Journal of International Economics 114 (2018) 44–62

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

Trade elasticities, heterogeneity, and optimal tariffs☆

Anson Soderbery

Purdue University, United States

ARTICLE INFO

Article history:
Received 19 April 2017
Received in revised form 13 April 2018
Accepted 26 April 2018
Available online 4 May 2018

Research data related to this submission: http://web.ics.purdue.edu/~asoderbe/Site/Elasticities.html

JEL classification:
F12
F14
F59
C13

Keywords:
Export supply
Import demand
Trade
Structural estimation
Optimal tariffs

ABSTRACT

We develop a structural estimator for heterogeneous supply and demand in the absence of instrumental variables. Using only readily available bilateral trade data we show how to leverage variation in prices and quantities across multiple markets in order to consistently estimate heterogeneous elasticities. Our elasticity estimates follow intuitive patterns of importer and exporter market power and produce believable distributions and magnitudes. To highlight the flexibility of the estimator, we extend the cornerstone theories of non-cooperative optimal trade policy to a setting where exporters have heterogeneous supply elasticities. Applying our estimates to trade and tariff data worldwide, we show that heterogeneous export supply elasticities provide new avenues for identification. We demonstrate statistically positive and persistent links between non-cooperative optimal tariffs and applied tariffs worldwide.

1. Introduction

At the root of empirical analysis of product markets are estimates of supply and demand elasticities. Here we develop a structural estimator of demand and supply that does not rely on instrumental variables and can identify variety by market specific heterogeneity in the elasticities. The estimator yields consistent estimates by leveraging variation in prices and quantities over time when products are sold in multiple markets.

These innovations are particularly important in the context of international trade where empirical analysis has a keen focus on the trade elasticities underlying our theoretical models. Trade economists generally observe many heterogeneous exporters selling varieties of many goods to multiple importers. The breadth of trade data thus implies estimating millions of elasticities across thousands of product markets worldwide. Consequently, constructing believable instruments for each market is a significant obstacle.

International trade, however, is particularly salient for the identification strategy of our estimator, as even publicly available data record countries exporting to and importing from many sources. We will show how price and quantity variation over time for the same good across export and import markets can be exploited to identify importer by exporter by product elasticities (i.e., heterogeneous elasticities). We thus develop a tractable model of international trade that follows the common assumptions from new trade theory.[1] Structurally estimating the model yields the first consistent estimates

☆ This paper has benefitted immensely from discussions with Bruce Blonigen, Matilde Bombardini, Rob Feenstra, Colin Hottman, David Hummels, Nuno Limão, Nick Sly, Alan Spearot, Bob Staiger and David Weinstein. I am additionally grateful to the seminar participants at Columbia University, Vanderbilt University, University of Oregon, Clemson University, McMaster University, Purdue University, Federal Reserve Bank of Philadelphia, Federal Reserve Bank of Dallas, Stanford Institute for Theoretical Economics and Empirical Investigations of International Trade and two anonymous referees. All errors and omissions are my own.

E-mail address: asoderbe@purdue.edu.

[1] Specifically, demand will follow from constant elasticity of substitution preferences and exporters will have heterogeneous upward sloping supply curves to every destination market.

https://doi.org/10.1016/j.jinteco.2018.04.008
0022-1996/

of heterogeneous export supply and import demand elasticities for every country pair and good traded worldwide.[2]

The key for the structural estimator is defining multiple markets with heteroskedastic differences in a time series of prices and quantities. While trade provides the perfect environment for defining multiple markets for a product (i.e., export and import flows), the methodology developed here could be more generally applied to any industry where the goal is to consistently estimate supply and demand elasticities simultaneously in the absence of believable instruments.

Applying the structural estimator to the universe of trade flows at the HS4 level from (publicly available) Comtrade data uncovers heterogeneity in export supply elasticities even at the most disaggregate level – across imported varieties of a particular good. Specifically, importer-product fixed effects explain 72% of the variation in our estimates. The remaining variation is explained by differences in export supply elasticities across exporters of particular imported goods. Evaluating the "reasonableness" of our elasticity estimates exploits the heterogeneity of the estimates and the interpretation of the inverse export supply elasticity as a measure of importer market power. We first establish intuitive relationships between product differentiation and market power. Importers are expected to possess more market power (i.e., large inverse export supply elasticities) for highly differentiated goods on average.[3] Our estimates confirm the intuition, and yield inverse export supply elasticities for differentiated goods that are around three times larger than for homogeneous goods, on average. Additionally, we expect differentiated goods to be less substitutable than homogeneous goods. We confirm this expectation as our estimated import demand elasticities are on average eight percent larger for homogeneous than differentiated goods.

Importers could also have varying degrees of market power across types of goods depending on the composition of their imports. To investigate, we regress our inverse export supply elasticities on importer by product differentiation fixed effects. The estimated fixed effects allow us to rank the median market power across importer good pairs. As we might expect, the largest importers in the data (Germany, US and China) possess the highest degree of market power across the goods they import, while the smaller less developed importers (Brazil, Caribbean and African Countries) have the least market power.

Notably, these market power rankings vary across types of goods. China for instance ranks first in market power across differentiated goods, but near last for non-differentiated goods. China is notably reliant on distant foreign suppliers of raw materials, supporting the result that they possess relatively weak market power over non-differentiated goods. Conversely, India ranks near last in differentiated goods, but third in market power for non-differentiated goods. India's imports of non-differentiated goods are predominantly supplied by regional trade partners, of which India is an important export destination. Our estimates thus support the idea that market power varies across types of goods and the composition of imports. These rich patterns suggest that heterogeneity in trade elasticities is important for understanding patterns of market power, and by association, the motives for tariffs set by importers.

We demonstrate the importance of our estimates by adapting the model to advance the optimal non-cooperative trade policy literature to allow for exporter supply heterogeneity. To be clear, we model importers setting tariffs taking into account exporter heterogeneity to maximize welfare without internalizing potential foreign retaliation. A large body of theoretical and empirical work argues that WTO member countries should not be expected to set non-cooperative optimal tariffs given the cooperative nature of the WTO. Ossa (2014) for instance contrasts the tariffs predicted by non-cooperative and cooperative optimal policies by simulating his model given moments in the data. Optimal tariffs here are derived in a non-cooperative setting, and estimated to be of similar magnitude to Ossa (2014)'s non-cooperative tariff estimates.

Our data cover all available trade flows and tariffs from 1991 to 2007. This period in trade policy highlights the difficulties of estimating a relationship between non-cooperative tariff motives and applied tariffs globally. From 1991 to 2007, according to the *CEPII*, the world saw 36 countries sign the GATT, 160 regional trade agreements spanning 1396 importer-exporter pairs implemented, and substantial reductions in applied tariffs. Additionally, WTO members during this period commonly restructured bound tariff rates to below the current applied tariffs suggesting that cooperative policy negotiations were becoming more and more binding.

Anecdotally these trends suggest a diminishing role for non-cooperative tariff motives as targeting terms of trade gains gives way to cooperative incentives. Our analysis is designed to demonstrate that even amidst growing cooperation between countries, applied tariffs still respond to non-cooperative tariff motives in the data. We argue that our estimates and methodology are uniquely able to identify terms of trade motives underlying applied tariffs. We employ our estimates to evaluate the model's efficacy in explaining applied tariffs worldwide from 1991 to 2007. Our estimates are shown to provide new insights into how trade policy is set in recent years.

Previous studies of optimal non-cooperative trade policy have assumed that every exporter of a particular good into a given destination exports under an *identical* supply elasticity. This assumption implies the optimal tariff that balances terms of trade gains and efficiency losses is the inverse of this common export supply elasticity. However, the heterogeneity uncovered by our estimates calls into question these simplifying assumptions. Intuitively, we are relaxing the assumption that, for instance, exported automobiles from Sweden to the US have the same export supply elasticity as automobiles originating from Japan. Our estimated export supply elasticities capture such things as differences in the production processes associated with these two varieties and the degree of market power that the US has for each variety. Both of these channels translate into differences in export supply elasticities for importer exporter pairs both within and across traded goods, which we demonstrate carries significant implications for our understanding of optimal non-cooperative trade policy.

With heterogeneity in export supply elasticities, we show that the optimal non-cooperative tariff set by an importer is no longer the inverse of a single elasticity. When varieties of a good are exported from countries with heterogeneous export supply elasticities, we demonstrate that the optimal non-cooperative tariff weights the relative contribution of each variety to terms of trade gains and efficiency losses resulting from the tariff. In essence, when the importer applies an identical tariff across multiple exporters with different export supply elasticities, each exporter yields a different terms of trade gain relative to its efficiency loss. The optimal tariff is therefore one that optimally weights each exporter's contribution to its total terms of trade gains and efficiency losses.

We employ our model and estimates to compare theoretically optimal non-cooperative trade policy under heterogeneity with

[2] There is a small but growing literature related methodologically (cf., Feenstra and Romalis, 2014; Feenstra and Weinstein, 2017 and Hottman et al., 2016). Additionally, there is a larger literature related topically on estimating trade elasticities (cf. Bas et al., 2017; Fernandes et al., 2015; Head et al., 2014; Tan, 2012). This literature is reviewed in more detail in Section 3, but it is worth noting that none of the methods can be applied to estimate heterogeneous export supply and import demand elasticities without believable instruments.

[3] Intuitively, if US demand for say German microscopes falls and Germany decreases prices globally, there will be relatively weak substitution by other countries toward German microscopes due to the differentiated nature of this industry. The equilibrium world price for German microscopes may fall substantially in response to the shock in the US market, which is our definition of importer market power.

applied tariffs in the data. Now that the terms of trade (i.e., importer market power) motives for optimal tariffs under heterogeneity theoretically depend on the composition of exporters that make up an importer's trade, exporter heterogeneity introduces a new channel whereby optimal tariffs respond to compositional changes in trade over time. This new source of variation in optimal tariffs is used to demonstrate a positive significant relationship between applied and optimal tariffs in modern trade data.

Notably, we are able to reconcile key features of the data that have been at odds with empirical analyses of optimal non-cooperative trade policy. For example, developed countries tend to import varieties with high estimated inverse export supply elasticities (i.e., have strong market power) yet set low tariffs. Our introduction of exporter heterogeneity yields optimal tariffs that depend not only on estimates of export supply elasticities but also on the composition of a country's imports.

In the data, our estimates do in fact suggest high degrees of market power but low predicted optimal tariffs for developed countries. It is important to be clear here that this is not to say we believe that countries do not set tariffs cooperatively in the data. Rather, we will demonstrate that across all importers, including countries that are almost surely setting tariffs cooperatively, optimal non-cooperative tariffs positively relate to applied tariffs in the data. Moreover, this relationship is borne out by our introduction of heterogeneity in export supply elasticities. Our results thus differ from studies such as Ludema and Mayda (2013) and Beshkar et al. (2015), who argue that similar patterns pertaining to importer market power do exist for WTO members, as we establish this relationship without appealing to a particular bargaining structure between importers and exporters.

Identifying importers in seemingly cooperative settings (e.g., WTO members) applying tariffs that target terms of trade motives is established by allowing for heterogeneity in export supply. Since importers' terms of trade motives depend on the composition of trade when facing heterogeneous export supply in our model, time series variation in tariffs can be exploited for identification. Even as countries join the WTO, they restructure tariffs in ways that still target terms of trade motives as the makeup of their trade relationships adjust with the new policies. For instance, our optimal tariff estimates are applied to subsets of countries that join the WTO during our sample. Using our new identification strategy to relate applied and optimal tariffs, we demonstrate that there is a positive relationship between applied tariffs and optimal tariffs post accession. However, this relationship weakens with WTO membership as in Bagwell and Staiger (2011). Specifically, our estimates suggest that before WTO accession a one standard deviation increase in our optimal tariff over time (i.e., within an importer-exporter-product) for the median importer corresponds to a 0.57 percentage point larger applied tariff. Post WTO accession, the same increase in the optimal tariff implies a 0.18 percentage point larger applied tariff. Finally, in the appendix we confirm that our estimates are robust to a number of theoretical and empirical extensions to non-cooperative optimal tariff theory including analyzing bound tariff rates and lobbying.

We proceed as follows. Section 2 develops our quantitative model of trade that we structurally estimate in Section 3. Section 4 evaluates the patterns of the resulting estimates of heterogeneous supply and demand elasticities. Section 5.1 adapts the prevailing model of optimal non-cooperative trade policy to allow for heterogeneity. Section 5.2 evaluates the relationship between the model's optimal tariffs and applied tariffs in the data. Section 6 concludes.

2. Estimation

The estimator developed here is tasked with identifying pairwise trade elasticities for thousands of goods traded by hundreds of countries. This amounts to estimating millions of elasticities without a single reliable instrument. Our predecessors are Feenstra (1994) and Broda et al. (2008), which rely on time series variation in prices and market shares of imported varieties of goods to identify trade elasticities. Given that their focus is solely on import data from various countries, ensuring identification requires the assumption that the import demand and export supply elasticities are identical across all imported varieties of a good. The estimator developed here will also leverage time series variation in prices and quantities, but will demonstrate that the Feenstra (1994) estimator can no longer identify import demand and export supply with heterogeneity. We show how to overcome these issues by combining time series variation from multiple markets for the same traded good.

Soderbery (2015) argues that homogeneous import demand elasticities are supported by trade data, but imposing a common export supply elasticity is not. This is intuitive to us as the import demand elasticity generated from trade flow data, which records country level quantities and average prices, is plausibly characterized by constant elasticity substitution patterns across these country level aggregates. Conversely export supply elasticities characterize many tradeoffs within the production and trade of the exported variety that result in heterogeneity across countries.[4]

After writing down a parsimonious model of trade with exporter heterogeneity, we highlight the identification issues introduced by allowing for heterogeneous elasticities. We then show how to achieve identification by exploiting the structural differences between import and export markets for varieties of traded goods. Here we argue that the time series variation in market shares *across* export destinations from a given origin differs from the variation in shares *within* an import market. We show that since these data series are generated from the same underlying supply process, their statistical differences can be used to identify unique pairwise export supply elasticities.

2.1. A quantitative model of international trade

We start by specifying a quantitative model of trade that incorporates key assumptions from new trade theory. The model is flexible enough to accommodate structural estimation, yet parsimonious enough to apply to a wide range of applications. In general, we are modeling a world with many importers and exporters trading a host of goods and varieties. Denote goods imported by market I as $g \in G^I$. In the following empirics, we define varieties of imported goods utilizing the Armington (1969) assumption. Explicitly, goods (g) will be defined by their HS4 product category, and varieties of these goods consumed by a given importer (i) will be determined by the origin of the exporter (j).

A representative consumer in each country maximizes her utility by choosing imports and domestic consumption. Following the standard in the literature, consumers aggregate over the composite domestic (D) and imported (X) goods. The subutility derived from the composite imported good will be given by a CES aggregation across imported varieties with a good-importer specific elasticity of substitution given by σ_g^I, where I denotes the import market. Imports of a particular variety of good g by country i are denoted by x_{ijg}^I. Import

[4] This is not to say that the implications of heterogeneous import demand elasticities are not of potential importance. However, constant elasticity of substitution (CES) preferences still dominate the international trade literature. We thus opt to focus on extending our quantitative model of trade and developing an estimator that accounts for heterogeneous export supply elasticities while maintaining the homogeneous CES import demand elasticity. Notably, the estimation methodology developed here could be applied to a setting with heterogeneous import demand elasticities as well. The costs associated with doing so are discussed in Section 3. Ultimately, it should be noted that the implications of modifying the demand system on our export supply elasticity estimates are not predictable ex ante. Our estimates are thus specific to the model of CES demand and export supply heterogeneity presented here.

demand is also augmented by a variety specific taste parameter b^{I}_{ijg}. To focus our analysis, we will assume that consumers in I purchase a numeraire, c^{I}_{0}, and aggregate their consumption through Cobb-Douglas.[5] Under these assumptions, we write the utility obtained by a consumer in any importing country as,

$$U^{I} = c^{I}_{0} + \xi^{I}_{X} \sum_{g \in G^{I}} \phi^{I}_{g} log\left(\left(\sum_{j \in J^{I}_{g}} \left(b^{I}_{ijg}\right)^{\frac{1}{\sigma^{I}_{g}}} \left(x^{I}_{ijg}\right)^{\frac{\sigma^{I}_{g}-1}{\sigma^{I}_{g}}}\right)^{\frac{\sigma^{I}_{g}}{\sigma^{I}_{g}}-1}\right) + \xi^{I}_{D} log(D^{I}). \quad (1)$$

The importer consumes fixed shares ξ^{I}_{D} and ξ^{I}_{X} of domestic and imported goods, respectively. Additionally, the imported composite is formed by consumption of fixed shares of each imported good given by Cobb-Douglas parameters ϕ^{I}_{g}. The separability of utility allows us to focus on prices and consumption of imported goods in the estimation.[6] This specification also implies that trade policy affects goods independently, which later will allow us to evaluate the efficacy of tariffs through their impact on consumption losses and terms of trade gains good by good.

Consumers maximize Eq. (1) subject to their budget constraint, which yields demand $\left(x^{I}_{ijg}\right)$ for any variety (v) of an imported good (g) as a function of its price $\left(p^{I}_{ijg}\right)$,

$$x^{I}_{ijg} = \xi^{I}_{X} \phi^{I}_{g} b^{I}_{ijg} \left(p^{I}_{ijg}\right)^{-\sigma^{I}_{g}} \left(\mathcal{P}^{I}_{g}\right)^{\sigma^{I}_{g}-1}, \quad (2)$$

where $\mathcal{P}^{I}_{g} = \left(\sum_{j} b^{I}_{ijg} \left(p^{I}_{ijg}\right)^{1-\sigma^{I}_{g}}\right)^{\frac{1}{1-\sigma^{I}_{g}}}$ is the standard CES price index. All prices thus far are delivered prices in the importing country.

Given the structure of demand, we now need to specify exporter supply. We want export supply to be tractable enough to estimate, yet general and flexible enough to apply to many studies. To do so, we assume that exporter supply curves are upward sloping of the form,

$$p^{I}_{ijg} = exp\left(\eta^{I}_{ijg}\right) \left(x^{I}_{ijg}\right)^{\omega^{I}_{ijg}}. \quad (3)$$

An upward sloping constant elasticity export supply curve of this nature was championed by Feenstra (1994), and has become standard with Broda and Weinstein (2006) and Broda et al. (2008) for structurally estimating import demand and export supply elasticities. Additionally, recent deviations from Feenstra (1994) by Feenstra and Weinstein (2017) and Hottman et al. (2016) model a tighter link between exporter cost functions and export supply, but effectively assume that export supply is isoelastic and upward sloping. Feenstra (2009) also provides a basis for this generic form of export supply curves. He shows that the equilibrium in Melitz (2003) follows a constant elasticity of transformation that to some extent parallels with an upward sloping constant elasticity export supply curve at the country level. Here we are following this literature regarding the shape of export supply, but are additionally allowing for unique supply curves for exporters both within and across countries. The importer-exporter-good-specific inverse export supply elasticity is thus ω^{I}_{ijg}. We also allow for unobservable variety specific supply shocks η^{I}_{ijg} to facilitate estimation.

3. Empirical strategy

Understanding the intuition of the estimator requires an explicit characterization of how our trade data are generated given the model. Supply and demand shocks fluctuate supply and demand curves. The well known issue with supply and demand estimation is that our market data only record equilibrium outcomes of prices and quantities, which translates into endogeneity from potential simultaneity in supply and demand. Since constructing feasible instruments to address this endogeneity is impossible given scope of the data (i.e., we would require a feasible instrument for each of the millions of elasticities to be estimated), we will utilize heteroskedasticity in supply and demand shocks to estimate the model. As with all of these so called heteroskedasticity supply and demand estimators, we will also require that demand (taste) and supply (productivity) shocks be independent from one another. Shocks combine with equilibrium conditions to form the observed data.

Leamer (1981) demonstrates that we can use the variation in observed price and quantity outcomes to bound the parameters that generate our observed equilibria. Feenstra (1994) argues that if multiple varieties exported to the same destination have identical elasticity parameters, we can structurally estimate the model's parameters using Leamer (1981)'s insights. Here we are unwilling to assume that varieties from different origin countries are exported with identical supply elasticities. Consequently, looking within a single market (e.g., imports) no longer identifies either the demand or supply elasticity. By defining the export market and combining it with the import market, we will show how to achieve identification with heterogeneity. Specifically, our estimator leverages heteroskedastic differences in prices and quantities *across* versus *within* markets to produce consistent estimates of heterogeneous elasticities.

The following shows how to consistently estimate heterogeneous export supply and import demand elasticities for every importer-exporter-good triplet in the data. To do so, the method structurally estimates the preceding supply and demand model using only publicly available trade data for identification. We show explicitly how to utilize price and quantity variation across multiple markets to overcome endogeneity of supply and demand and achieve identification under heterogeneity. Finally, we discuss the assumptions governing shocks to supply and demand required by the estimator to yield consistent elasticity estimates along with potential threats to identification in the data.

---

[5] The numeraire abstracts away from labor market effects in order to focus on trade flows and their elasticities. Our preferences also rule out income effects in the model. Both assumptions prevail in the literature (cf., Broda et al., 2008), but neither are necessary for estimation. The structural assumption facilitating estimation is the CES form of the imported variety nest.

[6] Notably, this assumption allows us to avoid explicitly specifying the form of the composite domestic good. Given the poor data on production and consumption of domestic goods for most (if not all) countries, it is impossible to respectably estimate supply and demand elasticities to construct the domestic composite (see Ardelean and Lugovskyy (2010) and Blonigen and Soderbery (2010) for a more thorough discussion).

3.1. Supply and demand

Specifying a tractable empirical technique will depend on available data. Trade flows of goods between all countries are drawn from Comtrade and span 1991–2007. For consistency with previous studies we aggregate these data to the HS4 level.[7] The data are distributed in two parts. The first part consists of the values and quantities of goods exported by an origin to a destination as reported by the importer. The second part consists of the values and quantities of trade of the same good as reported by the exporter. We take advantage of both surveys in order to identify the elasticities of interest. In the following, denote variables as reported by the importer with *I*, and those reported by the exporter with *J*.

The direction of trade will be important in what follows. For example, x_{ijg}^I denotes imports by *i* from *j*, and x_{ijg}^J are exports by *j* to *i*. To be clear, x_{ijg}^I and x_{ijg}^J are the same trade flow of the product. We will also rely on time-series variation for estimation, thus we denote a given year in the data by *t*.

Identification will exploit time series variation and statistical differences between import and export markets. To begin, take the perspective of the importer. The importer faces the delivered price p_{ijg}^I for a given variety. In the data, prices are unit values and quantities are generally given in kilograms. The estimator will address measurement error in unit values through optimal weighting. Additionally, as in Feenstra (1994), we follow Kemp (1962) and convert the data into market shares in order to alleviate measurement error in quantities. Call the market share of a variety *within* an import market in period *t*, s_{ijgt}^I. From the quantitative model in Section 2.1, we derive the exporter's market share in the importing country as,

$$s_{ijgt}^I \equiv \frac{p_{ijgt}^I x_{ijgt}^I}{\sum_j p_{ijgt}^I x_{ijgt}^I} = \left(\frac{p_{ijgt}^I}{\mathcal{P}_{gt}^I}\right)^{\sigma_g^I - 1} b_{ijgt}^I. \tag{4}$$

Begin by first-differencing to remove any time invariant importer specific effects.[8] Then to absorb good-by-time specific effects, select a reference variety *k* and difference. This yields market shares in logs,

$$\Delta^k log\left(s_{ijgt}^I\right) = -\left(\sigma_g^I - 1\right)\Delta^k log\left(p_{ijgt}^I\right) + \epsilon_{ijgt}^I, \tag{5}$$

where Δ denotes the first difference and superscript *k* denotes differencing by reference country *k* such that Δ^k denotes first- then referenced-differenced variables. ϵ_{ijgt}^I are first and reference differenced unobservable variety specific taste shocks. Eq. (5) is the demand curve of the importing country for the exported variety *v*.

Estimating the system also requires a supply curve for the variety delivered to *I*. Taking the same approach as above, we can write export supply in logs. After first- and reference-differencing we are left with,

$$\Delta^k log\left(s_{ijgt}^I\right) = \frac{\omega_{ijg}^I + 1}{\omega_{ijg}^I}\Delta log\left(p_{ijgt}^I\right) - \frac{\omega_{ikg}^I + 1}{\omega_{ikg}^I}\Delta log\left(p_{ikgt}^I\right) + \rho_{ijgt}^I, \tag{6}$$

where ρ_{ijg}^I are the unobserved differenced supply shocks. Notice that the export supply curve from any country is determined by the exporter's supply elasticity (ω_{ijg}^I) relative to its competitor's supply elasticity (ω_{ikg}^I). Consequently, applying Feenstra (1994) to the system cannot identify our trade elasticities with heterogeneity.[9]

Additional variation will be required to achieve identification. This variation can be found by taking the exporter's perspective. An exporter faces the same demand and supply elasticities in market *I* as specified above. However, the decisions made by an exporter *J* regarding goods shipped *across* destinations are substantively different than the outcomes *within* a destination. These differences are made apparent by calculating the share of total export supply from country *J* represented by country *I*, which is

$$s_{ijgt}^J = \frac{p_{ijgt}^J x_{ijgt}^J}{\sum_i p_{ijgt}^J x_{ijgt}^J} = \frac{\left(p_{ijgt}^J\right)^{\frac{\omega_{ijg}^I+1}{\omega_{ijg}^I}} exp\left(\frac{-\eta_{ijgt}^I}{\omega_{ijg}^I}\right)}{\sum_i \left(p_{ijgt}^J\right)^{\frac{\omega_{ijg}^I+1}{\omega_{ijg}^I}} exp\left(\frac{-\eta_{ijgt}^I}{\omega_{ijg}^I}\right)}. \tag{7}$$

[7] *ComTrade* data are generally available down to the more disaggregate HS6 product level. We have estimated, and make available estimates at the HS6 level. The themes of the paper are unaffected by the level of aggregation. Therefore, we opt to present the HS4 estimates as they are directly comparable to important results in the literature (e.g., Broda et al., 2008).

[8] Appendix A details precisely how the following equations are constructed, and how error terms are defined structurally.

[9] In Appendix B we demonstrate the inability of Feenstra (1994) to identify our trade elasticities graphically. Fundamentally, Feenstra (1994)'s method of mapping data to hyperbolae and searching for the intersection of the hyperbolae no longer identifies elasticities with heterogeneity, as each hyperbolae is generated from a different export supply process. Explicitly, heteroskedasticity of supply and demand shocks within an importer alone is not sufficient to identify heterogeneous elasticities.

Again we want to eliminate time and country specific effects. To do so, take logs, first difference, then choose a reference *destination* (v) and difference. This yields,

$$\Delta^v log\left(p_{ijgt}^J\right) = \frac{\omega_{ijg}^I}{\omega_{ijg}^I + 1}\Delta log\left(s_{ijgt}^J\right) - \frac{\omega_{vjg}^V}{\omega_{vjg}^V + 1}\Delta log\left(s_{vjgt}^J\right) + \rho_{ijgt}^J, \tag{8}$$

where ρ_{ijgt}^J are the unobserved double difference supply shocks. Eq. (8) is the supply curve for the exported variety *across* destinations. The supply curve again depends upon the export supply elasticity in I relative to the reference. However, the reference country is now defined as the same variety shipped by an exporter to a different destination.

Next we define the demand curve across the exporter's destinations. In logs and first- and reference-differences, demand is given by,

$$\Delta^v log\left(s_{ijgt}^J\right) = \left(1-\sigma_g^I\right)\Delta log\left(p_{ijgt}^J\right) - \left(1-\sigma_g^V\right)\Delta log\left(p_{vjgt}^J\right) + \epsilon_{ijgt}^J, \tag{9}$$

where ϵ_{ijgt}^J are the relative taste parameters across destinations. Relative prices and shares are thus determined by differences in the elasticity of substitution and market composition of each of the exporter's destination markets.

Here it is worth spelling out the intuition of the estimator so that the following explicit formulation is clear. First we need to acknowledge the data generating process in our trade data. Supply and demand shocks fluctuate our four supply and demand curves. Shocks then combine with equilibrium conditions to form the observed data. Leamer (1981) demonstrates that we can use the variation in shocks to construct hyperbolae that bound the parameters underlying our observed equilibria.[10] Feenstra (1994) argues that multiple varieties following the same underlying elasticity parameters allow us to estimate supply and demand elasticities by minimizing the distance between hyperbolae. The preceding makes evident that supply and demand for a variety in a single (import or export) market depends on the differences between the heterogeneous elasticities in that market. Consequently, minimizing the distance between the hyperbolae generated by multiple varieties within a market cannot identify elasticities with heterogeneity. The following formalizes how to combine price and quantity variation of a particular variety (i.e., hyperbolae) *across* multiple markets to identify our elasticities.

3.2. Identification and estimation

Notice that while market outcomes for exporters and importers are related, they are not identical. Market share *within* a destination captured by a particular exporter (Eq. (4)) depends on the composition of that particular market (e.g., other exporters to that market). Market share of a particular importer *across* an exporter's destinations (Eq. (7)) depends on the composition of the set of destination markets (e.g., other importers of the exported variety). Differential relationships between the outcomes of an exporter within markets versus across markets will allow us to identify heterogeneous export supply elasticities by jointly estimating Eqs. (5), (6), (8) and (9). In Leamer (1981) terms, differences in the hyperbolae generated by looking at fluctuations in prices and shares across markets versus the hyperbolae generated by looking at fluctuations in prices and shares within markets will be utilized to achieve identification. In other words, we require that the export and import hyperbolae for any origin destination pair are not asymptotically identical.

First, it is worthwhile to motivate differences between export markets and import markets in the data. In the raw data, variation in market shares realized across export destinations do in fact look sufficiently different from market shares captured within a destination. To provide a specific example, Canada exported about *$34Bill* of HS 8703 (automobiles) to the US in 2006. The share of the import market captured by Canada in the US was 27%. However, the share represented by the US for the Canadian export market was almost four times the size at 96%. Additionally, the raw correlation of these market shares over time from 1991 to 2007 was 0.52. An added source of variation in the data will come from the differences in reported shipped and delivered prices (unit values), which capture differences in prices received by exporters versus those faced by importers. The raw correlation between shipped and delivered prices for Canadian exports of autos to the US is 0.71. Variation of this sort is exactly what is required by the estimator.[11] Simply, as long as the fluctuations in prices and shares over time in the two markets are *not identical* the identification strategy is sound.

We begin by assuming, as each of our predecessors have done, that supply and demand shocks to a variety are uncorrelated over time. Specifically, $E\left[\epsilon_{ijgt}^I\rho_{ijgt}^I\right] = 0$ and $E\left[\epsilon_{ijgt}^J\rho_{ijgt}^J\right] = 0$. We can then multiply the residuals from the supply and demand equations for the importer

[10] Leamer (1981) demonstrates hyperbolae from a time series of price and quantity data are defined as:

$$(\beta - \frac{cov(p,x)}{var(p)})\left(\Theta - \frac{cov(p,x)}{var(p)}\right) = \left(\frac{cov(p,x)^2}{var(x)var(p)} - 1\right)\left(\frac{var(x)}{var(p)}\right)$$

where Θ is the true supply elasticity and β is the true demand elasticity. For those unfamiliar with Leamer (1981) hyperbolae, Soderbery (2015) surveys the methodology using actual data to construct hyperbolae and estimate import demand and export supply elasticities using Feenstra (1994)'s method.

[11] Theoretically, it is beneficial to exploit differences between shipped and delivered prices to further ensure that hyperbolae are not asymptotically identical. We will thus opt to utilize both importer and exporter reported trade values in Comtrade. Feenstra and Romalis (2014) also rely on both sections of Comtrade, which provides some measure of support. However, it is worth acknowledging the critiques levied by Hummels and Lugovskyy (2006) regarding the accuracy of the exporter reported data in particular. We do address the potential introduction of measurement error following Broda and Weinstein (2006). Specifically, we follow their intuitive weighting and the inclusion of a weighted constant term. As long as what remains after our control for measurement error is not correlated with price and quantity variances and covariances our strategy is sound. Finally, if we were still concerned about the accuracy of exporter reported data, it is reassuring that the estimator presented here does not require both importer and exporter reported data. One could easily calculate the import and export shares needed for estimation using only importer reported data as long as the researcher constructs the variables and estimator using the totality of world trade. The fundamental goal is defining multiple markets with price and market share variation that is sufficiently different but the underlying elasticities are the same. We expect the differences in import and export shares constructed using only import data would provide enough variation to generate unique hyperbolae even without exploiting the differences between shipped and delivered prices if the researcher deems exporter reported data unsuitable.

and exporter markets to generate consistent estimating equations. For the import market we multiply the residuals from Eqs. (5) and (6) to get,

$$\Delta^k log\left(p^I_{ijgt}\right)^2 = \frac{\omega^I_{ijg}}{\left(1+\omega^I_{ijg}\right)\left(\sigma^I_g-1\right)}\Delta^k log\left(s^I_{ijgt}\right)^2 + \frac{\omega^I_{ijg}}{1+\omega^I_{ijg}}\Delta^k log\left(s^I_{ijgt}\right)\Delta^k log\left(p^I_{ijgt}\right) - \frac{1}{\sigma^I_g-1}\Delta^k log\left(s^I_{ijgt}\right)\Delta log\left(p^I_{ijgt}\right)$$
$$+ \frac{\omega^I_{ijg}\left(1+\omega^I_{ikg}\right)}{\omega^I_{ikg}\left(1+\omega^I_{ijg}\right)\left(\sigma^I_g-1\right)}\Delta^k log\left(s^I_{ijgt}\right)\Delta log\left(p^I_{ikgt}\right) + \frac{\omega^I_{ijg}-\omega^I_{ikg}}{\omega^I_{ikg}\left(1+\omega^I_{ijg}\right)}\Delta^k log\left(p^I_{ijgt}\right)\Delta log\left(p^I_{ikgt}\right) + u^I_{ijgt}, \tag{10}$$

where the error term $u^I_{ijgt} = \frac{\omega^I_{ijg}\rho^I_{ijgt}\epsilon^I_{ijgt}}{\left(1+\omega^I_{ijg}\right)\left(\sigma^I_g-1\right)}$ is zero in expectation.[12] To estimate the model we will build on the methodology proposed by Feenstra (1994).[13] Taking exporter by exporter averages over time transforms Eq. (10) into a linear regression of price variances on share variances and price-share covariances (i.e., hyperbolae). However Eq. (10) is unidentified, so that applying Feenstra (1994)'s weighted least squared to Eq. (10) cannot consistently estimate our heterogeneous elasticities.

To overcome this identification problem, we produce a similar estimating equation for the export market. Multiplying our error terms from Eqs. (8) and (9) together yields,

$$\Delta^v log\left(p^J_{ijgt}\right)^2 = \frac{\omega^I_{ijg}}{\left(1+\omega^I_{ijg}\right)\left(\sigma^I_g-1\right)}\Delta log\left(s^J_{ijgt}\right)^2 + \frac{\omega^I_{ijg}\left(\sigma^I_g-2\right)-1}{\left(1+\omega^I_{ijg}\right)\left(\sigma^I_g-1\right)}\Delta log\left(s^J_{ijgt}\right)\Delta log\left(p^J_{ijgt}\right) + \frac{\omega^V_{vjg}}{\left(1+\omega^V_{vjg}\right)\left(\sigma^I_g-1\right)}\Delta log\left(s^J_{vjgt}\right)^2$$
$$+ \frac{1-\omega^V_{vjg}\left(\sigma^I_g-2\right)}{\left(1+\omega^V_{vjg}\right)\left(\sigma^I_g-1\right)}\Delta log\left(s^J_{vjgt}\right)\Delta log\left(p^J_{ijgt}\right) + \frac{1-\omega^V_{vjg}\left(\sigma^V_g-2\right)}{\left(1+\omega^V_{vjg}\right)\left(\sigma^I_g-1\right)}\Delta log\left(s^J_{vjgt}\right)\Delta log\left(p^J_{vjgt}\right) + \frac{\omega^I_{ijg}\left(\sigma^V_g-2\right)-1}{\left(1+\omega^I_{ijg}\right)\left(\sigma^I_g-1\right)}\Delta log\left(s^J_{ijgt}\right)\Delta log\left(p^J_{vjgt}\right)$$
$$- \frac{\omega^V_{vjg}\left(1+\omega^I_{ijg}\right)+\omega^I_{ijg}\left(1+\omega^V_{vjg}\right)}{\left(1+\omega^I_{ijg}\right)\left(1+\omega^V_{vjg}\right)\left(\sigma^I_g-1\right)}\Delta log\left(s^J_{ijgt}\right)\Delta log\left(s^J_{vjgt}\right) + \frac{\sigma^I_g-\sigma^V_g}{\sigma^I_g-1}\Delta log\left(p^J_{vjgt}\right)^2 + \frac{\sigma^V_g-\sigma^I_g}{\sigma^I_g-1}\Delta log(p^J_{vjgt})\Delta log\left(p^J_{ijgt}\right) + u^J_{ijgt}, \tag{11}$$

where the error term $u^J_{ijgt} = \frac{\rho^J_{ijgt}\epsilon^J_{ijgt}}{\sigma^I_g-1}$ is zero in expectation. Combining Eqs. (10) and (11) provides us the necessary variation to identify all heterogeneous elasticities. To estimate the full model, we apply Feenstra (1994)'s methodology to convert both equations to regression of price variances on share and price variances and covariances (roughly, we map Leamer (1981) hyperbolae into data). Our estimation then chooses the elasticities which jointly minimize the distance between Leamer (1981) hyperbolae across the import and export markets.

We apply the estimator looking within an importer-good pair across exported varieties over time. Multiple exporters jointly identify the import demand elasticity. Combining variation from import and export flows then pins down the pairwise export supply elasticities. In practice, we simultaneously estimate Eqs. (10) and (11) with a nonlinear seemingly unrelated-regression constrained to the feasible region where $\sigma > 1$ and $\omega > 0$.[14] Estimates of the elasticities are consistent provided the import and export hyperbolae for each importer-exporter-product triplet are not asymptotically proportional. [15]

Fig. 1 presents the intuition of the estimator graphically. Consider three countries denoted I,J and K trading good g. Country I imports varieties of g from J and K. We can then map prices and quantities following Eq. (10) into hyperbolae (denoted Imports (s^I_{ijg}) and Imports (s^I_{ikg})). Fig. 1 assumes we know for certain the import demand σ^I_g and export supply elasticities ω^I_{ijg} and ω^I_{ikg} in the market. Looking solely at the import hyperbolae, notice that we cannot identify the three elasticities with only the two import hyperbolae. Our identification strategy is to provide additional information (i.e., hyperbolae) from the export market. Realizing exporters J and K export to (many) other markets allows us to map Eq. (11) into the hyperbolae denoted Exports (s^I_{ijg}) and Exports $\left(s^I_{ikg}\right)$.

By combining importer and exporter hyperbolae, Fig. 1 displays our identification strategy.[16] Notice that the intersections of the importer and exporter hyperbolae in the import and export markets for multiple varieties can be combined to identify heterogeneous export supply elasticities. Notice that the intersections of the hyperbolae in the two markets jointly pin down the import demand elasticity since all exporters

[12] Note that Eq. (10) is identical to Feenstra (1994) if export supply elasticities are homogeneous (i.e. $\omega^I_{ijg} = \omega^I_g \quad \forall \quad v$).

[13] Feenstra (1994) demonstrates that applying WLS on country averages (or equivalently 2SLS with country indicators as instruments) to Eq. (10) yields consistent estimates of the supply and demand elasticities when both are homogeneous.

[14] One concern with existing trade data is measurement error in prices (unit values) and quantities (product weight). As mentioned above, the use of market shares alleviates much of our concern regarding the measurement of quantities. Using unit values in place of transaction prices necessitates a weighting scheme. Given the use of both import and export data and the context of the problem, we adopt the weights proposed by Broda and Weinstein (2006) and include an importer-exporter specific constant. Specifically, the data are weighted by $T^{\frac{3}{2}}_{ijg}\left(\frac{1}{x_{ijgt}} + \frac{1}{x_{gvt-1}}\right)^{-\frac{1}{2}}$, as we expect greater accuracy in the recorded data between large ($\uparrow x_{ijgt}$) and persistent ($\uparrow T_{ijg}$) trading partners. Broda and Weinstein (2006) demonstrate that these weights accompanied by a constant in estimation correct for random measurement error in prices along with measurement error that may be decreasing in the amount of total trade. Additionally, this weighting scheme matches our intuition derived from hyperbolae – we will give more weight to hyperbolae generated by large trading partners in the estimation. Soderbery (2015) discusses the appropriate weighting of hyperbolae in great detail. In the case of Feenstra (1994) LIML is the superior scheme. Here the problem is exponentially more computationally intensive and not amenable to LIML, thus we opt for the intuitive weighting scheme.

[15] Mathematically, this amounts to

$$\frac{\sigma^2_{\epsilon^I_{ijgt}} + \sigma^2_{\epsilon^I_{ikgt}}}{\sigma^2_{\epsilon^J_{jigt}} + \sigma^2_{\epsilon^J_{jkgt}}} \neq \frac{\sigma^2_{\rho^I_{ijgt}} + \sigma^2_{\rho^I_{ikgt}}}{\sigma^2_{\rho^J_{jigt}} + \sigma^2_{\rho^J_{jkgt}}}$$

for every variety $j \neq k$, where σ^2 is the variance of the subscripted shock process (e.g., $\sigma^2_{\epsilon^I_{ijgt}}$ is the variance of ϵ^I_{ijgt}, which are the demand shocks in country I for exports originating from country J). Intuitively, the above condition describes how shocks within a market must differ from shocks across markets to ensure that importer and exporter hyperbolae are not asymptotically proportional.

[16] Appendix B provides more detail on the construction of importer and exporter hyperbolae along with the identification issues introduced by heterogeneity.

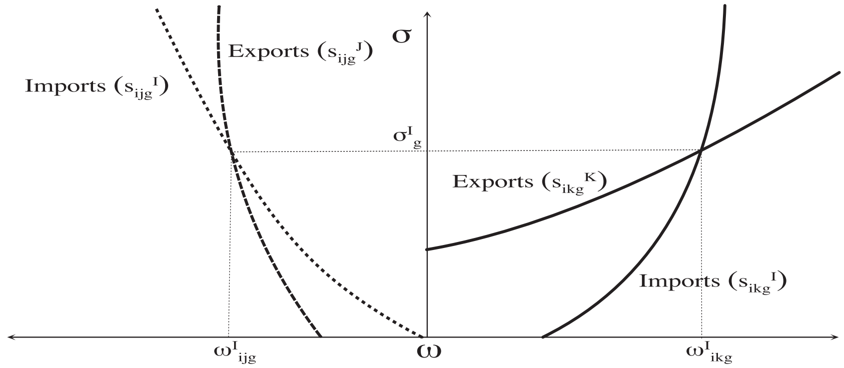

Fig. 1. Combining importer and exporter hyperbolae.

to I face the same value. With the demand elasticity in hand, our estimator can then simultaneously identify the heterogeneous export supply elasticities by combining hyperbolae from the import and export markets.[17]

Fig. 1 presents the ideal situation where all of the hyperbolae cross and line up directly at the import demand elasticity. The reality of the data is that this clean result will be rare – hyperbolae in general may cross multiple times at different values in the feasible region or not cross at all. This necessitates our estimation strategy where we jointly estimate the demand elasticity and use it to identify the heterogeneous export supply elasticities.

Our estimation strategy for import demand and export supply elasticities fits with a small but growing set of structural estimators in trade. Feenstra and Weinstein (2017) propose a structural estimator based on translog preferences to estimate exporter markups. Feenstra and Romalis (2014) estimate an importer specific quality distribution parameter. Notably, their estimators are not tasked with uncovering pairwise heterogeneity between importers and specific exporters, but rather potential distributional differences across exported varieties to the US. Using a similar methodology, Hottman et al. (2016) leverage detailed scanner data in the US to uncover quality, costs and markups across goods sold by particular US retailers. In comparison, we analyze the heterogeneous relationships both within and across all traded goods and countries.

A valuable difference between our estimator and the literature are the data required for estimation.[18] Our technique needs only readily available data on trade flows to be implemented. Feenstra and Weinstein (2017) require measures of product specific Herfindahl indexes that were specially constructed in the US for their study, and do not exist for most (if not all) other countries. Feenstra and Romalis (2014) rely on highly detailed trade cost and wage data across countries that are not readily available. Hottman et al. (2016) utilize proprietary scanner data as they require granular transaction level data, which presents significant hurdles to extend their estimator beyond the US or even across the full set of products imported by the US.

3.3. Threats to identification

Before presenting the estimates, it is worthwhile to discuss some potential correlation patterns across countries that will and will not call the estimator into question. The stalwart assumption that we (and the preceding literature) require is that supply and demand shocks from both the importer and exporter perspective are uncorrelated (i.e., $E\left[\epsilon^I_{ijgt}\rho^I_{ijgt}\right]=0$ and $E\left[\epsilon^J_{ijgt}\rho^J_{ijgt}\right]=0$). This assumption cannot be relaxed. While this assumption is a mainstay in the literature (i.e., Feenstra (1994) and extensions), it notably rules out models of endogenous quality choice where there is a mechanical correlation between importers' tastes and exporters' supply shocks.

However, this assumption does not rule out a host of other potential patterns in the data. Suppose for instance that Japan's productivity increases over time such that its supply curves to all destinations shift out systematically. This would imply that $E\left[\rho^J_{ijgt}\rho^I_{ijgt}\right]>0$ for the Japanese variety. As long as importers' taste for the Japanese variety is not increasing or decreasing *due to* the persistent supply shock, our estimator is consistent. Alternatively, suppose all importers' taste for say Germany's variety increase. This would imply that $E\left[\epsilon^J_{ijgt}\epsilon^I_{ijgt}\right]>0$ for German exports. Again, as long as German supply shocks are not increasing or decreasing due to this persistent demand shock our estimator is consistent.

One may argue that a persistent supply shock to a large country, such as Japan, will have feedback into importers' price indexes and total expenditures for a good (in fact, we will make this argument later when we analyze market power and optimal tariffs). This argument poses no threat to the estimating equation from the importer perspective, as the $E\left[\epsilon^I_{ijgt}\rho^I_{ijgt}\right]$ depends only on importer taste and exporter supply shocks, which we assume are fundamentally uncorrelated. However, from the exporter's perspective, ϵ^J_{ijgt} does contain importer price

[17] Theoretically, one could use the intuition of Fig. 1 to estimate a model with both heterogeneous import demand and export supply elasticities. However, given the results of Soderbery (2015), the prevalence of CES demand in international trade, and the tax this would place on the data, we focus on extending the estimator to heterogeneity in export supply elasticities.

[18] Topically related, there is a larger literature regarding estimating trade elasticities to measure welfare. Recently, Bas et al. (2017), Fernandes et al. (2015) Head et al. (2014) have developed gravity type estimators derived from the so called New Trade Theory models under CES preferences. Similarly, Simonovska and Waugh (2014) aim to compare estimates of macro elasticities of trade across models. In related work, Tan (2012) develops a similar type of estimator for a model with translog preferences. While related in spirit, this literature aim to estimate more macroeconomic elasticities, while the preceding is focused on estimating disaggregate product and market specific supply and demand elasticities.

indexes. We argue that even with large countries potentially impacting price indexes through their export supply shocks, $E\left[\epsilon^J_{ijgt}\rho^J_{ijgt}\right]$ still equals zero. This argument is based on observing that since ρ^J_{ijgt} is structurally the reference difference of first-differenced export supply shocks for the same exporter, it is zero in expectation over time. If we were unwilling to assume $E\left[\rho^J_{ijgt}\right]=0$, the structure of the model, data, and estimator do allow the practitioner to decompose ϵ^J_{ijgt} by including relative price indexes in the joint estimation rather than the error term. While this is a viable alternative, we are confident that the intuition of $E\left[\rho^J_{ijgt}\right]=0$, combined with our strategy to correct for measurement error, appropriately addresses these concerns.[19]

Fundamentally, our estimator relies on heteroskedasticity in supply and demand shocks. In Feenstra (1994), heteroskedasticity in shocks across exporters of a particular good is the key. He showed how this variation combined with homogeneous elasticities led to identification. Our estimator leverages this same variation in the data, but also requires heteroskedasticity in shocks across importers of the exported good. This additional source of variation, which we have highlighted is prevalent in the data, is how we achieve identification of the heterogeneous elasticities. The following presents our estimates, and provides some tests of their reasonableness given our intuition.

4. Elasticity estimates

The only data required for estimation are trade flows associated with country pairs across goods – here we rely on Comtrade data from 1991 to 2007, which are readily available. The data contain 1243 goods at the HS4 level and 192 importing and exporting countries.

Not all countries trade all goods with one another, but we are still tasked with estimating approximately 3 million export supply elasticities (the number of importer-exporter-goods in the data) and 200,000 (the number of importer-goods in the data) import demand elasticities. Simply, this is computationally infeasible.[20] To reduce the parameter space, assume small countries in the same region have identical supply technologies. Estimating exporter heterogeneity globally thus comes at the expense of the restrictive assumption in the estimation: large numbers of developing countries have identical export supply elasticities.[21] However, it is necessary for computational tractability, and still considerably more disaggregate than any preceding study. Table 1 lists the regions designated along with their total imports and exports.[22]

Applying the estimator requires imports from at least two countries that both export to at least one other destination for a minimum of three periods. These are relatively weak requirements conceptually, but in practice the intersection of trade data and the estimator meeting its minimum requirements reduces the sample of trade flows from around 1.25 million observations per year to 0.8 million. Reassuringly, the remaining data account for 94% of all trade by value globally.[23]

Table 2 summarizes our estimates across importers. The median absolute deviation (MAD) indicates that around 75% of our export supply elasticities lie between 0.197 and 1.183 with a long right tail.[24] Our demand elasticities have a similar long right tail and aninterquartile range of 2.41–3.34, which is very much in line with the literature.[25]

Additionally, our estimated elasticities vary across goods and importers and even within goods across exported varieties. To be specific, variation in our estimated inverse export supply elasticities explained by importer-product fixed effects is 72%. Remaining variation in the estimates comes from heterogeneity across exported varieties within an imported good. Notably, variation in homogeneous export supply elasticity estimates, such as those estimated by Broda and Weinstein (2006) and Broda et al. (2008), is fully explained by importer-product fixed effects.

It is also beneficial to discuss some features of the distribution of export supply elasticities within an importer-product pair. We can calculate the median of ω^I_{ijg} within every importer-product in the data. The median of this median is 0.712, which is similar to the median across all estimates. For this median importer-product, the standard deviation of the export supply elasticities is 0.625 and the MAD is 0.125.[26] These within importer-product distributions will provide key insights into furthering our understanding of importer-exporter-product relationships.

A useful interpretation of the inverse export supply elasticity ω^I_{ijg} is as a measure of importer market power. Since ω^I_{ijg} governs the degree of passthrough of a shock (e.g., tariffs) to delivered prices, large estimates correspond with a high degree of importer market power as more inelastically exported varieties only weakly pass through shock-induced price changes. Table 2 suggests that countries with the highest GDPs tend to have the most market power, as their imports have the largest median inverse supply elasticities. Of these countries, Japan (JPN), China (CHN), Germany (DEU) and the US (USA) seem to have the highest degree of market power on their median import. As we may expect, regions with the least market

[19] Additionally, we have estimated the alternative strategy discussed here by controlling for Sato-Vartia price indexes and expenditure in the estimation rather than the error term. For the products we have estimated, this alternative yields almost identical results. The within importer correlation of the export supply elasticities is 0.91 on average, and the median percentage difference is 0.000002% with an interquartile range of −0.00003% and 0.04%. The import demand elasticities are even closer to one another with 90% if the estimates lying between −0.07% and 0.02% within one another. These results are highly supportive of our assumptions and preferred methodology, since including the price index as a control has little to no effect on the estimates.

[20] There are 3 million importer-exporter-product triplets in the data, which implies 3 million export supply elasticities to be estimated. The following thus makes some assumptions to reduce the parameter space. After these assumptions, we still estimate over 1.2 million export supply and 125,000 import demand elasticities. The estimation routine requires two full weeks to complete running concurrently on three computers with 4 processors using StataMP.

[21] For example, the 43 African countries have the same destination-good export supply elasticities (see Table 1).

[22] We have tried to follow the United Nations Statistical Division's definitions of country regions, which can be found here https://unstats.un.org/unsd/methods/m49/m49regin.htm. The full listing of countries and their assigned regions is given by Appendix Table A8.

[23] Appendix E is devoted to comparing my estimates to those of Broda et al. (2008), but it is worth noting that their analysis covers data with 250,000 observations over 15 importers that only accounts for 8% of global trade.

[24] Robust standard errors for each estimate is available upon request. For the sake of space, further discussion of standard errors is undertaken in Appendix C.

[25] While the mean and median of the import demand elasticities estimated here are lower than those in previous studies (cf., Broda et al., 2008), they are very much in line with Soderbery (2015) who argues previous studies mask significant bias in the standard estimator and provides a correction. This paper is focused on the role of export supply elasticities. Yet, it is worth noting that there is a sense by which accounting for heterogeneity in export supply tames our estimates of import demand, as our estimates are not as volatile as previous studies and still follow intuitive patterns across products.

[26] There is also considerable variation in these statistics across importer-products. The 25th percentile of this within importer-product median is 0.353 with a standard deviation of 0.231 and a MAD of 0.023. Additionally, the 25th percentile of this within median is 1.485 with a standard deviation of 1.536 and MAD of 0.356.

Table 1
Regions and trade.

Country			GDP ($Billion)		Total trade ($Billion)	
Name	Code	Count	Average	Total	Imports	Exports
Australia	AUS	1	768.18	768.18	189.03	126.30
Brazil	BRA	1	1067.96	1067.96	154.55	153.13
Canada	CAN	1	1251.46	1251.46	527.79	467.53
China & Hong Kong	CHN	2	1574.23	2872.15	1791.36	2814.51
Germany	DEU	1	2906.68	2906.68	1473.64	2120.57
France	FRA	1	2230.72	2230.72	908.88	834.60
Great Britain	GBR	1	2345.02	2345.02	855.11	656.88
India	IND	1	906.27	906.27	238.88	195.04
Italy	ITA	1	1844.75	1844.75	713.24	818.41
Japan	JPN	1	4340.13	4340.13	789.13	1296.27
Mexico	MEX	1	839.18	839.18	442.62	322.61
Russia	RUS	1	986.94	986.94	265.06	259.23
United States	USA	1	13,201.82	13,201.82	2729.85	1775.20
Region						
African	AFR	43	83.39	941.65	304.65	281.94
Asian	ASA	38	248.54	2994.04	2179.07	2082.72
Caribbean	CAR	16	17.68	64.67	34.40	22.62
Northern/Western Europe	NWU	18	328.79	3278.56	2375.50	2354.46
Oceania	OCE	11	90.15	113.45	42.03	20.10
South American	SAM	20	98.64	962.40	310.57	275.90
Southern/Eastern Europe	SEU	22	292.89	2716.01	1591.13	1038.48

Notes: GDP is total gross domestic products reported by *CEPII* in US dollars for 2006. Imports and exports are in US dollars from ComTrade in 2006.

Table 2
Summary statistics: Elasticities and country size.

		σ_g^I			ω_{ijg}^I		
Importer	Obs	Mean	Median	MAD†	Mean	Median	MAD†
Australia	17,923	3.727	2.997	0.474	38.56	0.798	0.549
Brazil	15,664	3.372	2.847	0.492	31.42	0.734	0.562
Canada	17,896	3.702	3.051	0.621	52.07	1.158	0.734
China & Hong Kong	39,393	3.395	2.897	0.554	189.74	1.270	0.770
Germany	35,880	3.608	3.029	0.505	101.31	1.206	0.572
France	30,124	3.351	2.972	0.411	40.26	0.828	0.468
Great Britain	31,064	3.333	2.977	0.363	18.56	0.853	0.453
India	16,496	3.832	2.996	0.529	54.01	0.752	0.552
Italy	30,608	3.439	3.012	0.478	43.26	0.798	0.451
Japan	21,574	3.640	2.985	0.543	123.13	1.214	0.758
Mexico	15,715	3.625	2.925	0.526	29.63	0.890	0.697
Russia	20,310	3.621	3.007	0.563	204.18	0.773	0.524
United States	34,443	3.937	2.940	0.383	39.17	1.203	0.541
African	116,143	3.921	2.656	0.393	131.02	0.581	0.555
Asian	261,285	3.282	2.827	0.494	69.40	0.502	0.449
Caribbean	18,925	3.537	2.852	0.503	38.09	0.669	0.605
Northern/Western Europe	216,373	3.127	2.843	0.446	83.64	0.648	0.445
Oceania	16,915	3.829	2.991	0.642	59.88	0.732	0.638
South American	104,882	3.425	2.882	0.425	30.04	0.588	0.421
Southern/Eastern Europe	214,013	3.251	2.857	0.457	24.61	0.665	0.439
World	1,275,626	3.405	2.878	0.465	68.55	0.690	0.493

Notes: For exposition and comparability to previous studies, σ_g^I is truncated at 131.05, and the top and bottom 0.5% of ω_{ijg}^I are dropped. † MAD stands for the median absolute deviation.

power are the small Asian (ASA), South American (SAM) and African (AFR) countries. One surprising feature of the summary statistics is the relatively high median estimated market power for Oceanic (OCE) and Caribbean (CAR) countries.

We will explain these moments in the estimates by controlling for intuitive patterns of market power subsequently. Yet it is worth noting here that these patterns in the raw estimates are a direct consequence of compositional heterogeneity in the goods and partners that comprise importers' trade. This heterogeneity is precisely the feature of the data the estimator is tasked with uncovering.

4.1. Product differentiation and market power

Here we consider some tests, in the spirit of Broda et al. (2008), of our intuition regarding the relationships between importer-exporter and country-good characteristics that we expect to underlie market power. First, we explore the relationship between product differentiation and market power and the substitutability of goods. Second, we examine the role of shares of countries' trade in forming market power across goods. In the analysis we will highlight the role and importance of heterogeneous elasticities for our intuition.

Table 3
Trade elasticities and product differentiation.

	$log\left(\omega^I_{ijg}\right)$			$log\left(\sigma^I_g\right)$	
	(1)	(2)	(3)	(4)	(5)
Reference good	−1.816***	−1.830***	−0.665***	0.027***	0.027***
	(0.012)	(0.011)	(0.061)	(0.001)	(0.001)
Homogeneous good	−3.192***	−3.260***	−2.129***	0.080***	0.079***
	(0.027)	(0.027)	(0.065)	(0.002)	(0.002)
Importer FEs	No	Yes	Yes	No	Yes
Importer × Non-differentiated Good FEs	No	No	Yes	No	No
R^2	0.033	0.044	0.049	0.002	0.005
Obs.	1,229,710	1,229,710	1,229,710	1,229,710	1,229,710

Notes: Rauch (1999)'s conservative classification is used to indicate differentiated, reference priced, and homogeneous goods (the liberal classification yields identical results). Robust standard errors are in parentheses where, * p<0.10, ** p<0.05, *** p<0.01 indicate significance.

Conceptually, we believe importers have more market power over imports of differentiated goods since the equilibrium price of a differentiated good may be more affected by a single importer. For instance, if US demand for German microscopes (HS 9011) falls and Germany decreases prices globally, there will be relatively weak substitution by other countries toward German microscopes due to the differentiated nature of this industry. Consequently, the equilibrium world price for German microscopes may fall substantially in response to the shock in the US market, which is our definition of importer market power. Conversely, we would expect strong substitution by other countries with price fluctuations if the good is more homogeneous(e.g., steel sheet). Plainly, we expect differentiated goods to present with higher ω^I_{ijg}. A more direct comparison can be made with the demand elasticity. We expect homogeneous goods to be more substitutable, and thus present higher estimated import demand elasticities $\left(\sigma^I_g\right)$.[27]

Table 3 explores whether product differentiation relates to our estimated elasticities through various reduced-form regressions of our inverse export supply and import demand elasticities on indicators of product differentiation.[28] Considering the long right tail of the distributions of our elasticity estimates, our regressions are run in logs. Regardless of specification, Table 3 demonstrates that importers have lower market power on average in markets for less differentiated goods. Column (1) estimates that the average differentiated good is supplied with an inverse export supply elasticity triple the size of the average homogeneous good.

The heterogeneity in our estimates allows us to analyze whether certain countries have different degrees of market power across types of goods. We thus include various fixed effects to examine whether the relationship between product differentiation and trade elasticities varies across importers. Column (2) begins by adding importer fixed effects. Column (3) pushes the estimates further by exploring whether countries have differential market power across types of goods via importer fixed effects interacted with a product differentiation dummy. In all of these regressions countries have higher market power for differentiated goods on average. Columns (4)–(5) extend the exercise to the import demand elasticity estimates. We can see that our estimates of σ^I_g are highly correlated with product differentiation. As we would expect, differentiated goods have the lowest elasticities of substitution, and in increasing order, reference and homogeneous goods have higher elasticities.

The fixed effects in the preceding regression can be used to rank countries by their median inverse export supply elasticity across types of goods after controlling for product differentiation. Table 4 presents the estimated ordering of each region's market power by these fixed effects. We can see that Germany, the US and China have the highest degree of market power across all goods. Outside of the US, North American countries have noticeably low market power. This feature is likely due to the strong ties Canada and Mexico have with the US and an inability to exert market power on US exports.

For differentiated goods, market power rankings follow a similar ordering to that of all goods. One noticeable departure is market power in Asian (ASA) countries over differentiated goods. Overall, Asian countries rank seventeenth, but in terms of differentiated goods they have the twelfth highest degree of market power. This is quite intuitive as this region contains countries such as Taiwan and Korea which are significant importers of high-tech goods.

Table 4
Rankings of average market power.

	Market power ranking		
Importer	All goods	Differentiated goods	Non-differ goods
Germany	1	2	2
China & Hong Kong	2	1	8
United States	3	3	5
Japan	4	4	4
Italy	5	8	7
South/East Europe	6	5	17
Canada	7	7	10
France	8	10	6
Great Britain	9	9	9
North/West Europe	10	6	18
Russia	11	11	14
South American	12	13	15
Australia	13	15	11
India	14	18	3
Oceania	15	17	13
Mexico	16	16	16
Asian	17	12	20
Brazil	18	19	12
African	19	14	19
Caribbean	20	20	1

Notes: Rankings are determined by sorting the fixed effects in Table 3. *All goods* sorts the fixed effects from Column (2), while *Differentiated* and *Non-differ* sort the fixed effects from Column (3).

[27] Perpetuating the US-Germany example, we estimate a relatively large export supply elasticity (3.25) and small import demand elasticity (2.69) for US imports of German microscopes (HS 9011). Conversely, we estimate a relatively small export supply elasticity (0.09) and large import demand elasticity (4.51) for US imports of German steel sheet (HS 7208). These estimates thus support the intuition that the US has strong market power over imports of more differentiated products and more differentiated products tend to present lower demand elasticities.

[28] Product differentiation is defined according to Rauch (1999) who classifies products as either differentiated, reference priced or homogeneous.

Table 5
Inverse export supply elasticities and market size and share.

	$\log\left(\omega^{I}_{ijg}\right)$					
Controls	(1)	(2)	(3)	(4)	(5)	(6)
log(Importer GDP)	0.045***	0.096***				0.117***
	(0.003)	(0.009)				(0.009)
log(Importer remoteness)		0.012*				0.016**
		(0.007)				(0.007)
log(Distance)		−0.544***				−0.475***
		(0.006)				(0.006)
Importer's share of global imports of the product			4.894***			
			(0.129)			
Importer's share of total exporter's exports across products				3.907***		
				(0.073)		
Importer's share of exporter's exports within product					0.737***	
					(0.035)	
log(Exporter GDP)						-0.435***
						(0.010)
log(Exporter remoteness)						-0.446***
						(0.008)
R^2	0.067	0.079	0.072	0.073	0.071	0.084
F stat	310.99	4928.98	1438.65	2876.35	450.96	3565.22
N	952,384	952,384	1,042,824	1,042,824	1,042,824	93,6503

Notes: All regressions include industry, defined at the HS4 level, fixed effects. Remoteness for country i is defined as $^1/_{\sum_j {}^{GDP_j}/_{distance_{ij}}}$ using data from *CEPII* from 2006. Robust standard errors are in parentheses where, * p<0.10, ** p<0.05, *** p<0.01 indicate significance.

Additionally, we believe Asia's reliance on imported commodities leads to the lowest level of market power across non-differentiated goods. Exactly the opposite story is revealed for India. India has one of the lowest levels of market power for differentiated goods, but the third highest degree of market power for non-differentiated goods. India's imports of non-differentiated goods are predominantly supplied by regional trade partners, of which India is an important export destination. These regional trade linkages have been long associated with importer market power. While these rankings support intuitive patterns of market power varying across types of goods and the composition of a country's imports, a more detailed analysis is warranted.

The preceding looks across products to make cross country comparisons. Table 5 now looks within products by including product fixed effects in order to compare our market power estimates across country-pair relationships. Column (1) begins with the assertion that country size, measured by total GDP in US dollars, correlates with estimated market power.[29] We see that large importers have higher degrees of market power. Conditional on country size, we have asserted that importers in remote regions possess added market power over their nearby network of trade partners. Column (2) thus adds a measure of importer remoteness to capture regional market power.[30] We see a larger estimated coefficient between country size and our inverse elasticity as a result, as well as a positive relationship with importer remoteness. Column (2) also controls for physical distance to capture the connectedness of the origin and destination. We also estimate a statistically significant negative relationship between distance and importer market power, suggesting that exporters are less responsive to distant markets.

Next, Table 5 investigates whether pairwise relationships between countries within and across goods influence market power. Columns (3) and (4) consider covariates that capture various measures of the share of the imports. We expect that large importers of a particular good possess greater influence on the world market for that good. Column (3) includes the importing country's share of total world imports of the HS4 product. Our estimates suggest that going from China's average share of a product's imports (5%) to the US share (16%) implies on average 53% larger inverse export supply elasticities for a given good.

In Column (4) we include the importer's share of the exporter's total trade. This regression demonstrates that large importers of a particular exporter's goods have greater market power on average. The heterogeneity of our estimates allows us to look at the finest level of the importer-exporter relationship. Column (5) highlights that even at the most disaggregate level we find intuitive patterns by including the importer's share of the exporter's global exports of the HS4 product. We see that exporters shipping a greater share of a particular product to a particular destination have stronger responses to fluctuations in that market, as they face on average 7.3% larger inverse export supply elasticities for a 10% increase in market share. Lastly, our data allow us to look at characteristics of the exporter and how they may affect importer market power. In Column (6) we include size and remoteness of the exporting country. We see that larger exporters have lower inverse supply elasticities suggesting that exporter size counteracts importer market power to some extent.

In what follows, these patterns of market power and their implications for trade policy are used to describe how importers construct tariffs to maximize welfare by exploiting terms of trade gains as policy interacts with market power. The heterogeneity uncovered in the preceding will be shown to provide valuable insight into how importers may set tariffs both theoretically and empirically.

5. Optimal tariffs

Our analysis of optimal trade policy is related in spirit to Ossa (2014), which calibrates a new trade model to examine the differences between non-cooperative and cooperative trade policies. Notably, his simulations rely on homogeneous elasticity estimates (i.e., Feenstra, 1994). Here we estimate the model based on heterogeneity and then evaluate an adapted version of optimal trade policy. Our estimation strategy and evaluation of the theoretical results are perfectly compatible. Ossa (2014)'s analysis is centered around describing the channels of the model under different policy equilibria. Our analysis is data driven, and aimed at presenting a flexible model and estimation technique. The exercise will be to highlight the unique explanatory power provided by exporter heterogeneity on the relationship between applied and optimal non-cooperative tariffs in the data.

[29] Importer specific measures such as GDP are extracted from *CEPII*. Subsequent regressions are based on observations from 2006, which has the most coverage.

[30] Remoteness for country i is defined as $\frac{1}{\sum_j \frac{GDP_j}{distance_{ij}}}$.

Our comparison of optimal and actual trade policy is therefore most closely related to Broda et al. (2008), who relate optimal non-cooperative tariffs under homogeneity to applied tariffs for a select group of countries. Optimal tariffs under heterogeneity are shown to yield new avenues for identification and maintain a positive relationship with applied tariffs over time. Our estimation differs from Broda et al. (2008) through two main channels. First, our heterogeneous elasticities reconcile the regularity that large importers tend to have both strong market power and low applied tariffs. Export supply heterogeneity implies that optimal tariffs depend on the composition of trade. In contrast, under homogeneity the optimal non-cooperative tariff for a given good is the inverse of its (homogeneous) export supply elasticity. As a consequence, the optimal tariffs under homogeneity do not vary over time. Second, as optimal tariff estimates under heterogeneity vary over time with changes in the composition of trade, they respond to shifts in importers' trade patterns. We will demonstrate that introducing heterogeneity in export supply requires our model of optimal trade policy to internalize market concentration when constructing tariffs. In this way, our results are related to Ludema and Mayda (2013) who argue that when relating applied tariffs to export supply elasticities for WTO members, controlling for market concentration (they rely on HHIs) is key to identifying terms of trade motives. In our model, the importance of market concentration is borne directly by optimal non-cooperative trade policy (i.e., terms of trade) motives without appeal to a particular bargaining process as in Ludema and Mayda (2013).

5.1. Non-cooperative optimal tariff theory

A benevolent social planner will weigh efficiency losses against terms of trade gains when setting tariffs. Generally, the planner will maximize the sum of household income (Y^h) and consumer surpluses (ψ_g), which yields the social welfare function $W = \Sigma_h \left(Y^h + \Sigma_g \psi_g\right)$. Consumers in country I obtain utility through their consumption of a numeraire good, denoted c_0^h, and consumption of imported and domestic composites of goods, generally denoted as c_g^h. Utility is given by $U = c_0^h + \Sigma_g u_g\left(c_g^h\right)$. This quasilinear structure rules out income effects in the model so that we can focus on trade flows and trade policy. While many factors may influence the relationship between shipped and delivered prices (e.g., exchange rates), we will focus explicitly on the role of tariffs. As such, we can write the delivered price as the shipped price scaled by the tariff, $p_{ijg} = \left(1+\tau_g^I\right) p_{ijg}^J$. Here we will suppress the importer superscript on prices and quantities for convenience. For now, assume that imported varieties of a particular good are subject to an identical tariff. Specifically, $\tau_{ijg}^I = \tau_g^I \quad \forall \quad v$.This assumption is not critical for the results, and the model where importers can set multiple tariffs for a given good will be discussed subsequently.

We follow the assumptions laid out by Grossman and Helpman (1994, 1995) to focus on the analysis. There is a long and storied literature surrounding importer motives to set optimal trade policy dating back to Torrens (1833) and Mill (1874). For a more comprehensive discussion of the history of optimal trade policy see Irwin (1996). For now we focus on the terms of trade motives for policy. The numeraire is freely traded in a perfectly competitive market, produced according to constant returns to scale using only labor as an input, and its price is normalized to unity. These assumptions taken together imply unit wages in the economy. Additionally, domestic varieties are produced under constant returns to scale using labor and a specific factor earning quasi-rent π_d. Lastly, tariff revenues are good specific and redistributed uniformly as r_g. Individuals own one unit of labor and a subset of them own at most one unit of industry specific capital.

Normalizing the population to one in conjunction with our assumptions regarding the numeraire and utility from Eq. (1) yields the social welfare problem,

$$\arg\max_{\tau_g^I} W = 1 + \pi_d + \sum_g \underbrace{\sum_v \tau_g^I p_{ijg}^J x_{ijg}}_{r_g} + \underbrace{log\left(\left(\sum_v b_{ijg}^{\frac{1}{\sigma_g^I}} x_{ijg}^{\frac{\sigma_g^I-1}{\sigma_g^I}}\right)^{\frac{\sigma_g^I}{\sigma_g^I-1}}\right) - p_{ijg}x_{ijg}}_{\psi_g}. \tag{12}$$

Since worker rents are separable, the above problem amounts to choosing tariffs for each good that balance tariff revenues from terms of trade gains against changes in consumer surplus. Noting the change in consumer surplus with respect to the tariff is $\frac{\partial\psi_g}{\partial\tau_g^I} = x_{ijg}\frac{\partial p_{ijg}}{\partial\tau_g^I}$, by the envelope theorem, and $\frac{\partial p_{ijg}}{\partial\tau_g^I} = \left(1+\tau_g^I\right)\frac{\partial p_{ijg}^J}{\partial\tau_g^I} + p_{ijg}^J$, from the relation between shipped and delivered prices, the first order condition from Eq. (12) is derived for each good as,

$$\sum_v \left(\tau_g^I p_{ijg}^J \frac{\partial x_{ijg}}{\partial\tau_g^I} - x_{ijg}\frac{\partial p_{ijg}^J}{\partial\tau_g^I}\right) = 0. \tag{13}$$

The first term in Eq. (13) are the efficiency costs associated with the tariff, as consumers decrease their imports when exporters pass through the cost of the tariff to delivered prices. The second term represents importer terms of trade gains, as exporters partially absorb the tariff in their shipped price. It is intuitive to rewrite the preceding in terms of elasticities with respect to the tariff:

$$\sum_v \left(p_{ijg}^J x_{ijg}\frac{\partial x_{ijg}}{\partial\tau_g^I}\frac{\tau_g^I}{x_{ijg}} - p_{ijg}^J x_{ijg}\frac{\partial p_{ijg}^J}{\partial\tau_g^I}\frac{\tau_g^I}{p_{ijg}^J}\frac{1}{\tau_g^I}\right) = 0. \tag{14}$$

Next we simplify Eq. (14) given the assumptions of the preceding trade model.

Proposition 1. *The optimal tariff for good g* $\left(\tau_g^{I^*}\right)$ *with exporter heterogeneity is,*

$$\tau_g^{I^*} = \frac{\sum_j p_{ijg}^J x_{ijg}\frac{\omega_{ijg}^I}{1+\omega_{ijg}^I\sigma_g^I}}{\sum_j p_{ijg}^J x_{ijg}\frac{1}{1+\omega_{ijg}^I\sigma_g^I}} \tag{15}$$

given our quantitative trade model.

Proof. *Combining the elasticities of price and quantity with respect to the tariff in our quantitative trade model with Eq* (14) *yields the importer's welfare-maximizing first-order condition,*

$$\sum_v p_{ijg}^J x_{ijg}\left(\frac{-\left(1+\bar{\omega}_{ig}^I\sigma_g^I\right)}{\left(1+\bar{\omega}_{ig}^I\right)\left(1+\omega_{ijg}^I\sigma_g^I\right)}\frac{\tau_g^I}{1+\tau_g^I} + \frac{\omega_{ijg}^I\left(1+\bar{\omega}_{ig}^I\sigma_g^I\right)}{\left(1+\bar{\omega}_{ig}^I\right)\left(1+\omega_{ijg}^I\sigma_g^I\right)}\frac{\tau_g^I}{1+\tau_g^I}\frac{1}{\tau_g^I}\right) = 0.$$

Rearranging the first-order condition and solving for τ_g^I *, yield the welfare maximizing tariff for country I importing good g. The full derivation is relegated to Appendix D.*

In essence, the optimal tariff is a trade weighted average of variety-level responses to policy. The numerator represents the total terms of trade gains of the tariff, while the denominator is the total efficiency loss in the industry from the tariff. The social planner thus chooses the tariff that optimally weights the terms of trade gains relative to the efficiency losses by the importance of each of the importer's trading partners. In the canonical case, where the inverse export supply elasticities are homogeneous, the planner will set a tariff equal to the common inverse supply elasticity since all varieties yield an identical relationship between terms of trade gains and efficiency losses. Trivially, when $\omega^I_{ijg} = \omega^I_g$, Eq. (15) reduces to $\tau^{I*}_g = \omega^I_g$. Here, the planner takes into account the full distribution of export supply curves along with their interactions with the import demand curve as supply shifts for all varieties in response to a tariff. Put more plainly, when the importer applies an identical tariff across multiple exporters with different export supply elasticities, each exporter yields a different terms of trade gain relative to its efficiency loss. The optimal tariff therefore is the tariff that optimally weights each exporter's contribution to its total terms of trade gains and efficiency losses.

While we have assumed that the importer is only able to set a single tariff rate per good, the theory is flexible enough to accommodate any number of tariff rates. For example, suppose the importer divides the varieties of a good into two subsets ($\mathcal{A}$ and $\mathcal{B}$) and sets tariffs independently for each (e.g., MFN tariffs set by WTO members). Then there would be two optimal tariffs for the good:

$$\tau^{I*}_{g\mathcal{A}} = \frac{\sum_{j\in\mathcal{A}} p^J_{ijg} x_{ijg} \frac{\omega^I_{ijg}}{1+\omega^I_{ijg}\sigma^I_g}}{\sum_{j\in\mathcal{A}} p^J_{ijg} x_{ijg} \frac{1}{1+\omega^I_{ijg}\sigma^I_g}} \quad \forall \quad j \in \mathcal{A} \quad \text{and}$$

$$\tau^{I*}_{g\mathcal{B}} = \frac{\sum_{j\in\mathcal{B}} p^J_{ijg} x_{ijg} \frac{\omega^I_{ijg}}{1+\omega^I_{ijg}\sigma^I_g}}{\sum_{j\in\mathcal{B}} p^J_{ijg} x_{ijg} \frac{1}{1+\omega^I_{ijg}\sigma^I_g}} \quad \forall \quad j \in \mathcal{B}.$$

This analysis can be extended to any division of the varieties of goods in this model. Also notice the parallel with the homogeneous elasticity case. If the importer perfectly discriminates across varieties (i.e., sets a unique tariff for each imported variety), the optimal tariff is $\tau^{I*}_{ijg} = \frac{p^J_{ijg} x_{ijg} \frac{\omega^I_{ijg}}{1+\omega^I_{ijg}\sigma^I_g}}{p^J_{ijg} x_{ijg} \frac{1}{1+\omega^I_{ijg}\sigma^I_g}} = \omega^I_{ijg}$. Since we never see perfect discrimination in the data, we do not examine this case in the following. However, we will define the optimal tariff across exporters for each tariff rate we see in the data in the following empirical analysis. To provide a clear example, WTO members setting different tariffs for other WTO members and non-members will have two optimal tariffs per good across members and non-members. Notably, our empirical results are not sensitive to these assumptions.

In contrast to the literature, Ludema and Mayda (2013) argue that market concentration (they rely on HHIs) is important for identifying a relationship between applied tariffs and terms of trade motives when export supply elasticities are homogeneous. To some extent, our optimal tariff internalizes the importance of industry concentration described by Ludema and Mayda (2013). That is to say, introducing exporter heterogeneity bears a resemblance to the channels described in Ludema and Mayda (2013) as it optimally weights each exporter's contribution to terms of trade and efficiency from tariffs. The difference here is that exporter heterogeneity highlights the importance of market composition when setting tariffs even in a non-cooperative policy setting.

5.2. Empirical evaluation

We next investigate how incorporating export supply heterogeneity into optimal trade policy relates to applied tariffs. The preceding results warrant a fully inclusive analysis of the relationship between optimal and applied tariffs. Here we analyze the full sample of tariffs and trade flows across all countries over 1991–2007.[31] We combine tariff data with *CEPII*, which records when countries joined the WTO by signing GATT plus various country characteristics. Finally, we concord the tariff data to the HS4 level and combine them with our estimates and *Comtrade* to construct optimal tariffs. Our aim is to demonstrate how to leverage the ability of our optimal tariff estimates to adjust to changes in the composition of trade over time. This will allow us to uncover trade policy responses to the evolving terms of trade motives embodied by our estimates.

Our goal is to analyze whether, regardless of the state of trade policy (e.g., GATT and RTAs), importers' applied tariffs relate to optimal non-cooperative tariffs when we incorporate exporter heterogeneity. It is however worth reiterating that our sample period was a far reaching period of global cooperation in trade policy. From 1991 to 2007, according to the *CEPII*, the world saw 36 countries sign the GATT, 160 regional trade agreements spanning 1396 importer-exporter pairs implemented, and substantial reductions in applied tariffs. Additionally, WTO members during this period commonly restructured bound tariff rates to below current applied tariffs, suggesting that cooperative policy negotiations were becoming more and more binding.

Anecdotally these trends suggest a diminishing role for non-cooperative tariff motives, as targeting terms of trade gains give way to cooperative incentives. Our analysis is designed to ask whether applied tariffs still respond to non-cooperative tariff motives in modern data. Our goal thus precludes more in-depth analysis of some potentially interesting bilateral policies in recent years, as we are interested in understanding whether, even in light of complicated policy environments, importers' applied tariffs relate to terms of trade motives underlying non-cooperative policy theory.

Our empirical exercise will consider various regressions of applied tariffs on optimal tariffs and controls. Explicitly, our framework will estimate the relationship,

$$\tau^I_{ijgt} = \alpha \log\left(\tau^{I*}_{ijgt}\right) + \beta X_{ijgt} + \varphi_{ijg} + \varepsilon_{ijgt}. \tag{16}$$

Applied tariffs can be importer-exporter-product specific in the full sample, and are denoted τ^I_{ijgt}.[32] From our estimates we calculate the optimal non-cooperative tariff, τ^{I*}_{ijgt}, which we will include in logs in order to address the long right tail of the estimates. If importers apply tariffs for goods in response to non-cooperative tariff motives we expect to estimate a positive relationship $\alpha > 0$. Our goal is to examine whether, regardless of the type of policy setting in reality, applied tariffs follow non-cooperative (terms of trade) motives globally. As such, we will also consider trade policy controls (e.g., bilateral WTO and RTA indicators) denoted by X_{ijgt}. Finally, our identification strategy will leverage the unique ability of our optimal tariff to adjust over time, and our preferred specification will include importer-exporter-product fixed effects φ_{ijg}.

Table 6 presents summary statistics of the applied and optimal tariffs across regions in the data. Within sample rankings of the average tariffs are also displayed. Countries setting some of the

[31] Tariff data were generously furnished by Robert Feenstra and John Romalis. They are the same utilized by Feenstra and Romalis (2014), which include a detailed description in their Appendix B.

[32] However, note that for WTO members applied tariffs are generally importer-WTO member-product specific, which we will take into account when clustering standard errors in the following.

Table 6
Summary statistics: Applied and optimal tariffs.

Importer	Applied tariff†			Optimal tariff ($\tau_g^{I^*}$)		
	Mean	Median	Rank	Mean	Median	Rank
Australia	6.46%	4.66%	13	112.7%	75.6%	17
Brazil	15.55%	14.20%	5	114.1%	74.3%	16
Canada	6.45%	3.40%	14	206.0%	128.3%	7
China & Hong Kong	19.50%	14.15%	2	204.0%	123.7%	8
Germany	5.77%	3.62%	17	141.3%	114.7%	13
France	5.88%	3.75%	15	107.1%	81.5%	19
Great Britain	5.85%	3.68%	16	108.0%	81.3%	18
India	34.99%	30.77%	1	104.4%	70.0%	21
Italy	5.71%	3.62%	20	105.7%	77.6%	20
Japan	5.74%	2.74%	19	157.9%	118.6%	12
Mexico	15.74%	15.25%	4	217.4%	111.7%	6
Russia	11.42%	10.17%	7	122.5%	80.0%	15
United States	4.38%	2.63%	21	140.4%	114.5%	14
African	17.17%	14.23%	3	1070.1%	66.8%	1
Asian	11.10%	7.00%	8	161.4%	63.0%	11
Caribbean	13.84%	15.00%	6	438.5%	86.6%	3
Northern/Western Europe	5.76%	3.39%	18	183.0%	79.3%	10
Oceania	6.84%	4.67%	12	815.7%	78.4%	2
South American	10.66%	10.00%	9	276.4%	74.9%	4
Southern/Eastern Europe	8.81%	8.28%	11	192.6%	79.4%	9
World	10.19%	7.78%	10	262.3%	78.9%	5

Notes: †Applied tariffs are from Feenstra and Romalis (2014) over 1991–2007. For exposition, the top and bottom 0.5% of $\tau_g^{I^*}$ within each country are dropped. Mean is the average within and then across years. Median is the median in the full sample. Rank is the order of the mean from highest to lowest across the listed countries and regions.

Table 7
The relationship between applied and optimal tariffs.

Controls	Applied tariff					
	(1)	(2)	(3)	(4)	(5)	(6)
log(Optimal tariff)	0.227***	0.185***	0.219***	0.216***	0.175***	0.122***
	(0.013)	(0.028)	(0.013)	(0.013)	(0.026)	(0.026)
Exporter WTO						−2.534***
						(0.037)
Importer WTO						−3.825***
						(0.070)
Regional Trade Agreement						−1.771***
						(0.032)
Product FEs	Yes	No	Yes	Yes	No	No
Importer FEs	Yes	No	No	No	No	No
ImporterXYear FEs	No	No	Yes	Yes	No	No
ExporterXYear FEs	No	No	No	Yes	No	No
ProductXImporter FEs	No	Yes	No	No	No	No
ProductXImporterXExporter FEs	No	No	No	No	Yes	Yes
R^2	0.412	0.754	0.548	0.550	0.771	0.783
Obs.	18,426,802	18,426,569	18,426,794	18,426,738	18,289,462	16,739,732

Notes: Each column regresses applied tariffs on optimal tariffs under heterogeneity. Importer and exporter WTO are indicators denoting WTO membership while Regional Trade Agreement indicates a current trade agreement. Fixed effects are denoted as FEs, where product is the HS4 level. Standard errors clustered two ways at the importer-WTO member and exporter-product levels are in parentheses, where * p<0.10, ** p<0.05, *** p<0.01 indicate significance.

largest applied tariffs include China and Mexico along with African and Caribbean countries. These countries also tend to have large estimated optimal tariffs. More developed countries tend to set lower tariffs on average. The developed countries setting the lowest tariffs are Italy, Germany, Japan and the US. Surprisingly, even though these countries have the strongest unconditional market power in our estimates (i.e., large average inverse export supply elasticities as highlighted by Table 4), they actually generate some of the lowest estimated optimal tariffs.

This result is unique to the literature. Non-cooperative optimal tariffs estimated under homogeneity are unable to reconcile the low tariffs set by developed countries that also have large estimated average inverse export supply elasticities (i.e., strong market power) without incorporating alternative trade policy motives. Here, this feature of the data is readily explained by simple non-cooperative policy and our estimates through marked heterogeneity across export supply mixed with the composition of these varieties consumed by importers. Notably, our optimal tariff estimates for the full sample are quite similar to recent general equilibrium analysis of optimal tariffs (cf., Ossa, 2014).

Table 7 begins by examining the correlations between applied and optimal tariffs worldwide. Column (1) regresses applied tariffs on optimal tariffs with only importer and product fixed effects. We see that even in the full sample there is a statistically significant positive relationship between applied and optimal tariffs.[33] Column (2) pushes the data further by including importer-product fixed effects. This specification absorbs any importer-product specific effects that

[33] Standard errors are clustered two ways at the importer-WTO member and exporter-product levels to allow for possible correlation in residuals across most favored tariff rates within the WTO and over time for the exporter. We have investigated alternative clustering methods including at the level of the included fixed effects, but opt for the current method as it tends to yield larger standard errors and seems the most sensible given trade policies.

do not vary over time (e.g., constant endogenous lobbying for protection). The coefficient is identified within importers across exporters over time. The estimated relationship remains positive and significant at the 1% level, and is comparable in magnitude to Column (1). Additionally, the explanatory power of the model increases as the adjusted R^2 almost doubles to 0.754. Columns (3) and (4) control for country-year fixed effects. The variation in the data identifying our estimates is importer-exporter-product specific over time. We can see that the relationship between applied and optimal tariffs is still positive and statistically significant of a similar magnitude to Column (1). Additionally, the adjusted R^2 has fallen relative to Column (2) suggesting more explanatory power lies in importer-product fixed effects.

Our estimates produce variation over time in the terms of trade motives borne by export supply elasticities underlying optimal non-cooperative tariffs. Columns (5) and (6) further exploit this variation by controlling for importer by exporter by product fixed effects, such that our estimates are identified solely by variation over time. This strategy has the benefit of absorbing any trade policy targeted at a particular exporter of a given product. Column (5) yields similar results to Column (2) suggesting that applied tariffs are mainly determined within importer-product pairs. Still, isolating the variation of tariff changes over time yields a statistically significant positive relationship between applied and optimal tariffs.

Column (6) checks the robustness of Column (5) by including policy regime changes in the form of importers or exporters signing GATT or a regional trade agreement. These regime variables are identified by countries that switch status over time. As we should expect, importers and exporters signing the GATT (i.e., joining the WTO) or a regional trade agreement set lower tariffs. Even after controlling for these regime driven motives of tariff setting there is still a statistically significant positive relationship between applied and optimal tariffs.[34]

In summary, previous studies note the difficulty of identifying terms of trade motives driving tariffs set by WTO members (cf., Bagwell and Staiger (2011) who argue that tariff reductions by WTO members are geared to mitigate terms of trade motives). Ludema and Mayda (2013) develop a model of cooperative tariff negotiation. They introduce heterogeneity along the dimension of negotiating power. Consequently, they argue that market concentration is an important control when relating terms of trade motives to applied tariffs in the data. The bargaining structure in their model provides tools for uncovering that most favored nations tariffs follow patterns of importer market power. Beshkar et al. (2015) analyze that how countries optimally negotiate bound rates and how tariff overhang (i.e., the difference between bound and applied tariffs) depend on the characteristics of the importers and exporters (including importer market power). Broda et al. (2008) found that US market power influences trade policy only for trade barriers not set cooperatively under the WTO (e.g., non-tariff barriers). Here we find that exporter heterogeneity in the model transforms the standard measures of market power and terms of trade gains by incorporating the composition of trade. Time series variation in our estimates then allows us a new avenue for identification unavailable to these previous studies that rely on homogeneous elasticities. Controlling for any importer by exporter by good motives driving applied tariffs identifies a statistically significant positive relationship between market power, as embodied by optimal tariffs, and applied tariffs globally. This result persists even after controlling for structural changes in how policy is set through trade agreements such as the GATT and RTAs.

5.2.1. WTO status

We have shown that applied tariffs relate to optimal non-cooperative tariffs even when we expect importers to follow cooperative policies (e.g., via WTO membership). To further examine the role of joining the WTO, we can split our sample into subsets of countries according to their membership, and thereby our expectations of their propensity to follow cooperative policies. We might expect that our preceding results, which pool all countries, are inappropriate for countries that have been WTO members for the entirety of our sample. Table 8 divides the sample into countries that were either always WTO members, never WTO members, or switched status at some point in the sample, and repeats our analysis.[35]

Column (1) presents our preferred specification, Column (6) of Table 7, for reference. We see that countries who have been WTO members for the entire sample (e.g., the US) are less responsive to the terms of trade motives embodied by the optimal tariffs than the full sample. This result is not surprising as we expect these countries to set tariffs (more) cooperatively. Yet, we still find evidence that there is a significant positive relationship between applied and optimal tariffs even for these countries. Somewhat surprisingly, countries that are yet to join the WTO (e.g., Russia) only slightly set tariffs more aggressively in response to terms of trade motives. We suspect that this is partly due to strong regional agreements for these countries, as suggested by the large negative effect of RTAs on applied tariffs.

Countries that switch WTO status in the sample (e.g., China) have statistically larger positive relationship between applied and optimal tariffs than the previous groups. This result potentially yields insight into why these countries were targeted to join the WTO. Upon their membership, the coefficient on the interaction of WTO membership and the optimal tariff demonstrates that these switching countries restructure their applied tariffs away from the optimal tariff. These results thus reinforce the notion that the WTO helps to move importers toward a cooperative tariff equilibrium. However, even under the purview of the WTO, countries switching status still associate their applied and optimal tariffs. Notice that the net effect of the optimal tariff after they join the WTO is still statistically significant and positive. Our estimates imply that doubling the optimal tariff corresponds with 0.2% larger applied tariffs even after countries join the WTO.

Our results provide evidence that importers become less responsive to non-cooperative (i.e., terms of trade) motives when setting tariffs following WTO membership. However, they also highlight a positive relationship between non-cooperative optimal tariffs and applied tariffs even after WTO membership. This highlights that neglecting non-cooperative motives when evaluating policy even under seemingly cooperative regimes overlooks a substantial channel for tariff setting in the data.

[34] The Appendix presents a wide range of robustness checks for this result. Including, extending the model to control for Grossman and Helpman (1995) lobbying motives that may drive tariffs observed in the data. Our estimates continue to yield a positive relationship between our optimal and applied tariffs over time and worldwide. Additionally, we augment the data to include bound tariff rates for WTO members. We confirm the predictions of Beshkar et al. (2015) and uncover new patterns relating newly negotiated bound rates to terms of trade motives in the data. Additionally, these regressions use year to year variation in tariffs for identification. One may be concerned that tariffs do not adjust at this frequency in practice. As such, we have repeated our preferred specification in Column (6) at three and five year intervals and differences. The results are available upon request, but are qualitatively identical to the presented estimates.

[35] It should also be noted that one need not be concerned that our elasticity estimates are potentially biased due to extensive margin adjustments before and after WTO accession. Trade flows that exist only before or after an importer's accession make up a small fraction of the data. Additionally the elasticities estimated from these trade flows show no statistical differences from the entire sample (these results are available upon request).

Table 8
The relationship between applied and optimal tariffs: WTO subsets.

	Applied tariff				
WTO status:	Pooled	Always	Never	Switchers	
Controls	(1)	(2)	(3)	(4)	(5)
log(Optimal tariff)	0.122***	0.031***	0.042***	0.413***	0.541***
	(0.026)	(0.012)	(0.014)	(0.021)	(0.023)
log(Optimal tariff)*Importer WTO					−0.368***
					(0.008)
Exporter WTO	−2.534***	−3.347***	−1.988***	−0.087***	−0.106***
	(0.037)	(0.022)	(0.065)	(0.027)	(0.027)
Importer WTO	−3.825***			−4.349***	−4.425***
	(0.070)			(0.014)	(0.014)
Regional Trade Agreement	−1.771***	−2.462***	−3.214***	0.578***	0.581***
	(0.032)	(0.014)	(0.055)	(0.019)	(0.019)
ProductXImporterXExporter FEs	Yes	Yes	Yes	Yes	Yes
R^2	0.783	0.775	0.775	0.775	0.775
Obs.	16,739,732	11,319,645	2,119,617	3,300,470	3,300,470
Countries	173	99	38	36	36

Notes: Each column regresses applied tariffs on optimal tariffs under heterogeneity. Fixed effects are denoted as FEs, where product is the HS4 level. Standard errors clustered two ways at the importer-WTO member and exporter-product levels are in parentheses, where * p<0.10, ** p<0.05, *** p<0.01 indicate significance.

5.3. Endogeneity

Finally, we might be concerned about endogeneity in our preferred regression – Column (6) of Table 7 – from two sources. Both sources are a product of the scope of the analysis (i.e., including all importers regardless of their institutions) and policy trends at the time. Along with the abundance of new cooperative policy agreements, countries in existing agreements were also actively negotiation reductions in tariff barriers (e.g., WTO members lower tariff bounds below applied tariff levels). New and existing policies should thus be expected to affect trade and tariffs in ways not captured by our simple non-cooperative policy theory.

First, we expect reverse causality in the preceding regressions as our optimal tariffs are determined using trade share weights within an imported product. When we look within imports over time, changes in applied tariffs could be mechanically correlated with import values such that a decrease (increase) in the tariff leads to a larger decrease (increase) in the weights on more elastically supplied exports. This relationship is clear when we differentiate the log of the optimal tariff (our regressor) with respect to the applied tariff (our regressand) in our model:[36]

$$\frac{\partial log\left(\tau_g^{I*}\right)}{\partial\tau_g^I} = \frac{\sum_j \frac{\partial p_{ijg}^J x_{ijg}}{\partial\tau_g^I}\frac{1}{1+\omega_{ijg}\sigma_g}\left(\frac{\omega_{ijg}}{\tau_g^{I*}}-1\right)}{\sum_j p_{ijg}^J x_{ijg}\frac{1}{1+\omega_{ijg}\sigma_g}}.$$

This derivative is strictly positive given the assumptions of the model.[37] The implication for our estimation is thus reverse causality should bias our coefficient estimates upward. To control for the mechanical bias in our estimation, notice that, given our elasticity estimates, this derivative can be calculated in the data. Specifically, we can write the optimal tariff in the current period as a function of the optimal tariff in the previous period, the change in the optimal tariff due to changes in applied tariffs, and a residual effect such that (matching our notation from the regression model):

$$\text{log(Optimal Tariff)}_t = \text{log(Optimal Tariff)}_{t-1} + \underbrace{\frac{\Delta\text{log(Optimal Tariff)}}{\Delta\text{Applied Tariff}}}_{\frac{\partial log\left(\tau_g^{I*}\right)}{\partial\tau_g^I}*\Delta\tau_g^I} + \zeta_t.$$

Controlling for the lagged optimal tariff and the change in the optimal tariff due to changes in applied tariffs will leave the portion of the optimal tariff free from first-order reverse causality (i.e., ζ).

Our second endogeneity concern is the possibility that lower applied tariffs might be a product of countries aiming to reduce terms of trade externalities (cf., Bagwell and Staiger, 2011). Consequently, this implies a negative relationship between the optimal and applied tariffs, which would bias our coefficients down. We confront this endogeneity concern and the issue of reverse causality through an instrumental variables approach that borrows from the gravity literature to instrument for the trade weights.

Table 9 presents our IV results. Column (1) begins by displaying our preferred specification which directly regresses applied tariffs on the log of the optimal tariff including controls for trade agreements and importer-exporter-product fixed effects. In Columns (2)–(3), we use the model to control for the change in the optimal tariff due to the observed change in the applied tariff in the data. To do so, we estimate the first derivative of the optimal tariff with respect to the applied tariff. We can then multiply this derivative in period $t-1$ by the change in the applied tariff to period t. We call this estimate $\frac{\Delta\text{log(Optimal Tariff)}}{\Delta\text{Applied Tariff}}$, which according to the model is the portion of the optimal tariff in period t that resulted from any change in applied tariffs as we go from $t-1$ to t. In Column (2) we control for the derivative along with the lag of the optimal tariff directly in our preferred regression. The data suggest a weakly significant positive relationship between $\frac{\Delta\text{log(Optimal Tariff)}}{\Delta\text{Applied Tariff}}$ and applied tariffs. As expected, the estimated coefficient on optimal tariffs falls. However, we still estimate a statistically significant positive relationship between applied and optimal tariffs even after controlling for reverse causality.

[36] Recall the relationship between the value of trade $p_{ijg}x_{ijg}$ and applied tariffs is:

$$\frac{\partial p_{ijg}^J x_{ijg}}{\partial\tau_g^I} = \frac{-\left(1+\omega_{ijg}^I\right)\left(1+\bar{\omega}_{ig}^I\sigma_g^I\right)}{\left(1+\bar{\omega}_{ig}^I\right)\left(1+\omega_{ijg}^I\sigma_g^I\right)}\frac{p_{ijg}^J x_{ijg}}{1+\tau_g^I}$$

[37] Thanks are due to an anonymous referee for a number of valuable comments, with one specifically highlighting this derivation. The proof is available upon request.

Table 9
The relationship between applied and optimal tariffs: Addressing endogeneity.

	Applied tariff					
Controls	(1)	(2)	(3)	(4)	(5)	(6)
log(Optimal tariff)	0.122***	0.108***	0.106***	0.092**	0.347***	0.267***
	(0.026)	(0.006)	(0.020)	(0.040)	(0.041)	(0.038)
Exporter GDP ($Bill)						−0.001***
						(0.000)
Importer GDP ($Bill)						−0.001***
						(0.000)
$\frac{\Delta log(OptimalTariff)}{\Delta AppliedTariff}$		0.001				
		(0.001)				
$log(OptimalTariff)_{t-1}$		0.063***				
		(0.017)				
Exporter WTO	−2.511***	−2.427***	−2.427***	−2.534***	−2.569***	−1.628***
	(0.037)	(0.036)	(0.010)	(0.026)	(0.010)	(0.010)
Importer WTO	−3.819***	−3.583***	−3.585***	−3.826***	−3.863***	−3.193***
	(0.070)	(0.062)	(0.008)	(0.097)	(0.009)	(0.008)
Regional Trade Agreement	−1.767***	−1.699***	−1.700***	−1.772***	−1.720***	−1.391***
	(0.032)	(0.032)	(0.006)	(0.033)	(0.006)	(0.006)
ProductXImporterXExporter FEs	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.783	0.802	0.833	0.835	0.834	0.863
Obs.	16,739,732	14,641,405	14,641,405	16,739,732	16,579,932	16,505,471
First stage results:	Baseline	Derivative	Purged	Reweighted	GDP weight	GDP weight
Coefficient			0.988***	0.609***	0.758***	0.760***
			(0.005)	(0.008)	(0.012)	(0.012)
F-Stat			45,184	5921	3889	4014

Notes: Each column regresses applied tariffs on optimal tariffs under heterogeneity. Importer and Exporter WTO are indicators denoting WTO membership while Regional Trade Agreement indicates a current trade agreement. Fixed effects are denoted as FEs, where product is the HS4 level. Columns (2)–(3) control for the predicted effect of applied tariffs on optimal tariffs denoted by $\frac{\partial \log(\text{Optimal Tariff})}{\partial \text{Applied Tariff}}$. The final two columns apply an IV strategy where the optimal tariff weights are instrumented by importer and exporter GDP. Standard errors clustered two ways at the importer-WTO member and exporter-product levels are in parentheses, where * $p<0.10$, ** $p<0.05$, *** $p<0.01$ indicate significance.

Column (3) takes the intuition of Column (2) a step further. Our goal is to purge the optimal tariff of variation due to changes in applied tariffs. To do so, we regress the optimal tariff on $\frac{\Delta \log(\text{Optimal Tariff})}{\Delta \text{Applied Tariff}}$ plus $\log(\text{Optimal Tariff})_{t-1}$ as suggested by the theory. The predicted residuals from this regression are optimal tariffs less any changes due to changes in applied tariffs. We can then use these residual estimates as instruments for optimal tariffs. The resulting first stage coefficient is a statistically significant positive and presents a large F-statistic. The second stage coefficient is again smaller than our baseline, which reaffirms the upward bias from reverse causality. Nonetheless, the estimated coefficient remains positive and statistically significant.

One may be concerned that purging the optimal tariff in the preceding manner confounds non-linearities in changes in applied tariffs and the optimal tariff formula. Column (4) thus considers a related variant of purging the variation in the optimal tariff due to changes in applied tariffs. Similar to the preceding, we use the model to predict the trade shares in the optimal tariff formula less the change in trade flows due to changes in applied tariffs. Subtracting off the portion of the value of trade due to changes in tariffs given our estimates and the model, we can recalculate the optimal tariff using counterfactual trade weights. Column (4) then instruments the optimal tariff using this counterfactual optimal tariff. The results are qualitatively identical to purging the optimal tariff directly of variation due to changes in applied tariffs.

While filtering changes in applied tariffs through the model supports the assertion that reverse causality in the trade weights is biasing our estimates upward, we also consider an instrument outside of the model to simultaneously address reverse causality and other potential endogeneity concerns. The gravity literature posits a positive relationship between country sizes (i.e., importer and exporter GDP) and trade flows. Additionally, we expect that changes in an importer's applied tariff at our level of aggregation (HS4) are unlikely to have any measurable impact on either importers' or exporters' GDP. We can thus recalculate our optimal tariff replacing the trade weights with importer times exporter GDP to construct a valid instrument for our optimal tariff.[38] Columns (4)–(5) apply this IV approach. In both specifications the first stage yields a positive relationship between our instrument and the optimal tariff and large F-statistics.[39] In Column (4) we see that the estimated relationship between applied and optimal tariffs nearly triples, which suggests, even though reverse causality biases our OLS estimates upward, other unobserved mechanisms relating market power to applied tariffs in the data are attenuating our estimates to a greater degree.

While the gravity literature not only establishes the strong relationship between GDPs and trade, it also implicitly assumes that GDP and tariffs are uncorrelated. Nonetheless, Column (5) acknowledges the possibility that GDP may not be excludable from the second stage. In other words, while we have addressed reverse causality from tariffs to GDP, we might be concerned that country size (GDP) is correlated with applied tariffs over time. For instance, we might expect relatively large or fast growing countries to set lower tariffs in general, and against other large or fast growing countries in particular. Column (5) thus also controls for GDP in the second stage. We

[38] Explicitly, we instrument for our optimal tariff $\tau_g^{I*} = \frac{\sum_j p_{ijg}^J x_{ijg} \frac{\omega_{ijg}^I}{1+\omega_{ijg}^I \sigma_g^I}}{\sum_j p_{ijg}^J x_{ijg} \frac{1}{1+\omega_{ijg}^I \sigma_g^I}}$ with $\tau_g^{I^{Instrument}} = \frac{\sum_j GDP^I * GDP^J \frac{\omega_{ijg}^I}{1+\omega_{ijg}^I \sigma_g^I}}{\sum_j GDP^I * GDP^J \frac{1}{1+\omega_{ijg}^I \sigma_g^I}}$.

[39] Additionally, each of the IV regressions pass all standard tests. Specifically, Sargan-Hansen (overidentifying restrictions), GMM distance (endogeneity of regressor and instrument subsets), Stock-Wright (weak instruments) and Cragg-Donald (weak identification) all pass with p-values less than 0.00001.

do see evidence that large importers set lower tariffs on average, and even lower tariffs for their large partners. Even after controlling for this variation, our point estimate is positive and significant and rises from 0.122 in our baseline to 0.267.

6. Conclusion

After adapting a quantitative model of trade and trade policy to incorporate exporter and importer heterogeneity, we construct a flexible structural estimator of the underlying heterogeneous export supply and import demand elasticities. The estimates produce intuitive relationships with expected patterns of importer and exporter market power. Our estimates and theory imply a new measure for the optimal tariff set by an importer. These optimal tariffs are positively correlated with applied tariffs across a plethora of dimensions of the data. Additionally, the rich detail of the estimates and theory allow us to meaningfully decompose the channels of tariff setting across countries and products. Among other things, the data display intuitive patterns of importers targeting goods that generate pronounced terms of trade gains with higher tariff rates.

The degree of heterogeneity in the trade elasticity estimates produced here opens new avenues to deepening our understanding of a host of prominent theories. Supply elasticities are fundamental to evaluating core channels of globalization such as the degree of passthrough between countries (as in Goldberg and Knetter, 1997) and our estimates of trade indexes (restrictiveness as in Anderson and Neary, 1996 and prices as in Feenstra, 1994). The tractability of the estimator constructed in this paper lends itself naturally to extending these prevailing theories of inter-country relationships to include heterogeneity.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jinteco.2018.04.008.

References

Anderson, J., Neary, J., 1996. A new approach to evaluating trade policy. Rev. Econ. Stud.

Ardelean, A., Lugovskyy, V., 2010. Domestic productivity and variety gains from trade. J. Int. Econ. 80 (2), 280–291.

Armington, P.S., 1969. A theory of demand for products distinguished by place of production. Staff Papers - International Monetary Fund 16 (1), 159–178.

Bagwell, K., Staiger, R., 2011. What do trade negotiators negotiate about? Empirical evidence from the world trade organization. Am. Econ. Rev. 101 (4), 1238–1273.

Bas, M., Mayer, T., Thoenig, M., 2017. From micro to macro: demand, supply, and heterogeneity in the trade elasticity. J. Int. Econ.

Beshkar, M., Bond, E., Rho, Y., 2015. Tariff binding and overhang: theory and evidence. J. Int. Econ. 97, 1–13.

Blonigen, B.A., Soderbery, A., 2010. Measuring the benefits of foreign product variety with an accurate variety set. J. Int. Econ. 82 (2), 168–180.

Broda, C., Limão, N., Weinstein, D., 2008. Optimal tariffs and market power: the evidence. Am. Econ. Rev. 98 (5), 2032–2065.

Broda, C., Weinstein, D., 2006. Globalization and the gains from variety. Q. J. Econ. 121 (2), 541–585.

Feenstra, R., Romalis, J., 2014. International prices and endogenous quality. Q. J. Econ. 129 (2), 477–527.

Feenstra, R.C., 1994. New product varieties and the measurement of international prices. Am. Econ. Rev. 84 (1), 157–177.

Feenstra, R.C., 2009. Measuring the Gains from Trade under Monopolistic Competition. NBER Working Paper, , pp. 1–39.

Feenstra, R.C., Weinstein, D.E., 2017. Globalization, Markups, and US Welfare. J. Polit. Econ. 125 (4), 1040–1074. August.

Fernandes, A. M., Klenow, P. J., Meleshchuk, S., Pierola, M. D., 2015. The intensive margin in trade: moving beyond pareto. Mimeo.

Goldberg, P.K., Knetter, M., 1997. Goods prices and exchange rates: what have we learned? J. Econ. Lit. 35 (3), 1243–1272.

Grossman, G., Helpman, E., 1994. Protection for sale. Am. Econ. Rev. 84 (4), 833–850.

Grossman, G., Helpman, E., 1995. Trade wars and trade talks. J. Polit. Econ. 103 (4), 675–708.

Head, K., Mayer, T., Thoenig, M., 2014. Welfare and trade without pareto. Am. Econ. Rev. 104 (5), 310–316.

Hottman, C., Redding, S.J., Weinstein, D., 2016. Quantifying the sources of firm heterogeneity. Q. J. Econ. 131 (3), 1291–1364.

Hummels, D., Lugovskyy, V., 2006. Are matched partner trade statistics a usable measure of transportation costs? Rev. Int. Econ. 14 (1), 69–86.

Irwin, D., 1996. Against the Tide: An Intellectual History of Free Trade. Princeton University Press, Princeton.

Kemp, M., 1962. Errors of measurement and bias in estimates of import demand parameters. Econ. Rec. 38 (83), 369–372.

Leamer, E., 1981. Is it a demand curve, or is it a supply curve? Partial identification through inequality constraints. Rev. Econ. Stat. 63 (3), 319–327.

Ludema, R., Mayda, A., 2013. Do terms-of-trade effects matter for trade agreements? Theory and evidence from WTO countries*. Q. J. Econ. 128 (4), 1837–1893.

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

Mill, J., 1874. Essays on Some Unsettled Questions of Political Economy. 2 ed., Longman, London.

Ossa, R., 2014. Trade wars and trade talks with data †. Am. Econ. Rev. 104 (12), 4104–4146.

Rauch, J., 1999. Networks versus markets in international trade. J. Int. Econ. 48 (1), 7–35.

Simonovska, I., Waugh, M.E., 2014. The elasticity of trade: estimates and evidence. J. Int. Econ. 92 (1), 34–50.

Soderbery, A., 2015. Estimating import demand and supply elasticities: analysis and implications. J. Int. Econ. 96 (1), 1–17.

Tan, S., 2012. Structural estimation of a flexible Translog Gravity Model. Mimeo.

Torrens, R., 1833. Letters on Commercial Policy. Longman, London.