

Trade Models and Trade Elasticities

Ina Simonovska
University of California, Davis and NBER

Michael E. Waugh
New York University

This Version: February 2014

ABSTRACT

This paper shows that new trade models with different micro-level margins—estimated to fit the same moments in the data—imply lower trade elasticities and, hence, larger welfare gains from trade relative to models without these margins. The key feature of the estimation approach is to focus on common moments where new micro-level margins, such as an extensive margin or variable markups, alter the mapping from the data to the estimate of the trade elasticity. We find that the introduction of an extensive margin as in Eaton and Kortum (2002) or Melitz (2003) increases the welfare cost of autarky by up to 50 percent relative to Armington or Krugman (1980) which feature no extensive margin. Variable markups in Bernard, Eaton, Jensen, and Kortum (2003) further increase the welfare cost of autarky by 50 percent relative to Eaton and Kortum (2002) which features perfect competition.

JEL Classification: F10, F11, F14, F17

Keywords: elasticity of trade, bilateral, gravity, price dispersion, indirect inference

Email: inasimonovska@ucdavis.edu, mwaugh@stern.nyu.edu.

We thank Ariel Burstein and seminar participants at Columbia, Northwestern, Penn, New York Fed, Chicago, Houston, Arizona State, West Coast Trade Workshop at UCSC, Spring 2012 International Trade Workshop at FIU, AEA 2012, Princeton, Penn State, and NYU for their feedback.

1. Introduction

How has the development of new trade models with different micro-level margins changed our understanding of the welfare gains from trade? In this paper, we show that different models—estimated to fit the same moments in the data—imply different trade elasticities and, hence, different welfare gains from trade. The key insight is that the different margins of adjustment in new trade models, e.g., an extensive margin or variable markups, alter the mapping from the data to the estimate of the trade elasticity. Thus, while our approach uses moment conditions common across models, different models will have different trade elasticities and, hence, different welfare gains from trade.

The particular estimation approach that we use builds on Simonovska and Waugh (2011). In the present paper we show that the aforementioned methodology represents a common estimator for the trade elasticity that is applicable across models that feature within-country product differentiation. The basic idea behind our estimation strategy is to match moments about bilateral price differences at the product level relative to trade flows between the model and the data. What is critical to our approach is that all trade models have implications for these moments (in contrast to moments about size or value added across firms) and hence, our estimator is applicable across trade models.

Using disaggregate price and trade-flow data for the year 2004 from the thirty largest economies in the world, we apply our estimator to five canonical models of trade. We find that the introduction of an extensive margin as in Eaton and Kortum (2002) or Melitz (2003) increases the welfare cost of autarky by up to 50 percent relative to Armington or Krugman (1980) which feature no extensive margin. Variable markups in Bernard, Eaton, Jensen, and Kortum (2003) further increase the welfare cost of autarky by 50 percent relative to Eaton and Kortum (2002) which features perfect competition.

Given the stark welfare implications arising from our exercise, it is important to understand how our estimation method works and why we obtain estimates of the parameter that differ across models. Our estimation matches moments about the bilateral maximal price differences at the product level scaled by trade flows between the model and the data. Maximal price differences are meaningful because they reveal information about the size of the unobserved trade frictions. This information along with an assumption about the underlying model can then be used to readily identify the trade elasticity and, hence, the welfare gains from trade.

Differences in our estimates across models correspond exactly with the different micro-level margins introduced. For example, consider the ranking across Armington, Eaton and Kortum (2002), and Bernard, Eaton, Jensen, and Kortum (2003). In the Armington model, the maximal bilateral price difference essentially equals the unobserved trade friction. In Eaton and Kortum (2002), the maximal bilateral price difference always lies below trade frictions because of the

presence of endogenously non-traded goods, i.e. an extensive margin. Hence, when viewing maximal price differences through the lens of the Eaton and Kortum (2002) model, the unobserved trade frictions must be larger relative to Armington, and hence one needs a lower trade elasticity to rationalize the same amount of trade observed in the data.

Relative to these models, Bernard, Eaton, Jensen, and Kortum's (2003) introduces an additional margin through variable markups. In this model, producers are able to price at a markup over marginal costs depending on other latent competitors. Unlike in Eaton and Kortum (2002), the maximal bilateral price difference reflects both markups, trade frictions, and productivity differences (if the good is not traded between the pair). The crucial observation is that because mark-ups are negatively correlated to marginal costs, the maximal bilateral price difference lies even further below the trade friction relative to the Eaton and Kortum (2002) model. Hence, when viewing maximal price differences through the lens of the Bernard, Eaton, Jensen, and Kortum's (2003) model, the unobserved trade frictions must be larger relative to Eaton and Kortum (2002), and hence one needs a even lower trade elasticity to rationalize the same amount of trade observed in the data.

Finally, the distinction between Krugman (1980) and Melitz (2003) is because of firm heterogeneity and patterns of selection into export markets. In the Krugman (1980) model, the maximal bilateral price difference is likely to equal the unobserved trade friction (exactly like in the Armington model). In the Melitz (2003) model, only goods sold by firms that enter all markets are able to make the sample of prices that data collectors would consider. Moreover, the firms that enter all markets are likely to be from sources with low trade costs to export. This implies that the average maximal price difference across country pairs is likely to be below the average trade friction across country pairs. Thus one needs an even lower trade elasticity in Melitz (2003) relative to Krugman (1980) to rationalize the same amount of trade observed in the data.

Arkolakis, Costinot, and Rodriguez-Clare (2011) provide one answer to this question. They show that the trade elasticity is one of only two statistics needed to measure the welfare cost of autarky in a large and important class of trade models.¹ Hence, the authors argue that—for a given trade elasticity—new trade models yield the same welfare gains as in old trade models.

¹The class of models includes Armington, Krugman (1980), Eaton and Kortum (2002), Bernard, Eaton, Jensen, and Kortum (2003), and Melitz (2003) as articulated in Chaney (2008), which all generate log-linear relationships between bilateral trade flows and trade frictions.

2. Five Canonical Models of Trade

2.1. Armington

The simplest model of international trade that yields a gravity equation is the Armington model outlined in Anderson and van Wincoop (2003). The framework features N countries populated by consumers with constant elasticity of substitution (CES) preferences and tradable goods that are differentiated by the country of origin. Perfect competition among producers and product differentiation by origin imply that each good is purchased in each destination at a price that equals the marginal cost of production and delivery of the good there.

We consider a more empirically-relevant version of the Armington model that features product differentiation within and across countries.² Throughout the paper, let i denote the source country and n the destination. We allow each country to produce an exogenously given measure of tradable goods equal to $1/N$. We assume that products are differentiated. Within each country n , there is a measure of consumers L_n . Each consumer has one unit of time supplied inelastically in the domestic labor market and enjoys the consumption of a CES bundle of final tradable goods with elasticity of substitution $\rho > 1$,

$$U_n = \left[\sum_{i=1}^N \int_0^{1/N} x_{ni}(j)^{\frac{\rho-1}{\rho}} dj \right]^{\frac{\rho}{\rho-1}}.$$

To produce quantity $x_{ni}(j)$ in country i , a firm employs labor using a linear production function with productivity $T_i^{1/\theta}$, where T_i is a country-specific technology parameter and $\theta = \rho - 1$. The perfectly competitive firm from country i incurs a marginal cost to produce good j of $w_i/T_i^{1/\theta}$, where w_i is the wage rate in the economy. Shipping the good to a destination n further requires a per-unit iceberg trade cost of $\tau_{ni} > 1$ for $n \neq i$, with $\tau_{ii} = 1$. We assume that cross-border arbitrage forces effective geographic barriers to obey the triangle inequality: For any three countries i, k, n , $\tau_{ni} \leq \tau_{nk}\tau_{ki}$. We maintain the assumption on the existence of iceberg trade costs in all five models.

Perfect competition forces the price of good j from country i to destination n to be equal to the marginal cost of production and delivery

$$p_{ni}(j) = \frac{\tau_{ni}w_i}{T_i^{1/\theta}}.$$

Since goods are differentiated, consumers in destination n buy all products from all sources and pay $p_{ni}(j)$ for good j from i . Hence, all tradable goods are traded among the N markets and

²This model yields identical aggregate predictions to the simple Armington model, but it accommodates a greater number of goods than countries, which is an observation that holds true in the data.

consumers buy a unit measure of goods.

2.2. Eaton and Kortum (2002)

We now outline the environment of the multi-country Ricardian model of trade introduced by Eaton and Kortum (2002). As in the Armington model, there is a continuum of tradable goods indexed by $j \in [0, 1]$. Preferences are represented by the following utility function

$$V_n = \left[\int_0^1 x_n(j)^{\frac{\rho-1}{\rho}} dj \right]^{\frac{\rho}{\rho-1}}.$$

To produce quantity $x_i(j)$ in country i , a firm employs labor using a linear production function with productivity $z_i(j)$. Unlike the Armington model, country i 's productivity is the realization of a random variable (drawn independently for each j) from its country-specific Fréchet probability distribution

$$F_i(z_i) = \exp(-T_i z_i^{-\theta}).$$

The country-specific parameter $T_i > 0$ governs the location of the distribution; higher values of it imply that a high productivity draw for any good j is more likely. The parameter $\theta > 1$ is common across countries and, if higher, it generates less variability in productivity across goods.

Having drawn a particular productivity level, a perfectly competitive firm from country i incurs a marginal cost to produce good j of $w_i/z_i(j)$. Perfect competition forces the price of good j from country i to destination n to be equal to the marginal cost of production and delivery

$$p_{ni}(j) = \frac{\tau_{ni} w_i}{z_i(j)}.$$

So, consumers in destination n would pay $p_{ni}(j)$, should they decide to buy good j from i .

Consumers purchase good j from the low-cost supplier; thus, the actual price consumers in n pay for good j is the minimum price across all sources k

$$p_n(j) = \min_{k=1, \dots, N} \left\{ p_{nk}(j) \right\}.$$

Hence, the Eaton and Kortum (2002) framework introduces endogenous tradability into the Armington model outlined above. In particular, countries export only a subset of the unit measure of tradable goods for which they are the most efficient suppliers.

2.3. Bernard, Eaton, Jensen, and Kortum (2003)

Bernard, Eaton, Jensen, and Kortum (2003) introduce Bertrand competition into Eaton and Kortum's (2002) model. The most important implication from this extension is that individual good prices differ from the Eaton and Kortum (2002) model.

Let $c_{kni}(j) \equiv \tau_{ni}w_i/z_{ki}(j)$ be the cost that the k -th most efficient producer of good j in country i faces in order to deliver a unit of the good to destination n . With Bertrand competition, as with perfect competition, the low-cost supplier of each good serves the market. For good j in market n , this supplier has the following cost $c_{1n}(j) = \min_i \{c_{1ni}(j)\}$. This supplier is constrained not to charge more than the second-lowest cost of supplying the market, which is $c_{2n} = \min \{c_{2ni^*}(j), \min_{i \neq i^*} \{c_{1ni}(j)\}\}$, where i^* satisfies $c_{1ni^*}(j) = c_{1n}(j)$. Hence, the price of good j in market n is

$$p_n(j) = \min \{c_{2n}(j), \bar{m}c_{1n}(j)\},$$

where $\bar{m} = \rho/(\rho - 1)$ is the Dixit-Stiglitz constant mark-up.

Finally, for each country i , productivity, $z_{ki}(j)$ for $k = 1, 2$ is drawn from

$$G_i(z_1, z_2) = [1 + T_i(z_2^{-\theta} - z_1^{-\theta})] \exp(-T_i z_2^{-\theta}).$$

Hence, the Bernard, Eaton, Jensen, and Kortum (2003) model features a key additional component relative to the Eaton and Kortum (2002) framework: the existence of variable mark-ups. In particular, the most efficient suppliers in this model enjoy the highest mark-ups.

2.4. Krugman (1980)

We now depart from the Ricardian framework and outline a multi-country version of the Krugman (1980) model. Preferences are identical to the Armington model, but the measure of goods produced by any country i is endogenous,

$$W_n = \left[\sum_{i=1}^N \int_{J_i} x_{ni}(j)^{\frac{\rho-1}{\rho}} dj \right]^{\frac{\rho}{\rho-1}}.$$

In the above expression, J_i denotes the set of goods produced by country i .

In particular, we assume that a producer in country i has monopoly power over a blueprint to produce a single good j . In order to produce the good, a firm in i uses the following production technology

$$y_i(j) = T_i^{1/\theta} l_i(j) + f_i.$$

Thus a firm incurs a marginal cost of $w_i/T_i^{1/\theta}$ as well as a fixed cost f_i , both in labor units. There is an unbounded pool of potential producers and the mass of firms is pinned down via a zero average profit condition.

We make the assumption that fixed costs are proportional to the mass of workers in country i , L_i . Under the proportionality assumption on fixed costs, the mass of entrants is identical across countries and proportional to $1/(1 + \theta)$.

Conditional on entry, all consumers buy all products from all sources. Moreover, due to the proportionality assumption on entry costs, each country contributes a measure of $1/N$ toward the unit measure of consumed products as in the Armington model. In contrast to Armington, the price $p_{ni}(j)$ for good j from i in destination n is given by the product of the marginal cost of production and delivery and the Dixit-Stiglitz constant mark-up $\bar{m}$.

2.5. Melitz (2003)

Finally, we outline a variant of the Melitz (2003) model parameterized as in Chaney (2008). In the exposition, we follow closely Arkolakis, Demidova, Klenow, and Rodríguez-Clare (2008).

The preference relation as well as the market structure are identical to the Krugman (1980) model described above. There are two novel features. First, in order to sell its product to market n , a firm from any country i incurs an additional market access cost, a_n , which is proportional to the mass of workers in the destination, L_n . Second, marginal costs of production are not identical across producers within a country. In particular, upon paying the entry cost f_i , a firm from i draws a productivity realization for good j , $z_i(j)$, from the country-specific Pareto distribution with pdf

$$\mu_i(z) = T_i z^{-\theta}. \quad (1)$$

As in the Frechet distribution, the parameter θ in the Pareto distribution governs the variability of productivity among firms.

The two additional assumptions imply that there exist cost thresholds that limit the participation of certain firms to certain markets. In particular, if a firm obtains a cost draw above the threshold cost that characterizes the “most accessible” destination, it exits immediately without operating. Thus, because of free entry, in equilibrium expected profits of a firm must be equal to the fixed entry cost. In addition, more productive firms serve more markets and the “toughest” markets attract the most efficient producers from all sources.

As in the Krugman (1980) model, firms charge a constant mark-up over their marginal cost of production. However, unlike the Krugman (1980) model, the prices of goods offered by a given country differ according to the efficiency of the individual producer. More importantly, even

though entry costs are proportional to the mass of workers, the measure of firms that serve each market is no longer constant. Thus, selection among exporters alters the composition of the bundle of consumed goods.

2.6. Trade Flows, Aggregate Prices, and Welfare

In this section, we show that the five models described above produce identical aggregate outcomes, even though they feature different micro-level behavior. In particular, under the parametric assumption made above, the five models yield the same expressions for trade flows, price indices (up to a constant scalar), and welfare gains from trade. Proposition 1 summarizes the result.

Proposition 1 *Let i represent the country of origin and n denote the destination market. Given the functional forms for productivity $F_i(\cdot)$, $G_i(\cdot)$, and $\mu_i(\cdot)$ for all $i = 1, \dots, N$, and assuming that (i) fixed entry costs are proportional to the mass of workers in the country of origin, $f_i \propto L_i$ for all $i = 1, \dots, N$; (ii) market access costs are proportional to the mass of workers in the destination market, $a_n \propto L_n$ for all $n = 1, \dots, N$;*

a. The share of expenditures that n spends on goods from i , X_{ni}/X_n , predicted by all five models is

$$\frac{X_{ni}}{X_n} = \frac{T_i(\tau_{ni}w_i)^{-\theta}}{\sum_{k=1}^N T_k(\tau_{nk}w_k)^{-\theta}}. \quad (2)$$

b. The CES exact price index for destination n , P_n , predicted by all five models is

$$P_n \propto \Phi_n^{-\frac{1}{\theta}}, \quad \text{where} \quad \Phi_n = \sum_{k=1}^N T_k(\tau_{nk}w_k)^{-\theta}. \quad (3)$$

c. The percentage compensation that a representative consumer in n requires to move between two trading equilibria predicted by all five models is

$$\frac{P'_n}{P_n} - 1 = 1 - \left(\frac{X'_{nn}/X'_n}{X_{nn}/X_n} \right)^{\frac{1}{\theta}}. \quad (4)$$

We relegate the proof of the proposition to the Appendix and we focus on the welfare results in the remainder of this section. Across the models, the welfare gains from trade are essentially captured by changes in the CES price index that a representative consumer faces. Using the objects from the Proposition above, it is easy to relate the price indices to trade shares and the parameter θ . In particular, expressions (2) and (3) allow us to relate trade shares to trade costs

and the price indices of each trading partner via the following equation

$$\frac{X_{ni}/X_n}{X_{ii}/X_i} = \frac{\Phi_i}{\Phi_n} \tau_{ni}^{-\theta} = \left(\frac{P_i \tau_{ni}}{P_n} \right)^{-\theta}, \quad (5)$$

where $\frac{X_{ii}}{X_i}$ is country i 's expenditure share on goods from country i , or its home trade share. The welfare equation follows trivially from this expression.

2.7. The Elasticity of Trade

The key parameter determining trade flows (equation (5)) and welfare (equation (4)) is θ . To see the parameter's importance for trade flows, take logs of equation (5) yielding

$$\log \left(\frac{X_{ni}/X_n}{X_{ii}/X_i} \right) = -\theta [\log(\tau_{ni}) - \log(P_i) + \log(P_n)]. \quad (6)$$

As this expression makes clear, θ controls how a change in the bilateral trade costs, τ_{ni} , will change bilateral trade between two countries. This elasticity is important because if one wants to understand how a bilateral trade agreement will impact aggregate trade or to simply understand the magnitude of the trade friction between two countries, then a stand on this elasticity is necessary. This is what we mean by the elasticity of trade.

To see the parameter's importance for welfare, it is easy to demonstrate that (4) implies that θ represents the inverse of the elasticity of welfare with respect to domestic expenditure shares

$$\log(P_n) = -\frac{1}{\theta} \log \left(\frac{X_{nn}}{X_n} \right).$$

Hence, decreasing the domestic expenditure share by one percent generates a $(1/\theta)/100$ -percent increase in consumer welfare. Thus, in order to measure the impact of trade policy on welfare, it is sufficient to obtain data on realized domestic expenditures and an estimate of the elasticity of trade.

Given θ 's impact on trade flows and welfare, this elasticity is absolutely critical in any quantitative study of international trade.

3. Estimating θ

There are three ways to estimate the parameter θ in (5). One approach uses data on changes in bilateral trade flows and tariffs during trade liberalization episodes as in Head and Ries (2001) and Romalis (2007). The approach typically associates the entire change in trade flows during a trade liberalization with changes in tariffs. This results in high estimates of the trade elasticity, since changes in non-tariff barriers that occur during trade liberalizations are not accounted

for in the estimation. Moreover, instrumental-variable methods aimed to alleviate the omitted-variables bias in the estimation yield a wide range of estimates for the trade elasticity. Yet, a clear guide toward choosing the best control variables is not available because the estimation does not tailor to a particular structural model.

Equation (5) suggests a second approach to estimate θ , if one had data on trade shares, aggregate prices, and trade costs. The key issue is that trade costs are not observed. In what follows, we explain how Eaton and Kortum (2002) approximate trade costs and aggregate prices in order to estimate θ . Then, we outline our simulated-method of moments estimator, which builds on Simonovska and Waugh (2011), and uses moments from Eaton and Kortum's (2002) methodology in order to back out θ in each model.

3.1. Approximating Trade Costs

The main problem with estimating θ is that one must disentangle θ from trade costs, which are not observed. Eaton and Kortum (2002) propose approximating trade costs using *disaggregate* price information across countries. In particular, in the Eaton and Kortum (2002) model, the maximum price difference across goods between two countries bounds the bilateral trade cost, which solves the indeterminacy issue.

To illustrate this argument, suppose that we observe the price of good ℓ across locations, but we do not know its country of origin.³ We know that the price of good ℓ in country n relative to country i must satisfy the following inequality

$$\frac{p_n(\ell)}{p_i(\ell)} \leq \tau_{ni}. \quad (7)$$

That is, the relative price of good ℓ must be less than or equal to the trade friction. This inequality must hold because if it does not, then $p_n(\ell) > \tau_{ni}p_i(\ell)$ and an agent could import ℓ at a lower price. Thus, the inequality in (7) places a lower bound on the trade friction.⁴

Improvements on this bound are possible if we observe a sample of L goods across locations. This follows by noting that the *maximum* relative price must satisfy the same inequality

$$\max_{\ell \in L} \left\{ \frac{p_n(\ell)}{p_i(\ell)} \right\} \leq \tau_{ni}. \quad (8)$$

This suggests a way to exploit *disaggregate* price information across countries and to arrive at an estimate of τ_{ni} by taking the maximum of relative prices over goods. Thus, Eaton and Kortum

³This is the most common case, though Donaldson (2009) exploits a case where he knows the place of origin for one particular good, salt. He argues convincingly that in India, salt was produced in only a few locations and exported everywhere; thus, the relative price of salt across locations identifies the trade friction.

⁴We apply a similar argument to the remaining four models in the Appendix.

(2002) approximate τ_{ni} , in logs, by

$$\log \hat{\tau}_{ni}(L) = \max_{\ell \in L} \{ \log(p_n(\ell)) - \log(p_i(\ell)) \}, \quad (9)$$

where the “hat” denotes the approximated value of τ_{ni} and (L) indexes its dependence on the sample size of prices.

3.2. Estimating the Elasticity

Given the approximation of trade costs, Eaton and Kortum (2002) derive an econometric model that corresponds to (6). For a sample of L goods, they estimate a parameter, β , using a method of moments estimator, which takes the ratio of the average of the left-hand side of (6) to the average of the term in the square bracket of the right-hand side of (6), where the averages are computed across all country pairs. Mathematically, their estimator is

$$\hat{\beta} = - \frac{\sum_n \sum_i \log \left(\frac{X_{ni}/X_n}{X_{ii}/X_i} \right)}{\sum_n \sum_i \left(\log \hat{\tau}_{ni}(L) + \log \hat{P}_i - \log \hat{P}_n \right)}, \quad (10)$$

$$\text{where } \log \hat{\tau}_{ni}(L) = \max_{\ell \in L} \{ \log p_n(\ell) - \log p_i(\ell) \},$$

$$\text{and } \log \hat{P}_i = \frac{1}{L} \sum_{\ell=1}^L \log(p_i(\ell)).$$

The value of β is EK’s preferred estimate of the elasticity θ . Throughout, we will denote by $\hat{\beta}$ the estimator defined in equation (10) to distinguish it from the value θ . As discussed, the second line of expression (10) approximates the trade cost. The third line approximates the aggregate price indices. The top line represents a rule that combines these statistics, together with observed trade flows, in an attempt to estimate the elasticity of trade.

3.3. Simonovska and Waugh (2011) Estimator

Simonovska and Waugh (2011) show that EK’s estimator of the trade elasticity is biased in any finite sample of goods’ prices. In particular, assuming that price and trade data are generated from the Eaton and Kortum (2002) model, the expected maximum log price difference is biased below the true log trade cost

$$\mathbb{E} \left(\max_{\ell \in L} \{ \log p_n(\ell) - \log p_i(\ell) \} \right) < \tau_{ni}$$

The intuition for the result is as follows. Relative prices of goods are bounded above by trade barriers, so the maximum operator over a finite sample of prices underestimates the trade cost

with positive probability and overestimates the trade cost with zero probability. Consequently, the maximum price difference lies strictly below the true trade cost, in expectation. This necessarily implies that any estimate β obtained from the rule $\hat{\beta}$ lies strictly above the elasticity parameter θ , in expectation,

$$\mathbb{E}(\hat{\beta}) > \theta.$$

The authors then propose to use the moment β from Eaton and Kortum's (2002) approach and to recover the parameter value for θ via a simulation of the Eaton and Kortum (2002) model. The procedure successfully recovers the parameter value for θ because Eaton and Kortum's (2002) estimator is biased but monotone, and nearly proportional to θ . Thus, using this informative moment as a basis for estimation, the procedure involves the following steps:

- A.** Recover necessary parameters (except θ) from trade data.
- B.** Estimate $\beta(\theta)$ using artificial data and compare to β from real data.
- C.** Update θ until $\beta(\theta)$ is "close" to β .

This procedure naturally extends to the other four models outlined earlier. In particular, assuming that the price and trade flow data were generated from either of these models, one can repeat the three steps above and obtain estimates of the trade elasticity that are consistent with the models. In the sections that follow, we outline the simulation procedure as it applies to each model and we report estimates of the trade elasticity.

3.4. Simulation Approach

In this subsection, we follow the exposition in Simonovska and Waugh (2011), which applied to the Eaton and Kortum (2002) model, and we demonstrate how to recover all parameters of interest in the remaining models up to the unknown scalar θ from trade data only. Then we describe our simulation approach. This provides the foundation for the simulated method of moments estimator that we propose.

Step 1.—We estimate the parameters for the country-specific productivity distributions and trade costs from bilateral trade-flow data. We follow closely the methodologies proposed by Eaton and Kortum (2002) and Waugh (2010b). First, we derive the gravity equation from expression (2) by dividing the bilateral trade share by the importing country's home trade share,

$$\log \left(\frac{X_{ni}/X_n}{X_{nn}/X_n} \right) = S_i - S_n - \theta \log \tau_{ni}, \quad (11)$$

where S_i is defined as $\log [T_i w_i^{-\theta}]$. Note that (11) is a different equation than expression (5), which is derived by dividing the bilateral trade share by the exporting country's home trade share, and is used to estimate θ . S_i 's are recovered as the coefficients on country-specific dummy variables given the restrictions on how trade costs can covary across countries. Following the arguments of Waugh (2010b), trade costs take the following functional form

$$\log(\tau_{ni}) = d_k + b_{ni} + ex_i + \nu_{ni}. \quad (12)$$

Here, trade costs are a logarithmic function of distance, where d_k with $k = 1, 2, \dots, 6$ is the effect of distance between country i and n lying in the k -th distance interval.⁵ b_{ni} is the effect of a shared border in which $b_{ni} = 1$ if country i and n share a border and zero otherwise. The term ex_i is an exporter fixed effect and allows for the trade-cost level to vary depending upon the exporter. We assume that ν_{ni} reflects other factors and is orthogonal to the regressors and normally distributed with mean zero and standard deviation σ_ν . We use least squares to estimate equations (11) and (12).

Step 2.—The parameter estimates obtained from the first-stage gravity regression are sufficient to simulate trade flows and micro-level prices in each model up to a constant, θ .

The relationship is obvious in the estimation of trade barriers since $\log(\tau_{ni})$ is scaled by θ in (11).

Step 2a.— To simulate micro-level prices in the Eaton and Kortum (2002), see Simonovska and Waugh (2011).

Step 2b.— To see that we can simulate micro-level prices in the Bernard, Eaton, Jensen, and Kortum (2003) model as a function of θ only, we draw on an argument in Bernard, Eaton, Jensen, and Kortum (2003). To simulate their model, Bernard, Eaton, Jensen, and Kortum (2003) reformulate the model in terms of efficiency. In particular, given two productivity draws for good j in country i , $z_{1i}(j)$ and $z_{2i}(j)$, they define the following objects

$$\begin{aligned} u_{1i}(j) &= T_i z_{1i}(j)^{-\theta} \\ u_{2i}(j) &= T_i z_{2i}(j)^{-\theta} \end{aligned}$$

Bernard, Eaton, Jensen, and Kortum (2003) demonstrate that these objects are distributed according to

$$\begin{aligned} Pr[u_{1i} \leq u_1] &= 1 - \exp(-u_1) \\ Pr[u_{2i} \leq u_2 | u_{1i} = u_1] &= 1 - \exp(-u_2 + u_1) \end{aligned}$$

⁵Intervals are in miles: [0, 375); [375, 750); [750, 1500); [1500, 3000); [3000, 6000); and [6000, maximum]. An alternative to specifying a trade-cost function is to recover scaled trade costs as a residual using equation (5), trade data, and measures of aggregate prices as in Waugh (2010a).

To simulate trade flows and prices from the model, we define the following variables

$$v_{1i}(j) = \left(\frac{u_{1i}(j)}{T_i w_i^{-\theta}} \right)$$

$$v_{2i}(j) = \left(\frac{u_{2i}(j)}{T_i w_i^{-\theta}} \right)$$

Let $\tilde{S}_i = \exp \{S_i\}$, with $S_i = \log(T_i w_i^{-\theta})$ coming from gravity. Applying the pdf transformation rule, it is easy to demonstrate that $v_{1i}(j)$ is distributed according to

$$Pr[v_{1i} \leq v_1] = 1 - \exp(-\tilde{S}_i v_1) \quad (13)$$

Similarly,

$$Pr[v_{2i} \leq v_2 | v_{1i} = v_1] = 1 - \exp(-\tilde{S}_i v_2 + \tilde{S}_i v_1) \quad (14)$$

Thus, to simulate the model we draw minimum unit costs from (13), and conditional on these draws, we draw the second lowest unit costs from (14). Hence, having obtained the coefficients S_i from the first-stage gravity regression, we can simulate the inverse of marginal costs and prices.

To simulate the model, we assume that there are a large number (150,000) of potentially tradable goods. For each country, the inverse marginal costs are drawn from the country-specific distributions above and assigned to each good. Then, for each importing country and each good, the two lowest-cost suppliers across countries are found, realized prices are recorded, and aggregate bilateral trade shares are computed. From the realized prices, a subset of goods common to all countries is defined and the subsample of prices is recorded – i.e., we are acting as if we were collecting prices for the international organization that collects the data. The models of Eaton and Kortum (2002) and Bernard, Eaton, Jensen, and Kortum (2003) give a natural common basket of goods to be priced across countries. In these models, agents in all countries consume all goods that lie within a fixed interval, $[0, 1]$. Thus, we consider this common list in the simulated models and we randomly sample the prices of its goods across countries, in order to approximate trade barriers, much like it is done in the data.

Step 2c.— The simulation of the Melitz (2003) model is more intricate. In the Eaton and Kortum (2002) and Bernard, Eaton, Jensen, and Kortum (2003) models, a good is indexed by $j \in [0, 1]$ or an integer when discretized on the computer. In the Melitz (2003) model, different countries are consuming and producing different goods. To carry out the simulation, we discretize the continuum to a set of potentially produced goods, which equals to $300,000 \cdot N$. We proceed to determine the subset of goods produced domestically for each country. Then, we simulate inverse marginal costs for each good. Inverse marginal costs are determined following Eaton,

Kortum, and Kramarz (2011) who show how normalized inverse marginal costs can be sampled from a parameter free uniform distribution.

Then, we determine the set of exported varieties for all country-pairs and we compute their prices. We proceed to define the common set of goods whose prices will be sampled. The set includes the goods produced by firms with low enough cost draws so that they serve all N destinations. Let M_i represent the measure of goods from i that appear on the common list and let $M = \sum_m M_m$ be the common list. Then, the share of country i 's goods on the common list is

$$\frac{M_i}{M} = \frac{\min_l \tilde{S}_i \tau_{li}^{-\theta} \left[\sum_k \tilde{S}_k \tau_{lk}^{-\theta} \right]^{-1}}{\sum_m \min_l \tilde{S}_m \tau_{lm}^{-\theta} \left[\sum_k \tilde{S}_k \tau_{lk}^{-\theta} \right]^{-1}}. \quad (15)$$

Step 2d.—The simulation of the Armington and Krugman (1980) models is trivial because they do not feature micro-level heterogeneity. In these models, $\tilde{S}_i$ represents the common marginal cost of production across all producers in country i . Moreover, all produced goods appear on the common list, and each country contributes a fraction $1/N$ of them.

Step 3.—We added disturbances to the predicted trade shares with the disturbances drawn from a mean zero normal distribution with the standard deviation set equal to the standard deviation of the residuals from Step 1.

These steps then provide us with an artificial data set of micro-level prices and trade shares that mimic their analogs in the data. Given this artificial data set, we can then compute moments—as functions of θ —and compare them to the moments in the data.

3.5. Estimation

We perform an overidentified estimation with two moments. Below, we describe the moments we try to match and the details of our estimation procedure.

Moments. Define $\hat{\beta}_k$ as EK's method of moment estimator defined in (10) using the k th-order statistic over micro-level price differences. Then, the moments that we are interested in are

$$\hat{\beta}_k = - \frac{\sum_n \sum_i \log \left(\frac{X_{ni}/X_n}{X_{ii}/X_i} \right)}{\sum_n \sum_i \left(\log \hat{\tau}_{ni}^k(L) + \log \hat{P}_i - \log \hat{P}_n \right)}, \quad k = 1, 2 \quad (16)$$

where $\hat{\tau}_{ni}^k(L)$ is computed as the k th-order statistic over L micro-level price differences between countries n and i .

We denote the simulated moments by $\beta_1(\theta_m, u_s|m)$ and $\beta_2(\theta_m, u_s|m)$, which come from the analogous formula as in (16) and are estimated from artificial data generated from each model by

following **Steps 1-3** above. Note that these moments are a function of θ_m , where m denotes the possibility of a model-specific value, and depend upon a vector of random variables u_s associated with a particular simulation s . There are three components to this vector. First, there are the random productivity draws for production technologies for each good and each country. The second component is the set of goods sampled from all countries. The third component mimics the residuals ν_{ni} from equation (11), which are described in Section 3.4.

Stacking our data moments and averaged simulation moments gives us the following zero function

$$y(\theta_m) = \begin{bmatrix} \beta_1 - \frac{1}{S} \sum_{s=1}^S \beta_1(\theta_m, u_s|m) \\ \beta_2 - \frac{1}{S} \sum_{s=1}^S \beta_2(\theta_m, u_s|m) \end{bmatrix}. \quad (17)$$

Estimation Procedure. We base our estimation procedure on the moment condition

$$E[y(\theta_o)] = 0,$$

where θ_{om} is the true value of θ_m . Thus, our simulated method of moments estimator is

$$\hat{\theta}_m = \arg \min_{\theta_m} [y(\theta_m)' \mathbf{W} y(\theta_m)],$$

where $\mathbf{W}$ is a 2×2 weighting matrix that we discuss below.

The idea behind this moment condition is that, though $\hat{\beta}_1$ and $\hat{\beta}_2$ will be biased away from θ , the moments $\beta_1(\theta_m, u_s|m)$ and $\beta_2(\theta_m, u_s|m)$ will be biased by the same amount when evaluated at θ_{om} , in expectation. Viewed in this language, our moment condition is closely related to the estimation of bias functions discussed in MacKinnon and Smith (1998) and to indirect inference, as discussed in Smith (2008). The key issue in MacKinnon and Smith (1998) is how the bias function behaves. As we argued earlier the bias is monotonic in the parameter of interest. Furthermore, Figure 1 below shows that the bias is basically linear, so it is well behaved.

For the weighting matrix, we use the optimal weighting matrix suggested by Gouriéroux and Monfort (1996) for simulated method of moments estimators. Because the weighting matrix depends on our estimate of θ_m , we use a standard iterative procedure outlined in the next steps.

Step 4.—We make an initial guess of the weighting matrix $\mathbf{W}^0$ and solve for $\hat{\theta}_m^0$. Then, given this value we simulate the model to generate a new estimate of the weighting matrix.⁶ With the new estimate of the weighting matrix we solve for a new $\hat{\theta}_m^1$. We perform this iterative procedure

⁶The computation of this matrix is described in Gouriéroux and Monfort (1996).

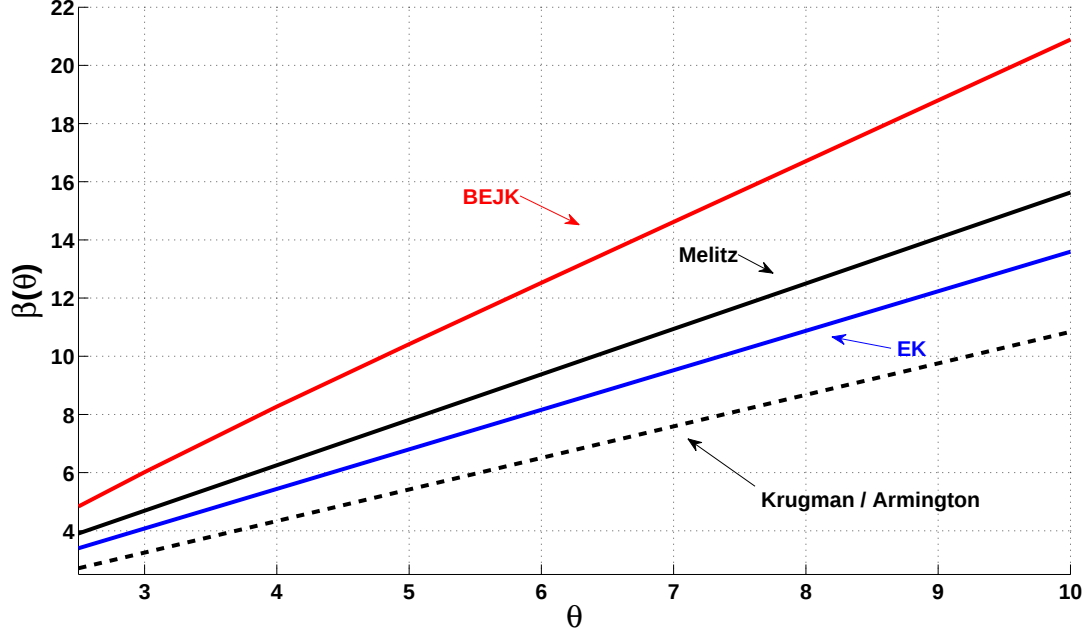


Figure 1: First Moment and θ

until our estimates of the weighting matrix and $\hat{\theta}_m$ converge. We explicitly consider simulation error because we utilize the weighting matrix suggested by Gouriéroux and Monfort (1996).

Step 5.—We compute standard errors using a parametric bootstrap technique. Assuming that the errors are normally distributed, we compute the variance of the normal distribution using the residuals from the data and the fitted values obtained using the estimates in (16). We proceed to simulate micro-level data, add error terms to the trade data, collect a sample of prices, and compute new estimates β_1^b and β_2^b . Next, we estimate the model using moments β_1^b and β_2^b and obtain θ^b . We repeat the procedure 100 times.

3.6. Performance on Simulated Data

Figure 1 in the previous section plots values for the moment $\beta_1(\theta_m)$ obtained from simulations of each model as we varied θ_m . It is clear that β_1 is a biased estimator for θ_m because the values do not lie on the 45° line. However, β_1 varies near linearly with θ_m . These observations motivated an estimation procedure that matches the data moments β_k to the moments $\beta_k(\theta_m)$ implied by the simulated model under a known θ_m .

To provide evidence that the estimation procedure recovers the underlying parameter for each model, we apply the methodology on data simulated by each model under a known θ_m . For exposition purposes, we let θ_m be the same across the models, and we set it equal to four as suggested by Simonovska and Waugh (2011). In order to apply our estimation procedure on

the Bernard, Eaton, Jensen, and Kortum (2003) model, a value for the elasticity of substitution is also needed. Throughout all exercises, we set $\rho = 1.5$.

Estimation Results With Artificial Data, Underlying $\theta = 4$			
	Estimate of θ	β_1	β_2
BEJK	3.98 [3.72, 4.19]	7.65	9.45
EK	3.96 [3.80, 4.18]	5.27	6.21
Melitz	4.03 [3.49, 4.76]	5.78	7.49
Armington/Krugman	4.00 [3.87, 4.35]	4.06	4.39

Note: Value is the mean estimate across simulations. In each simulation there are 18 countries and 100 simulations are performed. Values within brackets report 5 and 95 percentiles.

Table 1 reports the results from an overidentified estimation of θ applied to each model. Notice that the procedure successfully recovers the underlying parameter value for each model. However, the models yield different values for the moments of interest, β_1 and β_2 . The Bernard, Eaton, Jensen, and Kortum (2003) model yields the highest values for the two moments, while the Armington and Krugman (1980) models yield the lowest. These differences suggest that the five models will demand different values of the elasticity parameter in order to match the moments observed in the data.

4. Results Using ICP Data

In this section, we apply our estimation strategy to real data. Our sample contains the thirty largest countries in the world (in terms of absorption). We use trade flows and production data for the year 2004 to construct trade shares. The price data used to compute aggregate price indices and proxies for trade costs come from basic-heading-level data from the 2005 round of the International Comparison Programme (ICP). The dataset has been employed in a number of empirical studies. For example, Bradford (2003) and Bradford and Lawrence (2004) use the ICP price data in order to measure the degree of fragmentation, or the level of trade barriers, among OECD countries. In addition, the authors provide an excellent description of the data-collection process. Eaton and Kortum (2002) use a similar dataset for the year 1990 in their estimation.

The ICP collects price data on goods with identical characteristics across retail locations in the participating countries during the 2003-2005 period.⁷ The basic-heading level represents

⁷The ICP Methodological Handbook is available at <http://go.worldbank.org/MW520NNFK0>.

a narrowly-defined group of goods for which expenditure data are available. The data set contains a total of 129 basic headings, and we reduce the sample to 62 categories based on their correspondence with the trade data employed. Simonovska and Waugh (2011) provide a more detailed description of the ICP data.

The ICP provides a common list of “representative” goods whose prices are to be randomly sampled in each country over a certain period of time. A good is representative of a country if it comprises a significant share of a typical consumer’s bundle there. Thus, the ICP samples the prices of a common basket of goods across countries, where the goods have been pre-selected due to their highly informative content for the purpose of international comparisons.

Estimation Results				
	Estimate of θ	“J-statistic”	β_1	β_2
Data Moments	—	—	5.63	6.99
BEJK	2.81 [2.69, 2.92]	0.06 [0.80]	5.59	6.99
EK	4.21 [4.00, 4.37]	0.56 [0.79]	5.72	6.95
Melitz	3.41 [2.98, 3.86]	0.51 [0.82]	5.33	7.38
Armington/Krugman	5.21 [4.90, 5.46]	0.03 [0.72]	5.64	6.92

Note: The “J-statistic” reports the value $y(\hat{\theta})'W(\hat{\theta})y(\hat{\theta})$. Values within brackets report 5 and 95 percentiles.

4.1. Discussion of Results

The exercise resulted in different elasticity estimates across the three models. In particular, the Bernard, Eaton, Jensen, and Kortum (2003) model yields the lowest estimate of 2.81, while the Armington and Krugman (1980) models generate the highest estimate of 5.21. Since, welfare is inversely related to θ , the results imply that the Bernard, Eaton, Jensen, and Kortum (2003) model generates the highest, while the Armington and Krugman (1980) models yield the lowest gains from trade.

Given the stark welfare implications arising from our exercise, it is important to understand how our estimation method works and why we obtain estimates of the parameter that differ across models. Our estimation matches moments about the bilateral maximal price differences at the product level scaled by trade flows between the model and the data. Maximal price differences are meaningful because they reveal information about the size of the unobserved

trade frictions. This information along with an assumption about the underlying model can then be used to readily identify the trade elasticity and, hence, the welfare gains from trade.

Differences in our estimates across models correspond exactly with the different micro-level margins introduced. For example, consider the ranking across Armington, Eaton and Kortum (2002), and Bernard, Eaton, Jensen, and Kortum (2003). In the Armington model, the maximal bilateral price difference essentially equals the unobserved trade friction. In Eaton and Kortum (2002), the maximal bilateral price difference always lies below trade frictions because of the presence of endogenously non-traded goods, i.e. an extensive margin. Hence, when viewing maximal price differences through the lens of the Eaton and Kortum (2002) model, the unobserved trade frictions must be larger relative to Armington, and hence one needs a lower trade elasticity to rationalize the same amount of trade observed in the data.

Relative to these models, Bernard, Eaton, Jensen, and Kortum's (2003) introduces an additional margin through variable markups. In this model, producers are able to price at a markup over marginal costs depending on other latent competitors. Unlike in Eaton and Kortum (2002), the maximal bilateral price difference reflects both markups, trade frictions, and productivity differences (if the good is not traded between the pair). The crucial observation is that because mark-ups are negatively correlated to marginal costs, the maximal bilateral price difference lies even further below the trade friction relative to the Eaton and Kortum (2002) model. Hence, when viewing maximal price differences through the lens of the Bernard, Eaton, Jensen, and Kortum's (2003) model, the unobserved trade frictions must be larger relative to Eaton and Kortum (2002), and hence one needs a even lower trade elasticity to rationalize the same amount of trade observed in the data.

Figure 2 summarizes the latter argument. It plots the CDFs of the logged price differences generated by Bernard, Eaton, Jensen, and Kortum's (2003) and Eaton and Kortum's (2002) models. Recall that price differences in Eaton and Kortum's (2002) model correspond to cost differences, which are identical in the two models. Clearly, logged price (and cost) differences are bounded above by logged trade barriers. Moreover, the distribution of logged price differences in Eaton and Kortum's (2002) model lies below the distribution generated from Bernard, Eaton, Jensen, and Kortum's (2003) model, which suggests that the former stochastically dominates the latter.

The distinction between Krugman (1980) and Melitz (2003) is because of firm heterogeneity and patterns of selection into export markets. In the Krugman (1980) model, the maximal bilateral price difference is likely to equal the unobserved trade friction (exactly like in the Armington model). In the Melitz (2003) model, only goods sold by firms that enter all markets are able to make the sample of prices that data collectors would consider. Moreover, the firms that enter all markets are likely to be from sources with low trade costs to export. This implies that the average maximal price difference across country pairs is likely to be below the average trade

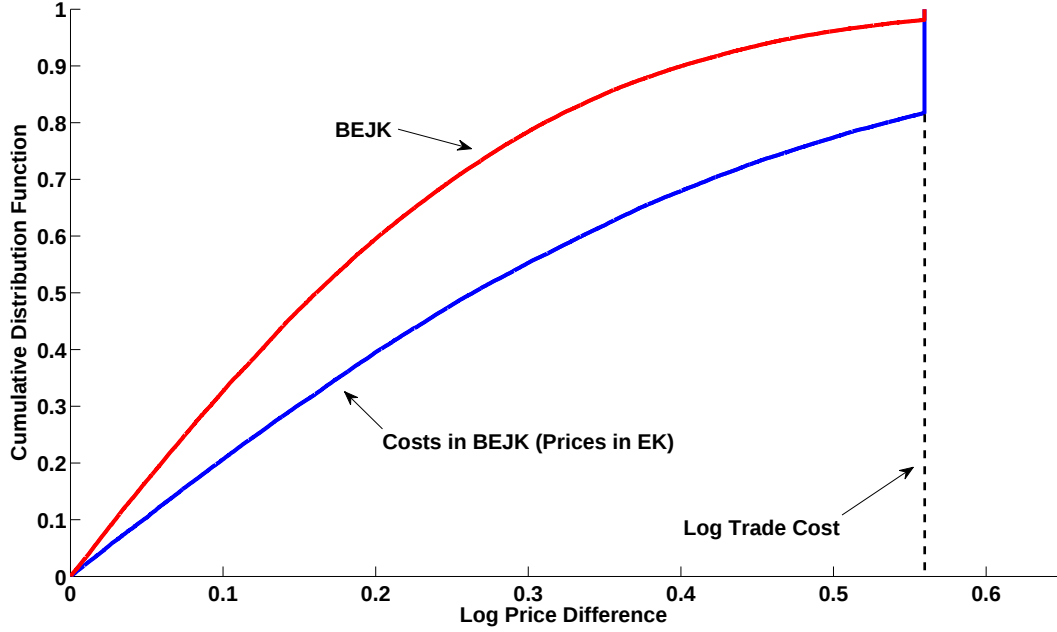


Figure 2: Distributions of logged price differences in Ricardian models

friction across country pairs. Thus one needs an even lower trade elasticity in Melitz (2003) relative to Krugman (1980) to rationalize the same amount of trade observed in the data.

In sum, the introduction of an extensive margin as in Eaton and Kortum (2002) or Melitz (2003) increases the welfare cost of autarky by up to 50 percent relative to Armington or Krugman (1980) which feature no extensive margin. Variable markups in Bernard, Eaton, Jensen, and Kortum (2003) further increase the welfare cost of autarky by 50 percent relative to Eaton and Kortum (2002) which features perfect competition.

5. Conclusion

This paper shows that new trade models with different micro-level margins—estimated to fit the same moments in the data—imply lower trade elasticities and, hence, larger welfare gains from trade relative to models without these margins. The key feature of the estimation approach is to focus on common moments where new micro-level margins, such as an extensive margin or variable markups, alter the mapping from the data to the estimate of the trade elasticity. We find that the introduction of an extensive margin as in Eaton and Kortum (2002) or Melitz (2003) increases the welfare cost of autarky by up to 50 percent relative to Armington or Krugman (1980) which feature no extensive margin. Variable markups in Bernard, Eaton, Jensen, and Kortum (2003) further increase the welfare cost of autarky by 50 percent relative to Eaton and Kortum (2002) which features perfect competition.

References

- ANDERSON, J., AND E. VAN WINCOOP (2003): "Gravity with Gravitas: A Solution to the Border Puzzle," *American Economic Review*, 93, 170–192.
- ARKOLAKIS, C., A. COSTINOT, AND A. RODRIGUEZ-CLARE (2011): "New Trade Models, Same Old Gains?," *American Economic Review* (forthcoming).
- ARKOLAKIS, C., S. DEMIDOVA, P. KLENOW, AND A. RODRÍGUEZ-CLARE (2008): "Endogenous Variety and the Gains from Trade," *The American Economic Review*, 98 (2)(2), 444–450.
- BERNARD, A., J. EATON, J. B. JENSEN, AND S. KORTUM (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93(4), 1268–1290.
- BRADFORD, S. (2003): "Paying the Price: Final Goods Protection in OECD Countries," *The Review of Economics and Statistics*, 85(1), 24–37.
- BRADFORD, S., AND R. Z. LAWRENCE (2004): *Has Globalization Gone Far Enough: The Costs of Fragmented Markets*, no. 349 in Peterson Institute Press: All Books. Peterson Institute for International Economics.
- CHANEY, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707–1721.
- DONALDSON, D. (2009): "Railroads of the Raj: Estimating the Economic Impact of Transportation Infrastructure," *unpublished paper, MIT*.
- EATON, J., AND S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.
- EATON, J., S. KORTUM, AND F. KRAMARZ (2011): "An Anatomy of International Trade: Evidence from French Firms," *Econometrica* (forthcoming).
- GOURIÉROUX, C., AND A. MONFORT (1996): *Simulation-based econometric methods*, CORE lectures. Oxford University Press.
- HEAD, K., AND J. RIES (2001): "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade," *American Economic Review*, 91(4), 858–876.
- KRUGMAN, P. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70(5), 950–959.
- MACKINNON, J. G., AND A. A. SMITH (1998): "Approximate bias correction in econometrics," *Journal of Econometrics*, 85(2), 205–230.

- MELITZ, M. J. (2003): “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(6), 1695–1725.
- ROMALIS, J. (2007): “NAFTA’s and CUSFTA’s Impact on International Trade,” *Review of Economics and Statistics*, 89(3), 416–35.
- SIMONOVSKA, I., AND M. E. WAUGH (2011): “The Elasticity of Trade: Estimates and Evidence,” NBER Working Papers 16796, National Bureau of Economic Research, Inc.
- SMITH, A. (2008): “Indirect inference,” *The New Palgrave Dictionary of Economics*, 2nd Edition (forthcoming).
- WAUGH, M. (2010a): “Bilateral Trade, Relative Prices, and Trade Costs,” *unpublished paper*, NYU.
- WAUGH, M. E. (2010b): “International Trade and Income Differences,” *American Economic Review*, 100(5), 2093–2124.

A. Proofs

The proof to Proposition 1 follows.

Proof

a. and b. We derive the trade shares and price indices for the versions of the Armington, Krugman (1980) and Melitz (2003) models that we consider in the present paper. We refer the reader to Eaton and Kortum (2002) and Bernard, Eaton, Jensen, and Kortum (2003) for derivations for these two models, since our exposition is identical to the one found in the papers.

For the Armington model, proof is to be completed.

For the Krugman (1980) model, maximizing W_i subject to a standard budget constraint yields the following CES demand function

$$x_{ni}(j) = \left(\frac{p_{ni}(j)}{P_n} \right)^{-\rho} \frac{w_n L_n}{P_n}, \quad (18)$$

$$P_n = \left[\sum_{i=1}^N \int_{J_i} p_{ni}(j)^{1-\rho} dj \right]^{\frac{1}{1-\rho}} \quad (19)$$

Moreover, maximizing firm profits and accounting for the demand in (18) yields

$$p_{ni}(j) = \frac{\bar{m} \tau_{ni} w_i}{T_i^{1/\theta}} \quad (20)$$

Free entry implies that the expected profit of a firm is zero, after paying the fixed entry cost. Using the firm's optimal pricing rule and quantity demanded, which must equal to the quantity produced, yields

$$\sum_{n=1}^N \frac{(p_{ni}(j))^{1-\rho}}{P_n^{1-\rho}} \frac{w_n L_n}{\rho} - w_i f_i = 0 \quad (21)$$

Let M_i denote the measure of entrants in country i , which corresponds to the measure of set J_i . In order to compute the measure of entrants, combine the free entry condition from (21) with the labor market clearing condition for country i ,

$$M_i \sum_{n=1}^N \frac{(p_{ni}(j))^{1-\rho}}{P_n^{1-\rho}} w_n L_n + f_i = L_i, \quad (22)$$

to obtain $M_i = L_i/(\rho f_i)$. Using the definition $\theta = \rho - 1$ and assuming that $L_i \propto f_i$ yields $M_i \propto 1/(1 + \theta)$. Relying on this result, use the equilibrium measure of entrants M_i and the optimal pricing rule (20) into (19) to obtain that $P_n \propto \Phi_n^{-1/\theta}$.

Finally, the share of expenditure that n spends on goods from i is given by

$$\frac{X_{ni}}{\sum_{k=1}^N X_{nk}} = \frac{M_i \frac{(p_{ni}(j))^{1-\rho}}{P_n^{1-\rho}} \frac{w_n L_n}{\rho}}{\sum_{k=1}^N M_k \frac{(p_{nk}(j))^{1-\rho}}{P_n^{1-\rho}} \frac{w_n L_n}{\rho}} \quad (23)$$

Using the definition $\theta = \rho - 1$, the equilibrium measure of entrants M_i under the proportionality assumption, the optimal pricing rule (20) and the exact price index (19) in (23) yields the expression in the text.

For the Melitz (2003) model, following the same steps as for the Krugman (1980) model obtains the demand function in (18) and the price index in (19). Let the marginal cost of a firm with efficiency draw $z_i(j)$ of producing good j in i and delivering it to n be

$$c_{ni}(j) = \frac{w_i \tau_{ni}}{z_i(j)}.$$

We will express all equilibrium objects in terms of cost draws in the remainder of the Appendix. As in the Krugman (1980) model, the optimal pricing rule is given by

$$p_{ni}(j) = \bar{m} c_{ni}(j) \quad (24)$$

Using the firm's optimal pricing rule and quantity demanded, which must equal to the quantity

produced, yields the firm's profit function

$$\pi_{ni}(c_{ni}(j)) = w_n L_n \left(\frac{\rho}{\rho-1} \right)^{-\rho} \frac{1}{\rho-1} \left(\frac{c_{ni}(j)}{P_n} \right)^{1-\rho} - w_n a_n$$

The cost cutoff for a firm from i to sell good j to n , $\bar{c}_{ni}(j)$, satisfies $\pi_{ni}(\bar{c}_{ni}(j)) = 0$ and is given by

$$\bar{c}_{ni}(j) = \left[\frac{a_n}{L_n} \frac{P_n^{1-\rho}}{\left(\frac{\rho}{\rho-1} \right)^{-\rho} \frac{1}{\rho-1}} \right]^{\frac{1}{1-\rho}}$$

Since $\bar{c}_{ni}(j) = \bar{c}_{nn}(j) \forall i$, denote cost cutoffs by the destination to which they apply, $\bar{c}_n(j)$.

To determine the measure of entrants, repeat the procedure applied to the Krugman (1980) model to obtain $M_i = L_i / f_i(\rho-1)/(\rho\theta)$. Only those firms with cost draws below $\bar{c}_i$ serve the domestic market and the remainder immediately exit. Using the Pareto distribution, the measure of domestic firms that sell j in country i is $M_{ii}(j) = M_i T_i w_i^{-\theta} (\bar{c}_i(j))^\theta$. Similarly, the measure of firms from i that sell j to n is $M_{ni}(j) = M_i T_i (w_i \tau_{ni})^{-\theta} (\bar{c}_n(j))^\theta$.

Furthermore, assuming that $L_i \propto f_i$ yields $M_i \propto (\rho-1)/(\rho\theta)$. Relying on this result, use the equilibrium measure of firms that serve destination n from any source i , M_{ni} , the optimal pricing rule (24), and the Pareto distribution into (19) to obtain that $P_n \propto \Phi_n^{-1/\theta}$.

Finally, to derive the trade share, use the optimal pricing rule (24), quantity demanded, which must equal to the quantity produced, the equilibrium measure of firms that serve each destination, and the Pareto distribution.

c. We refer the reader to Arkolakis, Costinot, and Rodriguez-Clare (2011) for derivations concerning welfare in all the models.

B. Trade Barriers and Prices

In this section we argue that bilateral trade barriers are bounded below by relative prices of goods in all five models.

The proof for the Armington model follows.

Proposition 2 *In the Armington model, $p_n(j)/p_i(j) \leq \tau_{ni}$ for any good j sold in destinations n and i .*

Proof In this model, goods are differentiated by the country of origin of the producer. Let a be the supplier of good j to all destinations. Then, the relative price of any good j between any

pair of destinations n and i is

$$\frac{p_{na}(j)}{p_{ia}(j)} = \frac{\frac{\tau_{na}w_a}{T_a^{1/\theta}}}{\frac{\tau_{ia}w_a}{T_a^{1/\theta}}} = \frac{\tau_{na}}{\tau_{ia}} \leq \tau_{ni}$$

by the assumed triangle inequality on trade barriers.

The proof for the Eaton and Kortum (2002) model can be found in Simonovska and Waugh (2011). The proof for the Bernard, Eaton, Jensen, and Kortum (2003) model (in logs) follows.

Proposition 3 *In the Bernard, Eaton, Jensen, and Kortum (2003) model, $\log p_n(j) - \log p_i(j) \leq \log \tau_{ni}$ for any good j sold in destinations n and i .*

Proof We prove the proposition for the general case in which country a supplies good j to n and country b supplies good j to i . The triangle inequality requires that $\tau_{nb}/\tau_{ib} = \tau_{na}/\tau_{ia} = \tau_{ni}$ in order for countries n and i to buy good j from a and b , respectively. Otherwise, both countries buy good j from the same source, in which case the extra condition holds trivially. The case in which both countries share a supplier is simply a sub-case of the general proof below.

The price of good j in the two destinations is

$$\begin{aligned} p_n(j) &= \min \{C_{2n}(j), \bar{m}C_{1na}(j)\} \\ p_i(j) &= \min \{C_{2i}(j), \bar{m}C_{1ib}(j)\} \end{aligned}$$

Since $\log$ is an increasing function,

$$\begin{aligned} \log p_n(j) &= \min \{\log(C_{2n}(j)), \log(\bar{m}C_{1na}(j))\} \\ \log p_i(j) &= \min \{\log(C_{2i}(j)), \log(\bar{m}C_{1ib}(j))\} \end{aligned}$$

It is useful to explicitly write the pricing rules as

$$\begin{aligned} \log p_n(j) &= \min \left\{ \min \left[\log(C_{2na}(j)), \min_{k \neq a} \{\log(C_{1nk}(j))\} \right], \log(\bar{m}C_{1na}(j)) \right\} \\ \log p_i(j) &= \min \left\{ \min \left[\log(C_{2ib}(j)), \min_{k \neq b} \{\log(C_{1ik}(j))\} \right], \log(\bar{m}C_{1ib}(j)) \right\} \end{aligned}$$

Broadly, there are four combinations of pricing rules to consider.

1. $\log(p_n(j)) = \log(\bar{m}C_{1na}(j))$ and $\log(p_i(j)) = \log(\bar{m}C_{1ib}(j))$. Then,

$$\begin{aligned} \log p_n(j) &= \log(\bar{m}) + \log(w_a \tau_{na}) - \log(Z_{1a}(j)) \\ \log p_i(j) &= \log(\bar{m}) + \log(w_b \tau_{ib}) - \log(Z_{1b}(j)) \end{aligned}$$

Since n imported j from a , it must be that $\log(w_a\tau_{na}) - \log(Z_{1a}(j)) \leq \log(w_b\tau_{nb}) - \log(Z_{1b}(j))$. Then,

$$\begin{aligned}
\log p_n(j) - \log p_i(j) &= \log(\bar{m}) + \log(w_a\tau_{na}) - \log(Z_{1a}(j)) - [\log(\bar{m}) + \log(w_b\tau_{ib}) - \log(Z_{1b}(j))] \\
&= \log(w_a\tau_{na}) - \log(Z_{1a}(j)) - [\log(w_b\tau_{ib}) - \log(Z_{1b}(j))] \\
&\leq \log(w_b\tau_{nb}) - \log(Z_{1b}(j)) - [\log(w_b\tau_{ib}) - \log(Z_{1b}(j))] \\
&= \log(\tau_{nb}) - \log(\tau_{ib}) \\
&\leq \log(\tau_{ni})
\end{aligned}$$

because $\tau_{nb} \leq \tau_{ni}\tau_{ib}$ by triangle inequality.

2. $\log(p_n(j)) = \log(C_{2n}(j))$ and $\log(p_i(j)) = \log(\bar{m}C_{1ib}(j))$.

Since $\log(p_n(j)) = \log(C_{2n}(j))$, it must be that $\log(C_{2n}(j)) \leq \log(\bar{m}C_{1na}(j))$. Then, the proof follows from 1.

3. $\log(p_n(j)) = \log(C_{2n}(j))$ and $\log(p_i(j)) = \log(C_{2i}(j))$. Then there are four sub-cases.

a. $\log(p_n(j)) = \log(C_{2na}(j))$ and $\log(p_i(j)) = \log(C_{2ib}(j))$.

Since, $\log(p_n(j)) = \log(C_{2na}(j))$, it must be that $\log(C_{2na}(j)) \leq \min_{k \neq a} \{\log(C_{1nk}(j))\}$ and in particular the inequality holds for $k = b$. Moreover, $\log(C_{1ib}(j)) \leq \log(C_{2ib}(j))$ by definition. Using these two inequalities yields,

$$\begin{aligned}
\log p_n(j) - \log p_i(j) &= \log(w_a\tau_{na}) - \log(Z_{2a}(j)) - [\log(w_b\tau_{ib}) - \log(Z_{2b}(j))] \\
&\leq \log(w_b\tau_{nb}) - \log(Z_{1b}(j)) - [\log(w_b\tau_{ib}) - \log(Z_{2b}(j))] \\
&\leq \log(w_b\tau_{nb}) - \log(Z_{2b}(j)) - [\log(w_b\tau_{ib}) - \log(Z_{2b}(j))] \\
&= \log(\tau_{nb}) - \log(\tau_{ib}) \\
&\leq \log(\tau_{ni})
\end{aligned}$$

because $\tau_{nb} \leq \tau_{ni}\tau_{ib}$ by triangle inequality.

b. $\log(p_n(j)) = \min_{k \neq a} \{\log(C_{1nk}(j))\}$ and $\log(p_i(j)) = \log(C_{2ib}(j))$.

Since $\log(p_n(j)) = \min_{k \neq a} \{\log(C_{1nk}(j))\}$, it must be that $\log(p_n(j)) \leq \log(C_{1nb}(j))$ (with equality if $k = b$). Moreover, $\log(C_{1ib}(j)) \leq \log(C_{2ib}(j))$ by definition. Using these two inequalities yields,

$$\begin{aligned}
\log p_n(j) - \log p_i(j) &\leq \log(w_b\tau_{nb}) - \log(Z_{1b}(j)) - [\log(w_b\tau_{ib}) - \log(Z_{2b}(j))] \\
&\leq \log(w_b\tau_{nb}) - \log(Z_{2b}(j)) - [\log(w_b\tau_{ib}) - \log(Z_{2b}(j))] \\
&= \log(\tau_{nb}) - \log(\tau_{ib}) \\
&\leq \log(\tau_{ni})
\end{aligned}$$

because $\tau_{nb} \leq \tau_{ni}\tau_{ib}$ by triangle inequality.

c. $\log(p_n(j)) = \min_{k \neq a} \{\log(C_{1nk}(j))\}$ and $\log(p_i(j)) = \min_{k \neq b} \{\log(C_{1ik}(j))\}$.

It suffices to prove that the result holds when the best competitor to the producer that supplies to country n comes from k' , while the best competitor to the producer that supplies to country i comes from k'' , since the case in which the best competitor to both producers comes from the same country is just a sub-case.

Since k' attains the minimum for destination n , it must be that $\log(w_{k'}\tau_{nk'}) - \log(Z_{1k'}(j)) \leq \log(w_{k''}\tau_{nk''}) - \log(Z_{1k''}(j))$. Then,

$$\begin{aligned} \log p_n(j) - \log p_i(j) &= \log(w_{k'}\tau_{nk'}) - \log(Z_{1k'}(j)) - [\log(w_{k''}\tau_{ik''}) - \log(Z_{1k''}(j))] \\ &\leq \log(w_{k''}\tau_{nk''}) - \log(Z_{1k''}(j)) - [\log(w_{k''}\tau_{ik''}) - \log(Z_{1k''}(j))] \\ &= \log(\tau_{nk''}) - \log(\tau_{ik''}) \\ &\leq \log(\tau_{ni}) \end{aligned}$$

because $\tau_{nk''} \leq \tau_{ni}\tau_{ik''}$ by triangle inequality.

d. $\log(p_n(j)) = \log(C_{2na}(j))$ and $\log(p_i(j)) = \min_{k \neq b} \{\log(C_{1ik}(j))\}$.

Since $\log(p_n(j)) = \log(C_{2na}(j))$ it must be that $\log(C_{2na}(j)) \leq \min_{k \neq a} \{\log(C_{1nk}(j))\}$. In particular, the inequality holds for the k'' that is the best competitor to the producer who supplies to country i . Then,

$$\begin{aligned} \log p_n(j) - \log p_i(j) &\leq \log(w_{k''}\tau_{nk''}) - \log(Z_{1k''}(j)) - [\log(w_{k''}\tau_{ik''}) - \log(Z_{1k''}(j))] \\ &= \log(\tau_{nk''}) - \log(\tau_{ik''}) \\ &\leq \log(\tau_{ni}) \end{aligned}$$

because $\tau_{nk''} \leq \tau_{ni}\tau_{ik''}$ by triangle inequality.

4. $\log(p_n(j)) = \log(\bar{m}C_{1na}(j))$ and $\log(p_i(j)) = \log(C_{2i}(j))$.

Since $\log(p_n(j)) = \log(\bar{m}C_{1na}(j))$, it must be that $\log(\bar{m}C_{1na}(j)) \leq \log(C_{2n}(j))$. This reduces to case 3 above, which completes the proof.

The proof for the Krugman (1980) model follows.

Proposition 4 *In the Krugman (1980) model, $p_n(j)/p_i(j) \leq \tau_{ni}$ for any good j sold in destinations n and i .*

Proof In this model, goods are differentiated by the country of origin of the producer. Let a be the supplier of good j to all destinations. Then, the relative price of any good j between any

pair of destinations n and i is

$$\frac{p_{na}(j)}{p_{ia}(j)} = \frac{\frac{\bar{m}\tau_{na}w_a}{T_a^{1/\theta}}}{\frac{\bar{m}\tau_{ia}w_a}{T_a^{1/\theta}}} = \frac{\tau_{na}}{\tau_{ia}} \leq \tau_{ni}$$

by the assumed triangle inequality on trade barriers.

The proof for the Melitz (2003) model follows.

Proposition 5 *In the Melitz (2003) model, $p_n(j)/p_i(j) \leq \tau_{ni}$ for any good j sold in destinations n and i .*

Proof In this model, goods are differentiated by the country of origin and the efficiency of the producer. Let the supplier of good j to all destinations originate from country a and have productivity draw $z_a(j)$. Then, the relative price of any good j between any pair of destinations n and i is

$$\frac{p_{na}(j)}{p_{ia}(j)} = \frac{\frac{\bar{m}\tau_{na}w_a}{z_a(j)}}{\frac{\bar{m}\tau_{ia}w_a}{z_a(j)}} = \frac{\tau_{na}}{\tau_{ia}} \leq \tau_{ni}$$

by the assumed triangle inequality on trade barriers.