent rate for logistics costs is minor for productive firms, but it is substantial for less productive firms.

Keywords: trade costs, heterogeneous firm model, welfare gains from trade

JEL classification: F12, F13, F14

*We are grateful to Michal Fabinger, Hidehiko Ichimura, Taiji Furusawa, Peter Schott, Sam Kortum, and the seminar participants at the Online Workshop on International Trade and FDI at Kobe University for insightful discussions.

†Graduate School of Public Policy, University of Tokyo. Email: kucher@pp.u-tokyo.ac.jp

‡Institute for Empirical Research in Organizational Economics, Waseda University. Email: roy1312@gmail.com

1 Introduction

The structure and magnitude of trade costs have important implications for various questions in international trade. To name a few: how firms make entry decisions, how much they sell in each destination, and what price they charge for the products exported. Answers to these questions, in their turn, have important implications for the magnitudes of welfare gains from trade. In this paper, we study how a richer trade costs structure introduced into a heterogeneous firms trade model *à la* Melitz (2003) affects our understanding of these questions. In the model, firms need to pay logistics costs to export, in addition to iceberg trade costs and fixed trade costs.We model the level of required logistics costs as a power function of the quantities traded. As a result, marginal trade costs depend on the exported amount of goods.

The first contribution of this paper is to derive the formula for welfare changes under a richer trade cost structure. Because in this model, iceberg trade cost is no longer the only variable trade cost, it is natural to expect the elasticity of logistics costs to play a significant role. In the symmetric two-country world, we show that the pass-through of zero-profit cutoff into welfare depends on the quantity elasticity of logistics costs and the demand elasticity, instead of a one-to-one pass-through in the case of no logistics costs. Notably, the sufficient statistics of gains from trade are composed of three elasticities, the trade elasticity, the demand elasticity, and the quantity elasticity of logistics costs. To see it more clearly, we further impose that firm-level productivity follows the Pareto distribution and derive the explicit formula for gains from trade. The resulted formula is simple and includes the ACR formula for gains from trade as a particular case. The result illustrates the importance of the microstructure of trade costs as three macro restrictions proposed in Arkolakis *et al.* (2012, ACR) are satisfied.

We use the extended economic order quantity model proposed in Fabinger and Weyl (2018) to micro-found the logistics costs through an inventory management problem. In this problem, each firm chooses an optimal frequency of shipment that balances the transaction costs and the inventory costs. Intuitively, if shipment frequency is high, the total transaction cost, which is the sum of transaction costs across shipments, will be increased.

At the same time, frequent shipment keeps the inventory level, and therefore storage costs, at a low level.

The micro foundation provides a straightforward way to estimate logistics cost elasticity via linear regression of log shipment frequency on log trade volume. Using the data on the universe of international trade transactions of Chinese exporters, we find the elasticity of logistics costs to be 0.6, which shows a clear rejection of it is equal to zero, a standard assumption in the literature. With the estimated trade costs elasticity, we compare the prediction of gains from trade, measured by the change in real income associated with moving to autarky, between the ACR formula and our formula. On average, we observe 11.5% lower welfare gains from trade.

To further evaluate the impact of a different trade cost structure, we calibrate the trade costs from observed trade flows assuming that productivity parameters across firms follow a log-normal distribution, using a method inspired by Fernandes *et al.* (2019) and Head *et al.* (2014). More specifically, we match the firm-level distributions of export sales and shipment frequency across origin-destination pairs. The distributions of firm-level shipment frequency help us to identify the parameter that governs the magnitude of logistics costs. Consistent with its increasing return to scale nature, we find that the logistics costs are small for the most productive firms, while it can account for more than 35% of revenue for least productive firms.

Using the calibrated model, we calculate the required level of iceberg costs for each firm at each destination, such that the logistics costs are zero and the operating profits remain at the same level. Then the difference between the implied and the calibrated iceberg trade costs can be viewed as the equivalent rate of logistics costs. In this way, we find that the impact of logistics costs are highly heterogeneous across firms. While the most productive firm incurs the logistics costs with an ad valorem equivalent rate of 1.9%, the least effective firm pays the logistics costs with an ad valorem rate of 174.9%.

Our analysis builds on Arkolakis *et al.* (2012, ACR) and Melitz and Redding (2015, MR). In their seminal paper, ACR shows that under some conditions, the welfare gains from trade for a large class of models can be summarized by two statistics, namely the expenditure share on domestic goods and the trade elasticity. In a paper closely related to

ACR, MR derives the formula of welfare gains from trade in a heterogeneous firm model under less restrictive assumptions. We extend their result under a different microstructure of trade costs.

Many authors have explored the implications of trade cost structure beyond the simple combination of iceberg trade costs and fixed trade costs. For example, Alessandria *et al.* (2010) use cost per shipment to explain lumpiness in international trade transactions. Hummels and Skiba (2004) use costs per unit to explain why high-quality goods tend to be shipped further away. Kropf and Sauré (2014) estimate the magnitude of fixed cost per shipment using an extension of the standard Melitz model. The effort of analyzing the implications of a richer cost structure, however, is limited by the difficulty that models quickly become intractable once additional costs are added. Our paper contributes to this literature by showing a way to increase the flexibility of trade cost structure while keeping the tractability.

The current paper is also related to the literature on inventory management (e.g., Alessandria *et al.* (2010); Blum *et al.* (2019)). A key difference is that we model the inventory problem on the exporter side instead of importers, emphasizing the productivity differences of producers. And to characterize the general equilibrium explicitly, we abstract from considerations of the uncertainty of demand and delay in transportation.

2 A Trade Model with Flexible Trade Costs

In this section we introduce the theoretical model. In order to make the intuition clearer, we first restrict the model to a closed economy in Section 2.1, then discuss the impact of a more flexible cost structure on welfare gains from trade in two symmetric countries case in Section 2.2. To simplify notation, we normalize wage to 1 in both sections. In Appendix A, the model is generalized to an environment with asymmetric countries.

2.1 Closed Economy

The setup of demand is the same as the Melitz (2003) model. The consumers in the economy are homogeneous, each maximizes a constant elasticity of substitution (CES) utility function by choosing consumption quantities across a variety of goods. This leads to the revenue function[1]

$$X_d(q) = Aq^{\nu_R}, \tag{1}$$

where q is the consumption quantity, $0 < \nu_R < 1$ is related to the elasticity of substitution between varieties, σ, through the relation $\nu_R = \frac{\sigma-1}{\sigma}$. Let L and P denote the total population and price index in the economy, the parameter A, defined as $A = P^{\nu_R}L^{1-\nu_R}$, summarizes the market condition.

Each firm takes am efficiency parameter a from a distribution $G(a)$. The production technology features constant return to scale, so that the total variable production costs for producing q units is aq. To deliver the product to the final consumer, the firm needs to incur logistics costs, which is equal to

$$C_{LT}(q) = B_d q^{\nu_{LT}}, \tag{2}$$

where B_d is a parameter determining the level of logistics costs in the domestic market, ν_{LT} determines how fast the logistics costs increase with quantity shipped. We make the following assumption about ν_{LT}:

Assumption 1. $0 \leq \nu_{LT} < \nu_R$.

The introduction of logistics costs is our main departure from the literature. Under Assumption 1, the marginal logistics costs is decreasing in quantity. Depending on the value of ν_{LT}, logistics costs capture the implications of different trade cost structures. When $\nu_{LT} = 0$, the logistics costs become constant, and can be merged with the fixed costs, therefore in that case the model reduces to the standard Melitz model. Because of this, and for the analytical tractability, we assume that fixed costs are zero for the theoretical discussion. When we perform quantitative analysis in Section 3, we remove the

[1] See Appendix A for a more detailed specification of preferences.

restriction of zero fixed costs to match the extensive margin of trade flows observed in the data. As a limiting case, when $\nu_{LT} = 1$, the logistics costs per unit are constant, thus it becomes additive trade costs emphasized by several papers in the literature of trade costs (e.g. Hummels and Skiba 2004, Irarrazabal *et al.* 2015). In Appendix C.1, we provide a micro-foundation of the logistic costs via an inventory management problem.

The profit function of a firm with efficiency equal to a is therefore

$$\pi = Aq^{\nu_R} - B_d q^{\nu_{LT}} - aq, \tag{3}$$

which implies a first-order condition

$$\frac{\partial \pi}{\partial q} = A\nu_R q^{\nu_R - 1} - B_d \nu_{LT} q^{\nu_{LT} - 1} - a = 0, \tag{4}$$

and a second-order condition

$$\frac{\partial^2 \pi}{\partial q^2} = A\nu_R \left(\nu_R - 1\right) q^{\nu_R - 2} - B_d \nu_{LT} \left(\nu_{LT} - 1\right) q^{\nu_{LT} - 2} < 0. \tag{5}$$

We can use the first-order condition (4) to express

$$a = A\nu_R q^{\nu_R - 1} - B_d \nu_{LT} q^{\nu_{LT} - 1}. \tag{6}$$

This expression implies that a can be thought of as a function of q, i.e. $a\left(q\right)$. This is one of the key insights that significantly simplifies the theoretical analysis that follows. Combining expression (6) for a together with the condition that firm should earn non-negative profits, gives

$$A\left(1 - \nu_R\right) q^{\nu_R} - B_d \left(1 - \nu_{LT}\right) q^{\nu_{LT}} \geq 0.$$

So that using Assumption 1, firm's profits are nonnegative if and only if $q \geq q_d^*$, where

$$q_d^* \equiv \left[\frac{B_d\left(1 - \nu_{LT}\right)}{A\left(1 - \nu_R\right)}\right]^{\frac{1}{\nu_R - \nu_{LT}}} \tag{7}$$

is the zero-profit quantity. Next, the second-order condition (5) holds if and only if

$$q^{\nu_R-\nu_{LT}} > \left(q_d^*\right)^{\nu_R-\nu_{LT}} \frac{\nu_{LT}}{\nu_R},$$

which holds for any $q \geq q_d^*$ under Assumption 1. Finally, it is straightforward to check that $a'(q) < 0$ if and only if the second-order condition (5) holds.

Hence, $a(q)$ is a monotone function defined for all $q \geq q_d^*$ that provides a one-to-one mapping between firm's marginal cost a and the optimal quantity q. This, in turn, implies that there is a unique profit-maximizing quantity q for each cost $a \in \left(0, a_d^*\right)$, where $a_d^* \equiv a\left(q_d^*\right)$. In addition, we can write the optimal quantity $q = q(a)$ without causing confusion.

The introduction of positive ν_{LT} impacts the relative prices between firms having different productivity. When $\nu_{LT} = 0$, it is easy to verify that the relative price is equal to the relative productivity. When $\nu_{LT} > 0$, the same proportion of advantage in marginal cost translates into a bigger difference in prices, due to the ability of taking advantage of the scale of economy in the transportation technology by productive firms. To see this more formally, consider two firms with $a_1 < a_2$, so that firm 1 has higher productivity. By Equation (6) and the monotonicity of $a(q)$,

$$\frac{a_1}{a_2} > \left(\frac{q_1}{q_2}\right)^{\nu_R-1} = \frac{p_1}{p_2},$$

where the equality follows from the definition of price, $p(q) = X_d(q)/q = Aq^{\nu_R-1}$.[2]

To solve the model, use the free entry condition

$$\int_0^{a_d^*} \pi_d(q(a))\, dG(a) = f^e, \tag{8}$$

[2] On the other hand, observe that the markup is equal to

$$\frac{p(q)}{VC'(q)} = \frac{1}{\nu_R},$$

where $VC(q)$ is the total variable cost that includes both logistics and production costs. This shows that the price has a constant markup over the marginal variable cost. The important difference from the standard model is that the marginal variable cost now depends on the quantity produced.

which states that the expected profit of entering the market should be equal to the entry cost f^e. Together with equations (6) and (7), this equation gives the solution to the zero-profit quantity q_d^*. The price index, thus the welfare level, is connected to the zero-profit quantity via the relation

$$P = \left[\frac{B_d \left(1 - \nu_{LT}\right)}{\left(q_d^*\right)^{\nu_R - \nu_{LT}} L^{1-\nu_R} \left(1 - \nu_R\right)} \right]^{\frac{1}{\nu_R}}. \tag{9}$$

2.2 Two Symmetric Countries

Next we discuss the implications of opening to trade on the welfare in a world with two symmetric countries. Since countries are symmetric, the market condition parameter, A, are equal in both countries, wages are equalized and normalized to 1. Again we assume that fixed costs of export are equal to zero. We use the subscript d to denote quantities related to the domestic market, and the subscript x to denote quantities related to the foreign market. Each firm can export to the foreign market subject to a iceberg cost τ, with the domestic iceberg trade cost normalized to 1, and the logistics cost parameter in the foreign market is B_x. We assume that the domestic market faces less friction in logistics, or $B_d < B_x$.

The goal in this subsection is to characterize the welfare changes from trade given any change in the trade costs, namely change in τ or B_x. Despite the additional complexity associated with the introduction of the logistics costs, the analysis largely remains tractable. In particular, we derive explicit formulas of welfare gains from trade. Moreover, the elasticity of trade with respect to logistics costs is shown to be equally informative about the gains from trade as the elasticity of trade with respect to iceberg trade costs.

Using similar argument as in the previous subsection, the first-order conditions of firm's profit maximization problem in each market define the one-one mappings between the marginal cost and the optimal quantity in each market as

$$a_d\left(q\right) = A\nu_R q^{\nu_R - 1} - B_d \nu_{LT} q^{\nu_{LT} - 1}, \tag{10}$$

$$a_x(q) = \frac{A\nu_R q^{\nu_R - 1} - B_x \nu_{LT} q^{\nu_{LT}-1}}{\tau}. \tag{11}$$

The inverse functions that define the mappings from the marginal costs to the optimal quantity are denoted as $q_d(a)$ and $q_x(a)$. The zero profit conditions for the domestic and the foreign market imply that

$$A = \frac{B_d(1-\nu_{LT})}{(q_d^*)^{\nu_R - \nu_{LT}}(1-\nu_R)} = \frac{B_x(1-\nu_{LT})}{(q_x^*)^{\nu_R - \nu_{LT}}(1-\nu_R)}, \tag{12}$$

where q_d^* is the zero-profit cutoff quantity in the domestic market, and q_x^* is the zero-profit cutoff in the foreign market. Plug in this relationship into equations (10), (11) to get

$$a_d^* = \left[\frac{(1-\nu_{LT})}{(1-\nu_R)}\nu_R - \nu_{LT}\right] B_d (q_d^*)^{\nu_{LT}-1}, \tag{13}$$

$$a_x^* = \frac{\left[\frac{(1-\nu_{LT})}{(1-\nu_R)}\nu_R - \nu_{LT}\right] B_x (q_x^*)^{\nu_{LT}-1}}{\tau}. \tag{14}$$

The second equality in Equation (12) shows that two cutoff quantities satisfy the relation

$$q_x^* = \left(\frac{B_x}{B_d}\right)^{\frac{1}{\nu_R - \nu_{LT}}} q_d^*, \tag{15}$$

which also shows that the foreign market has a higher entry threshold given Assumption 1 and $B_d < B_x$. Therefore the cutoff marginal costs satisfy the relation

$$a_x^* = \frac{1}{\tau}\left(\frac{B_d}{B_x}\right)^{\left(\frac{1-\nu_R}{\nu_R - \nu_{LT}}\right)} a_d^*. \tag{16}$$

Together with the free entry condition

$$\int_0^{a_d^*} \pi_d(q(a))\, dG(a) + \int_0^{a_x^*} \pi_x(q(a))\, dG(a) = f^e, \tag{17}$$

we can solve for the zero profit cutoffs a_d^* and a_x^*, and therefore the market condition parameter A. Other aggregated variables of interest can be expressed by the cutoffs. Using

Equation (12) and the definition $A = P^{\nu_R} L^{1-\nu_R}$, the real consumption can be derived as

$$W \equiv \frac{1}{P} = \left(\frac{B_d \left(1 - \nu_{LT}\right)}{\left(q_d^*\right)^{\nu_R - \nu_{LT}} \left(1 - \nu_R\right)} \right)^{-\frac{1}{\nu_R}} L^{\frac{1-\nu_R}{\nu_R}}. \tag{18}$$

Take logarithmic difference to this equation. Since we only consider the effects of changes in τ and B_x on the welfare, it reduces to

$$d \ln W = \left(\frac{\nu_R - \nu_{LT}}{\nu_R} \right) d \ln q_d^* = -\left(\frac{\nu_R - \nu_{LT}}{\nu_R \left(1 - \nu_{LT}\right)} \right) d \ln a_d^*. \tag{19}$$

The change in welfare is therefore summarized by the change in the domestic zero profit cutoff. Moreover, when $\nu_{LT} > 0$, the pass-through elasticity of change in the cutoff to the change in welfare is determined by two elasticities ν_R and ν_{LT}.

To build connection between welfare changes and data, we derive expressions for the key empirical moments proposed by ACR, namely the expenditure share on the domestic goods and the trade elasticity. The expenditure share on the domestic goods is

$$\begin{aligned} \lambda &= \frac{\int_0^{a_d^*} X_d \left(q\left(a\right)\right) dG\left(a\right)}{\int_0^{a_d^*} X_d \left(q\left(a\right)\right) dG\left(a\right) + \int_0^{a_x^*} X_x \left(q\left(a\right)\right) dG\left(a\right)} \\ &= \frac{\int_0^{a_d^*} q_d^{\nu_R}\left(a\right) dG\left(a\right)}{\int_0^{a_d^*} q_d^{\nu_R}\left(a\right) dG\left(a\right) + \int_0^{a_x^*} q_x^{\nu_R}\left(a\right) dG\left(a\right)}. \end{aligned} \tag{20}$$

When $\nu_{LT} > 0$, both the iceberg trade costs and the logistics costs serve as the variable trade costs, therefore we need to define both elasticities. Moreover, since most papers estimate the elasticities from the gravity equation, which exploits cross sectional variation in the bilateral frictions controlling for the origin and the destination fixed effects, we hold domestic environment, namely a_d^* as constant in our definition of elasticities. To be explicit, we treat the domestic expenditure share $\lambda = \lambda\left(a_d^*, a_x^*, \tau, B_x\right)$ as a function of two cutoff quantities and trade costs parameters, and the foreign market cutoff $a_x^* = a_x^*\left(a_d^*, \tau, B_x\right)$ as a function of the domestic cutoff and trade costs, as defined in Equation (16). Finally, the domestic cutoff $a_d^* = a_d^*\left(\tau, B_x\right)$ is a function of the trade costs parameters.

The elasticity with respect to τ is

$$
\begin{aligned}
\vartheta_\tau &\equiv -\left.\frac{d\ln\left(\frac{1-\lambda}{\lambda}\right)}{d\ln\tau}\right|_{a_d^*} \\
&= -\left[\frac{\partial\ln\left(\int_0^{a_x^*} q_x^{\nu_R}(a)\,dG(a)\right)}{\partial\ln\tau} + \frac{\partial\ln\left(\int_0^{a_x^*} q_x^{\nu_R}(a)\,dG(a)\right)}{\partial\ln a_x^*}\frac{\partial\ln a_x^*}{\partial\ln\tau}\right] \\
&= -\left[\frac{\partial\ln\left(\int_0^{a_x^*} q_x^{\nu_R}(a)\,dG(a)\right)}{\partial\ln\tau} + \frac{\partial\ln\left(\int_0^{a_x^*} q_x^{\nu_R}(a)\,dG(a)\right)}{\partial\ln a_x^*}(-1)\right], \qquad (21)
\end{aligned}
$$

where the first equality results from the restriction of q_d^* being constant, and the second equality uses Equation (16). Similarly, the elasticity with respect to B_x is

$$
\begin{aligned}
\vartheta_B &\equiv -\left.\frac{d\ln\left(\frac{1-\lambda}{\lambda}\right)}{d\ln B_x}\right|_{a_d^*} \\
&= -\left[\frac{\partial\ln\left(\int_0^{a_x^*} q_x^{\nu_R}(a)\,dG(a)\right)}{\partial\ln B_x} + \frac{\partial\ln\left(\int_0^{a_x^*} q_x^{\nu_R}(a)\,dG(a)\right)}{\partial\ln a_x^*}\frac{\partial\ln a_x^*}{\partial\ln B_x}\right] \\
&= -\left[\frac{\partial\ln\left(\int_0^{a_x^*} q_x^{\nu_R}(a)\,dG(a)\right)}{\partial\ln B_x} + \frac{\partial\ln\int_0^{a_x^*} q_x^{\nu_R}(a)\,dG(a)}{\partial\ln a_x^*}\left(-\frac{1-\nu_R}{\nu_R-\nu_{LT}}\right)\right], \qquad (22)
\end{aligned}
$$

where the second equality uses Equation (16) as well.

Using the definition of price index, we have

$$
P^{\frac{-\nu_R}{1-\nu_R}} = N^e A^{\frac{-\nu_R}{1-\nu_R}}\left(\int_0^{a_d^*} q_d^{\nu_R}(a)\,dG(a) + \int_0^{a_x^*} q_x^{\nu_R}(a)\,dG(a)\right), \qquad (23)
$$

where N^e is the mass of potential entrants. Using the definition of real consumption, and Equation (20), one can show that the real consumption is given by

$$
W = L^{-1}\left[N^e\frac{\int_0^{a_d^*} q_d^{\nu_R}(a)\,dG(a)}{\lambda}\right]^{\frac{1}{\nu_R}}. \qquad (24)
$$

Together with Equation (19), this shows that the change in real consumption can be expressed as

$$d\ln W = \frac{1}{\nu_R\left(1+\gamma\left(a_d^*\right)\left(\frac{1-\nu_{LT}}{\nu_R-\nu_{LT}}\right)\right)}\left(d\ln N^e - d\ln\lambda\right),$$

where $\gamma\left(x\right) \equiv \frac{d\ln\int_0^x q_d^{\nu_R}(a)dG(a)}{d\ln x}$. Because $\int_0^{a_d^*} q_d^{\nu_R}\left(a\right)dG\left(a\right)$ is proportional to the market share of the domestic firms, $\gamma\left(a_d^*\right)$ measures the sensitivity of market share with respect to the domestic zero profit cutoff.

Using the definitions of the elasticities in equations (21) and (22), the changes in welfare can be further expressed as either

$$d\ln W = \frac{\left(d\ln N^e - d\ln\lambda\right)}{\nu_R\left(1+\left[\gamma\left(a_d^*\right)-\gamma\left(a_x^*\right)+\vartheta_\tau+\frac{\partial\ln\left(\int_{q_x^*}^\infty q^{\nu_R}dG(a(q))\right)}{\partial\ln\tau}\right]\left(\frac{1-\nu_{LT}}{\nu_R-\nu_{LT}}\right)\right)}, \tag{25}$$

or

$$d\ln W = \frac{\left(d\ln N^e - d\ln\lambda\right)}{\nu_R\left(1+\left[\gamma\left(a_d^*\right)-\gamma\left(a_x^*\right)+\left(\frac{\nu_R-\nu_{LT}}{1-\nu_R}\right)\left[\vartheta_B+\frac{\partial\ln\left(\int_{q_x^*}^\infty q^{\nu_R}dG(a(q))\right)}{\partial\ln B_x}\right]\right]\left(\frac{1-\nu_{LT}}{\nu_R-\nu_{LT}}\right)\right)}. \tag{26}$$

Similar to the formula derived in MR, equations (25), (26) show that in the general case, the change in mass of firms could affect welfare, and the difference in the sensitivity $\gamma\left(a_d^*\right)-\gamma\left(a_x^*\right)$ also matters. Different from MR, we highlight that a different trade cost structure, as modeled through a positive ν_{LT}, could also affect the welfare changes. This effect is mediated via both a different transmission rate of changes in zero profit cutoffs as shown in Equation (19) and a different direct effects contained in the elasticities $\frac{\partial\ln\left(\int_{q_x^*}^\infty q^{\nu_R}dG(a(q))\right)}{\partial\ln\tau}$, $\frac{\partial\ln\left(\int_{q_x^*}^\infty q^{\nu_R}dG(a(q))\right)}{\partial\ln B_x}$. Moreover, the equations show that the elasticity with respect to logistics costs is equally informative about the welfare gains from trade, therefore suggesting a different way of quantifying the effect of trade liberalization.[3]

[3]When $\nu_{LT}=0$, it is straightforward to verify the the elasticity of trade respect to the fixed costs can also be used to calculate welfare gains from trade.

2.2.1 Solutions Under Pareto Distribution

We derive explicit solutions by imposing the assumption that the inverse of marginal cost follows Pareto distribution given by the cumulative density function

$$G(a) = \theta^{-1} (\kappa_G a)^{\theta}, \tag{27}$$

where $\theta > 0$ and $a \in \left[0, \kappa_G^{-1} \theta^{\frac{1}{\theta}}\right]$. Under this assumption, it can be shown that the aggregated optimal quantity of any power ν in the foreign market, subject to the condition $\frac{\nu + \nu_R \theta - \theta}{\nu_{LT} - \nu_R} > 0$, is given by

$$\int_0^{a_x^*} q_x^{\nu} dG(a) = \kappa_G^{\theta} \tau^{-\theta} (B_x)^{\theta} (q_x^*)^{\nu + \nu_{LT}\theta - \theta} \left(\frac{(1-\nu_{LT})}{(1-\nu_R)} \right)^{-\frac{\nu+\nu_{LT}\theta-\theta}{\nu_R - \nu_{LT}}} H(\nu, \nu_R, \nu_{LT}, \theta) \tag{28}$$

where $H(\nu, \nu_R, \nu_{LT}, \theta)$ is a positive-valued function that depends only on parameters of the model. The detailed derivation of this formula in any market can be found in Appendix B. In order to simplify exposition, we drop the later three arguments in the $H(\cdot)$ function and write $H(\nu)$ to represent $H(\nu, \nu_R, \nu_{LT}, \theta)$.

Using Equation (28), we see that[4]

$$\int_0^{a_d^*} q_d^{\nu_R}(a)\, dG(a) = \kappa_G^{\theta} (B_d)^{\theta} (q_d^*)^{\nu_R + \theta\nu_{LT} - \theta} H(\nu_R) \left(\frac{(1-\nu_{LT})}{(1-\nu_R)} \right)^{-\frac{\nu_R+\nu_{LT}\theta-\theta}{\nu_R - \nu_{LT}}},$$

$$\int_0^{a_x^*} q_x^{\nu_R}(a)\, dG(a) = \kappa_G^{\theta} (B_x)^{\theta} \tau^{-\theta} (q_x^*)^{\nu_R + \theta\nu_{LT} - \theta} H(\nu_R) \left(\frac{(1-\nu_{LT})}{(1-\nu_R)} \right)^{-\frac{\nu_R+\nu_{LT}\theta-\theta}{\nu_R - \nu_{LT}}},$$

from which we can calculate the elasticities as $\vartheta_\tau = \theta$, $\vartheta_B = \frac{\nu_R + \theta\nu_R - \theta}{\nu_{LT} - \nu_R}$, and [5]

$$d \ln W = -\frac{d \ln \lambda}{\nu_R \left(\frac{\nu_{LT} + \theta\nu_{LT} - \theta}{\nu_{LT} - \nu_R} \right)}. \tag{29}$$

The formula clearly illustrates that the trade elasticity ϑ_τ is no longer the sufficient statis-

[4]It is straightforward to apply the formula to the domestic market by replacing the corresponding market specific variables.

[5]Applying the formula to the labor market clearing condition, as shown in more detail in Appendix A, we show that the aggregate profits are a constant share of the aggregate revenue, and therefore $d \ln N^e = 0$.

tics for calculating welfare changes, even under an environment where all three macro restrictions proposed by ACR are satisfied. The key difference comes from the trade costs structure, or the introduction of ν_{LT}. Indeed, readers can easily verify that the formula collapses to the ACR formula when $\nu_{LT} = 0$.

3 Empirical analysis

3.1 Data

The empirical analysis mainly uses Chinese custom data. The data contains the universe of international trade transactions of Chinese firms for each month from 2000 to 2006. Each transaction is described in details by variables including the year-month when the transaction happens, the import/export dummy, 8-digit harmonized system (HS) code for product classification, 10-digit company identification number, the quantity and the value of the goods, the destination, the mode of transportation, and the mode of trade.[6] This data has been used for many papers, so we omit too much data description for the sake of brevity. For more information about the data, see for example, Bai *et al.* (2017). In order to measure the number of firms that sell domestically, we use the annual industrial survey data produced by the National Bureau of Statistics. Fore more information of this dataset, see Brandt *et al.* (2012). Table 1 shows number of exporters, products measured by HS8 codes, number of destinations, and total number of shipments in each year.

In our main sample, we use export transactions by manufacturing firms only. Intermediary firms are excluded as they do not incur production costs.[7] To reduce the noise from small export destinations, the sample only includes top 100 destinations in terms of export value, which accounts for more than 99% of the total export value.

[6]The mode of trade includes 18 categories, with the top 2 categories being the ordinary trade and the processing trade. The other variables include the unit of goods, the port and the route of the trade, the origin city, the zip code, and the type of the firm (with the main categories being state owned or private).

[7]We use the keywords of firm name provided in Ahn *et al.* (2011) to identify intermediary firms.

3.2 Welfare Gains from Trade

In this subsection, we quantify the gains from trade liberalization the formula (29) [8], and compare it with the gains derived using the ACR formula. As shown in more detail in Appendix C, using a simple extension of the standard Economic Order Quantity model proposed in Fabinger and Weyl (2018), we can link the logistics costs to the shipment frequency via an inventory management problem. More specifically, the inventory management problem gives us a relation between the optimal shipment frequency F from origin i to destination j and the total quantity q to be shipped as

$$F_{ij}(q) = \frac{1}{1-\nu_{LT}} \kappa_I B_{ij}^{-1} q^{(1-\nu_{LT})}, \tag{30}$$

where κ_I is the parameter related to inventory costs. Intuitively, in the model, in order to export a given quantity, the marginal cost of inventory is constant, while the marginal cost of coordination is decreasing in quantity. A firm chooses the optimal shipment frequency that equalizes the marginal cost of inventory and the marginal cost of coordination. The elasticity ν_{LT} measures how fast the coordination cost per shipment increases with the total quantity to be shipped.

Equation (30) allows us to estimate ν_{LT} by regressing the logarithm of shipment frequency to the logarithm of traded quantity. Given the detailed definition of a transaction, we use the number of transactions as the measure of the number of shipments. This idea is implemented in Appendix C.2, which gives us a very precise estimate of ν_{LT} close to 0.6. With the estimated ν_{LT}, we can compare gains from trade across models. In Table 2, we show the gains from trade that move from the autarky at the observed level of expenditure share on domestic goods for a number of countries. We compare the gains from trade under the assumption that $\nu_{LT} = 0$ and $\nu_{LT} = 0.6$. The former value is assumed in the standard trade models, while the latter value is supported by the estimation results using the shipment frequency data. ν_R is held to be 0.8, which is equivalent to elasticity of substitution equal to 5, consistent with the large literature estimating this parameter.

[8]The formula hold in an environment with asymmetric countries. For details, see Equation (59) in Appendix A.

The data of expenditure share comes from the WIOD in 2008 (see Timmer *et al.* 2015). The model under $\nu_{LT} = 0.6$ gives a lower welfare gains from trade. The difference is higher for countries like Hungary where the share of expenditure on domestic goods is lower. Overall, we observe a 11.5% lower welfare gains from trade on average.

3.3 Model Calibration

In this section, we show how to calibrate other model parameters, in particular those related to trade costs, using trade data. The goal of the calibration exercise is to quantify the significance of logistics costs in terms of rationalizing the observed trade flows.

3.3.1 Target Moments

This subsection presents the data moments that we use to calibrate trade costs parameters. For the calibration exercise, we aggregate the custom data to firm-destination level, and the sample restricts to sea transportation in order to avoid the potential heterogeneity in logistics costs across modes of transportation. For computational reason, we use firm level trade data from top 10 Chinese provinces[9] to top 8 destinations.[10] Since China has a broad territory, the distance at the country level could be misleading. For example, according to Google Map, while Dalian, a north-east city, is 1,645 km away from Tokyo, Chengdu, a mid-east city, has a distance about 3,353 km to Tokyo. To get a more precise distance measure for firm's exports, we split the firms according to their origin of province, and calculate the geographic distance between any province and destination using the latitude and gratitude information. More specifically, we calculate the distance between any cites, using the haversine formula, then calculate the population weighted average across all cities within any province-destination pair. The location and population data is from the website simplemap.[11]

[9]Provinces include Guangdong, Jiangsu, Shanghai, Zhejiang, Shandong, Fujian, Tianjin, Beijing, Liaoning, Hebei

[10]Destinations include the US, Japan, Korea, Germany, Netherlands, Singapore, UK, and Canada. We exclude Hong Kong since it is subject to a lot of re-export, whose real destinations are not clear.

[11]We use the basic database from https://simplemaps.com/data/world-cities

Figure 1 and Figure 2 show the empirical distributions of firm sales and shipment frequency from Guangdong province to its three popular destinations (the US, Japan, and Germany). The panels on the left show the histograms of firm-level sales and shipment frequency. As the export sales decreases, the export shipment frequency decreases as well. The panels on the right show the corresponding cumulative distribution functions (CDF), with the dashed lines being the CDFs of the normal random variables with the same mean and variance as the log export value or log shipment frequency. The log-normal distribution well explains the distribution of export value and has a reasonable fit for the distribution of shipment frequency, despite its discrete nature. The cross-destination variations in the firm distributions conditional on the origin province play a significant role in identification of trade costs. Within each province, the wage and productivity distribution is held constant, therefore the shift in export value and shipment frequency must be driven by changes in τ_{ij}, B_{ij} and the measure of market size A_j. While the last term A_j can be identified from the destination fixed effects as we include multiple origins.

Figure 3 shows the share of exporters that export to each destination from Guangdong. Note there are substantial variation across destinations. For example, abut 50% of exporters export to the US via sea, but only about 15% of exporters export to South Korea. This variation will allow us to calibrate the relative level of fixed export costs. Note that since we do not have a good estimate for the potential mass of entrants, the absolute level of fixed costs is not identified. To see this, if we observe a low number of exporters, it could be either due to a small mass of entrants, or a high level of fixed costs.

Finally, Table 3 shows gravity regressions using all provinces and all destinations in year 2006. The regressions put total export value X_{ij}, total number of firms N_{ij}, total shipment frequency F_{ij}, and value per firm, value per shipment on the left-hand side, the bilateral distance on the right-hand side after taking logarithmic transformation, controlling for the origin plus the destination fixed effects. We see that all the five margins decreases significantly with distance. As is widely used in the literature, the estimates of gravity equations provide us with information on the magnitude of the underlying trade costs.

3.3.2 Calibration Procedures

In order to match the target empirical patterns, we extend the model such that each firm takes a draw of fixed export cost in each destination independently. Namely the profit function now becomes

$$\pi_{ij}(q) = A_j q^{\nu_R} - w_i B_{ij} q^{\nu_{LT}} - a w_i \tau_{ij} q - w_i \xi_{fj} f_{ij}^x,$$

where $\ln \xi_{fj}$ follows a normal distribution with zero mean and variance σ_ξ^2:

$$P(x) = \Phi\left(\frac{\ln x}{\sigma_\xi}\right). \tag{31}$$

The introduction of heterogeneous fixed costs allows the model to rationalizes the empirical patterns that in the distant destinations, number of firms participating in export decreases dramatically while a significant number of small exporters are still present. In this extended model, even f_{ij}^x increases, which leads to the decrease of number of exporters, smaller exporter can survive as long as the firm-destination specific shock ξ_{fj} is low enough. Since fixed costs do not affect the optimal quantity either for the purpose of profit optimization nor logistics costs minimization, the firm's choice of quantity q_{fj} or shipment frequency F_{fj} do not change.

We assume that the marginal costs across firms within the same origin follow the log-normal distribution:

$$G_i(a) = \Phi\left(\frac{\ln a - \mu_{a,i}}{\sigma_a}\right), \tag{32}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal random variable. Although assuming the firm-level productivity follows a Pareto distribution would yield convenient closed-form solutions to the model, as shown in the Appendix D, this assumption implies that the elasticity of the intensive margin $\frac{X_{ij}}{N_{ij}}$ with respect to B_{ij} is positive given that $\nu_R > \nu_{LT}$. Together with a negative distance elasticity of $\frac{X_{ij}}{N_{ij}}$, the model would suggest that the logistics costs are decreasing in distance, a prediction that we find hard to justify. To avoid this problem, we follow the suggestion in Fernandes *et al.* (2019) and assume that firm-level productivity follow a log-normal distribution.

Since we don't have data on inventory costs, we follow the literature on inventory management (e.g. Alessandria *et al.* 2010) and set κ_I equal to 0.15. Since the inventory costs is equal to the product of the average inventory level and κ_I, this is equivalent to a annual inventory costs of 30% of the stored quantities. The demand side parameter ν_R is set to be 0.8, implying the elasticity of substitution is equal to 5. This value is consistent with the literature that estimates this parameter using different data and approaches (see e.g. Anderson and van Wincoop 2004). Same as previous section, we set $\nu_{LT} = 0.6$. Note that given these values of ν_R and ν_{LT}, we can solve the firm's profit maximization explicitly, which greatly speeds up the numerical calculations[12]. Wages in each province is identified by the GDP per capita that comes from National Statistics Bureau of China. We normalize the mean of wages across all provinces to be 1.

The remaining parameters to be calibrated are $\left(\sigma_a, \sigma_\xi, \mu_{a,i}, A_j, B_{ij}, \tau_{ij}, f_{ij}^x\right)$, so that if there are I origins and J destinations, we have $2 + I + J + 3IJ$ parameters to be calibrated. The parameters σ_a, σ_ξ that determine the variance of the productivity and fixed costs shocks are held to be constant across all origin and destinations. In order to identify these parameters, we exploit the following moment conditions. Firstly, the average export sales from origin i to destination j, conditional on firms that enter, is

$$\frac{X_{ij}}{N_{ij}} = \int \int_0^{a_{ij}^*(\xi)} A_j q_{ij}^{\nu_R} (a)\, d\frac{G(a)}{G\left(a_{ij}^*(\xi)\right)} dP(\xi), \tag{33}$$

and the variance of firm-level sales is

$$VarX_{ij} = \int \int_0^{a_{ij}^*(\xi)} \left[A_j q_{ij}^{\nu_R}(a) - \frac{X_{ij}}{N_{ij}}\right]^2 d\frac{G(a)}{G\left(a_{ij}^*(\xi)\right)} dP(\xi). \tag{34}$$

Secondly, the mean of shipment frequency conditional on firms that export, is

$$\frac{F_{ij}}{N_{ij}} = \int \int_0^{a_{ij}^*(\xi)} \frac{1}{1-\nu_{LT}} \kappa_I B_{ij}^{-1} q_{ij}^{(1-\nu_{LT})}(a)\, d\frac{G(a)}{G\left(a_{ij}^*(\xi)\right)} dP(\xi), \tag{35}$$

[12]More specifically, we use Equation (60) in Appendix A

and the variance of firm-level shipment frequency is

$$VarF_{ij} = \int \int_0^{a^*_{ij}(\xi)} \left[\frac{1}{1-\nu_{LT}} \kappa_I B_{ij}^{-1} q_{ij}^{(1-\nu_{LT})}(a) - \frac{F_{ij}}{N_{ij}} \right]^2 d \frac{G(a)}{G\left(a^*_{ij}(\xi)\right)} dP(\xi). \tag{36}$$

Finally, the number of firms that export from i to j can be expressed as:

$$N_{ij} = N_i^e \int \int_0^{a^*_{ij}(\xi)} dG(a)\, dP(\xi) = N_i^e \int G\left(a^*_{ij}(\xi)\right) dP(\xi),$$

where $\int G\left(a^*_{ij}(\xi)\right) dP(\xi)$ is the share of firms find it profitable to enter. Normalizing it with the number of firms that sell domestically, we get

$$\frac{N_{ij}}{N_{ii}} = \frac{\int G\left(a^*_{ij}(\xi)\right) dP(\xi)}{\int G\left(a^*_{ii}(\xi)\right) dP(\xi)}. \tag{37}$$

Therefore equations (33)-(37) define five moments for each origin-destination pair, and in total we observe $5IJ$ moment conditions from data. As a result, the number of the moment conditions will exceed the number of parameters we need to calibrate as long as the number of origin-destination pairs IJ is large enough.

The intuition of the identification works as follows. As can be seen from Equation (33), A_j shifts the mean of the sales distribution. Thus A_j can be identified from the average sales for each destination across origins. Moreover, since τ_{ij} affects the firm-level optimal quantity, it can be identified from the origin-destination specific change in average sales. Similarly, B_{ij} can be identified from the bilateral changes in average shipment frequency, as can be seen in Equation (35). Finally, f^x_{ij} is identified from the condition shown in Equation (37), since the average bilateral fixed costs affect the number of firms enter each market. The shape of the distributions, as measured in Equation (34), are informative about the variance of marginal costs σ_a^2, because within the province-destination, the marginal cost a is an important source of firm heterogeneity in sales or shipment frequency. The observed firm level sales and shipment frequency distribution represent mix of entrants across different levels of marginal cost, thus the shapes of the distributions are also informative about the variance σ_ξ^2. Importantly, the relative extensive margin only identifies levels of f_{ij}^x relative to the domestic fixed costs, which we normalize to 0. Similarly, we normalize the domestic iceberg trade costs to 1.

This model does not allow for a big firm in one destination to sell little in another destination. This model also does not allow for idiosyncratic shipment frequency across firms conditional on sales. That is, it predicts that a perfect correlation between sales and shipment frequency across firms within the same market. All these problems can be addressed by introducing more firm-destination specific shocks on demand and logistics costs. Since introducing such features do not contribute to our main goal of inferring province-destination level trade costs, we choose not to further complicate the model.

Simulation is performed and the parameters are chosen to match the simulated moments with the empirical moments. The simulation algorithm is summarized as follows:

- For each province, we simulate 10,000 firms, indexed by s. Each simulated firm takes a productivity draw $a_i(s)$, and for each destination a fixed export cost draw $\xi_{ij}(s)$.
- Given parameter values, solve the export decisions for each firm from each origin, according to the model discussed in the previous section. Shipment frequency is taken as the smallest integer greater or equal to the value given by Equation (30).
- Given other parameters, we solve for f_{ij}^x by solving Equation (37).
- In the outer loop, search over σ_ξ, σ_a, A_j, B_{ij} , τ_{ij} to match firm level sales and shipment frequency distribution for each origin-destination pair, by minimizing the distance defined by the moment conditions (33)-(37).

3.4 Calibration Results

The calibration result of σ_ξ and σ_a is shown in Table 4. The inferred σ_ξ is higher under model with $\nu_{LT} = 0.6$. It comes from the fact that productive firms have higher advantage when $\nu_{LT} > 0$, therefore in order to rationalize the presence of small firms we need more extreme shocks on fixed costs.

As an illustration of the model fits, Figure 4 and Figure 5 show the model fit of sales distribution and shipment frequency distribution respectively, for the exports from two biggest provinces, Guangdong and Jiangsu, to the US and Japan. The fits are good in general, and the fits to the sales distributions are better than the shipment frequency distribution. Figure 6 shows the model fit for the normalized extensive margin is perfect, as all the points of empirical and simulated normalized extensive margin lie on the 45 degree line. This is not surprising given that we solve for f_{ij}^x such that Equation (37) is satisfied exactly.

In order to check whether the calibrated parameters make sense, the following regressions are run:

$$\ln X_{ij} = \beta_0 + \beta_1 lnB_{ij} + \beta_2 \ln \tau_{ij} + \beta_3 \ln f_{ij}^x + \phi_i + \psi_j + u_{ij}, \tag{38}$$

where X_{ij} is the bilateral trade flow from province i to destination j, and B_{ij}, τ_{ij}, f_{ij}^x are trade costs parameters calibrated. We also replace the outcome with the logarithmic value of the bilateral shipment frequency $\ln F_{ij}$. All parameters are expected to be negatively correlated with trade flow and shipment frequency. As shown in Table 5, when regressed separately, all calibrated parameters are negatively correlated with both bilateral trade flows and shipment frequency. When including all explanatory variables, only $\ln \tau_{ij}$ and $\ln f_{ij}^x$ significantly impact the bilateral trade flow, while all variables negatively correlate with the bilateral shipment frequency. Interestingly, the estimated elasticity of τ_{ij} with respect to bilateral trade flow is comparable to the estimate of trade elasticity in the literature.

The first three columns of Table 6 report the distance elasticity of calibrated trade costs parameters, controlling for province and destination fixed effects. All parameters are increasing in distance, consistent with the intuition. B_{ij} has a lowest slope, driven by the slow change rate in shipment frequency, and f_{ij}^x has the highest slope, driven by the sharp decrease in the extensive margin of trade.

As measures of trade costs, for each firm with marginal cost a that exports from location i to location j, we calculate normalize logistics costs and iceberg trade costs using export revenue, namely the ratios $w_i B_{ij} q_{ij}^{\nu_{LT}} / A_j q^{\nu_R}$, and $\left(\tau_{ij} - 1\right) a w_i q_{ij} / A_j q^{\nu_R}$. As shown in Figure 7, the logistics costs are negligible for the most productive firms, but can be as high as 35% of revenue for the least productive firms. On the other hand, the iceberg trade costs play a relatively minor role for the less productive firms. This is in contrast with the case when $\nu_{LT} = 0$, under which the ratio of iceberg trade costs and revenue is constant.[13]

3.5 Ad Valorem Rate of The Logistics Costs

The calibrated parameters allow us to assess the significance of the logistics costs for each firm. More specifically, for each firm we ask what is the magnitude of iceberg costs required to generate the same operating profit (i.e., the difference between revenue and variables costs) if we shut down the logistics costs by adding the restriction $\nu_{LT} = 0$, $B_{ij} = 0$, keeping the market conditions unchanged. Let $O_{ij}(a)$ denote the the operating profit for a firm with marginal cost a that exports from location i to location j, and $\tau_{ij}^{Im}(a)$ denotes the iceberg costs required to generate the same operating profit. Since logistics costs are heterogeneous across firms, $\tau_{ij}^{Im}(a)$ depends on firm's productivity. The difference $\tau_{ij}^{Im}(a) - \tau_{ij}$ therefore captures the ad valorem rate of the logistics costs for each firm.

When $\nu_{LT} = 0$, the optimal quantity that maximizes the profit is given by

$$q = \left(\frac{A_j \nu_R}{a w_i \tau_{ij}}\right)^{\frac{1}{1-\nu_R}}. \tag{39}$$

And $O_{ij}(a)$ is given by

$$O_{ij}(a) = A_j^{\frac{1}{1-\nu_R}} \left(a w_i \tau_{ij}\right)^{-\frac{\nu_R}{1-\nu_R}} \left(\nu_R\right)^{\frac{1}{1-\nu_R}} \left[\nu_R^{-1} - 1\right]. \tag{40}$$

[13] When $\nu_{LT} = 0$, the first-order condition of profit maximization problem is

$$\frac{\partial \pi_{ij}}{\partial q} = A_j \nu_R q^{\nu_R - 1} - a w_i \tau_{ij} = 0,$$

so that $\frac{a w_i \left(\tau_{ij} - 1\right) q}{A_j q^{\nu_R}} = \nu_R \frac{\tau_{ij} - 1}{\tau_{ij}}$.

So that $\tau_{ij}^{Im}(a)$ is given by

$$\tau_{ij}^{Im}(a) = \frac{O_{ij}(a)^{-\frac{1-\nu_R}{\nu_R}}}{(\nu_R)^{-\frac{1}{\nu_R}}\left[\nu_R^{-1}-1\right]^{-\frac{1-\nu_R}{\nu_R}} A_j^{-\frac{1}{\nu_R}}(aw_i)}. \tag{41}$$

We take $O_{ij}(a)$ as well as the demand shifter A_j from the simulation in the above section to calculate $\tau_{ij}^{Im}(a)$ for each firm that exports from location i to location j. To calculate the ad valorem rate of logistics costs, $\tau_{ij}^{Im}(a) - \tau_{ij}$, we take τ_{ij} as the bilateral iceberg trade costs calibrated from the previous subsection.

The calculated ad valorem rate of logistics costs across firms that export from Guangdong to the US is plotted in Figure 8. As apparent from the figure, the effects of logistics costs are highly heterogeneous. While the most productive firm incurs the logistics costs that has a ad valorem equivalent rate of 1.9%, the least productive firm pays a ad valorem equivalent rate of 174.9%.

The two columns in Table 7 report the distance elasticity of the average implied iceberg trade costs τ_{ij}^{Im}, which is calculated by taking average across firms for each origin-destination pair, and the distance elasticity of the difference between the calibrated and the implied iceberg trade costs. The result in column 1 shows that after shutting down the logistics costs, distance elasticity of iceberg trade costs increases slightly. Lastly, the estimate in column 2 shows that the ad valorem equivalent rate of logistics costs differ substantially across space.

Conditional on the same operating profit, does the payment of logistics costs increase the total trade costs? To investigate this, we calculate the ratios between total trade costs and revenue, namely $\left(w_i B_{ij} q_{ij}^{\nu_{LT}} + (\tau_{ij}-1)\, aw_i q_{ij}\right)/A_j q^{\nu_R}$ and $\left(\tau_{ij}^{Im}-1\right) aw_i q_{ij}/A_j q^{\nu_R}$,for each simulated firm that exports from Guangdong to the US. The results are plotted in Figure 9. For the productive firms, it does not make much difference as the logistics costs are relatively low. But for the less productive firms, the total trade costs are higher when there are logistics costs.

4 Conclusion

In this paper we analyze the implications of a more flexible trade costs for heterogeneous firm trade model. Theoretically, we derive the formula for the welfare changes given trade costs changes under the assumption that firm-level productivity follows the Pareto distribution. Empirically, we quantify the significance of the logistics costs exploiting the patterns in the shipment frequency. Our results show that a more realistic trade cost structure can potentially impact the evaluation of trade policy. Not only the magnitude of gains from trade would be dependent on the trade costs structure, the response across firms can be highly heterogeneous.

References

AHN, J., KHANDELWAL, A. K. and WEI, S. J. (2011). The role of intermediaries in facilitating trade. *Journal of International Economics*, **84** (1), 73–85.

ALESSANDRIA, G., KABOSKI, J. P. and MIDRIGAN, V. (2010). Inventories, lumpy trade, and large devaluations. *American Economic Review*, **100** (5), 2304–2339.

ANDERSON, J. E. and VAN WINCOOP, E. (2004). Trade Costs. *Journal of Economic Literature*, **42** (3), 691–751.

ARKOLAKIS, C., COSTINOT, A. and RODRÍGUEZ-CLARE, A. (2012). New trade models, same old gains? *American Economic Review*, **102** (1), 94–130.

BAI, X., KRISHNA, K. and MA, H. (2017). How you export matters: Export mode, learning and productivity in China. *Journal of International Economics*, **104**, 122–137.

BLUM, B. S., CLARO, S., DASGUPTA, K. and HORSTMANN, I. J. (2019). Inventory management, product quality, and cross country income differences. *American Economic Journal: Macroeconomics*, **11** (1), 338–388.

BRANDT, L., VAN BIESEBROECK, J. and ZHANG, Y. (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics*, **97** (2), 339–351.

FABINGER, M. and WEYL, E. G. (2018). Functional Forms for Tractable Economic Models and the Cost Structure of International Trade. pp. 1–99.

FERNANDES, A., KLENOW, P., MELESHCHUK, S., DENISSE PIEROLA, M. and RODRIGUEZ-CLARE, A. (2019). The Intensive Margin in Trade. *IMF Working Papers*, **18** (259), 1.

HEAD, K., MAYER, T. and THOENIG, M. (2014). Welfare and trade without pareto. *American Economic Review*, **104** (5), 310–316.

HUMMELS, D. and SKIBA, A. (2004). Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture. *Journal of Political Economy*, **112** (6), 1384–1402.

IRARRAZABAL, A., MOXNES, A. and OPROMOLLA, L. D. (2015). The Tip of the Iceberg: A Quantitative Framework for Estimating Trade Costs. *Review of Economics and Statistics*, **97** (4), 777–792.

KROPF, A. and SAURÉ, P. (2014). Fixed costs per shipment. *Journal of International Economics*, **92**, 166–184.

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

— and REDDING, S. J. (2015). New Trade Models, New Welfare Implications. *American Economic Review*, **105** (3), 1105–1146.

TIMMER, M. P., DIETZENBACHER, E., LOS, B., STEHRER, R. and DE VRIES, G. J. (2015). An Illustrated User Guide to the World Input-Output Database: The Case of Global Automotive Production. *Review of International Economics*, **23** (3), 575–605.

year	No. of firm in trade data	No. of products	No. of destinations	No. of transactions
2000	62750	6734	222	5193140
2001	68459	6722	226	5900901
2002	78530	6889	228	7352879
2003	95610	7009	229	9215403
2004	120515	7014	229	11228228
2005	143895	7125	234	13674674
2006	208312	7416	236	25623829

Note: Summary statistics for each year. Products are defined by HS8 codes.

Table 1: Summary statistics

	Gains from Trade, %			Gains from Trade, %	
	$\nu_{LT} = 0$	$\nu_{LT} = 0.6$		$\nu_{LT} = 0$	$\nu_{LT} = 0.6$
Country	(1)	(2)	Country	(1)	(2)
AUS	2.32	2.08	IRL	8.04	7.21
AUT	5.65	5.06	ITA	2.89	2.58
BEL	7.49	6.72	JPN	1.69	1.51
BRA	1.50	1.34	KOR	4.30	3.85
CAN	3.77	3.37	MEX	3.30	2.95
CHN	2.65	2.37	NLD	6.16	5.52
CZE	6.00	5.37	POL	4.36	3.90
DEU	4.47	4.00	PRT	4.40	3.94
DNK	5.75	5.15	ROM	4.46	3.99
ESP	3.10	2.77	RUS	2.41	2.15
FIN	4.40	3.94	SVK	7.63	6.84
FRA	2.99	2.67	SVN	6.83	6.13
GBR	3.23	2.89	SWE	5.06	4.53
GRC	4.20	3.76	TUR	2.87	2.57
HUN	8.08	7.25	TWN	6.11	5.47
IDN	2.90	2.60	USA	1.77	1.58
IND	2.37	2.12	ROW	5.23	4.68
			Average	4.36	3.91

Notes: Average values in the last row are calculated based on the full set of countries.

Table 2: Gains from Trade

	$\ln X_{ij}$	$\ln N_{ij}$	$\ln F_{ij}$	$\ln \frac{X_{ij}}{N_{ij}}$	$\ln \frac{X_{ij}}{F_{ij}}$
$\ln Dist_{ij}$	-0.545^{***}	-0.324^{***}	-0.342^{***}	-0.221^{**}	-0.203^{**}
	(0.126)	(0.048)	(0.077)	(0.105)	(0.089)
No. Obs	2721	2721	2721	2721	2721
Adj. R^2	0.842	0.958	0.920	0.417	0.304

Note: estimation results from the gravity equation of shipment frequency. Origin and destination fixed effects are controlled. The standard errors are reported in the parenthesis. Statistical significance at 5%, 1% level, based on two-tailed tests, is indicated by **,***.

Table 3: Gravity estimates

σ_ξ	σ_a
3.92	0.71

Note: The table reported calibrated values for σ_ξ and σ_a . Estimates reported are the ones with lowest error from different random initial points.

Table 4: Estimates of σ_ξ, σ_a

	$\ln X_{ij}$	$\ln X_{ij}$	$\ln X_{ij}$	$\ln X_{ij}$	$\ln F_{ij}$	$\ln F_{ij}$	$\ln F_{ij}$	$\ln F_{ij}$
$\ln B_{ij}$			-3.133^{***}	-0.128			-4.062^{***}	-1.347^{***}
			(0.689)	(0.716)			(0.446)	(0.208)
$\ln \tau_{ij}$		-3.241^{***}		-3.072^{***}		-3.812^{***}		-3.156^{***}
		(0.442)		(0.518)		(0.183)		(0.150)
$\ln f_{ij}^x$	-0.329^{***}			-0.275^{***}	-0.216^{**}			-0.104^{***}
	(0.111)			(0.086)	(0.098)			(0.025)
No. Obs	80	80	80	80	80	80	80	80
Adj. R^2	0.871	0.921	0.890	0.932	0.891	0.985	0.950	0.994

Note: The table reports regression results of Equation (38). The standard errors are reported in the parenthesis. Statistical significance at 1% level, based on two-tailed tests, is indicated by ***.

Table 5: Correlation between calibrated trade costs and trade flows

	(1)	(2)	(3)
	$\ln \tau_{ij}^{LT}$	$\ln B_{ij}$	$\ln f_{ij}^{x}$
$\ln Dist_{ij}$	0.348^{***}	0.214^{***}	0.508
	(0.079)	(0.063)	(0.453)
No. Obs	80	80	80
Adj. R^2	0.798	0.819	0.851

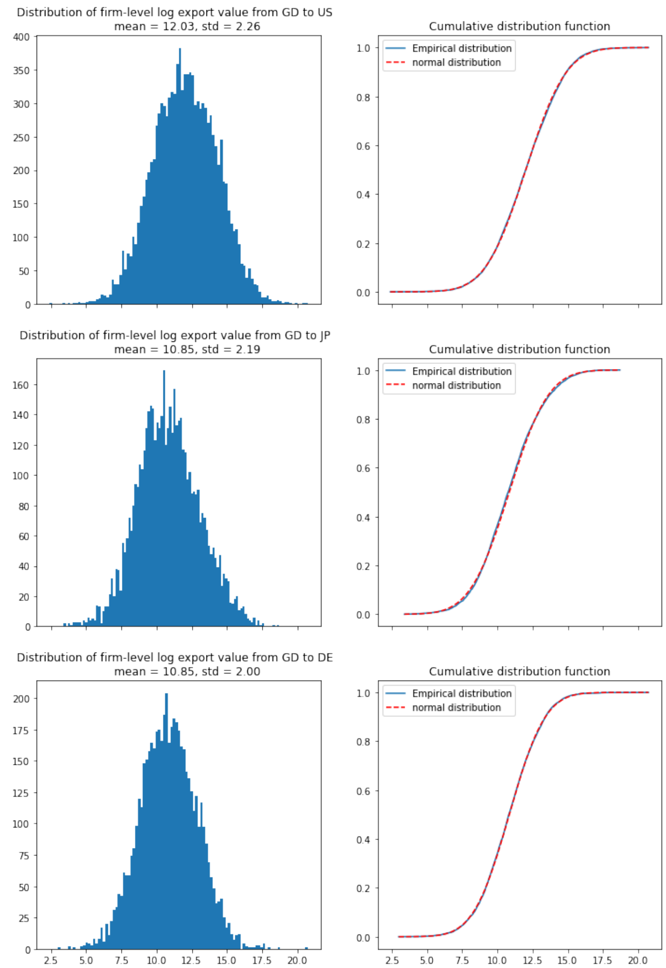
Note: Correlation between estimated trade costs and distance, controlling for province and destination fixed effects. The standard errors are reported in the parenthesis. Statistical significance at 1% level, based on two-tailed tests, is indicated by ***.

Table 6: Correlation between estimated trade costs and distance

	(1)	(2)
	$\ln \tau_{ij}^{Im}$	$\ln \left(\tau_{ij}^{Im} - \tau_{ij}^{LT}\right)$
$\ln Dist_{ij}$	0.404^{***}	0.750^{***}
	(0.094)	(0.169)
No. Obs	80	80
Adj. R^2	0.802	0.908

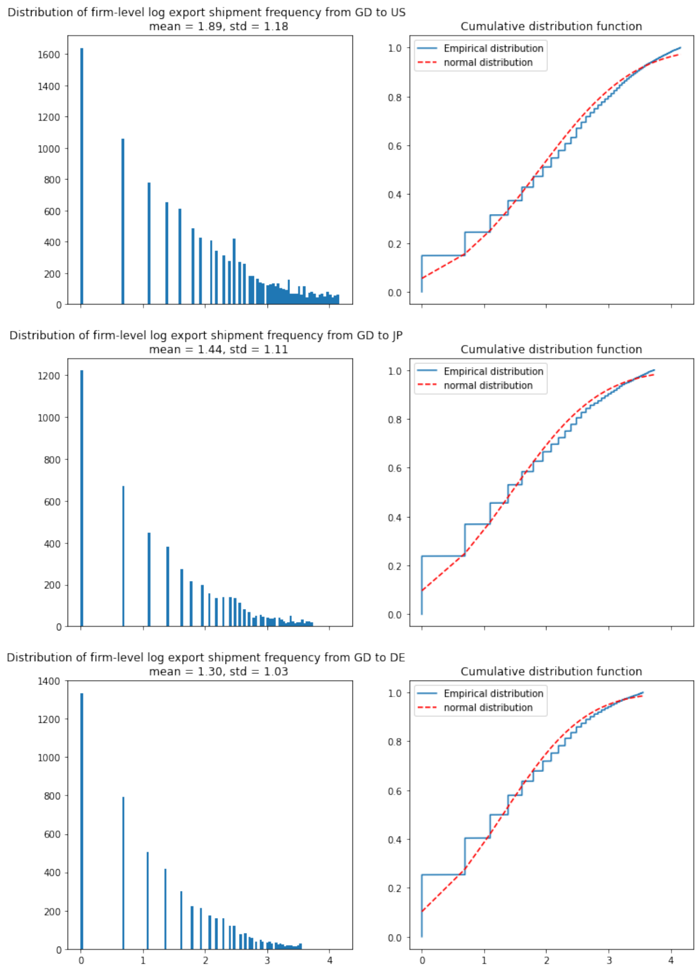
Note: Correlation between the implied iceberg trade costs, the ad-valorem equivalent rate of logistics costs, and distance, controlling for province and destination fixed effects. The standard errors are reported in the parenthesis. Statistical significance at 1% level, based on two-tailed tests, is indicated by ***.

Table 7: Correlation between implied iceberg trade costs and distance



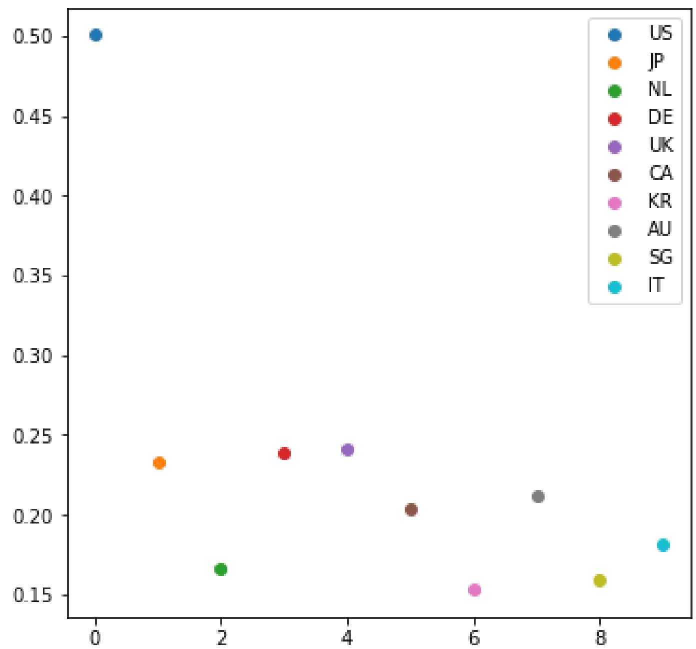
Note: The figures plot the empirical distributions of export sales from Guangdong province to its top three destinations. The dashed line shows the cumulative distribution function of a normal random variables with the same mean and variance as the empirical distribution of the log export value.

Figure 1: Empirical distributions of firm-level export value



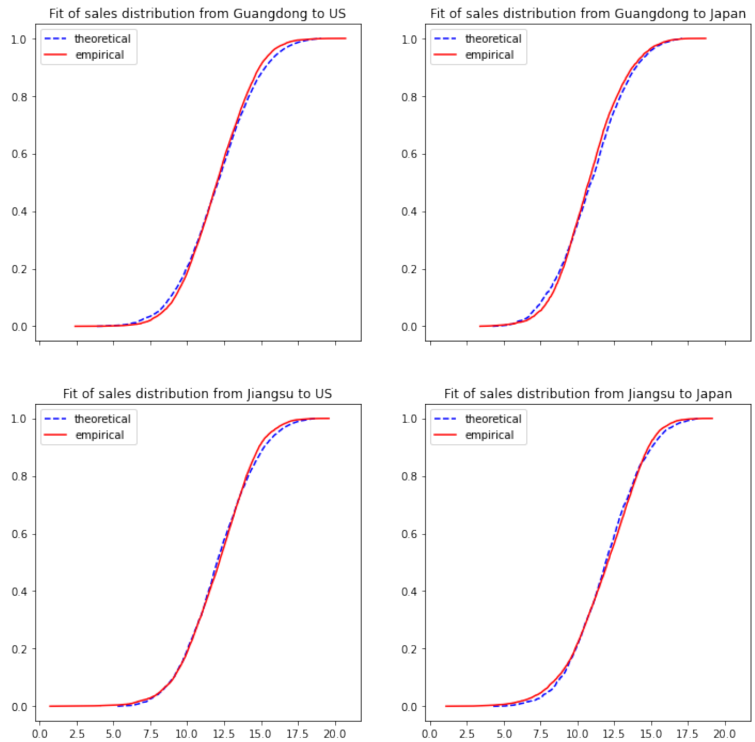
Note: The figures plot the empirical distributions of export shipment frequency from Guangdong province to its top three destinations. The dashed line shows the cumulative distribution function of a normal random variables with the same mean and variance as the empirical distribution of the log export shipment frequency.

Figure 2: Empirical distributions of firm-level export shipment frequency



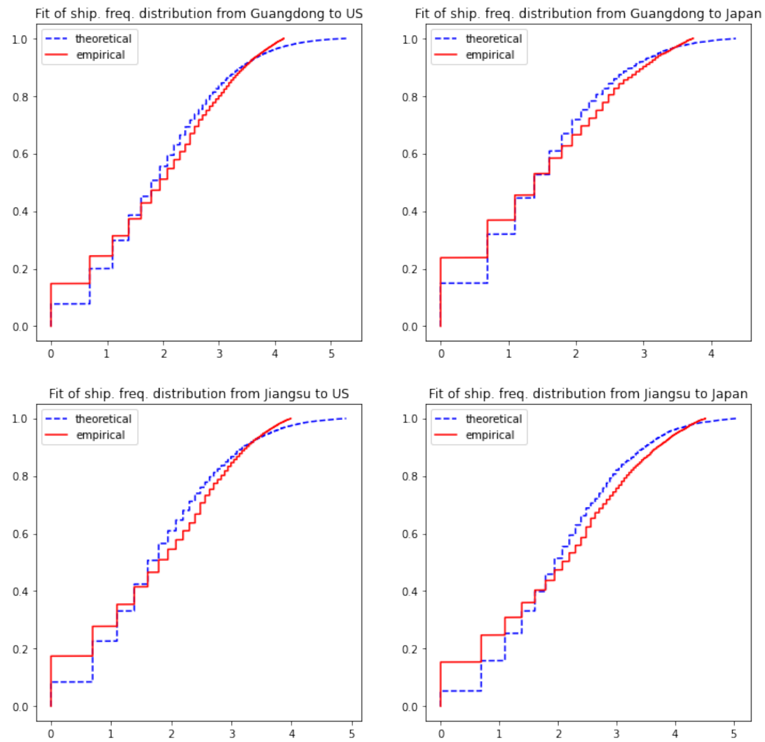
Note: The figure plots the share of firms from Guangdong province that export to each of the nine destinations.

Figure 3: Extensive margins



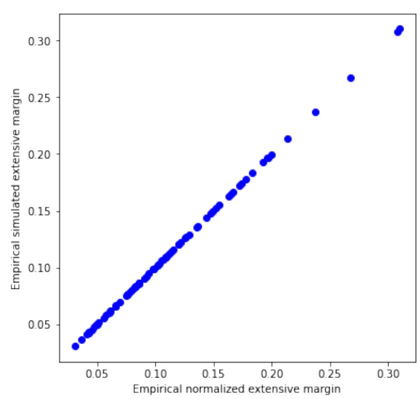
Note: The figures in first row show the fit of the distributions of log export sales from Guangdong Province to the US and Japan. The figures in the second row show the fit of the distributions of log export sales from Jiangsu Province to the US and Japan.

Figure 4: Model fit for sales ditributions



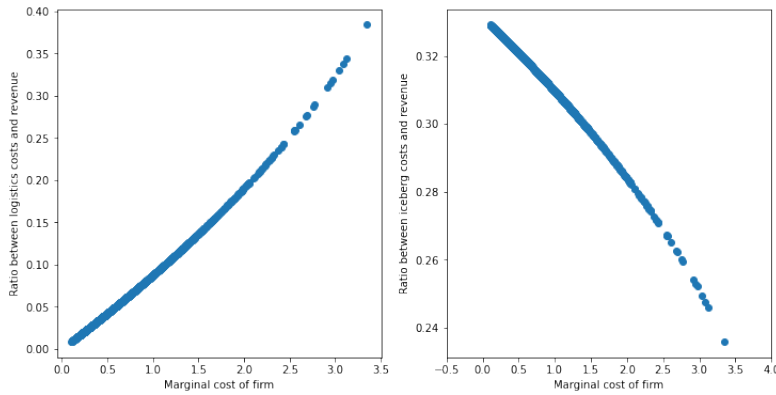
Note: The figures in first row show the fit of the distributions of log shipment frequency from Guangdong Province to the US and Japan. The figures in the second row show the fit of the distributions of log shipment frequency from Jiangsu Province to the US and Japan.

Figure 5: Model fit for sales distributions



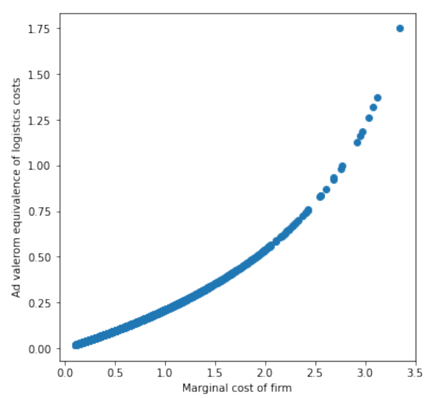
Note: The figure plots the empirical normalized extensive margin on the horizontal axis, and the simulated normalized extensive margin on the vertical axis.

Figure 6: Model fit for normalized extensive margin



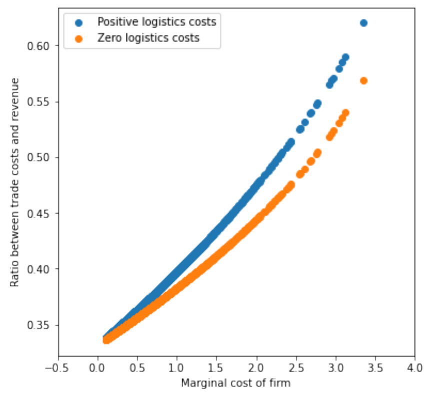
Note: The left figure plots the ratio between logistics costs and export revenue for simulated firms that export from Guangdong Province to the US. The right figure plots the ratio between iceberg costs and export revenue.

Figure 7: Ratios between different components of trade costs and revenue across firms



Note: The figure plots the ad valorem rate of logistics costs across simulated firms that export from Guangdong to the US.

Figure 8: Ad valorem equivalence of logistics costs across firms



Note: The left figure plots the ratio between total costs and export revenue for simulated firms that export from Guangdong Province to the US, with positive and zero logistics costs.

Figure 9: Ratio between total trade costs and revenue across firms

Appendix

A Asymmetric Countries

In this subsection, we extend the model to an environment with multiple asymmetric countries. The economy consists of J countries, indexed by i and j, each endowed with effective labor L_i. Labor is immobile across countries, but perfectly mobile between different uses within a country. Each country i can potentially produce a infinite set of varieties Ω_i indexed by ω. Only an endogenously determined subset Ω_{ij} of Ω_i is available in any destination country j. Utility of the representative consumer in country j is given by

$$U_j = \left(\sum_i \int_{\Omega_{ij}} q_{ij}\left(\omega\right)^{\nu_R} d\omega \right)^{\frac{1}{\nu_R}}, \tag{42}$$

where $q_{ij}\left(\omega\right)$ is the quantity of variety $\omega \in \Omega_{ij}$ from origin i consumed in destination j, and $0 < \nu_R < 1$ is related to the elasticity of substitution between varieties σ through the relation $\nu_R = \frac{\sigma-1}{\sigma}$. Utility maximization gives the inverse demand curve

$$p_{ij}\left(\omega\right) = A_j q_{ij}\left(\omega\right)^{\nu_R - 1}, \tag{43}$$

where the country-specific parameter $A_j = P_j^{\nu_R} I_j^{1-\nu_R}$ summarizes the market condition with I_j and P_j representing the total expenditure and price index in destination j. This gives the relationship between the value and quantity of each firm's export,

$$X_{ij}\left(\omega\right) = A_j q_{ij}\left(\omega\right)^{\nu_R}. \tag{44}$$

In what follows, in order to simplify notation, we also sometimes use the notation q_{fj} instead of $q_{ij}\left(\omega\right)$, with the understanding that firm f from origin i produces variety ω.

Technology of production of varieties features constant return to scale. Each firm in country i has a marginal cost of production $a > 0$ that is drawn from a known distribution

$G_i(a)$. The cost of serving any destination j has three components. The first component is the fixed costs: any firm f from origin i needs pay a cost of f_{ij}^x in terms of country i's labor in order to enter market j. The second component is the usual iceberg trade costs: delivering one unit of any variety from origin i to destination j requires shipping $\tau_{ij} \geq 1$ units of this variety. Finally, the third component constitutes our innovation relative to the standard Melitz model and captures the idea of the "logistic costs": in order to deliver q_{fj} units of its variety, firm f from origin i needs to pay the cost $B_{ij} q_{fj}^{\nu_{LT}}$ measured in units of country i's labor, where $\nu_{LT} > 0$ is the trade costs elasticity with respect to the quantity shipped.

Given the above specifications of technology of production and costs of serving markets, the profit country i's firm with marginal cost a from serving destination j is

$$\pi_{ij}(q) = A_j q^{\nu_R} - w_i B_{ij} q^{\nu_{LT}} - a w_i \tau_{ij} q - w_i f_{ij}^x. \tag{45}$$

For the analytical tractability, we assume that $f_{ij}^x = 0$ in this subsection as well. The first-order condition for the firm's profit maximization problem is

$$\frac{\partial \pi_{ij}}{\partial q} = A_j \nu_R q^{\nu_R - 1} - w_i B_{ij} \nu_{LT} q^{\nu_{LT} - 1} - a w_i \tau_{ij} = 0, \tag{46}$$

which defines a one-one mapping from the optimal quantity to marginal cost for any bilateral country pair as

$$a_{ij}(q) = \frac{A_j \nu_R q^{\nu_R - 1} - w_i B_{ij} \nu_{LT} q^{\nu_{LT} - 1}}{w_i \tau_{ij}}. \tag{47}$$

The inverse mapping from the marginal cost to the optimal quantity is thus $q_{ij}(a)$. The zero profit condition that defines the cutoff q_{ij}^* thus can be written as

$$A_j (1 - \nu_R) \left(q_{ij}^*\right)^{\nu_R} - w_i B_{ij} (1 - \nu_{LT}) \left(q_{ij}^*\right)^{\nu_{LT}} = 0. \tag{48}$$

The free entry condition is any market i is

$$\sum_{j=1}^{J} \int_{0}^{a_{ij}^{*}} \pi_{ij} \left(q_{ij} \left(a \right) \right) dG \left(a \right) = f^{e}. \tag{49}$$

For any market i, equations (48) and (49) therefore define a system of $J+1$ equations that can be used to solve for q_{ij}^{*} and A_i in terms of wage w_i.

The labor market clearing condition can be used to solve the total mass of potential entrants in country i, which is denoted as N_i^e. Labor is used for the following purposes: entry of firms, logistic costs, and production. The amount of labor used for entry is $L_i^e \equiv N_i^e f_i^e$,where f_i^e is the entry cost. Therefore the labor market clearing condition is

$$L_i = L_i^e + \sum_j L_{ij}^{lm}, \tag{50}$$

where L_{ij}^{lm} represents the amount of labor used in logistics and production for the exports to country j in origin i.

The wages, up to a scale, can be solved using the goods market clearing condition. More specifically, assuming that trade is balanced, the total output in country i is equal to the absorption of country i's production around the world:

$$w_i L_i = \sum_j \lambda_{ij} w_j L_j, \tag{51}$$

where λ_{ij} is the expenditure share of country i in the goods produced by country j. Since λ_{ij} is a function of wages, Equation (51) gives a system of equation that can be used to solve for wages in each country.

A.1 Solutions Under Pareto Distribution: Asymmetric Countries

In this subsection, we show closed form solutions assuming that the marginal cost across firms follows a Pareto distribution given by the cumulative density function

$$G_i \left(a \right) = \theta^{-1} \left(\kappa_{G,i} a \right)^{\theta}, \tag{52}$$

Under this assumption, as shown in Appendix B, we can express the aggregated quantity with any power ν, subject to the condition $\frac{\nu+\nu_R\theta-\theta}{\nu_{LT}-\nu_R} > 0$, as

$$\int_0^{a_{ij}^*} q^\nu dG_i\left(a\right) = \kappa_{G,i}^\theta \left(w_i\tau_{ij}\right)^{-\theta} A_j^{-\frac{\nu+\nu_{LT}\theta-\theta}{\nu_R-\nu_{LT}}} \left(w_i B_{ij}\right)^{\frac{\nu+\nu_R\theta-\theta}{\nu_R-\nu_{LT}}} H\left(\nu\right), \tag{53}$$

where $H\left(\nu\right)$ is a constant that depends on ν_R, ν_{LT} and θ. Using this formula, we can calculate the aggregate trade flow from country i to country j as

$$\begin{aligned} X_{ij} &= N_i^e A_j \int_0^{a_{ij}^*} q_{ij}^{\nu_R}\left(a\right) dG_i\left(a\right) \\ &= N_i^e \kappa_{G,i}^\theta \left(A_j\right)^{\frac{\nu_{LT}+\nu_{LT}\theta-\theta}{\nu_{LT}-\nu_R}} \left(w_i\tau_{ij}\right)^{-\theta} \left(w_i B_{ij}\right)^{-\frac{\nu_R+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}} H\left(\nu_R\right). \end{aligned} \tag{54}$$

The expenditure share of country j on country i's goods is then

$$\lambda_{ij} = \frac{X_{ij}}{\sum_l X_{lj}} = \frac{N_i^e \kappa_{G,i}^\theta \left(w_i\tau_{ij}\right)^{-\theta} \left(w_i B_{ij}\right)^{-\frac{\nu_R+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}}}{\sum_l N_l^e \kappa_{G,l}^\theta \left(w_l\tau_{lj}\right)^{-\theta} \left(w_l B_{lj}\right)^{-\frac{\nu_R+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}}}. \tag{55}$$

It shows that our model generates a gravity equation under a more flexible trade cost structure. The trade elasticity with respect to τ_{ij}, defined as $\frac{\partial \ln\left(X_{ij}/X_{jj}\right)}{\partial \ln \tau_{ij}}$, is given by the shape parameter of the firm's productivity distribution θ. And the trade elasticity with respect to B_{ij}, defined as $\frac{\partial \ln\left(X_{ij}/X_{jj}\right)}{\partial \ln B_{ij}}$, is determined by the demand elasticity ν_R, the elasticity of the logistics costs ν_{LT}, and the shape parameter θ.

Under the assumption of Pareto distribution, we show that the mass of firms N_i^e in each country is a fixed portion of L_i. Using first order condition (46) and the aggregation formula, L_{ij}^{lm} is equal to

$$L_{ij}^{lm} = N_i^e \int_0^{a_{ij}^*} \left[\left(1-\nu_{LT}\right) B_{ij} q_{ij}^{\nu_{LT}}\left(a\right) + \frac{\nu_R A_j q_{ij}^{\nu_R}\left(a\right)}{w_i}\right] dG_i\left(a\right)$$

$$= \nu_R \left[\left(\frac{1-\nu_{LT}}{\nu_R}\right) \frac{H\left(\nu_{LT}\right)}{H\left(\nu_R\right)} + 1\right] \frac{X_{ij}}{w_i}.$$

Then we can solve that

$$N_i^e = \left(1 - \nu_R - (1-\nu_{LT}) \frac{H(\nu_{LT})}{H(\nu_R)}\right) \frac{L_i}{f_i^e}, \tag{56}$$

which shows that N_i^e is a fixed portion of L_i as claimed. From the above calculations, we can also see that the aggregate profit is a constant share of the total revenue:

$$\Pi_i = \left(1 - \nu_R - (1-\nu_{LT}) \frac{H(\nu_{LT})}{H(\nu_R)}\right) \sum_j X_{ij}. \tag{57}$$

Given the gravity equation (55) and the fact that aggregate profit is a constant share of aggregate revenue, as shown in (57), it is clear that the model satisfies all the macro restrictions proposed in ACR. However, we next show that the the welfare gains from trade can not be inferred from the domestic expenditure share λ_{ii} and trade elasticity with respect to τ_{ij} only. Instead, the magnitudes of ν_{LT} and ν_R also play an important role.

Using the aggregation formula (53) and the definition $A_j = P_j^{\nu_R} I_j^{1-\nu_R}$, we can get

$$\left(P_j\right)^{\frac{-\nu_R(\nu_{LT}+\theta\nu_{LT}-\theta)}{\nu_{LT}-\nu_R}} = C_P \sum_i N_i^e \kappa_{G,i}^\theta \left(w_i \tau_{ij}\right)^{-\theta} \left(w_i B_{ij}\right)^{-\frac{\nu_R+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}}, \tag{58}$$

where

$$C_P = \left(I_j\right)^{-\nu_R + \frac{(1-\nu_R)(\theta\nu_{LT}+\nu_R-\theta)}{\nu_{LT}-\nu_R}} H(\nu_R),$$

and N_i^e has been solved above. Using (55) and (58), we can write the domestic expenditure share as

$$\lambda_{ii} = C_P N_i^e \kappa_{G,i}^\theta \left(w_i\right)^{-\theta} \left(w_i B_{ii}\right)^{-\frac{\nu_R+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}} \left(P_i\right)^{\frac{\nu_R(\nu_{LT}+\theta\nu_{LT}-\theta)}{\nu_{LT}-\nu_R}},$$

from which we can derive the expression of the real consumption as

$$W_i \equiv \frac{w_i}{P_i} = \left[\frac{\lambda_{ii}}{C_P N_i^e \kappa_{G,i}^\theta \left(B_{ii}\right)^{-\frac{\nu_R+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}}}\right]^{-\frac{\nu_{LT}-\nu_R}{\nu_R(\nu_{LT}+\theta\nu_{LT}-\theta)}}.$$

Because the mass of firms N_i^e is a fixed portion of the labor endowment, and under autarky, $\lambda_{ii}^A = 1$, the welfare gains from opening to trade in autarky is given by

$$\frac{W_i}{W_i^A} = \lambda_{ii}^{-\frac{\nu_{LT}-\nu_R}{\nu_R(\nu_{LT}+\theta\nu_{LT}-\theta)}}. \tag{59}$$

As a widely known formula shown in ACR, the expenditure share on the domestic goods and the trade elasticity are the only two sufficient statistics to calculate welfare gains from opening to trade. In contrast, our model gives a formula that combines trade elasticity, trade cost elasticity and demand elasticity. It is easy to verify that $-\frac{\nu_R(\nu_{LT}+\theta\nu_{LT}-\theta)}{\nu_{LT}-\nu_R}$ is equal to $-\theta$ when $\nu_{LT} = 0$, therefore this formula generalizes the results in ACR. As discussed in the previous section, the trade is balanced, the aggregate profit is a constant share of the total revenue, and the gravity equation is derived. Therefore our only departure from the class of models studied in ACR comes from a different micro structure, in particular the trade costs structure that we are considering.

A.2 Solution of Firm's Profit Maximization Problem

Equation (4) cannot be solved analytically for arbitrary values of ν_R and ν_{LT}. However, an empirically relevant way that gives an explicit solution to the firm's profit maximization problem is to impose $\nu_{LT} = 2\nu_R - 1$. With this assumption, the optimal quantity maximizing firm's profit is

$$q_{ij}(a) = \left[\frac{\left(\nu_R A_j + \sqrt{(\nu_R A_j)^2 - 4(aw_i\tau_{ij})(w_i\nu_{LT}B_{ij})}\right)}{2aw_i\tau_{ij}}\right]^{\frac{1}{1-\nu_R}}. \tag{60}$$

Each firm charges a price given by

$$p_{ij}(a) = \frac{2aw_i\tau_{ij}A_j}{\left(\nu_R A_j + \sqrt{(\nu_R A_j)^2 - 4(aw_i\tau_{ij})(w_i\nu_{LT}B_{ij})}\right)}. \tag{61}$$

From here we see that price depends not only on the marginal cost of production a, but also on the market conditions A_j and logistics costs parameter B_{ij}. As an example that the model has the potential to match the empirical regularities that call for a different cost structure from the standard model, we show that the model can generate Alchian-Allen effect.[14] Suppose that conditional on the marginal cost a, varieties are allowed to differentiate vertically, so that the demand shifter can differ across varieties, we can calculate the relative demand for varieties as

$$\frac{q_{ij}^{H}\left(a\right)}{q_{ij}^{L}\left(a\right)}=\left[\frac{\left(Q^{H}\nu_{R}A_{j}+\sqrt{\left(Q^{H}\nu_{R}A_{j}\right)^{2}-4\left(aw_{i}\tau_{ij}\right)\left(w_{i}\nu_{LT}B_{ij}\right)}\right)}{\left(Q^{L}\nu_{R}A_{j}+\sqrt{\left(Q^{L}\nu_{R}A_{j}\right)^{2}-4\left(aw_{i}\tau_{ij}\right)\left(w_{i}\nu_{LT}B_{ij}\right)}\right)}\right]^{\frac{1}{1-\nu_{R}}},$$

where $Q^{H}>Q^{L}$ represent high and low level of quality. The Alchian-Allen effect means that relative demand for high quality good is increasing in trade costs, which is true here because $\frac{\partial q_{ij}^{H}/q_{ij}^{L}}{\partial B_{ij}}>0$.

[14] For a definition of Alchian-Allen effect, see, for example, Hummels and Skiba (2004)

B Derivation of aggregation formula

Let us start with calculating the integral $\int_0^{a_{ij}^*} q^\nu dG_i(a)$ for some $\nu > 0$. Using expression (6) for a, and that $G_i(a) = \theta^{-1}(\kappa_{G,i}a)^\theta$, we can derive

$$\begin{aligned}
\int_0^{a_{ij}^*} q^\nu dG_i(a) &= -\int_{q_{ij}^*}^{\infty} q^\nu dG\left(\frac{\nu_R A_j q^{\nu_R-1} - \nu_{LT} w_i B_{ij} q^{\nu_{LT}-1}}{w_i\tau_{ij}}\right) \\
&= \int_{q_{ij}^*}^{\infty} \left\{ q^\nu \frac{(1-\nu_R)\,\nu_R A_j q^{\nu_R-2} - (1-\nu_{LT})\,\nu_{LT} w_i B_{ij} q^{\nu_{LT}-2}}{w_i\tau_{ij}} \right. \\
&\qquad \left. \times\, \kappa_{G,i}^\theta \left(\frac{\nu_R A_j q^{\nu_R-1} - \nu_{LT} w_i B_{ij} q^{\nu_{LT}-1}}{w_i\tau_{ij}}\right)^{\theta-1} \right\} dq \\
&= \kappa_{G,i}^\theta \left(w_i\tau_{ij}\right)^{-\theta}(1-\nu_R)\left(\nu_R A_j\right)^\theta \int_{q_{ij}^*}^{\infty} q^{\nu-1+\nu_R\theta-\theta}\left(1-\frac{\nu_{LT} w_i B_{ij}}{\nu_R A_j} q^{\nu_{LT}-\nu_R}\right)^{\theta-1} dq \\
&\quad - \left\{ \kappa_{G,i}^\theta \left(w_i\tau_{ij}\right)^{-\theta}(1-\nu_{LT})\,\nu_{LT}\left(\nu_R A_j\right)^{\theta-1} w_i B_{ij} \right. \\
&\qquad \left. \times \int_{q_{ij}^*}^{\infty} q^{\nu+\nu_{LT}-2+(\nu_R-1)(\theta-1)}\left(1-\frac{\nu_{LT} w_i B_{ij}}{\nu_R A_j} q^{\nu_{LT}-\nu_R}\right)^{\theta-1} dq \right\}.
\end{aligned}$$

Introduce the change of variables $x = \left(q/q_{ij}^*\right)^{\nu_{LT}-\nu_R}$. Then

$$\begin{aligned}
\int_0^{a_{ij}^*} q^\nu dG_i(a) = &\left\{ \frac{1-\nu_R}{\nu_R-\nu_{LT}} \kappa_{G,i}^\theta \left(w_i\tau_{ij}\right)^{-\theta}\left(\nu_R A_j\right)^\theta \left(q_{ij}^*\right)^{\nu+\nu_R\theta-\theta} \right. \\
&\qquad \left. \times \int_0^1 x^{\frac{\nu+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}-1}\left(1-\frac{\nu_{LT} w_i B_{ij}}{\nu_R A_j}\left(q_{ij}^*\right)^{\nu_{LT}-\nu_R} x\right)^{\theta-1} dx \right\} \\
&- \left\{ \frac{(1-\nu_{LT})\,\nu_{LT}}{\nu_R-\nu_{LT}} \kappa_{G,i}^\theta \left(w_i\tau_{ij}\right)^{-\theta}\left(\nu_R A_j\right)^{\theta-1} w_i B_{ij} \left(q_{ij}^*\right)^{\nu+\nu_{LT}-1+(\nu_R-1)(\theta-1)} \right. \\
&\qquad \left. \times \int_0^1 x^{\frac{\nu+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}}\left(1-\frac{\nu_{LT} w_i B_{ij}}{\nu_R A_j}\left(q_{ij}^*\right)^{\nu_{LT}-\nu_R} x\right)^{\theta-1} dx \right\}.
\end{aligned}$$

Using the fact that

$$q_{ij}^{*}=\left[\frac{w_iB_{ij}\left(1-\nu_{LT}\right)}{A_j\left(1-\nu_R\right)}\right]^{\frac{1}{\nu_R-\nu_{LT}}},$$

we get

$$\begin{aligned}\int_0^{a_{ij}^*}q^{\nu}dG_i\left(a\right)=&\left\{\frac{1-\nu_R}{\nu_R-\nu_{LT}}\kappa_{G,i}^{\theta}\left(w_i\tau_{ij}\right)^{-\theta}\left(\nu_RA_j\right)^{\theta}\left[q_{ij}^*\right]^{\nu+\nu_R\theta-\theta}\right.\\&\left.\times\int_0^1x^{\frac{\nu+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}-1}\left(1-\frac{\left(1-\nu_R\right)\nu_{LT}}{\left(1-\nu_{LT}\right)\nu_R}x\right)^{\theta-1}dx\right\}\\&-\left\{\frac{\left(1-\nu_{LT}\right)\nu_{LT}}{\nu_R-\nu_{LT}}\kappa_{G,i}^{\theta}\left(w_i\tau_{ij}\right)^{-\theta}\left(\nu_RA_j\right)^{\theta-1}w_iB_{ij}\left[\frac{w_iB_{ij}\left(1-\nu_{LT}\right)}{A_j\left(1-\nu_R\right)}\right]^{\frac{\nu+\nu_{LT}-1+\left(\nu_R-1\right)\left(\theta-1\right)}{\nu_R-\nu_{LT}}}\right.\\&\left.\times\int_0^1x^{\frac{\nu+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}}\left(1-\frac{\left(1-\nu_R\right)\nu_{LT}}{\left(1-\nu_{LT}\right)\nu_R}x\right)^{\theta-1}dx\right\}.\end{aligned}$$

At this point we are going to use the hyper-geometric function ${}_2F_1\left(a,b;c;z\right)$ defined by

$$B\left(b,c-b\right){}_2F_1\left(a,b;c;z\right)=\int_0^1x^{b-1}\left(1-x\right)^{c-b-1}\left(1-zx\right)^{-a}dx,$$

where $B\left(b,c-b\right)$ is the beta function. The integral in the expression for $B\left(b,c-b\right){}_2F_1\left(a,b;c;z\right)$ is defined only if $|z|<1$ and $c>b>0$. We have

$$\int_0^1x^{\frac{\nu+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}-1}\left(1-\frac{\left(1-\nu_R\right)\nu_{LT}}{\left(1-\nu_{LT}\right)\nu_R}x\right)^{\theta-1}dx=B\left(\gamma_1\left(\nu\right),1\right){}_2F_1\left(1-\theta,\gamma_1\left(\nu\right);\gamma_1\left(\nu\right)+1;\gamma_2\right),$$

where

$$\gamma_1\left(\nu\right)\equiv\frac{\nu+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}\quad\text{and}\quad\gamma_2\equiv\frac{\left(1-\nu_R\right)\nu_{LT}}{\left(1-\nu_{LT}\right)\nu_R},$$

and where we need to have $\gamma_1\left(\nu\right)>0$ and $\gamma_2<1$. The last inequality holds under our assumptions that $0<\nu_{LT}<\nu_R<1$.

We have

$$B\left(\gamma_1\left(\nu\right),1\right)=\int_0^1t^{\gamma_1\left(\nu\right)-1}dt=\frac{1}{\gamma_1\left(\nu\right)}t^{\gamma_1\left(\nu\right)}\Big|_0^1=\frac{1}{\gamma_1\left(\nu\right)}.$$

Then

$$\int_0^1 x^{\frac{\nu+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}-1}\left(1-\frac{(1-\nu_R)\,\nu_{LT}}{(1-\nu_{LT})\,\nu_R}x\right)^{\theta-1}dx=\frac{1}{\gamma_1(\nu)}\times{}_2F_1\left(1-\theta,\gamma_1(\nu);\gamma_1(\nu)+1;\gamma_2\right).$$

Next,

$$\begin{aligned}\int_0^1 x^{\frac{\nu+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}}\left(1-\frac{(1-\nu_R)\,\nu_{LT}}{(1-\nu_{LT})\,\nu_R}x\right)^{\theta-1}dx&=B\left(\gamma_1(\nu)+1,1\right){}_2F_1\left(1-\theta,\gamma_1(\nu)+1;\gamma_1(\nu)+2;\gamma_2\right)\\&=\frac{1}{\gamma_1(\nu)+1}\times{}_2F_1\left(1-\theta,\gamma_1(\nu)+1;\gamma_1(\nu)+2;\gamma_2\right).\end{aligned}$$

Thus, we get

$$\int_0^{a_{ij}^*} q^\nu dG_i(a)=\kappa_{G,i}^\theta\left(w_i\tau_{ij}\right)^{-\theta}A_j^{-\frac{\nu+\nu_{LT}\theta-\theta}{\nu_R-\nu_{LT}}}\left(w_iB_{ij}\right)^{\frac{\nu+\nu_R\theta-\theta}{\nu_R-\nu_{LT}}}H\left(\nu,\nu_R,\nu_{LT},\theta\right),$$

where

$$\begin{aligned}H\left(\nu,\nu_R,\nu_{LT},\theta\right)\equiv{}&\frac{1-\nu_R}{\nu_R-\nu_{LT}}\left[\nu_R\right]^\theta\left[\frac{1-\nu_{LT}}{1-\nu_R}\right]^{\frac{\nu+\nu_R\theta-\theta}{\nu_R-\nu_{LT}}}\frac{1}{\gamma_1(\nu)}\times{}_2F_1\left(1-\theta,\gamma_1(\nu);\gamma_1(\nu)+1;\gamma_2\right)\\&-\left\{\frac{(1-\nu_{LT})\,\nu_{LT}}{\nu_R-\nu_{LT}}\left[\nu_R\right]^{\theta-1}\left[\frac{1-\nu_{LT}}{1-\nu_R}\right]^{\frac{\nu+\nu_{LT}-1+(\nu_R-1)(\theta-1)}{\nu_R-\nu_{LT}}}\right.\\&\qquad\left.\times\frac{1}{\gamma_1(\nu)+1}\times{}_2F_1\left(1-\theta,\gamma_1(\nu)+1;\gamma_1(\nu)+2;\gamma_2\right)\right\}.\end{aligned}$$

Using the zero profit condition

$$A_j=\frac{w_iB_{ij}\left(1-\nu_{LT}\right)}{(1-\nu_R)\left(q_{ij}^*\right)^{\nu_R-\nu_{LT}}},$$

the formula can be written as

$$\int_0^{a_{ij}^*} q^\nu dG_i(a)=\kappa_{G,i}^\theta\left(w_i\tau_{ij}\right)^{-\theta}\left(w_iB_{ij}\right)^\theta\left(q_{ij}^*\right)^{\nu+\nu_{LT}\theta-\theta}\left(\frac{(1-\nu_{LT})}{(1-\nu_R)}\right)^{-\frac{\nu+\nu_{LT}\theta-\theta}{\nu_R-\nu_{LT}}}H\left(\nu,\nu_R,\nu_{LT},\theta\right),$$

C Using Shipment Frequency to Estimate ν_{LT}

C.1 Micro Foundation of Logistics Costs

In this section, we propose a simple logistics costs minimization problem to micro-found the logistics costs introduced into the profit function of firms. The demand of a variety during a given period of time is predicted without uncertainty. The problem is then how to deliver the required quantity q. The trade-off is to balance inventory cost and cost per shipment: since production is not instantaneous, the producer has to store the products and incur inventory costs.[15] Since the amount of inventory is proportional to the quantity per shipment, in order to save inventory cost, firms tend to ship frequently. But since there is a cost related to each shipment (e.g. paper work and other coordination with trade partner), too frequent shipments will be very costly. Therefore the firm optimally chooses a shipment frequency that balances these two costs.

Formally, we assume that firms choose a constant quantity per each shipment, which is denoted by q_s. The associated inventory cost is assumed to have the form $C_I(q_s) = \kappa_I q_s$, since the amount of goods that need to be stored is proportional to q_s. And for each shipment, firms need to incur coordination cost $C_T(q_s) = \kappa_T q_s^\alpha$, where α measures how fast the coordination cost changes with quantity. When $\alpha < 1$, there is a return to scale in coordination: marginal cost of coordination is decreasing in quantity per shipment. And if $\alpha < 0$, the coordination cost $C_T(q_s)$ decreases with quantity per shipment. Finally we add a fixed cost component f^x that does not depend on frequency of shipment. The logistics problem is therefore

$$\min_{q_s} C_I(q_s) + \frac{q}{q_s} C_T(q_s) + f^x.$$

Three special cases are worth mentioning. When $\alpha = 0$, the coordination cost is constant per shipment, the problem becomes the same as the economic order quantity model. When $\alpha = 1$, total coordination cost $\frac{q}{q_s} C_T(q_s)$ will be constant, it is therefore optimal to set

[15] As an alternative way to justify inventory cost, if we assume demand is uniformly distributed across time, the goods shipped but not consumed immediately must be stored, and the producer can be assumed to share this inventory costs.

q_s as small as possible, and the variable part of logistics cost will be $\kappa_T q$. When $\alpha = -\infty$, coordination cost is zero if q_s is slightly above 1. By choosing q_s near 1, the total logistics cost can be made arbitrarily close to $\kappa_I + f^x$, which is fixed regardless of trade quantity. The first-order condition of the above problem gives

$$q_s^{2-\alpha} = \frac{(1-\alpha)\,\kappa_T}{\kappa_I} q$$

With $\alpha < 1$, then second-order condition[16] will be positive. Plug in the above solution into the cost function, the optimized logistics cost has the form

$$C_{LT}(q) = (1-\alpha)^{-\frac{1-\alpha}{2-\alpha}} (2-\alpha)\, \kappa_I^{\frac{\alpha-1}{\alpha-2}} \kappa_T^{\frac{1}{2-\alpha}} q^{\frac{1}{2-\alpha}} + f^x = Bq^{\nu_{LT}} + f^x, \tag{62}$$

where for the ease of exposition, we let

$$\nu_{LT} = \frac{1}{2-\alpha},$$

$$B = \frac{1}{\nu_{LT}^{\nu_{LT}} (1-\nu_{LT})^{1-\nu_{LT}}} \kappa_I^{1-\nu_{LT}} \kappa_T^{\nu_{LT}}.$$

Since $\lim_{\nu_{LT}\to 0} Bq^{\nu_{LT}} = \kappa_I$, algebraically the variable part of logistic costs reduces to a constant. On the other hand, $\lim_{\nu_{LT}\to 1} Bq^{\nu_{LT}} = \kappa_T q$, in which case the marginal logistics cost is equal to κ_T. The model also gives an explicit expression for the optimal frequency of shipment:

$$F = \left(\frac{1}{\nu_{LT}} - 1\right)^{-\nu_{LT}} \kappa_I^{\nu_{LT}} \kappa_T^{-\nu_{LT}} q^{(1-\nu_{LT})} = \frac{1}{1-\nu_{LT}} \kappa_I B^{-1} q^{(1-\nu_{LT})}. \tag{63}$$

It is also worth emphasizing that the logistics problem is separated from other demand or supply side assumptions. The relationship shown in (63) also provides a straightforward way to estimate ν_{LT} using linear regression. The value of ν_{LT} is an easy test to differentiate different models. When $\nu_{LT} = 0$, as is often assumed in the standard heterogeneous firm model, the frequency of shipment is equal to export quantity, namely $F = q$. When the

[16] Second-order condition is $q\kappa_T (\alpha-1)(\alpha-2)\, q_s^{\alpha-3}$

coordination cost per shipment is a constant, $\nu_{LT} = \frac{1}{2}$. Finally, when $\nu_{LT} = 1$, coordination cost is proportional to the export quantity, the current formula does not apply, but the model predicts an infinitely frequent shipment to reduce inventory cost as low as possible.

C.2 Estimation of ν_{LT}

The inventory management problem provides a straightforward way to estimate ν_{LT}. After taking log, the equation (63) becomes:

$$\ln F = \ln \left(\frac{1}{\nu_{LT}} - 1 \right)^{-\nu_{LT}} \kappa_I^{\nu_{LT}} \kappa_T^{-\nu_{LT}} + (1 - \nu_{LT}) \ln q, \tag{64}$$

therefore we can estimate ν_{LT} by simply regressing log shipment frequency on log quantity. We first estimate ν_{LT} using the whole sample. Since the data contains a 7-year time series, we include quadratic function of the number of years of positive trade for each firm-product-destination to control for experience effect. More specifically, we run the regression

$$\ln F_{fojt} = \beta_0 + (1 - \nu_{LT}) \ln q_{fojt} + \beta_2 Exp_{fojt} + \beta_3 \left(Exp_{fojt} \right)^2 + \phi_{foj} + u_{fojt}, \tag{65}$$

where F_{fojt} is the number of shipment frequency firm f, product o, export to destination j in year t, q_{fojt} is trade volume, Exp_{fojt} is a measure of past trade experience, and ϕ_{foj} is the firm-product-destination fixed effects. We add the experience measure and fixed effects in order to control for the potential heterogeneity in $\kappa_I^{\nu_{LT}} \kappa_T^{-\nu_{LT}}$. As shown in Table 8, the estimate of ν_{LT} is close to 0.6. Importantly, the estimated value is significantly difference from either 0 or 1, showing that commonly assumed trade cost structure that imposes $\nu_{LT} = 0$ is not consistent with the data.

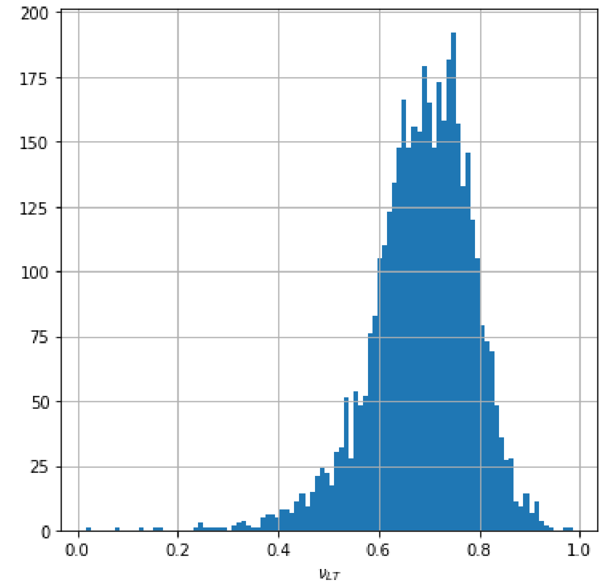
Variable	Estimate
ν_{LT}	0.6165^{***}
	(0.001)
β_2	0.0318^{***}
	(0.001)
β_3	-0.0163^{***}
	(0.000)
No. Obs.	11623454
R^2	0.495

Note: Estimate of ν_{LT} using the whole sample. Firm-product-destination fixed effects are controlled. Standard errors are reported in the brackets.

Table 8: Estimation of ν_{LT}

To have a better idea of potential heterogeneity of coordination costs across different sub-samples, using the year 2006 data, we run the frequency regression (64) for each product defined by the HS8 codes, controlling for the destination fixed effects. Figure 10 shows the distribution of estimated ν_{LT} across products, where we exclude products that have less than 30 observations. Note that consistent with the estimate using the whole sample, we clearly rejects the hypothesis that $\nu_{LT} = 0$. While $\nu_{LT} = 0.5$ or $\nu_{LT} = 1$ is closer to the data, the mean of this distribution is 0.66, with a small standard deviation equal to 0.09.[17]

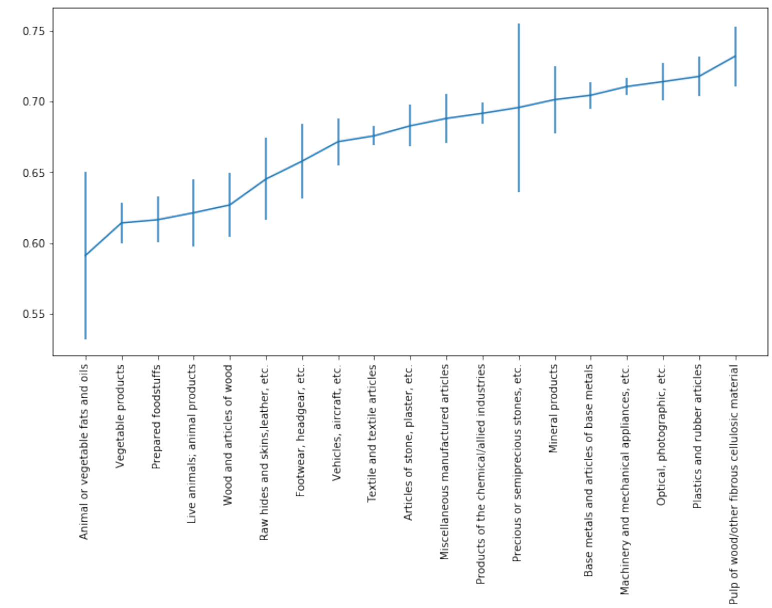
[17]Note that in contrast with theory, the frequency of shipment is measured as number of transactions, and therefore is a discrete variable in the data. To test whether this issue significantly affects our result, we run a Poisson regression instead of ordinary linear regression, and find that the results only change slightly.



Note: Distribution of estimated ν_{LT} across products. Products that have less than 30 observations are excluded. Mean of the estimates is 0.66, standard deviation is 0.09.

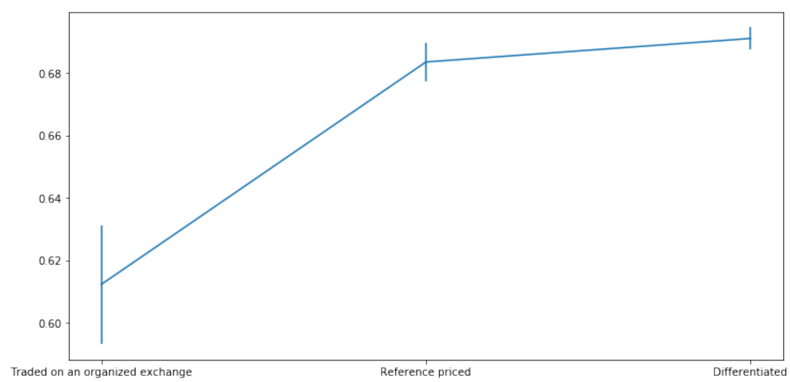
Figure 10: Distribution of estimated ν_{LT} across products

In order to understand what drives the heterogeneity of ν_{LT} across products, we regress the estimated product level ν_{LT} on the sector dummy. Figure 11 shows that the average value of ν_{LT} ranges from slightly above 0.55 to about 0.7. The vertical lines shows the 95 confidence interval of ν_{LT} in each sector. Sectors are ordered by the estimated value of ν_{LT}. Products that seem to be more homogeneous, like animal fat and prepared foodstuffs, tend to have lower ν_{LT} and therefore a lower logistics costs conditional on the quantity shipped. On the other hand, the more complex products like machinery tend to have a higher value of ν_{LT}. Consistent with this intuition, Figure 12 shows that when we regress estimated ν_{LT} across categories defined by the Rauch classification, a clear sorting pattern emerges: the differentiated products have a significantly higher value of ν_{LT} than the products having a reference price, which has ν_{LT} higher than the products traded in the organized exchange.



Note: Estimated ν_{LT} across sectors. The vertical lines shows the 95 confidence intervals of estimated ν_{LT} in each sector.

Figure 11: Estimated ν_{LT} across sectors of products



Note: Estimated ν_{LT} across Rauch categories. The vertical lines shows the 95 confidence intervals of estimated ν_{LT} in each category.

Figure 12: Estimated ν_{LT} across Rauch classifications

D Calibration under Pareto

In this section, we illustrate the idea of identifying trade costs parameters using flows under the assumption of $G(a)$ being Pareto distribution and $f_{ij}^{x} = 0$. Under these assumptions, we can apply the aggregation formula (28) to get the closed form solution for the bilateral trade flows. The number of firms that export from i to j is

$$N_{ij} = N_i^e \int_0^{a_{ij}^*} dG(a) = N_i^e G\left(a_{ij}^*\right)$$

$$= H(0)\left(A_j\right)^{\frac{\nu_{LT}\theta-\theta}{\nu_{LT}-\nu_R}} Z_{ij}\left(w_i B_{ij}\right)^{-\frac{\theta(1-\nu_R)}{\nu_R-\nu_{LT}}}, \tag{66}$$

where $Z_{ij} = N_i^e \kappa_{G,i}^{\theta}\left(w_i \tau_{ij}\right)^{-\theta}$. Therefore together with the expression on bilateral trade flow (54), the intensive margin is

$$\frac{X_{ij}}{N_{ij}} = \frac{H(\nu_R)}{H(0)}\left(A_j\right)^{\frac{\nu_{LT}}{\nu_{LT}-\nu_R}}\left(w_i B_{ij}\right)^{\frac{\nu_R}{\nu_R-\nu_{LT}}}. \tag{67}$$

The total frequency of shipments is

$$F_{ij} = N_i^e \int_0^{a_{ij}^*} F_{fj} dG(a) = N_i^e \int_0^{a_{ij}^*} \frac{1}{1-\nu_{LT}} \kappa_I B_{ij}^{-1} q_{fj}^{(1-\nu_{LT})} dG(a)$$

$$= \frac{H(1-\nu_{LT})}{1-\nu_{LT}} \kappa_I \left(A_j\right)^{\frac{(1-\nu_{LT})+\nu_{LT}\theta-\theta}{\nu_{LT}-\nu_R}} Z_{ij}\left(w_i B_{ij}\right)^{-\frac{1-\nu_R+\nu_R\theta-\theta}{\nu_{LT}-\nu_R}}.$$

Average shipment frequency is

$$\frac{F_{ij}}{N_{ij}} = \frac{H(1-\nu_{LT})}{H(0)(1-\nu_{LT})} \kappa_I \left(A_j\right)^{\frac{(1-\nu_{LT})}{\nu_{LT}-\nu_R}}\left(w_i B_{ij}\right)^{\frac{1-\nu_R}{\nu_R-\nu_{LT}}} \tag{68}$$

Given that we observe X_{ij}, N_{ij} and F_{ij} in the data, and ν_{LT} can be estimated separately, we can solve for A_j, B_{ij}, Z_{ij} up to a level of wage and the shape parameter of marginal cost distribution θ. Since we observe firm level sales, we can identify θ from the shape of the sales distribution.

However, despite the transparency in identification, the assumption of Pareto generates peculiar predictions regarding the intensive margins of trade. We could not think of a plausible theoretical reason for this to hold true. Fortunately, as suggested in Fernandes *et al.* (2019), when the marginal costs follow log-normal distribution, the Melitz model is able to match the patterns of intensive margin. Therefore in the next sub-section, we calibrate the model using the same logic but with the assumption that the marginal costs follow log-normal distribution.

E Auxiliary Empirical Analysis

In this subsection we use the sample of Guangdong province to show evidence that supports the idea how firms arrange shipment may reflect their underlying trade costs. Using data from a single province controls for the potential origin effects. Guangdong Province is the largest province in terms of trade value, accounting for about 30% of total trade value in China. We first examine the difference in shipment frequency between intermediaries and other firms. The literature studies intermediaries (e.g. Bai *et al.* (2017)) point out that firms export via intermediaries pay less entry cost, but incur higher variable cost, because producers have to pay a margin to intermediaries, and they also lose the chance of directly interact with buyers , thus are less likely to build long-term trust with buyers. This argument shows that trade via intermediaries is likely to have higher coordination cost and thus lower frequency of shipment. The regression we run to test this is

$$\log F_{foj} = \gamma_0 + (1 - \nu_{LT}) \log q_{foj} + \phi_{oj} + u_{foj},$$

where ϕ_{oj} is product-country fixed effect. The regressions are run for both the sub-sample of intermediary firms and non-intermediary firms. table 9 shows that the sub-sample of intermediaries not only has a lower value of intercept, but also has a higher estimated ν_{LT}. If the inventory costs are product specific and therefore on average the same across intermediaries and non-intermediaries, the lower intercept implies a higher coordination cost parameter κ_T for the intermediaries. Moreover, the higher estimated ν_{LT} shows that the logistics costs increase at a faster speed for the intermediaries.

	Intermediary	Non-intermediary
Intercept	-0.0657^{***}	0.0794^{***}
t-value	-153.959	133.821
ν_{LT}	0.838^{***}	0.787^{***}
t-value	704.982	752.8732
No. Obs	1321772	1022501
Adj. R^2	0.27	0.36

Note: Standard deviation is reported in the bracket, product-country fixed effect is controlled.

Table 9: Shipment frequency and intermediary

As another evidence of the associate between shipment frequency and trade costs, we run a gravity style regression and test whether the usual proxies of trade barriers are good predictors of shipment frequency. In the regression, we put annual number of shipment for each firm-product-destination on the left hand side, controls for value of shipment and gravity variables on the right hand side. For the gravity variables, besides GDP and GDP per ca pita, we also add proxy for trade costs including contiguous or not, sharing common language or not, and the cost of business startup in the destination country. The gravity variables are obtained from CEPII. To use the variation across destinations, firm-product fixed effect is controlled. Inference is clustered at firm level. Table 10 shows that contiguity and common language indeed associate higher frequency of shipment, and a higher business startup costs is associated with lower frequency of shipment.

	Estimation result	t-value
$\log X_{foj}$	0.305	214.057
$\log GDP_j$	0.055	78.143
$\log GDPpca_j$	0.016	14.858
$Contig_{ij}$	0.014	17.181
$Enthno_{ij}$	0.042	27.028
$BusiCost_{ij}$	-0.002	-1.956
No. Obs	2491115	
Adj. R^2	0.56	

Note: Firm-product fixed effect is controlled, standard deviation clustered at firm level.

Table 10: Shipment frequency and gravity variables