Trade Cost Elasticity: Estimates from a New Approach to Product-level Data

Juyoung Cheong*, Do Won Kwak†, and Kam Ki Tang‡

(Work in progress. Please do not cite.)

Abstract

This paper estimates trade cost (tariff) elasticities using bilateral tariff data at the HS2-digit level for 82 countries from 1996 to 2008. It extends the Helpman et al. (2008) model to incorporate firms' fixed costs of exporting that vary at the pair-product level. We apply a two-stage procedure as in Helpman et al. (2008) at the product level estimations to control for self-selection and firm heterogeneity using the signal from learning as exclusion restrictions. The empirical results show that there is substantial upward bias in the estimates of trade cost elasticities in the literature. Proper accounting of zero trade flows and firm heterogeneity using more disaggregated data yields significantly smaller estimates of trade cost elasticities (i.e. the magnitude decrease by half from -3.7 to -1.8), which imply much larger welfare gains from trade.

JEL Code: C13; C23, F10; F15

Keywords: Gravity Model; Firm Heterogeneity; Pseudo Poisson ML; Trade Cost Elasticity.

*Department of International Trade, Inha University, Korea; e-mail: jcheong@inha.ac.kr
†School of Economics, University of Queensland,QLD, Australia; e-mail: d.kwak@uq.edu.au
‡School of Economics, University of Queensland, Australia; e-mail: kk.tang@uq.edu.au

1 Introduction

The elasticity of trade flows with respect to trade costs is a key parameter to quantify welfare gains from trade. Arkolakis et al. (2012) note that the perfect competition models in Anderson and van Wincoop (2003) and Eaton and Kortum (2002) and the monopolistic competition models in Melitz (2003) are all in a broad class of models sharing Dixit–Stiglitz preferences, one factor of production, linear cost functions, complete specialization, and iceberg trade costs. Thus, if three macro-level restrictions hold, then all these models will share common parameters in measuring the gains from trade. They show that for a range of trade models, gains from trade can be estimated using two parameters: the import penetration ratio (or domestic share), and the trade elasticity with respect to variable trade cost obtained from a gravity equation. Bergstrand et al. (2013) also show that the gains from trade depends only on the import penetration ratio and a gravity equation based estimates of trade cost elasticity. For empirical estimation of the gains from trade, the import penetration ratio data can be obtained from national statistics and are publicly available across countries. On the contrary, trade cost elasticity is not directly observable and needs to be estimated. Furthermore, the estimation of trade cost elasticity is difficult because of issues such as omitted variable bias and sample selection bias.

In this paper, following Bergstrand et al. (2013) we estimate trade cost elasticity using an empirical gravity model. The trade cost elasticity is obtained from the response of trade flows to tariffs changes, which constitute changes in trade costs. The change in trade costs due to changes in tariffs is a representative of all sources of trade cost change, because for a given change in import prices, the source of the change should be irrelevant to the demand outcome. Furthermore, we focus on tariff changes because tariffs suffer less from measurement errors than other sources of variable trade costs such as transportation cost.

There are a fair number of empirical studies trade cost elasticity, including Anderson (1979), Harrigan (1993), Hummels (2001), Baier and Bergstrand (2001), Broda and Weinstein (2006) and Bergstrand et al. (2013). Anderson and van Wincoop (2004) have provided an excellent survey on this literature and show that trade cost elasticity estimates range from -4 to -10. Applying these estimates and the US's import penetration ratio of 0.07 as in year 2000 to the formula of Arkolakis et al. (2012), the gain from trade (from autarky) for the US will range between 0.7% and 1.8%. Given the importance of trade to modern economies, these welfare gain estimates seem to be too small to be true. Recent empirical studies such as Bergstrand et al. (2013) and Simonovska and Waugh (2014) find a similar level of estimates. This means that, with all the improvement in estimation techniques and data over time, elasticity estimates remain to be large and welfare gains to be small (Ossa, 2015).

Many studies including recent studies (e.g. Bergstrand et al. (2013)) use country-level data mostly due to data limitation. However, using aggregate data may not be as informative as we want. This is because two potentially serious biases could arise in the estimations of trade cost elasticities. Firstly, at any given time, tariffs vary substantially across products at least for some certain level, but data aggregation omits the information on substitution between these products, causing bias in the trade elasticity estimation.[1] Secondly, unfiltered bilateral trade data typically feature a large proportion of zero observations that have important implications to trade cost elasticity estimate, but information on zero trade flows at the product level is naturally lost with aggregate data, and firm's self-selection of exporting at the product level could not be addressed either.

A few studies provide trade elasticities estimates using product level data, e.g. Hummels (2001) and Baier and Bergstrand (2001). But they tend to cover only a single or at most a handful of exporting countries. Using data for a single exporting country may also be biased because it cannot account for unobserved heterogeneity at the exporter level such as the competing or complementary effects from exporters in other countries, or the broader multilateral resistance terms (MRTs) as shown in Anderson and van Wincoop

[1] Anderson and van Wincoop (2004) demonstrate with a numerical example that "the elasticity of substitution at the more aggregated level is entirely irrelevant." (p. 727). They advise that: "one should choose elasticities at a sufficiently disaggregated level at which firms truly compete."

(2003) (denoted as AvW (2003) hereafter).

Broda et al. (2008) and Kee et al. (2008) use fairly disaggregate data that cover many countries. Instead of using gravity models, they structurally estimate the elasticity using simplified demand and supply functions and GDP functions. Our paper distinguishes from them in that we use a gravity equation derived from a model incorporating firm characteristics, and thus our estimates do not much depend on the assumptions we make. Ossa (2015) uses a method similar to Broda et al. (2008) that incorporates sectoral linkages across industries, and shows that some industries having very small trade cost elasticities could in fact contribute to overall welfare gains significantly. Also, Ossa (2015)'s average trade elasticity estimate across industries (which is equivalent to aggregate trade cost elasticities), -3.9, is at the lower end of the spectrum of the previous literature. However, this estimate still implies a rather small welfare gain.

Comparing to the previous studies, our estimation shows that trade cost elasticities are about -1.9 in year 2000, which implies a welfare gain of close to 4% for the US – more than twice of the largest estimate in the previous literature. The sources of smaller elasticity estimates in our study come from accounting for zero trade flows as well as unobserved firm heterogeneity at the product levels, in line with the argument of the recent trade theory literature. Accounting for these two sources of bias simultaneously is challenging and, thus, successfully meeting this challenge will constitute a significant methodological contribution to the literature.

Helpman et al. (2008) (denoted as HMR hereafter) show that zero trade flows in aggregate data are the result of firms' self-selection out of foreign/export markets due to firm's heterogeneous productivity which depends on firms' fixed costs of exporting, and point out if we do not account for zero trade flows, the estimates could be upward biased. In order to account for bias from omitting zero trade flows, HMR suggest a two-stage procedure where export market entry decision is estimated in the first stage and volume decision conditional on entering the market is estimated in the second stage. However, the method is difficult to implement with product level data. This is because the exclusion restriction at the first stage estimation requires a variable that determines firms' export market entry decision but does not affect their volume of exports once they decide to enter a market.[2] An innovation of the current paper is that we suggest a series of valid exclusion restriction variables that allow we to extend the HMR approach to accounting for firms' fixed costs of exporting, to product level data. The justification of proposed exclusion restrictions rely upon the recent literature on search and learning on exporting markets, which are argued in Eaton et al. (2007), Eaton et al. (2014), Albornoz et al. (2012), Morales et al. (2011), Fernandes and Tang (2014) and others (details are provided in Section 5). The implication of our methodology is significant because it opens up opportunities to apply the highly influential HMR approach to test various trade theories using much richer product level data.

The rest of the paper is organized as follows. Section 2 extends the original HMR model for aggregate data to one for disaggregated, sector level data. Section 3 explains the new exclusion restriction variables we derive from the learning literature. Section 4 describes the data and provides the estimation results and Section 5 concludes.

2 The HMR Model for Sector Level Data

In this section, we extend the firms entry decision model used for aggregate data in HMR to a model for disaggregated, sector level data. We allow heterogeneity across sectors but assume independence between sectors.[3] Each equation in this section is sector-specific in the sense that parameters for the model of each sector vary by sector. Each sector is supposed to produce only one product so we use the terms "product" and "sector" inter-

[2] Recent theoretical studies like Chaney (2008) and Krautheim (2012) pay attention to the role of fixed costs in heterogeneous firms' entry decision in new export markets, which is empirically supported by Koenig et al. (2010).

[3] This assumption implies that we allow substitutions between sectors at the refined product level but not between sectors at the broader product level.

changeably.

Suppose exporting country j faces the following demand curve in destination i for product k under the monopolistic competition market:

$$q_{ijk} = Q_{ik} \left(\frac{c_{jk}\tau_{ijk}}{\alpha_k P_{ik}} \right)^{-\gamma} N_{jk} V_{ijk}$$

$$V_{ijk} = \frac{\theta a_{kL}^{\theta+\gamma}}{(\theta+\gamma)(a_{kH}^{\theta} - a_{kL}^{\theta})} W_{ijk},\ W_{ijk} = max\{(\frac{a_{ijk}}{a_{kL}})^{\theta+\gamma} - 1, 0\}$$

where Q_{ik} is the equilibrium market size in country i for product k; c_{jk} is a measure of average product-specific productivity in sector k of firms in j; τ_{ijk} is the variable trade cost of product k exported from j to i; P_{ik} is the price index for sector k in i, determined by domestic producers and existing exporters to country i; the inverse of a_k (i.e. $1/a_k$) represents firms' productivity in sector k. Productivity is heterogeneous across firms within a sector and $1/a_k$ determines firm-productivity cut-off of exporting in each sector. As in HMR, $G(a_k)$ has a truncated Pareto distribution with the support $[a_{kL,} a_{kH}]$, where a_{kH} (a_{kL}) implies the lowest (highest) productivity in sector k, so that $G(a_k) = (a_k^{\theta} - a_{kL}^{\theta})/(a_{kH}^{\theta} - a_{kL}^{\theta}), \theta > (\sigma_k - 1)$; N_{jk} is the number of firms from country j in the product market k; V_{ij} and W_{ijk} are a function of productivity cut-off which reflects the proportion of country j's exporting firms to country i for product k; and γ is the import demand elasticity (in absolute value). Firms in country j takes P_{ik} and Q_{ik} as given.[4]

Similar to HMR, we can write the volume of trade as follows. Using the sector independence and heterogeneity assumptions, we suppress k for the sake of simplicity.

$$ln(q_{ij}) = \beta_0 + \lambda_j + \xi_i + \delta_1 \mathbf{x}_{1ij} + \omega_{ij} + u_{ij} \tag{1}$$

where for each sector k, λ_j is exporter-country specific fixed effects (FEs) which sub-

[4]For brevity, we skip the parts to derive a trade flow equation (i.e.j's demand from i on product k) from a representative consumer's utility function in j. For the details of the model, see Helpman et al. (2008).

sume $ln(N_j)$ and $ln(c_j)$; ξ_i is destination-specific FEs which subsume $ln(Q_i)$ and $ln(P_i)$; $\mathbf{x}_{1ij}$ includes all observed variables that could capture trade costs, including pair-level gravity variables such as distance, cultural ties, and colonial relationship, and pair-sector level variables such as tariffs; $\omega_{ij}(= ln(W_{ij}))$ is a function of cut-off productivity that determines the fraction of firms in country j to destination i for each sector; and u_{ij} is an idiosyncratic error term. Effectively, the obtained equation for the volume of trade in eq.(1) is the same as the HMR. The only difference from the HMR is that as we allow sector heterogeneity such that the cut-off productivity differ by sector. As a result, we estimate eq.(1) sector by sector, and we need determinants of $\mathbf{x}_{1ij}$ and ω_{ij} that are specific to each sector for identification of sector specific parameters.

2.1 Model for entry decision of a firm

For each sector, the selection of country j's firms into a market i is determined by V_{ij}, which describes the cut-off productivity level for export market entry, a_{ij}. Now consider a latent variable Z_{ij} which is defined as

$$Z_{ij} = \frac{(1-\alpha)\left(\frac{c_j \tau_{ij}}{\alpha P_i}\right)^{-\gamma} Q_i a_{ij}^{-\gamma}}{c_j f_{ij}} \tag{2}$$

In eq.(2), the numerator is the operating revenue and the denominator is the fixed cost of exporting. So as long as $Z_{ij} > 1$, export accrues positive operating profits. The fixed exporting costs are stochastic due to unmeasured trade frictions v_{ij} that are assumed to be iid but correlated with errors (u_{ij}) in the second-stage estimation. So we assume the fixed costs of exporting are determined as follows:

$$f_{ij} = exp(\psi_j + \psi_i + \theta\sigma_{ij} - v_{ij})$$

where ψ_j subsumes inherent factors specific to exporter j that could affect their fixed costs of exporting, ψ_i is destination specific factors that could affect the fixed costs, σ_{ij} contains information on the fixed costs that are specific to both exporter j and destination i, and $v_{ij} \sim N(0, \phi_v^2)$ capture remaining unobserved factors. We take logarithm on eq.(2), then we have

$$z_{ij} \equiv ln(Z_{ij}) = \beta_0' + \eta_j + w_i + \delta \mathbf{x}_{ij} - \theta \sigma_{ij} + \epsilon_{1ij}$$

where $\mathbf{x}_{ij}$ represents typical observed pair variables (e.g. distance among others) included in the gravity model, η_j is exporter FEs, subsuming all j-specifc variables including c_j, w_i is importer FEs, subsuming all i-specifc variables including P_i and Q_i, σ_{ij} is information on the (sector specific) fixed costs for firms in j to export to i, and $\epsilon_{1ij} = \rho_0 u_{ij} + v_{ij} \sim N(0, \phi^2 + \phi_v^2)$ and assume $\rho_0 = 1$ for the sake of simplicity.

As σ_{ij} is not present in the eq.(1), it could be used as an exclusion restriction for identification of parameters in eq.(1). Implementation of two-stage estimation requires to have observed factors in σ_{ij} that vary at ij and affects on the fixed cost of exporting. We need exclusion restriction for σ_{ij}.

Using the probit model, we could obtain $\rho_{ij} = pr(q_{ij} > 0 | \eta_j, w_i, \mathbf{x}_{1ij}, \sigma_{ij})$ by:

$$\rho_{ij} = \Phi(\gamma_0^* + \eta_j^* + w_i^* + \delta^* \mathbf{x}_{1ij} + \theta^* \sigma_{ij}) \tag{3}$$

where $\Phi(\cdot)$ is a standard normal CDF. Let $\hat{\rho}_{ij}$ be the predicted probability from the probit estimation of eq.(3) and $\hat{z}_{ij}^* = \Phi^{-1}(\hat{\rho}_{ij})$ be the predicted value of $z_{ij}^* = \frac{z_{ij}}{\phi_v}$.

Similar to HMR, we can use the probit estimation of eq.(3) to obtain consistent estimates in the second stage by controlling for both endogenous number and self-selection of j's firms exporting to i (w_{ij}).

$$ln(q_{ij}) = \beta_0 + \lambda_j + \xi_i - \delta_1 \mathbf{x}_{1ij} + \omega_{ij} + u_{ij}$$

where ω_{ij} include factors that determine the fraction of firms exporting from j to i in sector k. Therefore, we need the estimates for both $E(\omega_{ij} | q_{ij} > 0, \mathbf{x}_{1ij}, \lambda_j, \xi_i)$ and $E(u_{ij} | q_{ij} >$

$0, \mathbf{x}_{1ij}, \lambda_j, \xi_i)$. Both terms depend on $\bar{v}^*_{ij} = E(v^*_{ij}|q_{ij} > 0, \eta_j, w_i, \mathbf{x}_{1ij}, \sigma_{ij})$ and note that $E(u_{ij}|q_{ij} > 0, \mathbf{x}_{1ij}, \lambda_j, \xi_i) = corr(u_{ij}, v_{ij}) \cdot \frac{\sigma_u}{\sigma_v}\bar{v}^*_{ij}$, $corr(u_{ij}, v_{ij}) \cdot \frac{\sigma_u}{\sigma_v} = \rho_1$ where $v^*_{ij} = \frac{v_{ij}}{\sigma_v}$. Also note that the estimate for $\bar{v}^*_{ij}$ could be obtained from the inverse Mills ratio (IMR), $\hat{\bar{v}}^*_{ij} = \frac{\phi(\hat{z}^*_{ij})}{\Phi(\hat{z}^*_{ij})}$. Furthermore, for the consistent estimate of $E(z_{ij}|q_{ij} > 0, \eta_j, w_i, \mathbf{x}_{1ij}, \sigma_{ij})$, we could use $\hat{z}^*_{ij} + \hat{\bar{v}}^*_{ij}$and $\hat{\bar{\omega}}^*_{ij} = ln[exp(\alpha(\hat{\bar{v}}^*_{ij} + \hat{z}^*_{ij})) - 1]$ for the consistent estimate for $E(\omega_{ij}|q_{ij} > 0, \mathbf{x}_{1ij}, \lambda_j, \xi_i)$.

Finally, we could estimate the second stage by using the following equation:

$$ln(q_{ij}) = \beta_0 + \lambda_j + \xi_i - \delta_1\mathbf{x}_{1ij} + ln[exp(\alpha(\hat{\bar{v}}^*_{ij} + \hat{z}^*_{ij})) - 1] + \rho_1\hat{\bar{v}}^*_{ij} + e_{ij} \tag{4}$$

where α is a function of γ as well as θ and in the eq.(4), $W_{ij} = Z^{\alpha}_{ij} - 1 = exp(\alpha z_{ij}) - 1$ is used to estimate ω_{ij} by taking log both sides of the equation. As long as exclusion restriction is available, we can easily implement eq.(3) to obtain $ln[exp(\delta(\hat{\bar{v}}^*_{ij} + \hat{z}^*_{ij})) - 1] + \rho_1\hat{\bar{v}}^*_{ij}$. Here $ln[exp(\delta(\hat{\bar{v}}^*_{ij} + \hat{z}^*_{ij})) - 1]$ and $\rho_1\hat{\bar{v}}^*_{ij}$ account for firm heterogeneity and self-selection of exporting at the sector level, respectively.

3 Exclusion restriction at sector level

3.1 Search and learning on exporting markets

In order to extend the HMR approach to sector/product level data, the estimation requires an exclusion restriction variables that affects firms' entrance into new export product markets but not their trade volumes in the markets they have already entered. These are factors that have affect on the fixed cost of exporting. However, finding a readily-available exclusion restriction that varies over pair-product-time is difficult not only due to data limitation, but also due to conceptual constraints.

In this section, we propose potential candidates for exclusion restriction variables based on information learnt from neighboring firms as in Fernandes and Tang (2014). We examine two main conditions for an exclusion restriction which are (i) that it should affect the fixed costs of exporting and; (ii) that it should not affect the volume of trade of product once an entrant becomes an incumbent in the export market. The first condition can be verified at the first stage of estimations using the Probit model and the F-test of partial correlation, but the second condition cannot be verified. Following the literature, we perform an overidentification test with multiple exclusion restrictions for an examination of the second condition. The overidentification test, however, is neither a sufficient, nor a necessary condition for instrument validity. It is to test whether all instruments identify the same vector of parameters, i.e. whether there is coherency amongst the instruments (Parente & Silva 2012).

Recent studies in the literature argue that firms can learn about their demand in a potential new export market from neighborhood countries' performance in that market. For instance, Fernandes and Tang (2014) examine which neighborhood countries' export performance to a specific market would be a signal to infer the potential market's demand and show that it positively affects a potential entrant's *entry* decision and initial export volume to the market. On the other hand, other studies examine learning from exporters' own experience in other markets to explain its export performance in a specific exporting market using firm-level data. For instance, Eaton et al. (2007) and Eaton et al. (2014) show that learning from its success in foreign markets affects a firm's incentive to search for more markets and also that a firms' geographic expansion paths depends on its initial destination market. Likewise, Albornoz et al. (2012) observe that a firm discovers its profitability as an exporter after actually engaging in exporting and adjusts quantities and decides whether to enter into, or exit from new markets. Furthermore, Morales et al. (2011) find that a firm's entry to a new destination is positively affected by its previous export experience in similar (geographically or economically) market.

The first stream of the literature implies that the larger number or faster export growth of an exporter's neighborhood could signal information about market demand as well as product demand. By definition of signal, once firms enter new export markets, information from exporter's neighborhood has no additional information on the demand for new entrant as entrant firms can directly observe their product demand.

3.2 Weak IV issue: Theoretical justification for the relevancy of exclusion restriction

For each sector k, we need observed variables that varies in both i as well as j and affects firms entry decision. Suppose that, before a country j exports to destination i, the fixed cost of exporting, σ_{ij}, is unknown. For brevity, we omit subscript k. Suppose σ_{ij} be decomposed into three parts:

$$\sigma_{ij} = \sigma_i + \sigma_j + \epsilon_{ij}$$

where σ_j is origin-specific but invariant across destination countries; σ_i is destination-specific; and ϵ_{ij} captures sum of factors that affect uncertainty of country i's product demand in country j. σ_j captures factors like product quality of source-country j, thus a low value of σ_j imply firm in country j to charge a higher price , and σ_i captures factors such as consumer preferences and destination specific trade barriers which apply to all countries exporting product k to market i.

σ_i, σ_j, and ϵ_{ij} are unknown to firms before they enter the market i. We assume that firms in country j decides whether to enter (or exit for an incumbent firm) market i by inferring $\sigma_i + \sigma_j + \epsilon_{ij}$ using information of their own performance in other destinations she already have been exported but market j (learning on σ_j) and that of firms from other countries exporting to market i (learning on σ_i) .

Suppose that the prior for σ_i, σ_j and ϵ_{ij} are given as follow:

$$\sigma_i \sim N(\mu_i, \phi_i^2)$$
$$\sigma_j \sim N(\mu_j, \phi_j^2)$$
$$\epsilon_{ij} \sim N(0, \phi_\epsilon^2)$$

where μ_i could be approximated by other exporters' performance in country i's market (in this paper, we use j's average performance with positive exports to countries shared the border with country *j*) and μ_j can be approximated by exporter j's average performance in similar countries/destinations. Similarly, ϕ_i^2 could be approximated by the variance of performance of other exporters in country i and ϕ_j^2 could be approximated by the variance of average performance of exporter j in other countries/destinations. ϕ_ϵ^2 could be approximated by the variance of interaction terms of country i's performance in other similar markets and average performance of similar countries at market j.

Suppose there are many exporters that exports to destination i. We assume that it is easier for j to learn about σ_i from its neighbouring countries than from those farther away. In this paper, we define j's neighbours as those that share borders with it, and use their average exports to i as information on σ_i. Likewise, learning σ_j also can be fairly precisely estimated from its own experience in all markets where they already export so we can essentially pin down σ_{ij}, which determines country i's export decisions to a specific market.

Suppose $\sigma_m = \sigma_j$ where m is a neighborhood country of country j, then

$$\begin{aligned}
E(\sigma_{im}|\sigma_{ij}) &= E(\sigma_i + \sigma_m|\sigma_{ij}) \\
&= E(\sigma_i + \sigma_j|\sigma_{ij}) \\
&= E(\sigma_{ij} - \epsilon_{ij}|\sigma_{ij}) \\
&= \sigma_{ij} - E(\epsilon_{ij}|\sigma_{ij}) \\
&= \frac{\phi_i^2 + \phi_j^2}{\phi_i^2 + \phi_j^2 + \phi_\epsilon^2}(\sigma_{ij} - \mu_i - \mu_j) + (\mu_i + \mu_j)
\end{aligned}$$

where the last equality is given by joint normality.

The above equation can be rewritten as:

$$E(\sigma_{im}|\sigma_{ij}) = \frac{\phi_\epsilon^2}{\phi_i^2 + \phi_j^2 + \phi_\epsilon^2}(\mu_i + \mu_j) + \frac{\phi_i^2 + \phi_j^2}{\phi_i^2 + \phi_j^2 + \phi_\epsilon^2}\sigma_{ij}$$

where σ_{ij} is obtained from information on other exporter j that is most close to m and already export to destination i. Similarly, expected value of σ_{mj} for known σ_{ij} can be determined where σ_{ij} is obtained from information on exporter j's other incumbent destination i that is closest to m.

3.3 Validity of Exclusion Restriction Variables

We propose to use information/signal used to infer market demand in potential new export market as exclusion restriction. Information that potential entrant j has in destination i is not complete before they enter the market. Thus, typically potential entrant j has to infer its product demand at destination i from two major sources. One is from their own experience for the same product from other destinations and the other is the market demand for neighbor's same product in the destination i. We note that the demand generally can vary by exporter, destination, product, and time. Although each firm cannot know her demand in new export market fully but they can infer their individual demand by using information from these two sources. We also note that these information firms use to infer their individual demand is no use for volume decision once they are in the market, as at that point they can observe the demand for their own product. Thus, after entering into new markets, the signals from neighbor's performance and their own experience in similar markets cannot have any additional value for the firms beyond their direct experience in the new market. In this sense, firms experience in other destinations and neighbor's performance has no additional information once we account for firms own performance in the particular market.[5]

In the construction of signals of market demand of product, we follow closely the specification of Fernandes and Tang (2014). Three variables are used as signal of new market demand: neighbor's average export growth, the number of incumbent neighbors, and the interaction between the two. We also expand variables for signal by including the log standard deviation of neighbor's exports. We consider learning in two dimension and define neighbor in two dimensions accordingly. The neighbor in the first dimension is exporter-dimension which includes exporters in the neighborhood who already export the same product at the same destination. On the other hand, the neighbor in the second dimension is destination-dimension which includes the same product from the same exporter in other destinations.

We estimate the following specification in firm's new entry market decision:

$$\begin{aligned}1(export_{ijkt} > 0) = \Phi(&\gamma_1 \triangle ln(export_{ikt}) + \gamma_2 ln(n_{ikt-1}) + \gamma_3 ln(n_{ikt-1}) \times \triangle ln(export_{ikt}) \\ &+ \beta_1 \triangle ln(export_{jkt}) + \beta_2 ln(n_{jkt-1}) + \beta_3 ln(n_{jkt-1}) \times \triangle ln(export_{jkt}) \\ &+ \mathbf{Z}_{ijt}\delta + \alpha_0 + u_{it} + v_{jt} + w_{kt} + e_{ijkt} \geq 0) \end{aligned} \tag{5}$$

[5]We need to use lagged dependent variable to account for it own experience. Suppose that a potential entrant try to infer the demand for her product by observing neighbor's performance in a potential export market. There is uncertainty about how close her demand would be to those of neighbor or there is uncertainty about how close her demand in other destination would be to the demand in potential new destination. However, once they get into the market, as they can see their own performance which is direct measure of the demand, information obtained from neighbor may not have no additional value in inferring demand.

where $\triangle ln(export_{ikt})$ is average export growth from $t-1$ to t at the destination i for product k among all exporters with positive export flows at $t-1$ and t ; n_{ikt-1} is the number of exporters with positive flows at the destination i for the product k at year $t-1$; $\triangle ln(export_{jkt})$ is the average export growth from $t-1$ to t of j's exports of product k to all destinations excluding i, and n_{jkt-1} is the number of destinations excluding i that j exports k to at year $t-1$. Thus, γ_s captures the effect of signal from neighbor firms in the potential new export markets while β_s captures the effect of signal from firm's own experience from other markets for the same product. $\mathbf{Z}_{ijt}$ includes gravity variables such as GDP of importer and exporter, GDP per capita of importer and exporter, log distance between importer and exporter, dummy variables for sharing border between importer and exporter, for sharing common legal origin, for sharing common colony and for sharing common language, and three preferential trade agreement dummy variables, partial scope agreement, free trade agreement, and custom union agreement, among others. Finally, $u_{it}+v_{jt}+w_{kt}$ are three types of unobserved heterogeneity that are accounted for by fixed effects.

We extend the specification of eq. (5) to also include a measure for the dispersion of signal as follows:

$$
\begin{aligned}
1(export_{ijkt} > 0) = \Phi(&\gamma_1 \triangle ln(export_{ikt}) + \gamma_2 ln(n_{ikt-1}) + \gamma_3 ln(n_{ikt-1}) \times \triangle ln(export_{ikt}) \\
&+ \beta_1 \triangle ln(export_{jkt}) + \beta_2 ln(n_{jkt-1}) + \beta_3 ln(n_{jkt-1}) \times \triangle ln(export_{jkt}) \\
&+ \gamma_4 V_1 + \gamma_5 V_1 \times \triangle ln(export_{ikt}) + \beta_4 V_2 + \beta_5 V_2 \times \triangle ln(export_{jkt}) \\
&+ \mathbf{Z}_{ijt}\delta + \alpha_0 + u_{it} + v_{jt} + w_{kt} + e_{ijkt} \geq 0)
\end{aligned} \tag{6}
$$

where V_1 is the log standard deviation of neighbor's exports at the destination i and V_2 is the log standard deviation of exports at destinations in the neighborhood for potential new exporter j.

4 Empirical Analysis

4.1 Data

The dependent variable is bilateral trade flows and the main explanatory variable is bilateral tariffs averaged at HS-2 digit from 1996 to 2008 for 82 countries. Our sample coverage for 82 countries and 13 years is determined mainly by tariff data availability. Trade flows at the HS2-digit level are obtained from the UNCOMTRADE and a time-variant bilateral tariffs at the HS2-digit level data are obtained from the World Integrated Trade Solution (WITS). We use applied tariff data and try to make our tariffs data as balanced as possible by using imputation. For instance, WITS allows us to impute tariff information for certain year with preceding adjacent available year. Thus, in the case of intermediate missing tariff, we use the data from the closest previous available year as a substitute. If it is not available in this manner, we use the data from the closest available year. Data on nominal GDP and GDP per capita are drawn from the Penn World Table (PWT) 7.0, and data on GDP deflator are drawn from the U.S. Department of Commerce's Bureau of Economic Analysis. PTA data are constructed from Regional Trade Agreements Information System (RTA-IS) of the World Trade Organization (WTO) and data on GATT/WTO membership are also drawn from the WTO website. Data on gravity variables such as distance, common language, common colony, common legal origin and adjacency are from CEPII.

4.2 Elasticity estimates from aggregate data

Table 4.2 shows that the proportion of zero trade flows for aggregate data in our sample which is as low as 15%. This is because we restrict our sample countries to 82 and to 13 years from 1996 to 2008. In other words, we exclude the majority of non- trading pairs

Table 1: Number of trade flows observations

	Aggregate, 1996-2008	HS2, 2002	HS2, 1996-2008
Positive value only	75,107	220,914	2,946,083
Zero + positive value	88,478	653,376	8,493,888
Proportion of zero	15.11%	66.19%	65.32 %

during our sample periods as we exclude countries with no tariff data. Therefore, most of pairs of countries already did trade at least one product during our sample period and the role for correction terms obtained from the first stage estimation should be limited.

With aggregate data, we estimate the following trade flow equation:

$$ln(T_{ijt}) = \beta_0+\lambda_{it}+\xi_{jt}+w_{ij}+\beta_1\cdot tariff_{ijt}+\mathbf{Z}_{ij}\delta+ln[exp(\delta(\hat{\bar{v}}^*_{ijt}+\hat{z}^*_{ijt}))-1]+\rho_1\hat{\bar{v}}^*_{ijt}+e_{ijt} \tag{7}$$

Here $T_{ijt} > 0$ and we approximate $ln[exp(\delta(\hat{\bar{v}}^*_{ijt}+\hat{z}^*_{ijt}))-1]+\rho_1\hat{\bar{v}}^*_{ijt}$ using $\rho_1\hat{\bar{\eta}}^*_{ijt}+\rho_2\hat{\bar{z}}^*_{ijt}+\rho_3\hat{\bar{z}}^{*2}_{ijt}+\rho_4\hat{\bar{z}}^{*3}_{ijt}$, where $\hat{\bar{\eta}}^*_{ij} = \frac{\varphi(\hat{z}^*_{ij})}{\Phi(\hat{z}^*_{ij})}$ is the IMR obtained from (6) to account for selection bias as well as firm heterogeneity. In the estimations with aggregate level data, the estimation model specification remains the same as the product level estimations. In the first stage estimation of eq. (6), exclusion restrictions are composed of 12 terms from two types of learning.[6] Note that with aggregate data, learning for the signal of market demand occurs at the country level , not the product-country level.

The estimation results in Table 4.2 confirm this conjecture, as the estimates for the elasticities of trade does not change much from OLS estimate by including additionally fixed effects and HMR correction terms for self-selection and firm heterogeneity. The qualitative results of the HMR estimation are the same as those of the OLS and FE esti-

[6]HMR in the aggregate data estimation first consider bilateral regulation measure as the exclusion restriction variable. Besides strong assumption of excludability in the model for trade volume, a practical limitation of this exclusion restriction is data availability, which prompts HMR to consider an alternative variable of an index of common religion (between any pair). Data on common religion has been available only for sporadic periods until Maoz and Henderson (2013) construct a world religion dataset for every five years from 1945 to 2010. While not to depreciate the value of this improved dataset on religion, we should be cautious about the measurement errors due to the challenging nature in measuring them at the first place.

mations.

4.3 Elasticity estimates from product level data

With product level data, our main estimation equation is the following:

$$ln(q_{ijkt}) = \beta_0+\lambda_{ijt}+u_{ikt}+v_{jkt}-\delta\mathbf{x}_{ijt}+\beta_1\cdot tariff_{ijkt}+\rho_1\hat{\bar{\eta}}^*_{ijkt}+\rho_2\hat{\bar{z}}^*_{ijkt}+\rho_3\hat{\bar{z}}^{*2}_{ijkt}+\rho_4\hat{\bar{z}}^{*3}_{ijkt}+e_{ijkt} \quad (8)$$

where we only use positive trade flows, $q_{ijkt} > 0$, $\mathbf{x}_{ijt}$ include gravity variables as in the previous section, and $\hat{\bar{\eta}}^*_{ijkt}, \hat{\bar{z}}^*_{ijkt}, \hat{\bar{z}}^{*2}_{ijkt}, \hat{\bar{z}}^{*3}_{ijkt}$ are obtained from the first stage estimation of eq. (6) using the Probit. It should be noted that eq. (8) has three distinct unobserved factors, $\lambda_{ijt} + u_{ikt} + v_{jkt}$ which account for the product-specific unobserved heterogeneity subsuming product-specific fixed costs and product level MRTs. As a result, we can perform regression analyses with product-level data while avoiding omitted variable bias (OVB) caused by product-level MRTs as emphasized in Anderson and Yotov (2011, 2012) and pair-product level unobserved fixed costs which arise from, for instance, information barriers, interest-group lobbying, and government red-tape as modeled in Chaney (2008) and Krautheim (2012).

Table 4.3 report the representative estimation results in year 2002 as a representative year. As year 1986 in the HMR, these results are not specific to year 2002. We report the estimation results for each of the years from 1997 to 2008 in the next table and obtain similar results. If anything, year 2002 data give the largest estimate of trade cost elasticity and thus the smallest welfare gain amongst all the other years we considered. In other words, the results for year 2002 are the most conservative one.

Unlike aggregate data, the proportion of zero trade flows for HS2 digit product level data in 2002 sample is about 66%. This is too large to be treated as negligible. We believe that accounting for zero trade flows from self-selection and firm heterogeneity could be

Table 2: Trade cost elasticity estimates based on aggregate data, 1996-2008

	(1)	(2)	(3)	(4)
Elasticity (1-σ)	-1.288***	-1.020***	-1.125**	-1.203**
	(0.348)	(0.285)	(0.500)	(0.526)
GDP_i	1.024***	2.052***		
	(0.013)	(0.200)		
GDP_j	1.198***	0.718***		
	(0.013)	(0.212)		
$log(distance)$	-1.203***			
	(0.037)			
PSA	0.121*	0.122		
	(0.067)	(0.102)		
FTA	0.513***	0.129***		
	(0.079)	(0.037)		
CU	0.543***	0.677*		
	(0.142)	(0.352)		
IMR				1.120***
				(0.526)
$hetergeneity$				1.533***
				(0.487)
$hetergeneity^2$				-0.420***
				(0.124)
$hetergeneity^3$				0.034***
				(0.011)
gravity vars	Yes	Yes	Yes	Yes
Pair (ij) FEs	No	Yes	Yes	Yes
MRTs (it, jt FEs)	No	No	Yes	Yes
R^2	0.719	0.915	0.922	0.923
Model		log-linear		HMR
No. of obs.		75,107		75,340

Notes: Gravity variables additionally include dummy variables for sharing border, common legal origin, common colony, and log of GDP per capita for importer as well as exporter. MRTs are accounted for by importer-year and exporter-year fixed effects. Cluster (pair) robust standard errors are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5%, 1% levels, respectively.

very important especially if the response from zero flows to positive flows is different from the response from one positive flows to another positive flows. Columns (1), (2) and (3) use the log-linear and they are different in terms of controlling for the fixed effects. The difference between column (2) and column (3) is to account for product level MRTs following Anderson and Yotov (2011). Compared to the estimate in column (2), additional control for the product level MRTs in column (3) reduces the trade elasticity from -3.6 to -2.8. This implies that the elasticity of trade costs is overestimated in absolute value by 29% (=$\frac{0.8}{2.8}$) if we ignore MRTs that are specific to product k. Further comparison to the estimates in column (4) shows that additionally taking into account the impact of self-selection and firm heterogeneity (due to the fixed costs of exporting) on prices of product k using the HMR approach, the trade elasticity is reduced in magnitude from -2.8 to -2.3. The difference between these two estimates is beyond sampling error and this implies that, in absolute value, ignoring self-selection and firm heterogeneity overestimates the elasticity of trade costs by 22% (=$\frac{0.5}{2.3}$). Overall, the trade elasticity estimate decreases by 57% as we further control for self-selection and firm heterogeneity as well as sector level MRTs.

To accommodate a number of zeros, the literature also estimates gravity models with an exponential mean using the PPML method as suggested by Silva and Tenreyro (2006):

$$q_{ijkt} = exp(\beta_0 + \lambda_{ijt} + u_{kt} - \delta \mathbf{x}_{ijt} + \beta_1 \cdot tariff_{ijkt}) \cdot e_{ijkt} \tag{9}$$

In Table 4.3, the estimate for β_1 from conditional PPML (CPPML) estimations is reported in column (5).[7] In the eq. (9), we can account for λ_{ijt} by applying conditional PPML method where the sum of trade flows over product, $\sum_{k=1}^{K} q_{ijkt}$, as conditioning argument. However, due to computational difficulty (i.e. convergency failed), we cannot control both v_{jkt} as well as u_{ikt}. Thus, we instead account for u_{kt} which is the best we can

[7] We also tried to estimate the trade elasticity using CPPML accounting for MRTs with the aggregate panel data, but failed to obtain an estimate due to convergence issue.

do with CPPML in terms of accounting for unobserved factors most comprehensive way. As shown in column (5), the estimate is -1.16 which is again smaller than the estimates from the log-linear.

Table 4.3 reports the robustness of the estimation results in Table 4.3. We estimate the same model using the sample from other years for which coverage for pair of importers and exporters as well as products remain the same. We report the results from the log-linear model estimations in the first two columns. The direction of overestimation is the same for all years as not accounting for unobserved heterogeneity and self-selection as well as firm heterogeneity. As shown from comparisons between (2) and (3), not accounting for HMR correction terms overestimate the magnitude of elasticity by 53% (1997, 2000, and 2003) while, as shown from comparisons between (1) and (2), not accounting for the product level unobserved heterogeneity overestimate the magnitude of elasticity by 43% (2004, and 2007). As we combine these two sets of comparisons, we can see that the magnitude of elasticity of trade costs reduced by 107% in 2007 by accounting for both product level unobserved factors and self-selection as well as firm heterogeneity.

4.4 Implication of small elasticity estimates for Welfare

Arkolakis et al. (2012) show that for a range of trade models, including the Armington model and new trade models with micro-foundation like Eaton and Kortum (2002) and Melitz and Ottaviano (2008), the welfare gains from trade (compared to autarky) can be simply measured using two statistics, the share of expenditure on domestic goods, λ and the elasticity of imports with respect to variable trade cost, ϕ. Consider the following example of US in 2000 as in Arkolakis et al. (2012). In year 2000, the share of expenditure devoted to domestic goods for US is 0.93 (i.e. $\lambda_{us} = 0.93$). Using the welfare change formula in Arkolakis et al. (2012) to evaluate the welfare change in US's year 2000 compared to Autarky, which is $(1 - \lambda^{-1/\phi})$ where $\lambda = 0.93$, they illustrate that the percentage change in real income needed to compensate a representative consumer for going back to

Table 3: Trade cost elasticity estimates based on HS2 disaggregate data, 2002

	(1)	(2)	(3)	(4)	(5)
	Disaggregate, HS2, 2002				
Elasticity (1-σ)	-2.828***	-3.641***	-2.814***	-2.299***	-1.155**
	(0.156)	(0.133)	(0.208)	(0.194)	(0.540)
GDP_i	0.671***				
	(0.011)				
GDP_j	0.830***				
	(0.012)				
$log(distance)$	-0.807***				
	(0.032)				
PSA	0.124**				
	(0.056)				
FTA	0.276***				
	(0.081)				
CU	0.608***				
	(0.085)				
IMR				0.604***	
				(0.057)	
$heterogeneity$				5.979***	
				(0.149)	
$heterogeneity^2$				-0.901***	
				(0.047)	
$heterogeneity^3$				0.059***	
				(0.005)	
gravity vars	Yes	Yes	Yes	Yes	Yes
Pair (ij) FEs	No	Yes	Yes	Yes	Yes
MRTs (ikt, jkt FEs)	No	No	Yes	Yes	
kt FEs					Yes
R^2	0.314	0.392	0.366	0.440	
Model		log-linear		HMR	CPPML
No. of obs.		220,914		220,256	561,542

Notes: 2002 is chosen just as a middle of year in the sample. Cluster (pair) robust standard errors are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5%, 1% levels, respectively.

Table 4: Trade cost elasticity estimates based on HS2 disaggregate data, 1997-2008

	(1)	(2)	(3)	(4)
	Disaggregate, HS2			
1997	-2.356***	-2.300***	-1.464***	-0.737**
	(0.110)	(0.230)	(0.222)	(0.333)
1998	-2.929***	-2.721***	-1.611***	-0.792**
	(0.129)	(0.257)	(0.240)	(0.386)
1999	-3.018***	-2.804***	-1.904***	-0.861**
	(0.116)	(0.224)	(0.219)	(0.384)
2000	-3.216***	-2.864***	-1.854***	-0.804*
	(0.122)	(0.202)	(0.191)	(0.425)
2001	-3.650***	-2.889***	-2.054***	-0.729
	(0.135)	(0.214)	(0.198)	(0.486)
2002	-3.641***	-2.814***	-2.299***	-1.155**
	(0.133)	(0.208)	(0.194)	(0.540)
2003	-3.493***	-2.908***	-1.941***	-0.999*
	(0.130)	(0.211)	(0.194)	(0.542)
2004	-3.677***	-2.618***	-2.057***	-1.117**
	(0.145)	(0.215)	(0.203)	(0.494)
2005	-3.531***	-2.687***	-2.290***	-1.102**
	(0.141)	(0.217)	(0.204)	(0.448)
2006	-3.932***	-3.051***	-2.259***	-1.404***
	(0.151)	(0.212)	(0.187)	(0.521)
2007	-3.722***	-2.610***	-1.829***	-1.612***
	(0.168)	(0.228)	(0.200)	(0.552)
2008	3.941***	-2.966***	-2.197***	-1.498**
	(0.168)	(0.234)	(0.213)	(0.623)
Pair (ij) FEs	Yes	Yes	Yes	Yes
MRTs (ikt,jkt FEs)	No	Yes	Yes	
kt FEs				Yes
HMR	No	No	Yes	Yes
Model	log-linear		HMR	CPPML

Notes: Cluster (pair) robust standard errors are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5%, 1% levels, respectively.

autarky is 0.7 percent to 1.4 percent depending if the trade elasticity are ranged from -10 to -5 as surveyed in Anderson and van Wincoop (2004). If we use the estimate obtained from non-zero observations as in column (1) for year 2000 in Table 4.3, US' gains from trade in year 2000 is implied to be 2.2 percent, slightly higher than the figure in Arkolakis et al. (2012) but still seems to be quite small. When we use the estimate obtained from data including zero observations with proper product-level controls as in column (3), the implied US's gains from trade in year 2000 increase to 3.8 percent.

In order to compute the total welfare gains with multiple sectors we need additional data on share of domestic expenditure, share of consumption and employment for each sector as well as sectoral trade elasticities (see section 5.1 in Arkolakis et al. (2012)). To deliver the main objective of this paper, however, we focus on trade elasticity average across industries and this simple numerical example serves the purpose that, to evaluate the welfare impact of trade liberalization, it is paramount to have an unbiased estimate of trade elasticity.

5 Conclusion

TBA

References

Albornoz, F., Pardo, H. F. C., Corcos, G., Ornelas, E., 2012. Sequential exporting. Journal of International Economics 88 (1), 17–31. 1, 3.1

Anderson, J. E., 1979. A theoretical foundation for the gravity equation. American Economic Review 69 (1), 106–116. 1

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

Anderson, J. E., van Wincoop, E., 2004. Trade costs. Journal of Economic Literature 42 (3), 691–751. 1, 1, 4.4

Anderson, J. E., Yotov, Y. V., Apr. 2011. Terms of Trade and Global Efficiency Effects of Free Trade Agreements, 1990-2002. NBER Working Papers 17003, National Bureau of Economic Research. 4.3

Anderson, J. E., Yotov, Y. V., Apr. 2012. Gold standard gravity. NBER Working Papers 17835, National Bureau of Economic Research. 4.3

Arkolakis, C., Costinot, A., Rodriguez-Clare, A., February 2012. New trade models, same old gains? American Economic Review 102 (1), 94–130. 1, 4.4

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

Bergstrand, J. H., Egger, P., Larch, M., 2013. Gravity redux: Estimation of gravity-equation coefficients, elasticities of substitution, and general equilibrium comparative statics under asymmetric bilateral trade costs. Journal of International Economics 89 (1), 110–121. 1

Broda, C., Limao, N., Weinstein, D. E., 2008. Optimal tariffs and market power: The evidence. American Economic Review 98 (5), 2032–65. 1

Broda, C., Weinstein, D. E., May 2006. Globalization and the gains from variety. The Quarterly Journal of Economics 121 (2), 541–585. 1

Chaney, T., 2008. Distorted gravity: The intensive and extensive margins of international trade. American Economic Review 98 (4), 1707–21. 2, 4.3

Eaton, J., Eslava, M., Krizan, C. J., Kugler, M., Tybout, J., 2014. A search and learning model of export dynamics. mimeo. 1, 3.1

Eaton, J., Eslava, M., Kugler, M., Tybout, J., 2007. Export dynamics in colombia: Firm-level evidence. Tech. rep., National Bureau of Economic Research. 1, 3.1

Eaton, J., Kortum, S., 2002. Consistent estimation from partially consistent observations. Econometrica 70, 1741–1779. 1, 4.4

Fernandes, A. P., Tang, H., 2014. Learning to export from neighbors. Journal of International Economics 94 (1), 67–84. 1, 3.1, 3.3

Harrigan, J., 1993. Oecd imports and trade barriers in 1983. Journal of International Economics 35 (1-2), 91–111. 1

Helpman, E., Melitz, M., Rubinstein, Y., 2008. Estimating trade flows: Trading partners and trading volumes. The Quarterly Journal of Economics 123 (2), 441–487. (document), 1, 4

Hummels, D., 2001. Toward a geography of trade costs. Tech. rep., Mimeo, Purdue University. 1

Koenig, P., Mayneris, F., Poncet, S., 2010. Local export spillovers in france. The European Economic Review 54 (4), 622–641. 2

Krautheim, S., 2012. Heterogeneous firms, exporter networks and the effect of distance on international trade. Journal of International Economics 87 (1), 27–35. 2, 4.3

Maoz, Z., Henderson, E. A., 2013. The world religion dataset, 1945-2010: Logic, estimates, and trends. International Interactions 39 (3), 265–291. 6

Melitz, M. J., November 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71 (6), 1695–1725. 1

Melitz, M. J., Ottaviano, G. I., 2008. Market size, trade, and productivity. The review of economic studies 75 (1), 295–316. 4.4

Morales, E., Sheu, G., Zahler, A., 2011. Gravity and extended gravity: Estimating a structural model of export entry. mimeo. 1, 3.1

Ossa, R., 2015. Why trade matters after all. Journal of International Economics 97 (2), 266–277. 1

Silva, J. S., Tenreyro, S., 2006. The log of gravity. The Review of Economics and Statistics 88 (4), 641–658. 4.3

Simonovska, I., Waugh, M. E., 2014. The elasticity of trade: Estimates and evidence. Journal of international Economics 92 (1), 34–50. 1