Crawford School of Public Policy

CAMA

Centre for Applied Macroeconomic Analysis

Trade Elasticity: Estimates From Product-Level Data

CAMA Working Paper 29/2020
March 2020

Juyoung Cheong
Department of Economics, Kyung Hee University, Korea

Do Won Kwak
Graduate School of International Studies, Korea University, Korea

Kam Ki Tang
School of Economics, University of Queensland
Centre for Applied Macroeconomic Analysis, ANU

Abstract

In a seminal paper, Helpman et al. (2008) demonstrate that accounting for self-selection and firm heterogeneity is crucial for obtaining unbiased estimates in the gravity equation. In this paper, we show how this insight helps solve the trade elasticity puzzle that, hitherto, trade elasticity estimates in the gravity equations have implied very small welfare gains from trade. To apply their approach to product-level data, we propose new exclusion restrictions based on the export learning literature. Our new estimates imply much larger welfare gains from trade than the previous literature suggested.

Keywords

Gravity Model, Firm Heterogeneity, Disaggregate Data, Trade Elasticity, Learning

JEL Classification

C13, C23, F10, F15

Address for correspondence:

(E) cama.admin@anu.edu.au

Trade Elasticity: Estimates From Product-Level Data

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

Abstract

Long version: In a seminal paper, Helpman et al. (2008) (HMR) demonstrate that, in the presence of a large proportion of zero bilateral trade observations, accounting for self-selection and firm heterogeneity is crucial for obtaining unbiased estimates in the gravity equation. In this paper, we show how this insight helps solve the trade elasticity puzzle that, hitherto, trade elasticity estimates in the gravity equations have implied very small welfare gains from trade. We estimate the trade elasticity using Harmonized System 2-digit product-level data for 63 importing and 135 exporting countries from 2001 to 2010. To apply the HMR approach, we propose new exclusion restrictions that vary across country-pair-product-time. The new exclusion restrictions are constructed based on the export market learning literature. Accounting for self-selection and firm heterogeneity substantially lowers the estimate of the trade elasticity from -2.6 to -1.4, which implies much larger welfare gains from trade than the previous literature suggested.

Short version: In a seminal paper, Helpman et al. (2008) demonstrate that accounting for self-selection and firm heterogeneity is crucial for obtaining unbiased estimates in the gravity equation. In this paper, we show how this insight helps solve the trade elasticity puzzle that, hitherto, trade elasticity estimates in the gravity equations have implied very small welfare gains from trade. To apply their approach to product-level data, we propose new exclusion restrictions based on the export learning literature. Our new estimates imply much larger welfare gains from trade than the previous literature suggested.

JEL Code: C13; C23, F10; F15

[*]Department of Economics, Kyung Hee University, Korea; e-mail: `jcheong@khu.ac.kr`

[†]Graduate School of International Studies, Korea University, Korea; e-mail: `dwkwak@korea.ac.kr`

[‡]School of Economics, University of Queensland, Australia; Centre for Applied Macroeconomic Analysis (CAMA), Australian National University, Australia; e-mail: `kk.tang@uq.edu.au`

[§]The authors would like to thank Peter Egger, Pao-Li Chang, James Lake, Benedikt Heid, Jason Query, Haishan Yuan, Youjin Hahn, and other participants at the WEAI conference in 2019, the Australian Trade Workshop in 2016 and 2017, the 2016 Asia-Pacific Trade Seminar, the 2018 Asian Meeting of the Econometric Society, and seminars at Kyung Hee University, Sungkyunkwan University, Yonsei University, and Korea University for their valuable suggestions. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A5A2A03037147).

1 Introduction

In a seminal paper, Helpman et al. (2008) (denoted as HMR hereafter) demonstrate that zero trade flows between countries are the result of heterogeneous firms' self-selection out of export markets, and show that not accounting for these heterogeneous firm characteristics could cause bias in the gravity model estimations. In this paper, we aim to show how this insight helps solve the trade elasticity puzzle that, hitherto, trade elasticity estimates in the gravity equations have implied very small welfare gains from trade. In particular, we develop a new method that extends the HMR approach for country-level data to product-level data and allows us to obtain unbiased estimates of trade elasticities. Our estimates imply much bigger welfare gains from trade than the previous literature suggested.

The price elasticity of trade or simply the trade elasticity is a key parameter to quantify welfare gains from trade.[1] As shown in Arkolakis et al. (2012) (denoted as ACR hereafter), for a broad range of trade models encompassing homogeneous and heterogeneous firm models, one can measure welfare gains from trade with just two parameters: the import penetration ratio (or domestic share out of total expenditure) and the trade elasticity with respect to variable trade costs.[2] The import penetration ratio can be readily obtained from national statistics, whereas the trade elasticity is not directly observable and needs to be estimated. However, estimates of the trade elasticity could vary significantly with model specifications and estimation methodologies, even after taking into account differences in the data coverage. For instance, a survey by Anderson and van Wincoop (2004) finds that elasticity estimates range from -5 to -10. This large range of trade elasticity

---

[1]Trade elasticity in this paper refers to micro-elasticity governing the substitution between varieties of imported goods. In the macro literature, elasticity, on the other hand, refers to the substitution between domestic and imported goods and its estimate is often concerned with small magnitude. See Feenstra et al. (2018) for a detailed discussion on this literature.

[2]Melitz and Redding (2015) show, however, that the trade elasticity is endogenously determined under heterogeneous firm models and that the additional adjustment margin in heterogeneous firm models is not captured by the trade elasticity. The implication is that the import penetration ratio and the trade elasticity are not sufficient to measure the welfare gains from trade in more general settings. Nevertheless, it is still one of the most important parameters in measuring the welfare gains from trade.

estimates implies sizable discrepancy in the measured trade frictions and welfare gains from trade.

Much work has been devoted to improving the estimation of trade elasticity over several decades (e.g., Anderson, 1979; Baier and Bergstrand, 2001; Broda and Weinstein, 2006; Harrigan, 1993; Imbs and Mejean, 2015; Ossa, 2015; Simonovska and Waugh, 2014a; Soderbery, 2018). ACR apply these estimates and the U.S.'s import penetration ratio of 0.07 for year 2000 to their formula and calculate the U.S.'s gains from trade (from autarky to the then trade regime) to lie between 0.7% and 1.4% of real income. As pointed out by Feenstra et al. (2018), these welfare gain estimates are rather small given the importance of trade to modern economies.[3]

Many early studies on trade elasticity relied on country-level data, but more recent studies are able to provide product-level estimates as disaggregate data have become more accessible. Simonovska and Waugh (2014a) and Caliendo and Parro (2014), who build on the Eaton and Kortum (2002) Ricardian model with geographic barriers, obtain sector-level trade elasticity estimates, but their estimates are not markedly different from those in the previous literature. Broda and Weinstein (2006) and Kee et al. (2008) conduct structural estimations with disaggregate data using simplified demand and supply functions or gross domestic product (GDP) functions. Although they provide much more detailed information on trade elasticity than earlier studies, their models do not taken into account the micro-level margins from firm heterogeneity in the new trade models. Ossa (2015), using a method that incorporated sectoral linkages across industries, show that a very small trade elasticity in just a few industries could contribute to significant overall welfare gains for the economy. Ossa's average or aggregate trade elasticity estimate across industries, -3.6, is at the lower end of the spectrum of the previous literature (in absolute terms, the same is true for the rest of the paper). However, this estimate still omits firm heterogeneity and, thus, implies a smaller welfare gain for the overall economy than otherwise would be the case. Soderbery (2018) introduces a structural estimator to estimate the export supply elasticity that incorporates exporter

---

[3]ACR formula is for a one-sector economy. Ossa (2015) shows that if one considers a multi-sector economy with sectoral linkages, the welfare gains will become larger.

heterogeneity but not at the firm level, and his mean estimates are not much different from those of Ossa (2015). Furthermore, Imbs and Mejean (2015) show that trade elasticity estimates from aggregate data would suffer from systematic downward bias (e.g., true estimate is -6 but biased estimate is -3). This indicates that the welfare gains from trade should be even smaller if we consider products heterogeneity using disaggregate data. As Ossa (2015) stressed, despite all the improvement in estimation techniques and data quality over time, trade elasticity estimates remain large and welfare gains small.

The objective of this paper is to improve the trade elasticity estimate using disaggregate tariff (as a variable cost) and trade flow data. As the proportion of zero trade flows observations is much larger in disaggregate data than in aggregate data, the insight of HMR suggests that accounting for self-selection and firm heterogeneity would be even more important. Indeed, by accounting for the causes of zero trade flows, we obtain a smaller trade elasticity and, thus, larger welfare gains from trade. Our results echo the findings of Simonovska and Waugh (2014b) that the extensive margin observed in the new trade models reduces the trade elasticity.[4]

HMR suggest a two-stage procedure to correct for the bias from omitting zero trade flows. The export market entry decision is modeled in the first stage and the volume decision conditional on entering the market in the second stage. Their method has been widely adopted in the gravity model literature (e.g., Baier et al., 2014; Cheong et al., 2015; Dutt et al., 2013). The method requires exogenous variations, known as the exclusion restrictions (ERs), that affect firms' entry decisions, but not their performance once they have entered a foreign market.[5] Prominent examples of ERs for aggregate data are entry regulation (e.g., Djankov et al., 2002) and religion proximity (e.g., Helpman et al., 2008; Cheong et al., 2015; Dutt et al., 2013), but its theoretical foundations may be

---

[4]Using year 2004 cross-sectional data covering 30 countries, Simonovska and Waugh (2014b) find that the trade elasticity estimate in the Melitz model is 30% smaller than in the Krugman model. The result is attributed to the presence of the extensive margin of trade in the former, but not in the latter.

[5]Recent theoretical studies such as those by Chaney (2008) and Krautheim (2012) have emphasized the role of fixed costs in heterogeneous firms' decisions to enter new export markets. Their conclusions are empirically supported by Koenig et al. (2010).

questioned. Furthermore, these candidates of ERs are not applicable to product-level data because they are subsumed by country-pair-time fixed effects. Finding an ER that works at the product-level data is challenging because very few country-pair-product-time-varying variables are available in practice. To the best of our knowledge, tariff and trade flows are the only two variables of this kind for which the data are publicly available. However, tariff and trade flows cannot be directly used as ERs in our estimations because they are the key explanatory and dependent variables, respectively, in the volume (i.e., second-stage) equation.

In this paper, we propose new ERs that allow us to extend the HMR approach to product-level data. The new ERs are derived based on the recent literature on learning in exporting markets such as Albornoz et al. (2012), Eaton et al. (2007), Eaton et al. (2014), Fernandes and Tang (2014), Holloway (2017), and Morales et al. (2011). In these studies, firms learn about their chance in a prospective export market from two sources: i) the performance of other countries/firms in the same market for the same product, and ii) their own performance in other destinations for the same product. Such learning about a prospective market's demand affects a firm's entry decision. However, once a firm has entered a new market, it can directly observe the demand for its product and therefore does not need to infer from other firms' experience or its own experience in other markets to decide how much it should export to the new market. We show that these learning variables pass standard tests for ERs after controlling for various unobserved factors. This methodological innovation is important because it opens up opportunities to apply the highly influential HMR approach to testing various trade theories using product-level data.

Building on the HMR model, we propose a new method to estimate an empirical gravity model at the sector level using Harmonized System (HS) 2-digit data from 2001 to 2010 for 63 importing and 135 exporting countries. The trade elasticity is obtained from the response of trade flows to tariff changes, which constitute changes in variable trade costs. For a given change in import prices, the source of the price change should be irrelevant to the demand outcome. However, focusing on tariffs has the advantage that tariff data suffer less from measurement errors than other sources of

variable trade costs such as transportation costs and information costs.

Our main findings are as follows. For our HS 2-digit data, where zero trade flows are around 63%, the elasticity estimate decreases to -1.4 (from -4 to -10 as in the previous literature) when we apply the two-stage HMR approach with the new ERs. We apply the new elasticity estimates to the ACR formula for welfare gains from trade, and show that for the U.S. the gains are almost double that of the log-linear estimate and nearly triple the upper-end estimate in the literature. The results imply that in evaluating the welfare effects of trade, it is paramount to control the sources of zero trade flows.

We further find heterogeneity of the trade elasticity based on the income level of country-pairs. Our results show that industrial countries have smaller trade elasticities than developing countries, indicating larger welfare gains from trade for high-income countries. We also find that the trade elasticity increases during the global financial crisis, implying that buyers are more price-sensitive during economic hardship.

The rest of the paper is organized as follows. Section 2 extends the HMR model to a model with disaggregate data and explains the new ERs that we derived from the export learning literature. Section 3 describes the data. Section 4 reports the main results, and it also provides the sensitivity tests including heterogeneity analysis across the income level of country-pairs, time, and sector. Section 5 provides simple welfare implications, and the last section concludes.

2   The HMR Model for Sector Level

Firm-level heterogeneity has received attention in the recent international trade literature, including from Bernard et al. (2003); Kugler and Verhoogen (2012); Manova and Zhang (2012); Melitz (2003) among others.[6] A key features of these papers is that the extensive margin of trade is de-

---

[6]Firm heterogeneity due to various sources such as preference factors, production cost factors, quality factors and export fixed cost factors has been studied in the past. In this paper, we do not distinguish the source of firm heterogeneity but focus on accounting for firm heterogeneity to obtain a consistent trade elasticity estimation.

termined by firm heterogeneity. As trade barriers decrease, more firms can enter export markets, and this endogenous selection of firms into export markets can explain observed patterns of trade flows. Building on the Melitz (2003) model of heterogeneous firms under monopolistic competition, HMR develop an estimable equation using the assumption that firm productivity has a truncated Pareto distribution and use this assumption to account for zero trade flows at the aggregate (i.e., national) level. For the estimation strategy, they suggest a Heckman (1979) type two-stage method, in which an inverse Mills ratio (IMR) and related variables obtained from the first-stage Probit estimation are used to account for firm heterogeneity in the second-stage estimation of a gravity equation.

In this section, we extend the firm's entry decision model used for aggregate data in HMR to a model for disaggregated sector-level data. Without loss of generality, we assume products are distinct across sectors, and within each sector each firm produces one slightly different variety in a monopolistic competition environment. Therefore, there is a one-to-one mapping between varieties and firms.

Suppose that an exporting firm in sector k from country j faces the following demand for its product variety in destination i, q_{ijk}, under the monopolistic competition condition:

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

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

where Q_{ik} is the equilibrium market size of importing country i for products in sector k; c_{jk} is a measure of average product-specific productivity in sector k of firms in country j; τ_{ijk} is the variable trade cost of firms in sector k exporting from j to i; P_{ik} is the price index for sector k in importing country i, determined by domestic producers and existing exporters selling in country i; and 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 firms' productivity cutoff, which determines the proportion of country j's firms in sector k exporting to country i. Unlike HMR, we allow the productivity cutoff to differ across sectors. However, as in HMR, we assume that firm productivity, $G(a_k)$, has a truncated Pareto distribution, where a_{kH} (a_{kL}) represents the lowest (highest) productivity in sector k, so that $G(a_k) = (a_k^\theta - a_{kL}^\theta)/(a_{kH}^\theta - a_{kL}^\theta)$, $\theta > \gamma_k$, where θ is the shape parameter of the Pareto distribution for firm productivity. N_{jk} is the number of firms from country j in sector k; γ_k is the import demand elasticity; and V_{ijk} and W_{ijk} are a function of the productivity cutoff. Firms in country j take P_{ik} and Q_{ik} as given.[7] Notice that the cutoff productivity and demand elasticity are sector-specific, and these variables are functions of sector-level trade barriers.

Similar to HMR, we can write the volume of trade as follows. Under sector independence assumptions, we suppress k for the sake of simplicity. Thus, for each sector k, we have the following estimable equation:

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

where λ_j is exporter-specific fixed effects (FEs), which subsume $ln(N_j)$ and $ln(c_j)$; ξ_i is destination FEs, which subsume $ln(Q_i)$ and $ln(P_i)$; $\mathbf{x}_{1ij}$ includes all observed variables that could capture trade cost factors, including pair gravity variables, such as distance, cultural ties, and colonial relationship, and pair-sector variables such as tariffs; $w_{ij}(= ln(W_{ij}))$ is a function of the productivity cutoff; and u_{ij} is an idiosyncratic error term. Effectively, the obtained equation for the volume of trade in eq. (1) is the same as HMR except that the productivity cutoff arising from sector-specific trade barriers differs by sector. As a result, we estimate eq. (1) using sector-level data and, thus, we need information on $\mathbf{x}_{1ij}$ and w_{ij}, which are sector-specific. However, we do not observe w_{ij}; and without a proper control for w_{ij}, δ_1 could be severely biased. In the next section, we explain

---

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

our strategy to account for w_{ij}.

2.1 Model for the entry decision of a firm

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

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

where the numerator is the operating revenue and the denominator is the fixed costs of exporting. As long as $Z_{ij} > 1$, export accrues positive operating profits. We assume that for each sector 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 its fixed costs of exporting; ψ_i is destination-specific factors that could affect the fixed costs (e.g., government red tape); σ_{ij} contains information on the fixed costs that are specific to both exporter j and destination i; and $v_{ij} \sim N(0, \phi_v^2)$ captures remaining unobserved factors. The fixed costs of exporting are stochastic due to unobserved trade frictions, v_{ij}, that are assumed to be independent and identically distributed (i.i.d.) but correlated with the errors (u_{ij}) in the second-stage estimation. We take the logarithm of eq. (2) to obtain

$$z_{ij} \equiv ln(Z_{ij}) = \gamma_0 + \eta_j + \omega_i + \mathbf{x}_{ij}\delta - \theta \sigma_{ij} + \epsilon_{1ij}$$

where $\mathbf{x}_{ij}$ represents typical observed country-pair variables included in the gravity model; η_j is exporter FEs, subsuming all j-specific variables including c_j and ψ_j; ω_i is importer FEs, subsuming

10

all i-specific variables including P_i, Q_i, and ψ_i; σ_{ij} is information on the (sector-specific) fixed costs for firms in j to export to i; and $\epsilon_{1ij} = \rho_0 u_{ij} + v_{ij} \sim N(0, \phi_u^2 + \phi_v^2)$, and it is assumed that $\rho_0 = 1$ for the sake of simplicity.

σ_{ij}, which captures country-pair-specific fixed cost factors, is not present in eq. (1). There-fore, if we could find some proxy variables for σ_{ij}, the proxy variables could be used as ERs for the identification of parameters in eq. (1). Implementation of the two-stage estimation requires *observed* factors in σ_{ij} that vary with ij and affect the fixed costs of exporting.

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

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

where $\Phi(\cdot)$ is a standard normal cumulative distribution function (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 the endogenous number of exporters and self-selection of j's firms exporting to i as in:

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

where ω_{ij} includes factors that determine the fraction of firms in sector k exporting from j to i. Therefore, we need the estimates for both $E(w_{ij}|q_{ij} > 0, \mathbf{x}_{1ij}, \lambda_j, \xi_i)$ and $E(u_{ij}|q_{ij} > 0, \mathbf{x}_{1ij}, \lambda_j, \xi_i)$. Both terms depend on $\bar{v}_{ij}^* = E(v_{ij}^*|q_{ij} > 0, \eta_j, \omega_i, \mathbf{x}_{1ij}, \sigma_{ij})$. It should be noted that $E(u_{ij}|q_{ij} > 0, \mathbf{x}_{1ij}, \lambda_j, \xi_i) = corr(u_{ij}, v_{ij}) \cdot \frac{\sigma_u}{\sigma_v}\bar{v}_{ij}^*$, and $corr(u_{ij}, v_{ij}) \cdot \frac{\sigma_u}{\sigma_v} = \rho_1$ where $v_{ij}^* = \frac{v_{ij}}{\sigma_v}$. Also, the esti-mate for $\bar{v}_{ij}^*$ could be obtained from the IMR, $\hat{\bar{v}}_{ij}^* = \frac{\phi(\hat{z}_{ij}^*)}{\Phi(\hat{z}_{ij}^*)}$. Furthermore, for the consistent estimate of $E(z_{ij}|q_{ij} > 0, \eta_j, \omega_i, \mathbf{x}_{1ij}, \sigma_{ij})$, we could use $\hat{z}_{ij}^* + \hat{\bar{v}}_{ij}^*$ and $\hat{\bar{w}}_{ij}^* = ln[exp(\alpha(\hat{\bar{v}}_{ij}^* + \hat{z}_{ij}^*)) - 1]$ for the consistent estimate for $E(w_{ij}|q_{ij} > 0, \mathbf{x}_{1ij}, \lambda_j, \xi_i)$.

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

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

where α is a function of γ as well as θ; and $W_{ij} = Z_{ij}^\alpha - 1 = exp(\alpha z_{ij}) - 1$ is used to estimate w_{ij} by taking the logarithm of both sides of the equation. As long as an ER at the sector level is available, we can 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 $\rho_1 \hat{\bar{v}}_{ij}^*$ are used to account for firm heterogeneity and self-selection of exporting, respectively.

2.2 Exclusion restrictions at the sector level: Learning in exporting markets

The HMR approach requires one or more ERs that: (i) affect the fixed costs of exporting and, more broadly, the market entry decision; and (ii) do not affect the volume of trade of product once the entrant has become an incumbent in the export market.[8] Finding variables that meet these strict requirements at the aggregate level is already difficult; doing it at the sector level is understandably far more challenging. In particular, at the sector level, ERs have to be varying over country-pair-product-time. We are not aware of any readily available variables that are country-pair-product-time variant and satisfy the two criteria of ERs at the same time. To tackle this problem, we draw upon on the recent literature on export learning to develop new ER variables.

Fernandes and Tang (2014) argue that a larger number or a faster growth of a country's neighboring exporters in a specific market could provide information about the market's demand for the country concerned. They show that the information affects the country's entry decision. On the other hand, Eaton et al. (2007) and Eaton et al. (2014) show that a firm's previous export success affects its incentive to search for new export destinations, and a firm's geographic expansion path depends on its initial destination markets. Morales et al. (2011) also find that a firm's entry into a new destination is positively affected by its previous export experience in geographically or eco-

---

[8]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.

nomically similar markets. Likewise, Albornoz et al. (2012) and Holloway (2017) observe that a firm discovers its profitability as an exporter after actually engaging in exporting and then decides whether to enter into new markets.

Based on this export learning literature, we postulate that firms can learn about the potential demand for their products in a new export market from two sources: i) their own experience in other markets with the same product; and ii) the performance of firms from other countries in the same market with the same product. As regards ER criterion (i), the information firms extract from these sources can affect their entry decision into a potential export market as suggested by the aforementioned literature. As regards criterion (ii), once a firm enters a new market, it can directly observe the demand for its product in the new market, and therefore the information it obtained from other markets, or firms from other countries, become irrelevant to its decision on the supply quantity for this market.[9]

To learn about the market demand for a product, we follow the specification of Fernandes and Tang (2014) to focus on the average export volumes and the number of incumbents (or destinations). As such, we use the following information to generate four "foundation" ER variables:

i. $\overline{q}_{jkt}$: the average GDP-weighted export volume of product k across all destinations excluding i by country j;[10]

ii. N_{jkt}: the number of destinations excluding i for country j and product k;

iii. $\overline{q}_{ikt}$: the average export volume of product k by all countries excluding j to destination i;

---

[9]Even though the signals become irrelevant for firms' decisions on the export volumes, there is a possibility that learning variables may still affect the extensive margin after firms enter the market. For instance, $\overline{q}_{ikt}$, the average export volume of product k in i's market, may affect the intensive margin even after firms from country j enter i's market. This scenario is plausible if j's exports significantly increase the competition in i's market. In addition, N_{ikt}, the number of countries exporting product k in i's market, may also affect j's export volumes. This is because under monopolistic competition, the price index of country i's market depends on the number of firms in the market, and a significant change in the price index, which determines the demand, may affect country j's export volumes for product k in the market. However, these cannot be verified, as mentioned in footnote 8. We use lagged values to help mitigate this problem.

[10]We first obtain the ratio of export volume of product k by firm j to a given destination h as a proportion of the 1999 GDP of destination h, and then compute the average of the export-to-GDP ratio across all destinations excluding i.

and

    iv. N_{ikt}: the number of countries excluding j exporting product k to destination i.

The first two variables, $\overline{q}_{jkt}$ and N_{jkt}, are proxies for the knowledge firms from country j can learn from themselves in other markets to which they are exporting product k to at time t. The last two variables, $\overline{q}_{ikt}$ and N_{ikt}, are proxies for the knowledge firms from country j can learn from other firms that are exporting product k to destination i at time t.

To incorporate these learning variables as ERs at the first stage, we estimate the following specification in a firm's new entry market decision:

$$1(q_{ijkt} > 0) = \Phi(\beta_1(lnN_{ikt-1} \cdot lnN_{jkt-1}) + \beta_2(ln\overline{q}_{ikt-1} \cdot ln\overline{q}_{jkt-1}) + \mathbf{Z}_{ijkt}\delta + \alpha_0 + e_{ijkt} \geq 0) \quad (5)$$

where $\mathbf{Z}_{ijt}$ is a set of gravity equation variables and e_{ijkt} is an error term. β_1 and β_1 capture the effects of learning at time $t-1$.

Our benchmark ERs are two interaction terms, each of which is a product of one ikt-varying and one jkt-varying learning variables. We use the interaction terms to make our ERs $ijtk$-varying like the dependent variable. We need to emphasize that our ERs are just two out of numerous possible linear and non-linear combinations of the four foundation ER variables. In the following section, we check the robustness of our key results using a different set of ER variables. Furthermore, lagged terms of the ERs are used to avoid potential reverse causality and other concerns discussed in footnote 9. We also check the robustness of our key findings using current ER.

It is also worth mentioning that the expected signs of the coefficient of each learning variable are not always positive. For example, an increase in the number of firms from other countries already exporting in a potential destination country (N_{ikt}) implies more competition in the market. It may reduce the probability of a firm entering this market. In addition, firms are likely to enter the market where they can earn most profits first. Therefore, an increase in the number of markets to which a firm has already exported (N_{jkt}) may imply a lower probability of entering another new

market because the profits there are expected to be lower.

3 Data

The dependent variable in our empirical analysis is bilateral trade flows, and the main explanatory variable is bilateral tariffs, both of which are averaged at the HS 2-digit level. Our dataset covers 2001 to 2010[11] for 63 importing and 135 exporting countries, all of which are World Trade Organization (WTO) members. The sample coverage, especially for importing countries, is determined mainly by tariff data availability. We use the HS 2-digit data despite the availability of the HS 6-digit data for several reasons. First, there are computational problems caused by using high dimensional FEs accounting for unobserved heterogeneity. Second, including all zero trade flows for all pairs (63×135) and products (around 5,000) induces too many zeroes in the trade flows variable. Finally, we can show that, even with the HS 2-digit data, the proportion of zero trade flows is already about two-thirds, and it is sufficient to demonstrate the importance of accounting for a firm's entry and exit decisions concerning export markets.

Trade flows are obtained from the UNCOMTRADE database, and time-variant bilateral tariffs are obtained from the World Integrated Trade Solution. The sources of original tariff data are the UNCTAD Trade Analysis Information System (TRAINS) and the WTO's Integrated Data Base and Consolidated Tariff Schedules database. We use applied tariff data for entries with positive flows. However, for entries with zero trade flows, we use preferential tariffs from the UNCTAD TRAINS if available and Most Favoured Nation (MFN) tariffs for the rest. By using multiple sources of tariff data, we can maximise the number of observations.

Data on nominal GDP and GDP per capita are drawn from the World Bank's World Development Indicators (WDI). Regional Trade Agreement (RTA) data are obtained from the WTO's Re-

---

[11]We use data from year 2000 when generating the lagged terms of learning variables included in the first-stage estimations so that data coverage in the second-stage estimations would remain the same from year 2001 to 2010.

gional Trade Agreements Information System (RTA-IS). Data on gravity equation variables such as contiguity, distance, common language and common colony are sourced from the CEPII.

Table 1 shows that the proportion of zero trade flows in our sample is about 63%.

Table 1: Sample statistics: Positive and zero trade flows

	HS 2-digit, 2001-2010
Positive value only (imports)	2,576,869
Zero + positive value	6,924,419
Proportion of zero	62.79%

Table 2 presents basic statistics for a few key variables. Two indicator variables, free trade agreements (FTA) and customs union (CU) are used to capture trade liberalization of varying depth. The first two columns show the unconditional mean and standard deviation, while the last two columns show the mean and standard deviation of variables conditional on positive trade flows. The statistics indicate that country-pairs with positive flows tend to have lower tariff rates and are more likely to form trade agreements.

Table 2: Basic statistics: Trade flows, tariffs, and trade agreements

	Mean	SD	Mean|Flow>0	SD|Flow>0
Trade flows (imports)	9602.2	254759.9	29562.9	446351.0
Tariffs	0.076	0.128	0.069	0.118
Free trade agreements (FTA)	0.09	0.23	0.15	0.35
Customs union (CU)	0.05	0.23	0.13	0.34

Trade flows are in 1000 USD.

4 Estimations

4.1 Estimations with product-level data

Before we estimate using product-level data, we do so with country-level data to check that our dataset generates results similar to the previous literature. The results from first- and second-stage

estimations are reported in the appendix. The coefficient estimates for the trade elasticity from the log-linear and HMR models are not statistically different from each other, and they are also comparable to recent studies including Ossa (2015).

Because our dataset does not exhibit any abnormality in the country-level analysis, we can proceed to the product-level analysis. The proportion of zero trade flows in our HS 2-digit data is about 63%. Therefore, accounting for zero trade flows is crucial if the response at the extensive margin (i.e., from zero flows to positive flows) due to tariff change is different from that at the intensive margin (i.e., from one volume of positive flows to another volume of positive flows).

With product-level panel data, our main estimation equation for the pooled data (i.e., pooling over all sectors) is as follows:

$$ln(q_{ijkt}) = \beta_0 + \lambda_{ijt} + \mathbf{x}_{ijt}\delta + \beta_1 ln(1 + tariff_{ijkt}) + \rho_1 \hat{\bar{\eta}}^*_{ijkt} + \rho_2 \hat{\bar{z}}^*_{ijkt} + \rho_3 \hat{\bar{z}}^{*2}_{ijkt} + \rho_4 \hat{\bar{z}}^{*3}_{ijkt} + \rho_5 \hat{\bar{z}}^{*4}_{ijkt} + u_{ijkt}$$

(6)

where we use only positive trade flows, i.e. $q_{ijkt} > 0$; $\mathbf{x}_{ijt}$ includes gravity equation variables as in the previous section but is subsumed when we include country-pair-time FEs, λ_{ijt}; and $\hat{\bar{\eta}}^*_{ijkt}, \hat{\bar{z}}^*_{ijkt}, \hat{\bar{z}}^{*2}_{ijkt}, \hat{\bar{z}}^{*3}_{ijkt}$ and $\hat{\bar{z}}^{*4}_{ijkt}$ are terms to approximate $ln[exp(\delta(\hat{\bar{v}}^*_{ij} + \hat{z}^*_{ij})) - 1] + \rho_1 \hat{\bar{v}}^*_{ij}$ in eq. (4), and are obtained from the first-stage estimation of eq. (5) using the learning variables defined in Section 2.

In eq. (6), there is a distinct unobserved factor, λ_{ijt}. It accounts for not only country-level multilateral resistance terms (MRTs) and country-pair FEs, but also any unobserved country-pair-time varying heterogeneity. Thus, heterogeneity factors that vary over ijt are controlled for by FEs, but heterogeneity factors that vary over $ijkt$ are controlled for by the HMR terms, $\hat{\bar{z}}^*_{ijkt}, \hat{\bar{z}}^{*2}_{ijkt}, \hat{\bar{z}}^{*3}_{ijkt}$, and $\hat{\bar{z}}^{*4}_{ijkt}$.

Even though we use country-pair-time FEs to subsume country-pair FEs and country-time FEs as emphasized in the literature, the country-pair-time FEs may not completely account for unob-

served heterogeneity at the sectoral level. On one hand, when the tariff on a product is reduced, the trade effect on this product may be smaller if tariff reduction also occurs with substitutable products. On the other hand, if the product concerned is an intermediate good, the trade effect caused by a tariff change may also depend on the tariff changes in the associated final products, and vice versa. These two issues may cause bias in the trade elasticity estimations. To partly address these issues of substitutability across sectors and intersectoral linkages, we additionally control for product FEs in our product-level estimations.

Table 3 presents the results from the first-stage estimation. In column (1), current learning variables are used as ERs, and in column (2), their lagged values are used instead. The coefficients of all ERs in both estimations are statistically significant at the 1% level. The results confirm that the learning variables affect the probability of trade for specific products between countries.

The estimation results from the second-stage are reported in Table 4. Column (1) is the log-linear estimation, and columns (2) and (3) are the HMR approach with the learning variables as ERs. All the estimations include ijt FEs and k FEs. The trade elasticity estimate decreases from -2.6 in column (1) to -1.4 in columns (2) and (3). The estimates of the IMR and the first and second HMR terms are statistically significant at the 1% significance level, confirming that accounting for self-selection and firm heterogeneity are important in modeling product-level trade flows. Column (3) uses the lagged terms of learning variables as ERs, and it is not statistically different from its column (2) counterpart. It is also worth mentioning that the R^2 statistics are larger when we use the HMR method.

The results from columns (2) and (3) are not statistically different. However, we prefer the specification of column (3) because it could additionally avoid potential bias from reserve causality. Thus, for brevity hereafter we report the results based on lagged learning variables only.

Table 3: HS 2-digit data: 1st stage

	(1) Learning	(2) Lag. Learning
$Contiguity$	0.726***	0.718***
	(0.006)	(0.006)
$Language$	0.665***	0.645***
	(0.002)	(0.002)
$Common\_Colony$	0.189***	0.187***
	(0.004)	(0.004)
$lnDistance$	-0.459***	-0.454***
	(0.001)	(0.001)
$lnGDP_i$	0.184***	0.179***
	(0.000)	(0.000)
$lnGDP_j$	0.200***	0.199***
	(0.000)	(0.000)
$lnGDPPC_i$	0.038***	0.048***
	(0.001)	(0.001)
$lnGDPPC_j$	-0.008***	-0.011***
	(0.001)	(0.001)
FTA	0.047***	0.051***
	(0.002)	(0.002)
CU	0.030***	0.045***
	(0.004)	(0.004)
$lnN_i \cdot lnN_j$	0.201***	
	(0.000)	
$ln\bar{q}_i \cdot ln\bar{q}_j$	0.005***	
	(0.000)	
L. $lnN_i \cdot lnN_j$		0.206***
		(0.000)
L.$ln\bar{q}_i \cdot ln\bar{q}_j$		0.005***
		(0.000)
Num of Obs	6924419	6924419

Time fixed effects are used. Cluster (pair) robust standard errors are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: HS 2-digit data: 2nd stage

	(1) Log-linear	(2) HMR:Learning	(3) HMR: Lag. learning
$ln(1+Tariff)$	-2.643***	-1.447***	-1.407***
	(0.130)	(0.093)	(0.094)
$\hat{\bar{\eta}}_{ijkt}^{*}$		0.402***	0.605***
		(0.064)	(0.064)
$\hat{\bar{z}}_{ijkt}^{*}$		10.154***	11.049***
		(0.494)	(0.479)
$\hat{\bar{z}}_{ijkt}^{*2}$		-0.852***	-1.455***
		(0.299)	(0.287)
$\hat{\bar{z}}_{ijkt}^{*3}$		-0.049	0.096
		(0.076)	(0.072)
$\hat{\bar{z}}_{ijkt}^{*4}$		0.017**	0.005
		(0.007)	(0.006)
Fixed Effects	ijt , k	ijt , k	ijt , k
Num of Obs	2576869	2576869	2562036
R^2	0.541	0.660	0.659

Cluster (pair) robust standard errors are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

20

4.2 Sensitivity analysis: Alternative exclusion restrictions

In this section, we perform a sensitivity analysis by applying an alternative set of ERs. Instead of using only interaction terms of the foundation learning variables, which are $ijkt$-varying, we also use each of the four foundation learning variables as an additional ER.

Table 5 reports the results of the first-stage estimation. Column (1) uses the (lagged) interaction terms only and is therefore the same as column (2) of Table 3, while column (2) uses six ERs including the four foundation learning variables. Both sets of ERs are statistically significant at the 1% significance level. The second-stage estimation results are presented in Table 6. The qualitative results from both columns remain the same. Both coefficient estimates of the trade elasticity are not statistically different, while they are statistically different from the log-linear estimate presented in columns (1) of Table 4.

Table 5: Sensitivity test, HS 2-digit data: First-stage

	(1) Lag. learning: Interaction	(2) Lag. learning: Interaction and individual
$Contiguity$	0.718***	0.822***
	(0.006)	(0.006)
$Language$	0.645***	0.633***
	(0.002)	(0.002)
$Common\_Colony$	0.187***	0.236***
	(0.004)	(0.004)
$lnDistance$	-0.454***	-0.470***
	(0.001)	(0.001)
$lnGDP_i$	0.179***	0.175***
	(0.000)	(0.001)
$lnGDP_j$	0.199***	0.231***
	(0.000)	(0.001)
$lnGDPPC_i$	0.048***	0.033***
	(0.001)	(0.001)
$lnGDPPC_j$	-0.011***	-0.010***
	(0.001)	(0.001)
FTA	0.051***	0.073***
	(0.002)	(0.002)
CU	0.045***	0.088***
	(0.004)	(0.004)
L. $lnN_i \cdot lnN_j$	0.206***	-0.015***
	(0.000)	(0.002)
L.$ln\bar{q}_i \cdot ln\bar{q}_j$	0.005***	0.001***
	(0.000)	(0.000)
L. lnN_i		1.029***
		(0.005)
L. lnN_j		0.697***
		(0.006)
L. $ln\bar{q}_i$		-0.035***
		(0.001)
L. $ln\bar{q}_j$		0.074***
		(0.001)
Num of Obs	6924419	6924419

Time fixed effects are used. Cluster (pair) robust standard errors are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Sensitivity Analyses, HS2-digit data: 2nd stage

	(1) Lag. learning: Interaction	(2) Lag. learning: Interaction and individual
$ln(1 + Tariff)$	-1.407***	-1.364***
	(0.094)	(0.091)
$\hat{\bar{\eta}}^*_{ijkt}$	0.605***	0.213***
	(0.064)	(0.068)
$\hat{\bar{z}}^*_{ijkt}$	11.049***	8.606***
	(0.479)	(0.468)
$\hat{\bar{z}}^{*2}_{ijkt}$	-1.455***	0.028
	(0.287)	(0.285)
$\hat{\bar{z}}^{*3}_{ijkt}$	0.096	-0.264***
	(0.072)	(0.073)
$\hat{\bar{z}}^{*4}_{ijkt}$	0.005	0.035***
	(0.006)	(0.006)
Fixed Effects	ijt , k	ijt , k
Num of Obs	2562036	2562036
R^2	0.659	0.664

Cluster (pair) robust standard errors are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.3 Robustness tests

To determine whether our results remain robust with different sub-samples, we first divide our sample into four country-pair groups based on the OECD membership in year 2010. In doing so, we also investigate whether the trade elasticity is homogenous across developing and industrial countries.

Table 7 shows the second-stage estimation results. The upper panel shows the results from the log-linear estimations and the lower panel from the HMR approach with learning. As for the group labels, the first term denotes the income category of the importing country, and the second term denotes that of the exporting country. For example, Ind-Dev in column (3) indicates that the importing country is an industrial country and the exporting country is a developing country. The proportion of zero trade flows is substantial in all four groups, but it also varies greatly across groups. It is highest between developing countries and lowest between industrial countries.

As found in the previous section, the trade elasticity is overestimated (in absolute terms) with the log-linear model for all groups. According to the results from the HMR approach, the trade elasticity point estimate is smaller when the importing country is an industrial country as in columns (1) and (3), and larger when it is a developing country as in columns (2) and (4). The smaller trade elasticities of the industrial importing countries imply that trade incurs the larger welfare gains from trade for industrial countries than for developing countries.

A comparison of the HMR point estimates in columns (1) and (3) also suggests that industrial countries are less sensitive to tariff-induced price changes for the products from industrial countries. A possible explanation is that, because tariffs in industrial countries are already low in our sample period, consumers in industrial countries may care more about product quality than small price changes.

As another robustness check, we estimate the trade elasticity for each of our sample years. In these year-by-year estimations, we use ij FEs and k FEs. The results are presented in Figure 1 and

the details are provided in the appendix. Again, the coefficient estimate from the HMR approach is smaller than that from the log-linear estimator for every year, and each of the differences is statistically significant.

Figure 1 also shows that the trade elasticity was around -1.2 between years 2001 and 2006 with small fluctuations, but it soared during the global financial crisis around 2008 and started to reverse after the crisis. This suggests that consumers are more sensitive to import price during an economic downturn.

In the last robustness test, we conduct section-by-section estimations. In principle we can estimate the trade elasticity for all 97 HS 2-digit sectors, which are known as 'chapters'. However, due to the burden of presenting 97 sets of results, we opt to use the category of 'section' instead. For the HS data, products are categorized into 20 sections, meaning that one section may include multiple chapters. A simple description of the 20 sections is provided in the appendix.

Figure 2 reports the results.[12] All sectional estimations yield negative trade elasticities, but there is large heterogeneity across the sections. The estimates from the log-linear range from -0.31 to -8.61, and those from the HMR method range from -0.24 to -5.45. Again, the key qualitative result remains the same that the point estimates from the HMR approach are smaller than those from the log-linear for most sections.

---

[12]Because three sections (Sections III, XIV, and XIX) include only one chapter, which represents "product k" in our estimations, we do not obtain estimates for them after accounting for k FEs. We also do not use it FEs in these estimations due to multicollinearity. The detailed explanation is provided in the appendix.

Table 7: Heterogeneity by income between country-pairs: HS 2-digit, second-stage

Log-liner	(1) Ind-Ind	(2) Dev-Ind	(3) Ind-Dev	(4) Dev-Dev
$ln(1+Tariff)$	-1.243***	-2.851***	-1.361***	-2.077***
	(0.244)	(0.240)	(0.293)	(0.217)
R^2	0.596	0.575	0.488	0.471
HMR: Lag. learning				
$ln(1+Tariff)$	-0.521**	-1.675***	-0.826***	-1.473***
	(0.206)	(0.164)	(0.224)	(0.164)
R^2	0.690	0.648	0.634	0.571
Portion of Zeros	22.95%	81.85%	72.14%	47.13%
Fixed Effects	ijt, k	ijt, k	ijt, k	ijt, k
Num of Obs	533233	489489	836344	537122

The first term denotes the income category of the importing country and the second term that of the exporting country.

Cluster (pair) robust standard errors are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

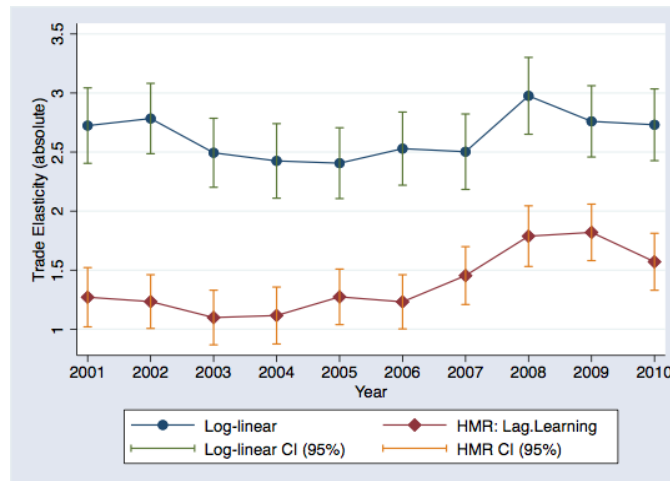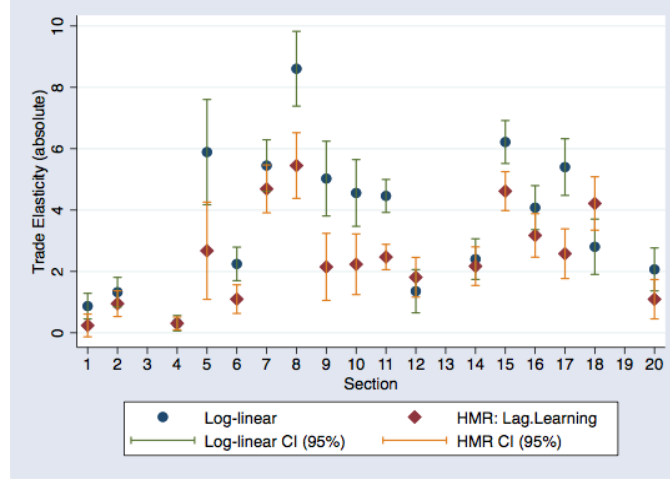Figure 1: Trade elasticity by years (abs.)

Figure 2: Trade elasticity by sections (abs.)

5 Welfare Implication

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 computed as $(1 - \lambda^{-1/\phi})$, where λ is the share of expenditure on domestic products and ϕ is the elasticity of imports with respect to variable trade costs. According to Anderson and van Wincoop (2004), the trade elasticity estimates in the previous literature range from -5 to -10. More recent studies using disaggregate data such as Ossa (2015) and Simonovska and Waugh (2014a) find the trade elasticity to be close to -4, slightly below the lower end of those in the previous literature.

In Table 8, we provide rough estimations of the welfare gains from trade for the U.S. in year 2000 following Arkolakis et al. (2012). In year 2000, the share of expenditure devoted to domestic products for the U.S. was 0.93. Using this value for λ, Arkolakis et al. (2012) illustrate that the percentage change in real income needed to compensate a representative consumer in the U.S. for going back to autarky is between 0.7% and 1.8%, based on the trade elasticity estimates found in the literature.

When we use the trade elasticity estimates obtained from our log-linear estimation for the HS 2-digit data as in column (1) of Table 4, the U.S.'s gains from trade increase to 2.8%. However, when we use our HMR estimation from column (3) of Table 4, the estimated welfare gains from trade increase much further to 5.1%, nearly three times the upper-end estimate in the past. This result may help resolve the puzzle of small estimates for gains from international trade.

It should be pointed out that in order to compute the total welfare gains with multiple sectors, we need additional data on the share of domestic expenditure, the share of consumption, and the share of employment for each sector, as well as the sectoral trade elasticity (see section 5.1 in Arkolakis et al. (2012)). To meet the main objective of this paper, we, however, focus on the average trade elasticity across sectors. The simplified numerical example in Table 8 serves to bring home the message that, to evaluate the welfare impact of trade liberalization, it is paramount to have an unbiased estimate of the trade elasticity. Including zero trade flows and accounting for self-selection and firm heterogeneity reduce the trade elasticity estimate substantially and therefore lead to a much larger estimate of welfare gains from trade.

Table 8: Welfare calculation: Gains from trade

Source	Trade elasticity	Gains from trade (2000 vs. autarky)
Literature lower end	-10	0.74%
Literature upper end	-4	1.83%
HS 2-digit with log-linear	-2.6	2.8%
HS 2-digit with HMR	-1.4	5.1%

Note: In year 2000, λ was 0.93 for the U.S..

6 Conclusion

In this paper, we estimate the trade elasticity using the product-level data for 64 importing and 135 exporting countries from 2001 to 2010. We adopt the HMR approach to account for two sources of biases–self-selection and firm heterogeneity–in the presence of a large portion of zero

bilateral trade flow observations. To implement the HMR approach at the sector level, we introduce new ERs that are constructed using trade data that vary over country-pair-time-product to extract information from export learning as in Fernandes and Tang (2014) and other search and learning literature. Importantly, the new method opens up the possibility of using the influential HMR approach to study various trade topics using disaggregate data.

We find upward bias in the estimates of the trade elasticity if only the positive trade flows are used. Properly accounting for self-selection and firm heterogeneity at the product level with large proportion of zero trade flows yields substantially smaller estimates of the trade elasticity (the magnitude decrease from -2.6 to -1.4), which implies much larger welfare gains from trade.

The previous literature usually finds the trade elasticity estimates ranging from -4 to -10. Based on these figures, the welfare gains from trade for the U.S. were estimated to be between 0.7% and 1.8% of real income. Given the importance of trade to modern economies, these estimates are puzzlingly small. On the contrary, our estimate of welfare gains for the U.S. is several times larger at over 5% of real income.

We also investigate the heterogeneity of the trade elasticity across country-pair groups, sectors, and time. Our estimates show that the bias from ignoring zero trade flows is substantial for every country-pair group regardless of whether the importer and exporter are a developing or an industrial country. However, our results also suggest that the trade elasticity is smaller for industrial countries, implying larger welfare gains from trade for industrial countries as compared to developing countries. Last, we find that trade elasticity increases (in absolute terms) during severe economic downturns like the global financial crisis.

References

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

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

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

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

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.

Baier, S. L., Bergstrand, J. H., Feng, M., 2014. Economic integration agreements and the margins of international trade. Journal of International Economics 93 (2), 339–350.

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

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

Caliendo, L., Parro, F., 2014. Estimates of the trade and welfare effects of nafta. The Review of Economic Studies, rdu035.

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

Cheong, J., Kwak, D. W., Tang, K. K., 05 2015. Can Trade Agreements Curtail Trade Creation and Prevent Trade Diversion? Review of International Economics 23 (2), 221–238.

Djankov, S., La Porta, R., Lopez-de Silanes, F., Shleifer, A., 2002. The regulation of entry. The quarterly Journal of economics 117 (1), 1–37.

Dutt, P., Mihov, I., Van Zandt, T., 2013. The effect of WTO on the extensive and the intensive margins of trade. Journal of International Economics 91 (2), 204–219.

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

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

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

Feenstra, R. C., Luck, P., Obstfeld, M., Russ, K. N., 2018. In search of the armington elasticity. Review of Economics and Statistics 100 (1), 135–150.

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

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

Heckman, J. J., 1979. Sample selection bias as a specification error. Econometrica 47 (1), 153–161.

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

Holloway, I. R., 2017. Learning via sequential market entry: Evidence from international release of u.s. movies. Journal of International Economics 104 (1), 104–121.

Imbs, J., Mejean, I., July 2015. Elasticity Optimism. American Economic Journal: Macroeconomics 7 (3), 43–83.

Kee, H. L., Nicita, A., Olarreaga, M., 2008. Import demand elasticities and trade distortions. The Review of Economics and Statistics 90 (4), 666–682.

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

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

Kugler, M., Verhoogen, E., 2012. Prices, plant size, and product quality. The Review of Economic Studies 79 (1), 307.

Manova, K., Zhang, Z., 2012. Export prices across firms and destinations. The Quarterly Journal of Economics 127 (1), 379–436.

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

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

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

Melitz, M. J., Redding, S. J., 2015. New trade models, new welfare implications. The American Economic Review 105 (3), 1105–1146.

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

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

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

Simonovska, I., Waugh, M. E., Sep. 2014b. Trade Models, Trade Elasticities, and the Gains from Trade. NBER Working Papers 20495, National Bureau of Economic Research, Inc.

Soderbery, A., 2018. Trade elasticities, heterogeneity, and optimal tariffs. Journal of International Economics 114, 44–62.

Appendix

Tables A1 and A2 report, respectively, the first- and second-stage estimation results using country-level data. Regarding the ER for country-level data, HMR first consider the bilateral regulation costs. However, besides a strong assumption of excludability in the model for trade volume, data availability is a problem. As a result, HMR resort to an index of bilateral religion proximity, which is more readily available, as a proxy for fixed trade costs. Following HMR, to construct an ER for the aggregate data, we focus on the four most popular religions–Christianity, Judaism, Islam, and Buddhism–using the data obtained from Maoz and Henderson (2013). Because the religion data are available only for every 5 years, we use the data from the closest previous year for a year where data are unavailable. For example, we use year 2000 data for years 2001–2004. In addition, we use ERs obtained from the interaction of the foundation learning variables at the country-level.

For the first-stage results reported in Table A1, column (1) uses the religion proximity, and columns (2) and (3) use, respectively, the current and lagged interaction terms of the foundation learning variables. For the second-stage results reported in Table A2, column (1) uses the log-linear estimator, and columns (2), (3), and (4) use the HMR method with the religion proximity and learning variable ERs, respectively. In all four estimations we control for country-pair and exporter-year FEs but not importer-year FEs. This is because most changes of bilateral tariffs at the country-level come from changes in MFN tariffs, which technically are set unilaterally by importing countries. Thus, the tariff variable loses most of its explanatory power when importer-year FEs are accounted for. The trade elasticity estimates of the four models are not statistically different from each other and are around -3.5 to -3.7, similar to those from the recent literature.

Table A1: Aggregate data: first-stage

	(1) Religion	(2) Learning	(3) Lag. learning
$Contiguity$	-0.179	0.530*	-0.063
	(0.127)	(0.272)	(0.113)
$Langauge$	0.377***	0.444***	0.390***
	(0.030)	(0.032)	(0.027)
$Common\_Colony$	-0.028	-0.011	0.037
	(0.033)	(0.034)	(0.030)
$lnDistance$	-0.371***	-0.475***	-0.311***
	(0.016)	(0.017)	(0.014)
$lnGDP_i$	0.318***	0.300***	0.331***
	(0.006)	(0.008)	(0.006)
$lnGDP_j$	0.391***	0.197***	0.329***
	(0.006)	(0.010)	(0.007)
$lnGDPPC_i$	0.173***	0.062***	0.092***
	(0.007)	(0.007)	(0.006)
$lnGDPPC_j$	0.017***	0.024***	0.025***
	(0.006)	(0.006)	(0.006)
FTA	0.436***	0.333***	0.317***
	(0.048)	(0.053)	(0.042)
CU	1.012***	0.894***	0.791***
	(0.183)	(0.221)	(0.130)
$Religion$	0.119***		
	(0.028)		
$lnN_i \cdot lnN_j$		0.709***	
		(0.012)	
$ln\overline{q}_i \cdot ln\overline{q}_j$		0.001	
		(0.001)	
$L.lnN_i \cdot lnN_j$			0.241***
			(0.005)
$L.ln\overline{q}_i \cdot ln\overline{q}_j$			-0.003***
			(0.001)
Num of Obs	69330	84420	84420

Time fixed effects are used. Cluster (pair) robust standard errors are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Aggregate data: 2nd stage

	(1) Log-linear	(2) Religion	(3) Learning	(4) Lag. learning
$ln(1 + Tariff)$	-3.738***	-3.690***	-3.838***	-3.672***
	(0.536)	(0.538)	(0.535)	(0.534)
FTA	-0.081*	-3.006	-0.273***	-0.082
	(0.048)	(2.496)	(0.056)	(0.053)
CU	-0.084	-6.858	-0.512***	-0.063
	(0.068)	(5.789)	(0.095)	(0.082)
$\hat{\hat{\eta}}^*_{ijt}$		-4.701	1.048	-0.108
		(5.877)	(0.644)	(0.306)
$\hat{\hat{z}}^*_{ijt}$		12.096*	2.157	-0.212
		(6.200)	(1.414)	(0.452)
$\hat{\hat{z}}^{*2}_{ijt}$		-2.292**	-0.425	0.133
		(1.088)	(0.566)	(0.275)
$\hat{\hat{z}}^{*3}_{ijt}$		0.418**	0.044	-0.036
		(0.195)	(0.096)	(0.064)
$\hat{\hat{z}}^{*4}_{ijt}$		-0.028**	-0.002	0.003
		(0.013)	(0.006)	(0.005)
Fixed Effects	ij, it, jt	ij, it, jt	ij, it, jt	ij, it, jt
Num of Obs	70205	63601	63601	63601
R^2	0.928	0.930	0.930	0.930

Cluster (pair) robust standard errors are reported in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: Heterogeneity by years: HS 2-digit, second-stage

Log-liner	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
$ln(1+Tariff)$	-2.724***	-2.784***	-2.494***	-2.425***	-2.406***	-2.529***	-2.503***	-2.976***	-2.760***	-2.731***
	(0.163)	(0.152)	(0.149)	(0.161)	(0.153)	(0.158)	(0.163)	(0.166)	(0.154)	(0.155)
R^2	0.527	0.534	0.536	0.541	0.544	0.544	0.548	0.549	0.544	0.546
HMR: Lag.learning										
$ln(1+Tariff)$	-1.271***	-1.234***	-1.099***	-1.116***	-1.274***	-1.232***	-1.454***	-1.788***	-1.820***	-1.571***
	(0.128)	(0.116)	(0.118)	(0.123)	(0.120)	(0.117)	(0.125)	(0.131)	(0.122)	(0.123)
R^2	0.655	0.659	0.661	0.663	0.661	0.657	0.661	0.661	0.658	0.660
Portion of Zeros	69.66%	68.88%	68.56%	67.31%	66.50%	66.19%	66.66%	66.74%	66.02%	65.97%
Fixed Effects	ij,k	ij,k	ij,k	ij,k	ij,k	ij,k	ij,k	ij,k	ij,k	ij,k
Num of Obs	236503	242373	245086	254883	263650	266232	262466	262232	260046	268565

Table A4: HS classification by section

HS sections	HS 2-digit codes	Simple descriptions
Sections I	01-05	Animal Products
Sections II	06-14	Vegetable Products
Sections III	15	Animal or Vegetable Fats and Oils
Section IV	16-24	Prepared Foodstuffs
Section V	25-27	Mineral Products
Sections VI	28-38	Chemical Products
Sections VII	39-40	Plastic, Rubber
Sections VIII	41-43	Raw Hides, Skins, Leather
Sections IX	44-46	Wood
Sections X	47-49	Pulp of Wood, Paper
Section XI	50-63	Textiles
Sections XII	64-67	Footwear
Sections XIII	68-70	Plaster, Glass
Sections XIV	71	Pearls, Precious Stones
Sections XV	72-83	Base Metals
Section XVI	84-85	Machinery, Appliances
Section XVII	86-89	Vehicles
Section XVIII	90-92	Optical, Watches
Section XIX	93	Arms
Section XX	94-96	Furniture, Toys