In Search of the Armington Elasticity*

Robert C. Feenstra   Philip Luck

Maurice Obstfeld   Katheryn N. Russ

May 19, 2017

Abstract

How big is the elasticity of substitution between goods from different countries-the Armington elasticity? Estimates of the "macro" elasticity between home and imported goods are often smaller than the "micro" elasticity between foreign sources of imports. Using new, highly disaggregate U.S. production data matched to imports and simulated data from a Melitz-style model with nested CES preferences, we explore estimation techniques for the two elasticities. For about three-quarters of sample goods there is no significant difference between the macro and micro elasticities, but for the rest the micro elasticity is significantly higher, even at the same level of disaggregation.

JEL Code: F14, F17, F40

---

*UC Davis and NBER, CU Denver, UC Berkeley NBER, and CEPR, and UC Davis and NBER. We thank Tom Becker, Zhiyuan Li, José Antonio Rodríguez-López, Anson Soderbery, and Greg Wright for sterling research assistance. We have benefitted from comments made by seminar participants at the Dartmouth College, Federal Reserve Bank of San Francisco, the Federal Reserve Bank of New York, Yale University, and the International Monetary Fund. We especially thank discussants James Harrigan and Menzie Chinn, as well as Andrés Rodríguez-Clare. Financial support was provided by the International Growth Centre at the London School of Economics, the Coleman Fung Risk Research Center at UC Berkeley, the National Science Foundation and the Sloan Foundation.

1 Introduction

The elasticity of substitution between goods produced in different countries – or the Armington (1969) elasticity – has long been one of the key parameters in international economics. Because it governs the strength of the relative demand response to relative international prices, this elasticity is central to understanding many features of the global economy. These include (with recent references) the role of international prices in trade balance adjustment (Imbs and Méjean, 2015) , the international transmission of business cycles (Heathcote and Perri, 2002), the effects of regional trade agreements (Romalis, 2007), and the welfare benefits of expanding world trade (Costinot and Rodríguez-Clare, 2014), and the possibility of multiple versus unique equilibria (Kucheryavyy, Lyn and Rodríguez-Clare, 2016).

Since at least the 1940s, economists have used both aggregate and disaggregate trade data in attempts to estimate the responsiveness of demand to international prices. Periodic comprehensive surveys by Machlup (1950), Orcutt (1950), Cheng (1959), Leamer and Stern (1970), Magee (1975), Stern, Francis, and Schumacher (1976), Goldstein and Khan (1985), Shiells, Stern, and Deardorff (1986), Marquez (2002), and McDaniel and Balistreri (2003), among others, document the growth over time in the supply of econometric studies on larger and increasingly detailed data sets. Yet despite an ever-expanding body of empirical study, there remains substantial uncertainty about the appropriate elasticity values to apply to different research and policy questions.

That uncertainty is reflected, in particular, in the elasticities that are used in computable general equilibrium models (CGE). Traditionally (for example, Harrison, Rutherford and Tarr 1997; Balstreri and Rutherford 2013; Hillberry and Hummels 2013), CGE models applied to international trade have used a nested CES structure on preferences, with an upper-level "macro" elasticity governing the substitution between home a foreign goods, and a lower-level "micro" elasticity governing the substitution between varieties of foreign goods. The calibrated values of the macro elasticity were lower than those of the micro elasticity, as justified by the differing elasticities estimated from data at various levels of aggregation.[1] Recently, however, the work of Dekle, Eaton and Kortum (2007, 2008) has spawned a new generation of computable models that *do not* allow for any difference between the macro and micro elasticities, but have a single elasticity in preferences between all product varieties, home or foreign. Calibration of these models instead relies on "trade costs" that differ for international versus domestic sales. We believe the absence of any difference between the micro and macro elasticities in this new generation of models can lead to substantial differences in results compared to those of the earlier CGE models.[2]

To evaluate the difference between the elasticity of substitution between home and foreign goods, and between varieties of foreign goods, requires two ingredients: a model that allows for such a nested CES structure and a dataset that has both home and foreign supplies at exactly the same level of disaggregation. We provide both these ingredients here. We build upon the general-equilibrium trade model growing out of work by Melitz (2003) and Chaney (2008), while allowing the Armington substitution elasticity between *domestic and foreign suppliers* to differ from that between *alternative foreign suppliers*, using a nested CES preference structure.[3]

[1]Goldstein and Khan (1985) survey a large body of research and endorse an earlier judgment by Harberger (1957) that for a typical country the price elasticity of import demand "lies in or above the range of −0.5 to −1.0...." More recent macro studies such as Heathcote and Perri (2002) and Bergin (2006) estimate aggregate substitution elasticities around unity. In apparently sharp contrast, recent studies of individual product groups such as Feenstra (1994), Lai and Trefler (2002), Broda and Weinstein (2006), Romalis (2007), and Imbs and Méjean (2015) tend to identify much stronger price responses.

[2]For example, contrast the results from the CGE models of China's growth and trade with the rest of the world in Tokarick (2012) and di Giovanni, Levchenko, and Zhang (2012).

[3]Ardelean and Lugovskyy (2010) and Mattoo, Mishra and Subramanian (2017) also introduce a nested CES structure, but do not allow for heteregeneous firms. See Rauch and Trindade (2002, 2003) for a micro-founded approach and Cosar et al. (2015) for industry-specific empirical evidence. Blonigen and Wilson (1999) identify other factors.

Section 2 develops the disaggregate (by good and country) import demand equation implied by the model. Endogeneity of the terms in this equation, along with measurement error due to the use of unit values rather than ideal price indexes in the estimation, introduces statistical biases. In section 3 we draw on Feenstra (1994) to propose a generalized method of moments (GMM) estimation strategy that – in large samples at least – corrects for the biases implied by our model. The moment condition we rely on is that the demand error is uncorrelated with the supply error for each country and good. Soderbury (2010, 2012) has recently identified small-sample biases in this estimator using simulated data.[4] That concern is addressed here by adding additional moment conditions. Our nested CES framework allows for a natural condition by aggregating the demand equation over foreign countries, to obtain an alternative equation for total imports relative to domestic demand, which only involves the macro Armington elasticity. We naturally refer to this as the "macro" demand equation. The corresponding moment condition is that the error of the macro demand equation is uncorrelated with the error of the macro supply equation for each good. As we discuss in section 3, this moment condition adds new information.

Estimation is performed in section 4 using simulated as well as U.S. data. The simulated dataset incorporates shocks to tastes and technologies that are calibrated from the literature. The U.S. dataset matches data on imports (by source country) and exports for about 100 goods with product-level data on U.S. production, and therefore implied apparent consumption. The U.S. production data are obtained from *Current Industrial Reports*, and our estimation is the first time that such data have been matched to the highly-disaggregate (Harmonized System, or HS) level for imports. The key feature of our data and results is that the macro and micro elasticities are estimated at the same level of disaggregation, that is, close to the same HS level. We find that in up to one-quarter of goods the micro elasticity is significantly higher than the macro elasticity, while in three-quarters of goods there is no significant difference between the macro and micro

[4]The potential for small-sample bias in the GMM estimator applied to other datasets was found earlier by Altonji and Segal (1996).

elasticities, when they are estimated at the same level of disaggregation. For about four-fifths of the goods, we fail to reject the hypothesis that the micro elasticity is twice as large as the macro elasticity.

In section 5 we briefly explore the implications of having different micro and macro elasticities for the gains from trade. Using our U.S. dataset we quantify the gains from trade first under the assumption that there is a single Armington elasticity and then letting the macro and micro elasticities differ, as implied by our estimates in section 4. When we relax the assumption that the macro and micro elasticity are the same our estimated gains from trade increase between three and six-fold depending on parameter values. Further conclusions are given in section 6 and various technical results are gathered in two on-line Appendices.

2 The Model

2.1 Preferences and Prices

There are J countries in the world and a fixed number G of different goods. Each country produces a range of distinct varieties of each good $g \in \{1, ..., G\}$, the set of varieties produced to be determined endogenously within our model.

In the classic Armington (1969) model, goods are differentiated not only by inherent differences in their characteristics, but also by their place of production. In the "home country" j, the representative consumer has a comprehensive consumption index given by

$$C^j = \left[\sum_{g=1}^{G} \left(\alpha_g^j\right)^{\frac{1}{\eta}} \left(C_g^j\right)^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}, \tag{1}$$

where the weights $\left\{\alpha_g^j\right\}_{g=1}^G$ are random preference shocks, with $\sum_g \alpha_g^j = 1$, and η is the elasticity of substitution between different goods. Since each good is produced in all countries, the Armington assumption does not become relevant until one defines the good-specific consumption sub-indexes $\left\{C_g^j\right\}_{g=1}^G$.

A general Armington setup differentiates products not only by their domestic or foreign origin, but also by the specific foreign country of origin. Define β_g^j as a random preference weight that country-j residents attach to domestically produced units of good g. We assume that

$$C_g^j = \left[\left(\beta_g^j\right)^{\frac{1}{\omega_g}} \left(C_g^{jj}\right)^{\frac{\omega_g - 1}{\omega_g}} + \left(1 - \beta_g^j\right)^{\frac{1}{\omega_g}} \left(C_g^{Fj}\right)^{\frac{\omega_g - 1}{\omega_g}} \right]^{\frac{\omega_g}{\omega_g - 1}}, \tag{2}$$

where C_g^{jj} denotes the consumption index of varieties of good g produced at home, C_g^{Fj} denotes the consumption aggregate of varieties of good g produced abroad, and ω_g is the substitution elasticity between home and foreign varieties of good g.

In turn, the country j foreign consumption index C_g^{Fj} depends on consumption from all possible sources of imports $i \neq j$, with random country-of-origin weights $\left\{\kappa_g^{ij}\right\}_{i \neq j}$, $\sum_{i \neq j} \kappa_g^{ij} = 1$:

$$C_g^{Fj} = \left[\sum_{i=1, i \neq j}^{J} \left(\kappa_g^{ij}\right)^{\frac{1}{\sigma_g}} \left(C_g^{ij}\right)^{\frac{\sigma_g - 1}{\sigma_g}} \right]^{\frac{\sigma_g}{\sigma_g - 1}}.$$

Here, σ_g is the elasticity of substitution between baskets of good g varieties originating in different potential exporters to country j, and we assume that this elasticity also applies *within* the consumption index C_g^{ij} of good g varieties imported from country i.

Denote the measure of varieties of good g that country j imports from country i by N_g^{ij}. (Country j will itself produce a measure N_g^{jj} of varieties for home consumption.) In our model, each set of measure N_g^{ij} is determined endogenously by a country-pair-specific fixed cost of trade and other factors to be described in detail below. Because σ_g also denotes the elasticity of substitution between different varieties φ of good g produced by a particular country i, then for all $i \in \{1, ..., J\}$,

$$C_g^{ij} = \left[\int_{N_g^{ij}} \left(c_g^{ij}(\varphi)^{\frac{\sigma_g - 1}{\sigma_g}} \mathrm{d}\varphi \right) \right]^{\frac{\sigma_g}{\sigma_g - 1}},$$

where the notation indicates that integration is done over a set of varieties that we indicate by its measure, N_g^{ij}.

The preceding preference setup defines a structure of canonical cost-of-living indexes and sub-indexes. The comprehensive consumer price index (CPI) for country j is

$$P^j = \left[\sum_{g=1}^{G} \alpha_g^j \left(P_g^j \right)^{1-\eta} \right]^{\frac{1}{1-\eta}}.$$

Corresponding to the consumption aggregator (2) for country j residents is a price index P_g^{jj} for varieties of good g produced at home and an index P_g^{Fj} for the aggregate of imported varieties. For example, the price index for imported goods P_g^{Fj} is given by

$$P_g^{Fj} = \left[\sum_{\substack{i=1 \\ i \neq j}}^{J} \kappa_g^{ij} \left(P_g^{ij} \right)^{1-\sigma_g} \right]^{\frac{1}{1-\sigma_g}} \tag{3}$$

Let us assume that when good g is shipped from i to j, only a fraction $1/\tau_g^{ij} \leq 1$ arrives in j. Thus, the model makes a distinction between c.i.f and f.o.b. prices. If p_g^i denotes the f.o.b. price of a variety of good g produced in country i, the (c.i.f.) price faced by country j consumers who import the good from country i is $\tau_g^{ij} p_g^i$. If P_g^{ij} denotes the price index for varieties of good g that country j imports from i, then the good-by-good components of the country j CPI, $\left\{ P_g^j \right\}_{g=1}^{G}$ are given by

$$P_g^j = \left\{ \beta_g^j \left(P_g^{jj} \right)^{1-\omega_g} + \left(1 - \beta_g^j \right) \left(P_g^{Fj} \right)^{1-\omega_g} \right\}^{\frac{1}{1-\omega_g}}.$$

2.2 Productivity and Production

Recall that in each country i and for each good g, N_g^{ij} represents the measure of goods exported to country j. Let φ denote a producer-specific productivity factor. In our model, N_g^{ij} will be the size of an interval of producer-specific productivity factors and firms can be indexed by φ in that interval. For a firm φ in i that exports the amount $y_g^{ij}(\varphi)$ to country j, the unit labor requirement is

$$\ell_g^{ij}(\varphi) = \frac{y_g^{ij}(\varphi)}{A_g A^i \varphi} + f_g^{ij},$$

where A_g is a global good-specific productivity shock, A^i is a country-specific productivity shock, and f_g^{ij} is a fixed labor cost of exporting g from i to j.

The distribution of producer-specific productivity factors φ among varieties follows the cumulative distribution function $H_g^i(\varphi)$. With a continuum of firms the law of large numbers applies and the measure of potential varieties produced at a firm-specific productivity exceeding φ is $1- H_g^i(\varphi)$. We will determine an endogenous *cutoff productivity level* $\hat{\varphi}_g^{ij}$ below which country i producers of varieties of g will find it unprofitable to ship to j's market. Under this notation, if the distribution of productivity levels is unbounded from above, country i producers with $\varphi \in \left[\hat{\varphi}_g^{ij}, \infty\right)$ export to j and the measure of varieties of g exported from i to j is given by $N_g^{ij} = 1- H_g^i\left(\hat{\varphi}_g^{ij}\right)$.

Let W^i be country i's wage denominated in some global numeraire. Then the price of a variety of good g "exported" to the same country i in which it is produced (its f.o.b. price) is

$$p_g^i(\varphi) = \frac{\sigma_g}{\sigma_g - 1}\left(\frac{W^i}{A_g A^i \varphi}\right). \tag{4}$$

In the presence of trade costs, as we have seen, higher (c.i.f.) prices $\tau_g^{ij} p_g^i(\varphi)$ will prevail in the countries j that import this product from i.

Exporter revenues less variable costs on shipments of g from i to j are given by $\pi_g^{ij}(\varphi) = p_g^i(\varphi)\, y_g^{ij}(\varphi)/\sigma_g$. Invoking the standard demand functions implied by CES utility, we therefore define the cutoff productivity level for exports from i to j by:

$$\pi_g^{ij}\left(\hat{\varphi}_g^{ij}\right) = W^i f_g^{ij}. \tag{5}$$

The first line above follows because, due to shipping costs, exporter production $y_g^{ij}(\varphi)$ must equal τ_g^{ij} times the number of units that actually end up being consumed by importers in country j. Equation (4) allows one to solve condition (5) explicitly for $\hat{\varphi}_g^{ij}$ as a function of variables exogenous to the firm. The cutoff productivity $\hat{\varphi}_g^{jj}$ for country j "imports" from ("exports" to) itself is found by replacing the product $\kappa_g^{ij}(1 - \beta_g^j)$ by β_g^j in equation (5), setting $i = j$ (where $\tau_g^{jj} = 1$), and replacing P_g^{Fj} by $P_g^{jj} = \left[\int_{N_g^{jj}} p_g^j(\varphi)^{1-\sigma_g}\,\mathrm{d}\varphi\right]^{\frac{1}{1-\sigma_g}}$. Notice that P_g^{ij}, the price index for varieties of g imported by

j from i, is given by

$$\begin{aligned} P_g^{ij} &= \left[\int_{N_g^{ij}} \left(\tau_g^{ij} p_g^i(\varphi) \right)^{1-\sigma_g} \mathrm{d}\varphi \right]^{\frac{1}{1-\sigma_g}} \\ &= \left(N_g^{ij} \mathbb{E} \left\{ \left(\tau_g^{ij} p_g^i(\varphi) \right)^{1-\sigma_g} | \varphi \geq \hat{\varphi}_g^{ij} \right\} \right)^{\frac{1}{1-\sigma_g}}. \end{aligned} \tag{6}$$

If the labor supply in each country j, L^j, is fixed, imposing labor-market clearing conditions for each country yields the equilibrium allocation. In Appendix A we show how to solve for this equilibrium under the assumption that the distribution of variety-specific productivity shocks is Pareto:

$$H_g^i(\varphi) = 1 - \varphi^{-\gamma_g^i}.$$

Under this specification, the price index for varieties of g imported by j from i (including $i = j$) becomes

$$P_g^{ij} = \left(\frac{\sigma_g}{\sigma_g - 1} \right) \left[\frac{\gamma_g^i}{\gamma_g^i - (\sigma_g - 1)} \right] \frac{\tau_g^{ij} W^i}{A^i A_g} (N_g^{ij})^{\frac{-[\gamma_g^i - (\sigma_g - 1)]}{\gamma_g^i (\sigma_g - 1)}}, \tag{7}$$

where the standard assumption that $\gamma_g^i > \sigma_g - 1$ is needed for this price index to be well defined.

2.3 Import Demand

It is helpful to add a time subscript now to all variables, where we are supposing that the data available are a panel of one destination country j, multiple source countries $i = 1, ..., J, i \neq j$, and multiple time periods $t = 1, ..., T$.[5] We allow the random taste and productivity parameters to vary over time, which implies that all endogenous variables are time-varying, too.

The assumptions on preferences imply that we can express the value of country j's

[5]Having multiple destination countries is straightforward in the theory, but since that is not the case in our U.S. data, we do not pursue that generalization here.

imports of good g from country $i \neq j$ (covering all varieties N_{gt}^{ij}) as:

$$V_{gt}^{ij} = \alpha_{gt}^{j} \kappa_{gt}^{ij} (1 - \beta_{gt}^{j}) \left(\frac{P_{gt}^{ij}}{P_{gt}^{Fj}} \right)^{1-\sigma_g} \left(\frac{P_{gt}^{Fj}}{P_{gt}^{j}} \right)^{1-\omega_g} \left(\frac{P_{gt}^{j}}{P_{t}^{j}} \right)^{1-\eta} P_t^j C_t^j. \tag{8}$$

Spending on good g from home supply is:

$$V_{gt}^{jj} = \alpha_{gt}^{j} \beta_{gt}^{j} \left(\frac{P_{gt}^{jj}}{P_{gt}^{j}} \right)^{1-\omega_g} \left(\frac{P_{gt}^{j}}{P_{t}^{j}} \right)^{1-\eta} P_t^j C_t^j. \tag{9}$$

Dividing (8) and (9) we obtain imports from country i relative to home demand,

$$\frac{V_{gt}^{ij}}{V_{gt}^{jj}} = \kappa_{gt}^{ij} \left(\frac{1 - \beta_{gt}^{j}}{\beta_{gt}^{j}} \right) \left(\frac{P_{gt}^{ij}}{P_{gt}^{Fj}} \right)^{1-\sigma_g} \left(\frac{P_{gt}^{Fj}}{P_{gt}^{jj}} \right)^{1-\omega_g}. \tag{10}$$

Notice that this import demand equation includes the multilateral import price index relative to the home price, P_{gt}^{Fj}/P_{gt}^{jj}, on the right, from which the elasticity ω_g is identified, whereas the elasticity σ_g is identified from the relative bilateral import price, P_{gt}^{ij}/P_{gt}^{Fj}.[6]

This import demand equation differs from the form in which it would be estimated, however, because the CES price index, P_{gt}^{ij}, is rarely if ever measured in practice by official statistical agencies. As it is specified in (6), P_{gt}^{ij} will fall whenever there is an expansion in the set of varieties N_{gt}^{ij}, because such an expansion provides a utility gain for consumers and therefore lowers the "true" price index. This negative relationship between P_{gt}^{ij} and N_{gt}^{ij} can be seen from (7), for example. Price indexes used in practice, such as the Laspeyres import and export prices used by the Bureau of Labor Statistics (BLS), do not make such a correction for variety. The same is true for unit values, which we shall use in our empirical application and which are in fact *adversely* affected by expansions in variety.

The unit value for good g sold by country i to j is defined as a consumption-weighted

[6]The role of the multilateral price index is analogous to the "multilteral resistance" effect highlighted by Anderson and van Wincoop (2004).

average of prices:

$$
\begin{aligned}
UV_{gt}^{ij} &= \int_{\hat{\varphi}_{gt}^{ij}}^{\infty} \tau_{gt}^{ij} p_{gt}^{i}(\varphi) \left[\frac{c_{gt}^{ij}(\varphi)}{\int_{\hat{\varphi}_{gt}^{ij}}^{\infty} c_{gt}^{ij}(\varphi)\, dH_{g}^{i}(\varphi)} \right] dH_{g}^{i}(\varphi) \qquad (11) \\
&= \frac{\left(\gamma_{g}^{i} - \sigma_{g}\right)}{\left(\gamma_{g}^{i} - \sigma_{g} + 1\right)} \tau_{gt}^{ij} p_{gt}^{i}\left(\hat{\varphi}_{gt}^{ij}\right).
\end{aligned}
$$

Taking the ratio of unit values in the current and previous periods, we are left with:

$$
\frac{UV_{gt}^{ij}}{UV_{gt-1}^{ij}} = \left(\frac{\tau_{gt}^{ij} W_{t}^{i} / A_{gt} A_{t}^{i}}{\tau_{gt-1}^{ij} W_{t-1}^{i} / A_{gt-1} A_{t-1}^{i}} \right) \left(\frac{N_{g}^{ij}}{N_{gt-1}^{ij}} \right)^{1/\gamma_{g}^{i}} = \frac{P_{gt}^{ij}}{P_{gt-1}^{ij}} \left(\frac{N_{gt}^{ij}}{N_{gt-1}^{ij}} \right)^{1/(\sigma_{g}-1)}, \qquad (12)
$$

where the first equality makes use of the prices in (4) and $N_{gt}^{ij} = 1 - H_{g}\left(\hat{\varphi}_{gt}^{ij}\right) = \left(\hat{\varphi}_{gt}^{ij}\right)^{-\gamma_{g}}$, and the second equality follows from (7). It is apparent from (12) that the unit value is *positively* associated with an increase in product variety N_{gt}^{ij}, in contrast to the CES price index in (7). Another way to state this result is that product variety N_{gt}^{ij} is the *measurement error* in the unit value as compared to the exact price index. The reason for this is that an expansion of demand in country j for the goods from i will lead to entry in country i, thereby driving *up* the average price as less efficient firms enter. The rate at which the average price rises as compared to the relative wage depends on the inverse of the Pareto parameter, $1/\gamma_{g}^{i}$, which appears in (12). Note that this expression holds equally well for the home county j unit value UV_{gt}^{jj}.

The true import demand equation involves the overall import price index P_{gt}^{Fj}, which is a CES function of the underlying bilateral prices P_{gt}^{ij} according to equation (3). The intertemporal ratio of CES import price indexes can be measured by the exact index due to Sato (1976) and Vartia (1976). In this case the taste coefficients appearing in (3) are random, so they also need to be included in the Sato-Vartia index, which is:[7]

$$\frac{P_{gt}^{Fj}}{P_{gt-1}^{Fj}} = \prod_{i=1, i\neq j}^{J} \left[\left(\frac{\kappa_{gt}^{ij}}{\kappa_{gt-1}^{ij}} \right)^{\frac{1}{1-\sigma_g}} \frac{P_{gt}^{ij}}{P_{gt-1}^{ij}} \right]^{w_{gt}^{ij}}, \tag{13}$$

$$w_{gt}^{ij} \equiv \frac{\left(\frac{s_{gt}^{ij} - s_{gt-1}^{ij}}{\ln s_{gt}^{ij} - \ln s_{gt-1}^{ij}} \right)}{\sum_{i=1, i\neq j}^{J} \left(\frac{s_{gt}^{ij} - s_{gt-1}^{ij}}{\ln s_{gt}^{ij} - \ln s_{gt-1}^{ij}} \right)}.$$

If we use the unit values $UV_{gt}^{ij}/UV_{gt-1}^{ij}$ instead of the CES prices on the right of (13), then we obtain a multilateral *unit-value index* as we shall use in our empirical work:

$$\frac{UV_{gt}^{Fj}}{UV_{gt-1}^{Fj}} \equiv \prod_{i=1, i\neq j}^{J} \left(\frac{UV_{gt}^{ij}}{UV_{gt-1}^{ij}} \right)^{w_{gt}^{ij}}.$$

Because the unit-value index does not properly correct for variety and taste shocks, it will differ from the true CES multilateral index by an aggregate of variety and taste-shock terms,

$$\frac{UV_{gt}^{Fj}}{UV_{gt-1}^{Fj}} = \left(\frac{P_{gt}^{Fj}}{P_{gt-1}^{Fj}} \right) \left(\frac{\kappa_{gt}^{Fj} N_{gt}^{Fj}}{\kappa_{gt-1}^{Fj} N_{gt-1}^{Fj}} \right)^{\frac{1}{(\sigma_g - 1)}} \tag{14}$$

as is obtained by using (7) and (13), where $\frac{N_{gt}^{Fj}}{N_{gt-1}^{Fj}} = \prod_{i=1, i\neq j}^{J} \left(\frac{N_{gt}^{ij}}{N_{gt-1}^{ij}} \right)^{w_{gt}^{ij}}$ and $\frac{\kappa_{gt}^{Fj}}{\kappa_{gt-1}^{Fj}} = \prod_{i=1, i\neq j}^{J} \left(\frac{\kappa_{gt}^{ij}}{\kappa_{gt-1}^{ij}} \right)^{w_{gt}^{ij}}$.

We can now specify the import demand equation in a form that we shall estimate. Using Δ to denote the first difference, from (10), (12) and (14) we have,

$$\Delta \ln \left(\frac{V_{gt}^{ij}}{V_{gt}^{jj}} \right) = -(\sigma_g - 1) \Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) + (1 - \omega_g) \Delta \ln \left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}} \right) + \varepsilon_{gt}^{ij}, \; i = 1, ..., J, i \neq j, \tag{15}$$

[7] To prove (13), note that $s_g^{ij} = \kappa_g^{ij} \left(P_g^{ij}/P_g^{Fj} \right)^{1-\sigma_g}$, so that $P_g^{Fj} = (\kappa_g^{ij}/s_g^{ij})^{\frac{1}{1-\sigma_g}} P_g^{ij}$. Then take the ratio with respect to the base period and the geometric mean using the weights w_g^{ij}. It is readily confirmed that the shares $s_g^{ij}/s_{g,0}^{ij}$ have a weighted geometric mean of unity (since the natural log of this mean sums to zero), leaving (13).

for time periods $t = 2, ..., T$ (allowing for the first difference), with the error term

$$\varepsilon_{gt}^{ij} \equiv \Delta \ln \left(\frac{\kappa_{gt}^{ij}}{\kappa_{gt}^{Fj}} \right) + \Delta \ln \left(\frac{N_{gt}^{ij}}{N_{gt}^{Fj}} \right) + \Delta \ln \frac{(1-\beta_{gt}^{j})}{\beta_{gt}^{j}} - \frac{(1-\omega_g)}{(\sigma_g - 1)} \Delta \ln \left(\frac{\kappa_{gt}^{Fj} N_{gt}^{Fj}}{N_{gt}^{jj}} \right), \quad (16)$$

which reflects exogenous taste shocks and endogenous changes to product variety at several levels of aggregation.

3 Estimating with Moment Conditions

We have every reason to expect that the error term is correlated with the relative prices that appear on the right of (15). For example, a taste shock toward goods from foreign country i (a rise in κ_g^{ij}) would raise imports V_{gt}^{ij} but would also tend to raise the unit value UV_{gt}^{ij}, because wages in country i would increase. This correlation will tend to create a downward bias in the OLS estimates of the price elasticity σ_g.[8] A further bias occurs because the unit values measure the true price indexes with error, so that the error term incorporates relative variety, which is itself changing endogenously. Similar sources of bias due to taste shocks and due to variety occur in the OLS estimates of the macro elasticity ω_g.

To offset these biases we employ Generalized Method of Moment (GMM) estimation. The nested CES preferences allow us to estimate the micro and macro elasticities from several inter-related moment conditions. The first of these conditions is obtained from *relative demand for imports*, together with a reduced-form supply equation. Feenstra (1994) assumed that the errors in these import demand and supply equations are uncorrelated, which gives a set of moment conditions – one for each source of imports – that is used to obtain the micro elasticity σ_g. This procedure is described as "step 1" below.

A second step, very similar to the first, is to work with the *total import demand relative to home demand,* in what we call "step 2." That demand equation only involves the macro Armington elasticity — not the micro elasticity from step 1. That moment condition

[8] Such downward bias is demonstrated using simulated data in our working paper (Feenstra, Luck, Obstfeld and Russ, 2014).

becomes particularly useful when we combine it with the *country-specific demand for imports relative to home demand*, in “step 3”. That country-specific import demand equation involves both the micro and macro elasticities, and likewise, we work with a reduced-form supply equation that has two supply-side parameters. We will find that this corresponding reduced-form supply equation involves the supply parameters estimated in steps 1 and 2, i.e. there is a cross-equation restriction between the various moment conditions. That cross-equation restriction gives us the extra identifying information we need to estimate the macro elasticity.

3.1 Step 1: The Micro Elasticities

For the micro elasticity, we can start with a simplified demand equation that only relies on import data. Sum (15) across foreign countries $i \neq j$ using the Sato-Vartia weights w_{gt}^{ij}, and then take the difference between (15) and the resulting equation. The terms involving home demand V_{gt}^{jj} and unit value UV_{gt}^{jj} cancel out when we take the difference, and we obtain the simple import demand equation:

$$\Delta \ln \left(\frac{V_{gt}^{ij}}{V_{gt}^{Fj}} \right) = -(\sigma_g - 1) \Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) + \varepsilon_{gt}^{iF},$$

with, the error term, ε_{gt}^{iF}, equal to $\Delta \ln \left(\frac{\kappa_{gt}^{ij}}{\kappa_{gt}^{Fj}} \right) + \Delta \ln \left(\frac{N_{gt}^{ij}}{N_{gt}^{Fj}} \right)$.

Shifting the unit-values to the left and dividing by $(\sigma_g - 1)$, we obtain:

$$\Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) = \frac{-\Delta \ln(V_{gt}^{ij}/V_{gt}^{Fj})}{(\sigma_g - 1)} + \frac{\varepsilon_{gt}^{iF}}{(\sigma_g - 1)}. \tag{17}$$

The error on the right are shocks to the relative demand for imports due to changes in tastes or variety – both of which appear in the error term ε_{gt}^{iF}. We expect that V_{gt}^{ij}/V_{gt}^{Fj} will increase with a positive shock to ε_{g}^{iF}, thereby *dampening* the response of the relative import unit value. Accordingly, we will suppose that a linear projection of the relative import unit value on the demand shocks pooled over all import sources i takes the form,

$$\Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) = \rho_{1g} \frac{\varepsilon_{gt}^{iF}}{(\sigma_g - 1)} + \delta_{gt}^{iF}, \tag{18}$$

where δ_{gt}^{iF} is an error term and ρ_{1g} denotes the OLS coefficient of the demand error ε_{gt}^{iF} which we expect to be between 0 and 1, as a result of the dampening discussed above. The presumed positive sign for this coefficient means that we interpret (18) as a reduced-form supply curve. We now recognize that the supply and demand curves for each country $i = 1, ..., J, i \neq j$, may run over different time periods, denoted by $t = 1, ..., T_g^i$.

As it is stated, equation (18) is without loss of generality and the residual δ_{gt}^{iF} in this supply curve is uncorrelated with the variables on the right of (18) over all import sources i by construction. That is, conditional on the data on the right of (18), the OLS coefficient ρ_{1g} is chosen so that the following condition holds:

$$\sum_t \sum_{i \neq j} \varepsilon_{gt}^{iF} \delta_{gt}^{iF} = 0.$$

We will, however, make the stronger assumption that this moment condition holds in expectation *for each individual* source country i:

Assumption 1: $E\left(\sum_t \varepsilon_{gt}^{iF} \delta_{gt}^{iF}\right) = 0$ for $i = 1, \ldots, J,\ i \neq j$.

The same assumption was made in Feenstra (1994) in a simpler system that used a partial equilibrium supply curve. The motivation for Assumption 1 in Feenstra (1994) was that different factors are shifting demand and supply, and it is these unmeasured factors that are entering the error terms in each equation, making it reasonable to treat these errors as uncorrelated.[9] In this paper we do not assume such a partial equilibrium supply curve because we are using the Melitz model, where wages and prices are determined in general equilibrium. So the reduced form (18) plays the role of a supply curve.

Assumption 1 gives us $J - 1$ moment conditions for each good that we can use to estimate the two parameters σ_g and ρ_{1g}. To see this, we proceed as in Feenstra (1994), by isolating the error terms in the demand equation (17) and supply equation (18):

[9]This logic does not go through, however, when an unmeasured factor influences *both* demand and supply, such as for unmeasured quality. The solution in that case is to explicitly model the choice of quality by firms, and introduce that variable into both the demand and supply equations, as in Feenstra and Romalis (2014).

$$
\begin{aligned}
\varepsilon_{gt}^{iF} &= \Delta \ln \left(\frac{V_{gt}^{ij}}{V_{gt}^{Fj}} \right) + (\sigma_g - 1) \Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right), \\
\delta_{gt}^{iF} &= (1 - \rho_{1g}) \Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) - \frac{\rho_{1g}}{(\sigma_{1g} - 1)} \Delta \ln \left(\frac{V_{gt}^{ij}}{V_{gt}^{Fj}} \right),
\end{aligned}
$$

where the final line follows by substituting from the first equation. Multiplying these two equations together and dividing by $(1 - \rho_{1g})(\sigma_g - 1)$, we obtain the estimating equation:

$$
Y_{gt}^{iF} = \theta_{1g} X_{1gt}^{iF} + \theta_{2g} X_{2gt}^{iF} + u_{gt}^{iF}, \tag{19}
$$

for $i = 1, ..., J, i \neq j$, and $t = 2, ..., T_g^i$, where

$$
\begin{aligned}
Y_{gt}^{iF} &= [\Delta \ln(UV_{gt}^{ij}/UV_{gt}^{Fj})]^2, \quad X_{1gt}^{iF} = [\Delta \ln(V_{gt}^{ij}/V_{gt}^{Fj})]^2, \\
X_{2gt}^{iF} &= [\Delta \ln(UV_{gt}^{ij}/UV_{gt}^{Fj})][\Delta \ln(V_{gt}^{ij}/V_{gt}^{Fj})],
\end{aligned}
$$

with

$$
\theta_{1g} = \frac{\rho_{1g}}{(\sigma_g - 1)^2 (1 - \rho_{1g})}, \theta_{2g} = \frac{(2\rho_{1g} - 1)}{(\sigma_g - 1)(1 - \rho_{1g})} \tag{20}
$$

and the error term, u_{gt}^{iF} equal to $\frac{\varepsilon_{gt}^{iF} \delta_{gt}^{iF}}{(\sigma_g - 1)(1 - \rho_{1g})}$.

Summing over time, the expectation of the error term u_{gt}^{iF} is zero from Assumption 1, so that gives us $J - 1$ moment conditions that we can use for estimation. Formally, we can proceed by using source-country indicator variables as instrumental variables (IV) in non-linear estimation. The inner-product of the error term with the indicator variable for country i is just the average value of u_{gt}^{iF} over time, for that i. From Assumption 1, this magnitude has expected value of zero, so that the source-country indicator variables are not correlated with the error term and are therefore valid IV, so that (19) can be estimated with two-stage least squares (TSLS).

To see the effect of using these IV in practice, suppose that country i appears for T_g^i periods in (19). If we regress the left and right-hand side variables on country indicators, we obtain the following equation:

$$\overline{Y}_g^{iF} = \theta_{1g}\overline{X}_{1g}^{iF} + \theta_{2g}\overline{X}_{2g}^{iF} + \overline{u}_g^{iF},\ i = 1, ..., J, i \neq j, \tag{21}$$

where the bar indicates the *average* value of each variable over time. Because country i appeared for T_g^i periods in the original equation (19), then likewise that country appears for T_g^i periods in (21), i.e. just repeating the averaged observation for that country T_g^i times. So the TSLS procedure is equivalent to estimating (21) over countries while weighting each country-observation by T_g^i (i.e. multiplying each country-observation in (21) by $\sqrt{T_g^i}$). We will still refer to these estimates as TSLS, even when the averaging and weighting of the country-observations is done manually.[10]

In order for the source-country indicators to be valid instruments, we also need to check that a rank condition holds: namely, that the matrix of right-hand side variables regressed on the instruments has full rank. Feenstra (1994) argues that the rank condition holds in this system if and only if there is sufficient heteroskedasticity in the error terms for demand and supply.[11] That requirement is an example of "identification through heteroskedasticity," as labeled by Rigobon (2003), and applied here in a panel context.[12]

Using initial estimates of $\hat{\theta}_{1g}$ and $\hat{\theta}_{2g}$, we obtain $\hat{\sigma}_g$, along with $\hat{\rho}_{1g}$ by a quadratic equation arising from (20) (see Feenstra 1994). We can improve efficiency by using weighting estimates, obtained from the errors:

[10]Under conventional TSLS, we would estimate θ_{1g} and θ_{2g}, and then solve a quadratic equation to obtain σ_g and ρ_{1g}. One reason to prefer the manual approach is that we can instead use nonlinear least squares (NLS) to obtain σ_g and ρ_{1g} directly. The manual approach using NLS will be useful when estimating several equations with cross-equation restrictions, as done later.

[11]Feenstra (1994) treats those variances as constant over time. See Feenstra (2010, appendix to Chapter 2) for the derivation of this result.

[12]Rigobon (2003) gives many examples of identification through heterscedasticity in finance and other fields.

$$\widehat{u}_{gt}^{iF} = Y_{gt}^{iF} - \hat{\theta}_{1g} X_{1gt}^{iF} - \hat{\theta}_{2g} X_{2gt}^{iF}.$$

We weight the estimating equation (19) by the inverse of the variance of these errors over time, for each country i, and then re-estimate (19) to obtain efficient estimates of σ_g and ρ_{1g}. This procedure corresponds to the weighting matrix that is optimally used in 2-step GMM estimation, which can be programmed manually or automatically in STATA (Baum, Schaffer and Stillman 2007).

There are two other GMM estimators available from STATA. The first is limited information maximum likelihood (LIML). In that case, the variance of the true errors $u_{gt}^{iF} = Y_{gt}^{iF} - \theta_{1g} X_{1gt}^{iF} - \theta_{2g} X_{2gt}^{iF}$ is computed over all exporting countries i and time periods. Denoting that variance by σ_{gu}^{2F}, LIML is equivalent to manually weighting the estimating equation (21) by T_g^i / σ_{gu}^{2F} and minimizing the sum of squared residuals over countries. This estimation is nonlinear because σ_{gu}^{2F} itself depends on the parameters θ_{1g} and θ_{2g}. It can be interpreted as maximum likelihood only if the true errors are normally distributed and homoskedastic across countries (which we have already ruled out for identification). Another estimator, introduced by Hansen, Heaton and Yaron (1996), is the "continously-updated estimator" (CUE). In that estimator, the standard deviation of the true errors u_{gt}^{2iF} is computed over time *for each* exporting country i, therefore allowing for heteroscedasticity. Denoting that variance by σ_{gu}^{2iF}, CUE is equivalent to manually weighting the estimating equation (21) by $T_g^i / \sigma_{gu}^{2iF}$ and minimizing the sum of squared residuals over countries. This estimation is again nonlinear because σ_{gu}^{2iF} depends on the parameters θ_{1g} and θ_{2g}, and as we shall find, has more difficulty converging than LIML.

3.2 Step 2: The Macro Elasticity

A similar moment condition to that described above – but now at the macro level – can be obtained by aggregating across countries. To achieve this, we start from (8), which gives country j's spending on imports of good g from country i. For convenience, we assume the macro Armington elasticities are the same across a set of goods:

$$\omega_g = \omega \text{ for } g = 1, ..., G.$$

Summing over all trade partners $i \neq j$ yields country j spending on imports of good g from all foreign sources, denoted V_{gt}^{Fj}:

$$V_{gt}^{Fj} = \sum_{i \neq j} V_{gt}^{ij} = \alpha_{gt}^{j} \left(1 - \beta_{gt}^{j}\right) \left(\frac{P_{gt}^{Fj}}{P_{gt}^{j}}\right)^{1-\omega} \left(\frac{P_{gt}^{j}}{P_{t}^{j}}\right)^{1-\eta} P_t^j C_t^j.$$

The last line follows from definition (3) and we also impose $\omega_g = \omega$. Combining the foregoing expression with the demand V_{gt}^{jj} as computed from equation (9), we obtain:

$$\ln\left(\frac{V_{gt}^{Fj}}{V_{gt}^{jj}}\right) = (1-\omega)\ln\left(\frac{P_{gt}^{Fj}}{P_{gt}^{jj}}\right) + \ln\left(\frac{1-\beta_{gt}^{j}}{\beta_{gt}^{j}}\right),$$

so that the home-foreign Armington elasticity ω can be identified from a multilaterally aggregated equation for imports of good $g = 1, ..., G$.[13]

Of course, when estimating this equation we use unit value rather than the true price indexes, as shown in (12) (for $i = j$) and (14). It follows that the aggregate equation for estimation is:

$$\Delta \ln\left(\frac{V_{g}^{Fj}}{V_{g}^{jj}}\right) = (1-\omega)\Delta \ln\left(\frac{UV_{g}^{Fj}}{UV_{g}^{jj}}\right) + \varepsilon_{g}^{Fj},$$

with the error term,

$$\varepsilon_{gt}^{Fj} \equiv \Delta \ln \frac{(1-\beta_t^j)}{\beta_t^j} + \frac{(\omega-1)}{(\sigma_g-1)}\left[\Delta \ln \kappa_{gt}^{Fj} + \Delta \ln\left(\frac{N_{gt}^{Fj}}{N_{gt}^{jj}}\right)\right]. \tag{22}$$

The properties of the macro demand equation are quite similar to those of the disaggregate equation (15), and we can similarly adapt the technique of Feenstra (1994) to estimate the macro elasticity. We begin by re-writing the demand equation slightly as:

[13] This estimating equation is closely related to those that Reinert and Roland-Holst (1992), Blonigen and Wilson (1999), and Gallaway, McDaniel, and Rivera (2003) use.

$$\Delta \ln \left(\frac{UV_g^{Fj}}{UV_g^{jj}} \right) = -\frac{1}{(\omega - 1)} \Delta \ln \left(\frac{V_g^{Fj}}{V_g^{jj}} \right) + \frac{1}{(\omega - 1)} \varepsilon_g^{Fj}.$$

Again, we expect that V_{gt}^{Fj}/V_{gt}^{jj} will increase with a positive shock to ε_g^{Fj}, thereby *dampening* the response of the relative import unit value. Accordingly, we take a linear projection across goods and time of the relative unit-value on the error term to obtain:

$$\Delta \ln \left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}} \right) = \rho^F \frac{\varepsilon_{gt}^{Fj}}{(\omega - 1)} + \delta_{gt}^{Fj}, \tag{23}$$

for $g = 1, ..., G$, and $t = 1, ..., T_g$. The coefficient ρ^F denote the impact of the demand error ε_{gt}^{Fj} on the relative unit-value, and we expect that $0 < \rho^F < 1$, so that (23) is interpreted as a reduced-form "macro" supply curve.

By construction, the supply error δ_{gt}^{Fj} is uncorrelated with the demand error ε_{gt}^{Fj} in (23) when taken over all observations $g = 1, ..., G$, and $t = 1, ..., T_g$. We make the stronger assumption that these errors are uncorrelated for each good:

Assumption 2: $E\left(\sum_t \varepsilon_{gt}^{Fj} \delta_{gt}^{Fj}\right) = 0$ for $g = 1, ..., G$.

To make use of the macro demand and supply equations, we now proceed in the same manner as with the estimating equation (19) for the micro elasticities. Multiplying the errors in the macro demand and supply equations, we can obtain:

$$Y_{gt}^{Fj} = \phi_1 X_{1gt}^{Fj} + \phi_2 X_{2gt}^{Fj} + u_{gt}^{Fj}, \tag{24}$$

for $g = 1, ..., G$ and $t = 2, ..., T_g$, where

$$\begin{aligned} Y_{gt}^{Fj} &= [\Delta \ln(UV_{gt}^{Fj}/UV_{gt}^{jj})]^2, \quad X_{1gt}^{Fj} = [\Delta \ln(V_{gt}^{Fj}/V_{gt}^{jj})]^2, \\ X_{2gt}^{Fj} &= [\Delta \ln(UV_{gt}^{Fj}/UV_{gt}^{jj})][\Delta \ln(V_{gt}^{Fj}/V_{gt}^{jj})], \end{aligned}$$

with the coefficients ϕ_1 equal to $\frac{\rho^F}{(\omega-1)^2(1-\rho^F)}$ and ϕ_2 equal to $\frac{(2\rho^F-1)}{(\omega-1)(1-\rho^F)}$ and the error term u_{gt}^{Fj} equal to $\frac{\varepsilon_{gt}^{Fj}\delta_{gt}^{Fj}}{(\omega-1)(1-\rho^F)}$

If we average (24) over time, then we obtain additional moment conditions (one condition for each good) that involve the macro elasticity ω and a new supply parameter ρ^F. We have found in preliminary results, however, that GMM estimates from this equation alone do not perform that well: in simulated data the estimates converge to the true value of the macro elasticity very slowly, and on US data we do not obtain convergence for some industries. We therefore add another moment condition, obtained from the nested-CES demand equation, that involves both the micro and macro elasticities.[14]

3.3 Step 3: Combining the Micro and Macro Elasticities

We now consider the nested-CES demand equation (15), which we re-write slightly so that the unit values appear on the left and the import values on the right:

$$\Delta \ln\left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}}\right) = \frac{-\Delta \ln(V_{gt}^{ij}/V_{gt}^{jj})}{(\sigma_g - 1)} - \frac{(\omega - 1)}{(\sigma_g - 1)}\Delta \ln\left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}}\right) + \frac{\varepsilon_{gt}^{ij}}{(\sigma_g - 1)}, \qquad (25)$$

where we continue to use $\omega_g = \omega$ for $g = 1, ..., G$. This specification unpacks (17), as it has the same dependent variable on the left but now has two variables on the right, including the multilateral unit value UV_{gt}^{Fj} relative to the home unit value, UV_{gt}^{jj}. Similar to our discussion of (17), the error term on the right includes shocks to the relative demand for imports due to changes in tastes or variety. If there were no response at all in relative demand V_{gt}^{ij}/V_{gt}^{jj}, then the relative import unit value on the left-hand side would rise by the full amount of the term involving $UV_{gt}^{Fj}/UV_{gt}^{jj}$. We once again expect that V_{gt}^{ij}/V_{gt}^{jj} will increase with a positive shock to ε_g^{ij}, thereby *dampening* the response of the relative import unit value. The amount of dampening could very well depend on the source of the shock, however. Accordingly, we will suppose that the relative import unit values are related to the demand shocks by the reduced-form supply equation,

[14] We also estimate the parameters using steps 1 and 3, but without the macro import demand in step 2; see our working paper (Feenstra, Luck, Obstfeld and Russ, 2014). We found that the results on simulated data did not perform as well as with using all three steps.

$$\Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) = \rho_{1g} \frac{\varepsilon_{gt}^{ij}}{(\sigma_g - 1)} - \rho_{2g} \frac{(\omega - 1)}{(\sigma_g - 1)} \Delta \ln \left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}} \right) + \delta_{gt}^{ij}, \tag{26}$$

for $i = 1, ..., J, i \neq j$, and $t = 1, ..., T$.

We regard (26) as the natural extension of the reduced-form supply curve (18), one that now incorporates the relative multilateral unit value $UV_{gt}^{Fj}/UV_{gt}^{jj}$, appearing on the right. We expect that $\rho_{1g}, \rho_{2g} > 0$ represent the possibly dampened impact of the demand shock $\varepsilon_{gt}^{ij}/(\sigma_g - 1)$ and a shock to the relative multilateral unit value, respectively. As before, we assume that error term δ_{gt}^{ij} in this supply equation is uncorrelated with the error term in demand for each source country:

Assumption 3: $E\left(\sum_t \varepsilon_{gt}^{ij} \delta_{gt}^{ij}\right) = 0$ for $i = 1, \ldots, J$, $i \neq j$.

Whereas Assumption 1 referred to the correlation of errors in a supply and demand system that differences out home demand, Assumption 3 now refers to the correlation of errors in a system that retains home demand and its unit value. It again give us $J - 1$ moment conditions that can be used to estimate the model parameters, which now are σ_g, ω_g, ρ_{1g} and ρ_{2g}. To see this, isolate the errors terms ε_{gt}^{ij} and δ_{gt}^{ij} in (25) and (26), respectively, then multiply these two equations together and divide by $(1 - \rho_{1g})(\sigma_g - 1)$. We obtain an estimating equation expressed in the convenient form:

$$Y_{gt}^{iF} = \sum_{n=1}^{2} \theta_{ng} X_{ngt}^{ij} + \sum_{n=3}^{4} (\omega_g - 1)\theta_{ng} X_{ngt}^{ij} + (\omega_g - 1)^2 \theta_{5g} X_{ngt}^{j} + u_{gt}^{ij}, \tag{27}$$

for $i = 1, ..., J, i \neq j$, and $t = 1, ..., T_g^i$, where

$$\begin{aligned}
Y_{gt}^{iF} &= [\Delta \ln(UV_{gt}^{ij}/UV_{gt}^{Fj})]^2, \quad X_{1gt}^{ij} = [\Delta \ln(V_{gt}^{ij}/V_{gt}^{jj})]^2, \\
X_{2gt}^{ij} &= [\Delta \ln(UV_{gt}^{ij}/UV_{gt}^{Fj})][\Delta \ln(V_{gt}^{ij}/V_{gt}^{jj})], \quad X_{3gt}^{ij} = [\Delta \ln(UV_{gt}^{Fj}/UV_{gt}^{jj})][\Delta \ln(UV_{gt}^{ij}/UV_{gt}^{Fj})], \\
X_{4gt}^{ij} &= [\Delta \ln(UV_{gt}^{Fj}/UV_{gt}^{jj})][\Delta \ln(V_{gt}^{ij}/V_{gt}^{jj})], \quad X_{5gt}^{j} = [\Delta \ln(UV_{gt}^{Fj}/UV_{gt}^{jj})]^2,
\end{aligned}$$

with the coefficients θ_{1g} and θ_{2g} defined as in (20), and also:

$$\theta_{3g} = \frac{-(1+\rho_{2g}-2\rho_{1g})}{(\sigma_g-1)(1-\rho_{1g})}, \theta_{4g} = \frac{-(\rho_{2g}-2\rho_{1g})}{(\sigma_g-1)^2(1-\rho_{1g})}, \theta_{5g} = \frac{-(\rho_{2g}-\rho_{1g})}{(\sigma_g-1)^2(1-\rho_{1g})}. \tag{28}$$

Notice that the dependent variable Y_{gt}^{iF} in (19) and (27) are the same but the X variables are different. Nevertheless, the coefficients θ_{1g} and θ_{2g} are the same in the two systems, and then the extended system (27) also has three additional dependent variables with coefficients that depend on the micro demand elasticity σ_g and supply elasticity ρ_{1g}, as well as the macro demand elasticity ω and supply elasticity ρ_{2g}.

As in the earlier estimation of (19), we can use source-country indicator variables as IV, in which case the dependent variables are averaged over time. Because of the nonlinear structure of the coefficients in (27), in practice we perform IV estimation manually by averaging (27) over time for each source country and good. The averaged-over-time equation appears T_g^i times for each source country and good. Therefore, we weight the averaged-over-time equation by T_g^i and then apply NLS to estimate the macro demand ω and supply ρ_2 elasticities. These are referred to as the TSLS estimates. Given these initial estimates, we can construct more efficient estimates by weighting (27) by the inverse of the variance of the residuals computed over time, for each country i and good g, and then re-estimating (27) to obtain efficient estimates of ω and ρ_2. Equivalently, we weight the averaged-over-time equation by T_g^i divided by the variance of the residuals, and then apply NLS to obtain the 2-step GMM estimates. In both cases, we include a constant in (27) to control for measurement error in the relative unit-value within Y_{gt}^{iF}.

We claim that this new set of equations adds information to the micro and macro equations from step 1 and step 2, respectively. To see this, notice that the error term ε_{gt}^{ij} in equation (25) is equal to $\varepsilon_{gt}^{iF} + \varepsilon_{gt}^{Fj}$, the sum of the micro and macro errors defined above. Using (23) the reduced-form supply relation (18) can be re-written as:

$$\Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) = \rho_{1g} \frac{\varepsilon_{gt}^{ij}}{(\sigma_g - 1)} - \rho_{2g} \left(\frac{\omega - 1}{\sigma_g - 1} \right) \Delta \ln \left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}} \right) + \delta_{gt}^{ij}, \tag{29}$$

with $\rho_{2g} \equiv \frac{\rho_{1g}}{\rho^F}$ and $\delta_{gt}^{ij} \equiv \rho_{2g} \left(\frac{\omega-1}{\sigma_g-1} \right) \delta_{gt}^{Fj} + \delta_{gt}^{iF}$. The new equation (29) justifies our use of (26) by making use of the macro supply equation (23), which allows us to deduce the cross-equation restrictions $\rho_{2g} = \rho_{1g}/\rho_F$. This additional cross-equation restriction is an advantage for estimation purposes. Imposition of the restriction $\rho_{2g} = \rho_{1g}/\rho_F$ is one way that the macro equation adds information to our estimating system

We further note that the macro Assumption 2 adds extra information to Assumption 3. To see this, recall that the demand shock ε_{gt}^{Fj} equals $\sum_{i\neq j} w_{gt}^{ij}\varepsilon_{gt}^{ij}$, and the reduced-form supply shock $\rho_{2g}\left(\frac{\omega-1}{\sigma_g-1}\right)\delta_{gt}^{Fj}$ equals $\sum_{i\neq j} w_{gt}^{ij}\delta_{gt}^{ij}$, which follows from the definitions of δ_{gt}^{ij} and ρ_{2g} and $\sum_{i\neq j} w_{gt}^{ij}\delta_{gt}^{iF} = 0$. For $\rho_{2g} \neq 0$ and $\omega \neq 1$, we can re-write the expectation in Assumption 2 as:

$$0 = E\left(\sum_t \sum_{i\neq j} (w_{gt}^{ij})^2 \varepsilon_{gt}^{ij}\delta_{gt}^{ij}\right) + E\left(\sum_t \sum_{i\neq j} \sum_{k\neq i,j} w_{gt}^{ij} w_{gt}^{kj} \varepsilon_{gt}^{ij}\delta_{gt}^{kj}\right). \tag{30}$$

The first summation in (30) equals zero in its unweighted form from Assumption 3, and we could expect it to be close to zero even in its weighted form. But the second summation involves the (weighted) cross-correlation of the terms ε_{gt}^{ij} and δ_{gt}^{kj}, which refer to different source countries $i \neq k$. Assumption 2 imposes that the sum of these two complex summations equals zero. For that reason, the moment condition for the macro demand equation in Assumption 2 is adding information to what we have used in Assumption 3 for the nested CES import demand equation. Accordingly, we will exploit both moment conditions, combined with that in Assumption 1 for the micro elasticity, in the estimation.

4 Data and estimation

4.1 U.S. Data

The import data at 10-digit Harmonized System (HS) level are readily available, along with the associated unit values, but it is difficult to match these imports to the associated U.S. supply. We make use of a unique data source called *Current Industrial Reports* (CIR), which is published by the U.S. Bureau of the Census and reports imports, exports and U.S. production at a disaggregate "product code" level. Recent years are available online,[15] and past years were obtained from an online archive, so the dataset spans 1992-2007. The data are in readable PDF or similar format, so we laboriously transcribed these to machine-readable datasets. After a careful process of matching and pooling the NAICS- and SIC-based product codes with HS product codes (see Appendix for details), we end up with 109 goods, roughly 80 based on a single 10-digit HS code and the rest based on two or three, making the dataset novel in its highly disaggregate nature.

4.2 Simulated Data

We also simulate a small-scale example of our theoretical model to check the robustness of our GMM estimation. To minimize computational requirements, we use five countries and ten goods in the simulation. The goods differ in the way we calibrate their relevant productivity and taste shocks, as described in this section, which is intended to satisfy the assumptions necessary for identification, specified in section 4. To reflect the short sample for most goods in the actual data, there are 10 observations for each simulated country-pair-good combination. We calibrate the model using variances of shocks taken from plant-level and macroeconomic studies. The Appendix shows how we solve for the model's equilibrium and provides detailed descriptions of the calibration.

4.3 Step 1: Estimating the Micro Elasticities in Simulated and U.S. Data

The estimation of (19) can sometimes lead to values for $\hat{\sigma}_g$ less than unity. Broda and Weinstein (2006) implemented a grid search procedure to avoid that outcome. We do not implement that procedure here, because we are interested in comparing the estimates of the micro and macro elasticities without constraining either estimate. Instead, we allow the estimates of $\hat{\sigma}_g$ to be less than unity. We isolated a small number of goods in our dataset, however, where the estimates of $\hat{\sigma}_g$ are most frequently negative, i.e. in more than 75% of the bootstrap estimates, so that we conclude that these data are faulty or incompatible with our model. There are 6 such goods out of the 109 used in

[15] http://www.census.gov/manufacturing/cir/index.html.

the OLS estimates, and another 5 goods had imaginary point estimates $\hat{\sigma}_g$ when solving the quadratic equation arising from (20), so in the GMM estimation we work with the remaining 98 goods.

In the simulated dataset with a true value of $\sigma = 3$, applying the above TSLS and GMM procedures to equation (19) gives the median results shown in Table 1 using 1,000 simulations.[16] We find that both TSLS and 2-step GMM result in estimates of the micro elasticity that are about $10 - 15\%$ *above* the true value of $\sigma = 3$. That bias is surprisingly persistent as the number of time periods T in the simulation is increased 10 to 50 to 100. (Recall that the number of years for each product in our first-differenced U.S. data is at most 15, but often much less.) We can compare these results to Soderbery (2010), who performs a Monte Carlo analysis on the estimation of the micro elasticity, where the data generating process uses CES demand and a partial equilibrium supply curve. In the presence of measurement error in prices, he finds that the upward bias fall from about $10 - 15\%$ with $T = 10$, like we have found, to about 5% with $T = 50$ or 100. While we do not see this fall in bias from $T = 10$ to 50, both his results and ours have a quite persistent bias from $T = 50$ to 100.

The median estimates from 1,000 simulations using LIML and CUE are shown in the final two columns of Table 1. We find that both estimates are *downward* biased by roughly 20%. Once again, we find no evidence that this bias is reduced when longer time periods are used in the simulation. In contrast, Soderbury (2010) does not find a downward bias when using LIML in his Monte Carlo analysis (he did not estimate with CUE). But he does find a large downward difference – often of 50% or more – from using LIML as compared to the TSLS estimates for the five U.S. imported products that he analyzes. Likewise, we shall find that the LIML and CUE estimates in our U.S. data – reported below – tend to be lower that the TSLS and 2-step GMM estimates. So we believe that the downward bias of LIML and CUE reported in Table 2 are broadly similar

[16]In all estimators, we include a constant term in (19). Feenstra (1994) argued that this term can control for classical measurement error – that is, error uncorrelated with other variables – in the unit values.

Table 1: Median GMM Estimates of the Micro Elasticity Using Simulated and U.S. Data

	Sigma estimated from Eq. (20)			
	TSLS	2-step GMM	LIML	CUE
Data	(1)	(2)	(3)	(4)
Simulated	3.36	3.72	2.66	2.75
$T = 10$	(2.15, 3.56)	(3.42, 3.86)	(2.25, 3.11)	(2.70, 2.78)
Simulated	3.27	3.70	2.60	2.66
$T = 50$	(3.03, 3.49)	(3.66, 3.72)	(1.84, 3.25)	(2.63, 2.67)
Simulated	3.42	3.69	2.60	2.65
$T = 100$	(3.27, 3.57)	(3.57, 3.76)	(1.57, 3.38)	(2.60, 2.69)
U.S.	3.22	4.05	1.54	1.91
$T \leq 15$	(-2.98, 4.77)	(2.73, 6.05)	(0.96, 2.98)	(0.35, 4.58)

Notes: The first three rows report the median estimate from 1,000 simulated datasets, where the true value of σ is 3, the true value of ω is 2, and T indicates the length of the time series. For the simulated data, the confidence intervals shown in parentheses are computed by bootstrapping the simulated dataset corresponding to the median estimate. The final row reports the median estimate over goods in the U.S. data, and the confidence intervals shown in parentheses are likewise computed by bootstrapping the U.S. dataset corresponding to the median good.

to the results from our U.S. data and those of Soderbery.

Turning to the U.S. data, the median estimate is 3.24 from the TSLS estimates, as reported at the bottom of Table 1. That median is close to the median estimate of 3.1 from Broda and Weinstein (2006), computed over some 10,000 HS categories of imports, so our much more limited sample of 98 goods is similar in this respect. To obtain confidence intervals on the GMM estimates we perform a nested bootstrap procedure. That is, we randomly re-draw observations with replacement 500 times and re-estimate (19) to generate bootstrapped point estimates. We then randomly re-draw observations from each bootstrap dataset 100 times and re-estimate (19) once again to generate bootstrapped standard errors for each of our 500 bootstrapped point estimates. Performing a nested bootstrap allows us to construct confidence intervals using an asymptotic refinement and to construct test statistics that are pivotal and therefore consistent even in the event of bias in the underlying point estimates (see MacKinnon, 2006). The kernel density of the TSLS estimates for σ_g over the 98 goods with the lower and upper 95% confidence bounds are graphed in Figure 1A.

Turning to other results from the U.S. data, the median 2-step GMM estimate is 4.12, which is somewhat *higher* than the median TSLS estimate. That higher estimate from 2-step GMM is not found consistently in our simulated data, but occurs for 62% of our U.S. products and is persistent in this sense. The kernel density of the 2-step GMM estimates along with the lower and upper 95% confidence bounds are graphed in Figure 1B. The median LIML and CUE estimates, also reported at the bottom of Table 1, are *lower* than found for TSLS and 2-step GMM. These medians are taken over slightly different sets of goods because LIML and especially CUE fail to converge in some cases.[17] If we focus on the 85 goods where convergence is always achieved, then in 66 of these cases the LIML estimate is less than TSLS, and in 49 cases the CUE estimate is less than TSLS. We conclude that there is a quite persistent downward bias in the LIML and CUE estimates, as we also found using our simulated data. Understanding the source of this downward

[17]Going forward we focus only on the 98 goods for which both TSLS and 2-step GMM both provide real estimates.

bias in LIML and CUE when estimating the micro Armington elasticity is a topic for further research.[18] For the remainder of the paper we shall focus on TSLS and 2-step GMM as the preferred estimation methods, in part due to the persistent downward bias in LIML and CUE, and also because the equations will become nonlinear in the coefficients below when we use multiple moment conditions in our estimation.

4.4 Steps 2 and 3: Estimating the Macro Elasticity in Simulated and U.S. Data

The micro elasticities for disaggregate goods are always estimated according to equation (19), in step 1 of our procedure, though we will repeat step 1 many times with bootstrapped data to generate standard errors for the micro and the macro elasticities. The macro elasticity is estimated at a more aggregate level using equations (24) and (27) simultaneously, in steps 2 and 3 of our procedure. We begin by reporting results for the macro elasticity from simulated data in Table 2.

The results from simultaneous estimation of (24) and (27) on simulated data are shown in columns (1) and (2) of Table 2. Once again, we show median estimates of the macro elasticity from 1,000 simulations of our model, together with the 95% confidence intervals obtained by bootstrapping the data for that median estimate. The estimation proceeds by first obtaining micro elasticity estimates $\hat{\sigma}_g$ from (19), or step 1, and substituting these into (24) and (27), so that steps 2 and 3 are estimated as a system. Standard errors are obtained by bootstrapping the entire system, i.e. equations (19), (24) and (27). For the TSLS estimates in column (1), we see that the median point estimates of the macro elasticity are not that close to their true value of 2, unless the sample size in increased to $T = 100$. When applying 2-step GMM, shown in column (3) of Table 2, the estimates are closer to their true value. For $T = 100$, the median estimate of the macro elasticity is

[18] LIML and CUE both have the desirable property that the estimates are independent of the normalization of the estimating equation, i.e. which variable in (21) is used on the left-hand side. It is surprising that a persistent downward bias is found despite this desirable property.

1.93, only slightly lower than its true value of $\omega = 2$. For lower values of T the downward bias is more noticeable, but the confidence intervals either include 2 (for $T = 50$) or nearly so (for $T = 10$).

Turning to the U.S. data, our estimates of the macro elasticity obtained from simultaneous estimation of the macro equation (24) and the nested CES equation (27) are shown in columns (1) and (2) of Table 3. The point estimates of the macro elasticity are significantly less than unity in only one sector (metal products) and are significantly greater than unity in two other sectors (electronics, as well as primary metals for the TSLS estimate). The finding that the confidence intervals are quite large– including the values of unity in four out of eight cases and values of 2 or higher in most other cases – reflects the fact that we are bootstrapping the standard errors using all three estimating equations, so there is a high degree of coefficient variation across bootstrapped samples. Still, a judicious interpretation of the results is that the macro elasticity is often found to be greater than unity.

To investigate the size of the macro elasticity more formally, in Table 4, we report the results of tests for the null hypothesis that $\omega \geq \sigma_g$. To perform these one-sided tests we use the bootstrapped data from the three estimating equations: (19), (24) and (27).[19] Each bootstrap results in estimates of the micro elasticity $\hat{\sigma}_g$ for each of the goods in a sector and the single macro elasticity $\hat{\omega}$ for that sector. Using our nested bootstrap we estimate both point estimates and standard errors for each of our 500 bootstrap samples allowing us to construct a t-statistic for each bootstrap. By counting the number of cases where the bootstrap t-statisic exceeds that of the original estimate we are able to construct a p-value for the one-sided test of $\omega \geq \sigma_g$. With a 5% significance level, if there are fewer than 25 bootstrap samples where the bootstrap t-statisic exceeds that of the original estimate then we reject the null hypothesis that $\omega \geq \sigma_g$. From this result, we would conclude that the macro elasticity ω is significantly less than the micro elasticity

[19]The use of boostrapping to test hypotheses is discussed in MacKinnon (2006). Following MacKinnon we compute a bootstrap test statistic that is pivotal by constructing a t-statistic for each bootstrapped dataset.

Table 2: Median System GMM Estimates of the Macro Elasticity Using Simulated Data

T	Omega from Eqs. (25) and (28)	
	TSLS	2-step GMM
	(1)	(2)
10	1.44	1.52
	(1.22, 1.59)	(1.20, 1.90)
50	1.62	1.70
	(1.38, 1.83)	(1.37, 2.22)
100	1.78	1.93
	(1.56, 2.33)	(1.32, 3.48)

Notes: The true value of ω is 2. This table reports estimate of ω obtaining by running TSLS or 2-step GMM on equations (25) and (28) jointly, where the instruments are indicator variables by country and good within each sector. The estimates of σ_g and ρ_{1g} used are obtained from first-stage estimation of (20). Reported in parentheses are the 95% confidence intervals obtained by bootstrapping the entire system.

Table 3: System GMM Estimates of the Macro Elasticity Using U.S. Data

Sector	Number of Goods	Omega from Eqs. (25) and (28) TSLS (1)	2-step GMM (2)
Food Products	6	4.08	3.12
		(-4.27, 12.52)	(-4.20, 9.94)
Apparel Manufacturing	13	2.51	3.60
		(1.85, 3.17)	(3.47, 3.75)
Rubber, Stone, & Misc Metal	5	1.38	1.65
		(0.61, 2.14)	(1.24, 1.89)
Chemical Manufacturing	6	2.10	1.46
		(-3.27, 7.52)	(-6.04, 15.25)
Primary Metals	20	2.064	1.16
		(1.28, 2.85)	(-98.4, 4.89)
Metal Products	9	0.87	0.88
		(0.55, 1.17)	(0.78, 1.14)
Machinery	15	2.01	2.36
		(1.24, 2.78)	(1.75, 2.92)
Electronics	24	2.40	3.48
		(1.97, 2.83)	(3.15, 3.77)

Notes: Same as Table 2, except that the Table 2 assumed a value of $\omega = 2$ that does not apply in this table which is based on estimates from U.S. data.

Table 4: Testing that the Macro Elasticity is less than the Micro Elasticity

Sigma estimated from Eq. (20) and Omega from Eqs. (25) and (28)							
	Number	σ_g from TSLS			σ_g from 2-step GMM		
	of	ω from TSLS			ω from 2-step GMM		
Sector	Goods	Number of Goods with:			Number of Goods with:		
		Count	Rejection of null		Count	Rejection of null	
		$\omega < \sigma_g$	$\omega \geq \sigma_g$	$2\omega = \sigma_g$	$\omega < \sigma_g$	$\omega \geq \sigma_g$	$2\omega = \sigma_g$
		(1)	(2)	(3)	(4)	(5)	(6)
Food Products	6	4	0	4	5	4	2
Apparel Manufacturing	13	12	4	10	11	10	13
Rubber, Stone & Misc Metal	5	0	3	3	3	3	4
Chemical Manufacturing	6	4	0	1	6	1	0
Primary Metals	20	16	1	5	20	12	7
Metal Products	9	9	3	5	9	7	6
Machinery	15	12	0	3	9	5	6
Electronics	24	11	3	15	10	7	2
Total	98	73	34	11	73	26	20

Notes: This table reports the number of goods in each sector for which $\omega < \sigma_g$ in the point estimates in columns (1) and (4). Columns (2) and (5) report the results of the one-sided test for $\omega \geq \sigma_g$ constructed from a bootstrap technique explained in the text. We report the number of goods for which that test is rejected at the 5% level. Columns (3) and (6) report the results of the two-sided test for $2\omega = \sigma_g$ constructed from the same bootstrap technique. We report the number of goods for which that test is rejected at the 5% level.

σ_g.

To give further guidance on the relative size of the elasticities, we note that some researchers have employed an *ad hoc* assumption known as the "Rule of Two", which states that the macro elasticity should be roughly one half the micro elasticity (see Hillberry and Hummels 2013 for a discussion). We test how well the "Rule of Two" fits our data using the same nested bootstrap procedure and a two-sided test of the null hypothesis $2\omega = \sigma_g$ for each of our 98 goods. Within Table 4, we report hypothesis tests for both TSLS and 2-step GMM estimates. We report the results of hypothesis tests using our TSLS estimates for the micro and macro elasticities in columns (2) and (3). In columns (5)-(6) we report the results of tests using our 2-step GMM estimates.[20]

Looking first at the hypothesis test using TSLS estimates, in column (1) we show how many goods in each sector have estimates of the macro elasticity $\hat{\omega}$ less than that of the micro elasticities $\hat{\sigma}_g$. Out of 98 goods, fully 73 of then have $\hat{\omega} < \hat{\sigma}_g$ in the point estimates. Then in column (2) we report the number of goods for which we reject the null hypothesis that $\omega \geq \sigma_g$. Looking at the total reported in the final row of column (2), there are 34 out of the 98 goods where we reject the null hypothesis that the macro elasticity is equal to or exceeds the micro elasticity. In other words, for two thirds of the cases we fail to reject that null hypothesis, meaning that we cannot distinguish the size of the macro and micro elasticities. In column (3) we report the number of goods for which we can reject the null hypothesis that $2\omega = \sigma_g$. We can reject this hypothesis in only 11 cases. These results reflect somewhat large standard errors in the TSLS estimates, and

[20]We found earlier in Table 1 that the 2-step GMM estimate of the micro elasticity are noticeably higher than the TSLS estimates, and that both elasticities are upward biased in the simulations. In contrast, the median macro elasticities reported in Table 2 are always downward biased in the simulations. So we might be concerned that these biases will make it more likely that $\hat{\omega} < \hat{\sigma}_g$ in the bootstrap samples. However, because our test statistic is based on a pivotal object, the t-statistic rather than the point estimates of $\hat{\omega}$ and $\hat{\sigma}_g$, our test is unbiased even in the event of bias our point estimates (see MacKinnon (2006) for details).

we obtain somewhat sharper results using 2-step GMM estimates.

In column (4)-(6) where we use the 2-step GMM estimates of both elasticities. We find that 73 goods have the estimated macro elasticity $\hat{\omega}$ less than the estimated micro elasticities $\hat{\sigma}_g$ (column 4), the same as the count with TSLS estimates. But now when we test the null hypothesis that $\omega \geq \sigma_g$, we are able to reject it for 26 of the goods (column 5). In other words, for one-quarter of the total number of goods or one-third of cases where our point estimate satisfy $\omega < \sigma_g$, we find that the macro elasticity is significantly *below* the micro elasticity. When estimating both σ_g and ω with 2-step GMM we can reject the "Rule of Two" for only 20 of our 98 goods.

While our results are somewhat sensitive to using the TSLS versus 2-step GMM estimates, if we focus on the latter estimates then we conclude that for *roughly one-quarter* of the goods the macro elasticity is significantly less that the micro elasticities, but not as low as the value of unity sometimes found using macro time-series methods. We reject the hypothesis that the micro elasticity is twice as large as the macro elasticity in only one-fifth of the cases.

5 Gains from Trade

Despite the nuanced nature of our results we believe they have important implications for many issues in international trade and macroeconomics, and discuss here just one application, to the quantitative welfare assessments of trade policy. In their comprehensive recent survey, Costinot and Rodríguez-Clare (2014) argue within a simple gravity model that the "trade elasticity" ε relevant for welfare analysis the elasticity of substitution between foreign and domestic goods. However, they note that the formula for ε is more complicated in some models, including ours. In our model the elasticity of trade (elasticity of the aggregate import/domestic consumption ratio with respect to relative trade costs for imports) is:[21]

$$\varepsilon_g = \frac{\gamma_g(\omega - 1)}{(\sigma_g - \omega)\left(\frac{\gamma_g}{\sigma_g - 1} - 1\right) + \sigma_g - 1}, \tag{31}$$

which applies in each sector. For simplicity, we will assume that the utility function across goods is Cobb-Douglas rather than CES, that is, we assume the $\eta = 1$ in (1), and we continue to use a single value for ω within each broad sector. If σ_g is equal to ω in (31), then the trade elasticity ε_g is simply the Pareto shape parameter γ_g, as in the Melitz-Chaney model.[22] However, if $\sigma_g \neq \omega$, as we found in a number of cases, then the trade elasticity ε_g is a function of σ_g, ω, and γ_g, and is increasing in ω. Moreover, we find from (31) that $\varepsilon_g < \gamma_g$ if $\omega < \sigma_g$. Since our results are often consistent with $\omega < \sigma_g$ and the trade gains are inversely proportional to ε, our results imply quantitatively bigger trade gains than one might have expected based on the typical parameter assumption that $\sigma_g = \omega$.

Given the above expression for the elasticity of trade, the gains from trade relative to autarky are measured by the weighted geometric mean of the domestic expenditure shares, with share weights α_g^i multiplied by $-1/\varepsilon_g$ as follows,

$$Gains = \prod_g \left(\lambda_g^{ii\prime}\right)^{\frac{-\alpha_g^i}{\gamma_g(\omega-1)}\left[(\sigma_g - \omega)\left(\frac{\gamma_g}{\sigma_g - 1} - 1\right) + \sigma_g - 1\right]},$$

where $\lambda_g^{ii\prime}$ denotes the domestic expenditure share in a trade equilibrium, measured relative to $\lambda_g^{ii} = 1$ in autarky. In an effort to evaluate the importance of distinguishing between the micro and macro Armington elasticities we evaluate the gains from trade using our estimates of σ_g and ω and compare them to the gain under the assumption that $\sigma_g = \omega$. As noted before, in the latter case the trade elasticity is simply equal to the Pareto shape parameter γ_g, which in our model can be expressed as $\gamma_g = \zeta(\sigma_g - 1)$,

[21]For a complete derivation of this result see the online Appendix. We are grateful to Andrés Rodríguez-Clare for assistance in deriving this expression.

[22]There is a corresponding result for the Fréchet parameter θ in the Ricardian model of Eaton and Kortum (2002).

where ζ is the Pareto shape parameter for the firm size distribution. Using estimates of ζ in the literature, Table 5 shows that when we relax the assumption that the macro and micro elasticity are the same, our estimated gains from trade nearly triple when $\zeta = 1.1$. As ζ increases the overall estimated gains decline, but the relative importance of allowing the micro and macro elasticities to differ increases: for $\zeta = 1.6$ the estimated gains from trade when the macro and micro elasticity are five times larger then when they are the same, and for $\zeta = 2.1$ they are seven times larger. (See Appendix for full description of estimation).

6 Conclusions

In this paper, we distinguish between the substitution elasticity between alternative foreign import sources and the substitution elasticity between domestic and foreign import sources. These two elasticities are conceptually quite distinct. We show that they are in some cases empirically quite distinct, too. We find evidence that the former elasticity – which we call the "micro" Armington elasticity – is larger than the latter elasticity – the "macro" Armington elasticity. Our median estimates of the micro elasticity across individual industries are 3.24 and 4.12 for TSLS and 2-step GMM respectively, whereas the macro elasticities are lower in three-quarters of the goods we analyze and significantly lower in up to one-quarter of the goods, as in the approach to calibration in traditional CGE policy analysis. On the other hand, the fact that the micro and macro elasticities are not significantly different from each other in many of these cases offers some limited support for the newer generation of computable structural models, which do not allow for any difference between them.

Our finding that the micro elasticity is often larger than the macro elasticity has important implications for quantitative welfare assessments of trade policy. We show that when the macro and micro elasticities are different then the "trade elasticity" relevant for welfare analysis, ε, is a function of both elasticities and the Pareto shape parameter. Furthermore, the trade elasticity ε is less than the Pareto parameter when the macro elasticity is less than the micro elasticity. Since our results are often consistent with a smaller macro elasticity, and the trade gains are inversely proportional to ε, we find bigger trade gains than one might have expected based on ε equal to the Pareto parameter (as holds when the two elasticities are equal). This result is shown to be qualitatively important as well: the estimated gains from trade are between three and seven times greater once we relax the assumption that the micro and macro elasticities are equal.

We also find point estimates for the macro elasticity that exceed unity in almost all sectors. Values around unity are common in the various studies of substitution between domestic and imported goods carried out over decades by researchers who generally applied OLS to datasets more highly aggregated than ours. In contrast to these earlier works, ours is the first to estimate the micro and macro elasticities simultaneously at a disaggregate level for a number of products. Our econometric methodology, based on Feenstra (1994), corrects for potential biases in OLS estimation, including the errors introduced by reliance on unit-value price indexes rather than the exact indexes implied by theory. We frame the empirical analysis within a theoretical general-equilibrium trade model, based on Melitz (2003) and Chaney (2008), as a guide to both econometric specification and simulation analysis of alternative estimation approaches.

Our empirical findings raise the question of why substitution between home goods and imports is often lower than substitution between different foreign supply sources. Blonigen and Wilson (1999) documented several factors influencing the size of macro Armington elasticities across sectors, but to our knowledge there has been no corresponding study comparing macro to micro elasticities. One theoretical answer might come from the theory of discrete choice under uncertainty. Drawing on Anderson, de Palma, and Thisse (1992), a smaller elasticity for the macro Armington elasticity could be obtained if the variance of the random utility component between home and foreign goods in general is greater than the variance of the random utility component between two foreign (or home) varieties. Rauch and Trindade (2002, 2003) offer a micro-foundation for such an assumption, with empirical support, with additional empirical support coming from Cosar, Grieco and Tintelnot (2015).

Alternatively, low existing estimates of the macro elasticity may very well be due to differences between short-run and long-run elasticities. Gallaway, McDaniel, and Rivera (2003) have recently estimated short-run U.S. macro elasticities on monthly data that average 0.95, but long-run elasticities that are twice as large on average, in some cases up to five times larger. Our estimation is performed on annual data, so the elasticity estimates are not exactly short-run; but because we do not introduce lags in the adjustment of demand, we might view our estimates as applying to the "medium run." Introducing such an adjustment process in the theory and the estimation is an important avenue for future work.[23]

We close by emphasizing that while the macro Armington elasticity, which we have labeled ω, is the prime determinant of the aggregate *import* response to a terms of trade change, the overall *trade balance* sensitivity may depend powerfully on the micro elasticity governing substitution between alternative foreign suppliers. Once we move beyond the unrealistic assumption of a two-country world, it is evident that the *export* response to a terms of trade change depends not only on ω, but also on the foreign-foreign substitution elasticities that we labeled σ above.

As an example, suppose that the Korean won depreciates against all trading-partner currencies. Three things will happen. First, Korean residents will switch consumption from imports to domestic import-competing firms with elasticity ω. Second, consumers and firms outside Korea will switch from domestic goods competing with Korean exports to Korean exports with elasticity ω. But third, consumers and firms outside Korea will switch their demand from Korea's export competitors to Korea with elasticity σ. (For example, United States residents will import more ships and steel from Korea, less from China.) Thus, the overall effect of currency depreciation on Korea's net exports depends on both σ and ω. Because σ could be quite a bit larger than ω, there may be grounds for some degree of "elasticity optimism" after all.

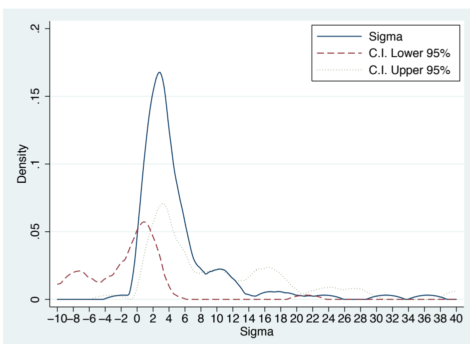
[23]Ruhl (2008) and Kehoe and Ruhl (2009) argue that due to supply-side responses, the expected permanence of a tariff cut gives it a much greater impact on trade flows than an equivalent temporary change in exchange rates.

Table 5: Gains from Trade

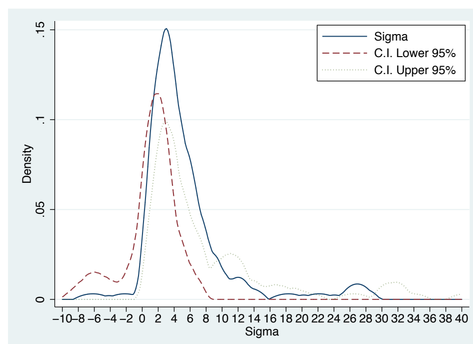
# of goods	ζ	$\sigma_g = \omega$	$\sigma_g \neq \omega$
89	1.1	1.169	1.585
		(1.165, 1.170)	(1.367, 1.740)
84	1.6	1.100	1.512
		(1.098, 1.101)	(1.309 1.654)
79	2.1	1.067	1.481
		(1.066, 1.068)	(1.227, 1.613)

Note: We report the estimated welfare gains from trade for both the naive case, where $\sigma_g = \omega$, as well as when we allow $\sigma_g \neq \omega$ using our 2-step GMM estimates of the micro and macro elasticities. For both cases we set γ_g equal to $\zeta(\sigma_g - 1)$, where ζ is the shape parameter of the firm size distribution, which is allowed to vary between 1.1 and 2.1. To calculate both the point estimates and the confidence intervals for the gains from trade we only consider estimates of σ_g and ω that imply positive estimates of ε. We construct 95% confidence intervals by estimating the gains from trade using our 500 bootstrapped estimates of σ_g and ω, and reporting the 2.5 and 97.5 percentiles in parentheses.

Figure 1: TSLS and GMM Results for Sigma



Panel A: TSLS



Panel B: 2-step GMM

References

Altonji, Joseph G., and Lewis M. Segal, "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics* 14(1996), 353–366.

Anderson, James E., and Eric van Wincoop, "Trade Costs," *Journal of Economic Literature* 42(2004), 691–751.

Anderson, Simon P., André de Palma, and Jacques-François Thisse, *Discrete Choice Theory of Product Differentiation* (Cambridge, MA: MIT Press, 1992).

Ardelean, Adina, and Volodymyr Lugovskyy, "Domestic Productivity and Variety Gains from Trade," *Journal of International Economics* 80(2010), 280–291.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?" *American Economic Review* 102(2012), 94–130.

Armington, Paul S., "A Theory of Demand for Products Distinguished by Place of Production," *IMF Staff Papers* 16(1969), 159–178.

Balistreri, Edward J., and Thomas F. Rutherford, "Computing General Equilibrium Theories of Monopolistic Competition and Heterogeneous Firms," In Peter B. Dixon, and Dale W. Jorgenson, eds., *Handbook of Computable General Equilibrium Modeling*, vol 1B (Amsterdam: Elsevier, 2013).

Baum, Christopher F., Mark E. Schaffer, and Steven Stillman, "Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing," *Stata Journal* 7(2007), 465–506.

Bergin, Paul, "How Well Can the New Open Economy Macroeconomics Explain the Exchange Rate and Current Account?" *Journal of International Money and Finance* 25(2006), 675–701.

Blonigen, Bruce A., and Wesley W. Wilson, "Explaining Armington: What Determines Substitutability between Home and Foreign Goods?" *Canadian Journal of Economics* 32(1999), 1–21.

Broda, Christian, and David E. Weinstein, "Globalization and the Gains from Variety," *Quarterly Journal of Economics* 121(2006), 541–585.

Chaney, Thomas, "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review* 98(2008), 1707–1721.

Cheng, H. S, "Statistical Estimates of Elasticities and Propensities in International Trade," *IMF Staff Papers* 7(1959), 107–158.

Cosar, A. Kerem, Paul L. E. Grieco, and Felix Tintelnot, "Borders, Geography, and Oligopoly: Evidence from the Wind Turbine Industry," *Review of Economics and Statistics* 107(2015), 623–637.

Costinot, Arnaud, and Andrés Rodríguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," In Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Economics,* vol. 4, (Amsterdam: Elsevier, 2014)

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Global Rebalancing with Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers* 55(2008), 511–540.

Di Giovanni, Julian, Andrei Levchenko, and Jing Zhang, "The Global Welfare Impact of China: Trade Integration and Technological Change," *American Economic Journal: Macroeconomics* 6(2014), 153–183.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica* 70(2002), 1741–1779.

Feenstra, Robert C, "New Product Varieties and the Measurement of International Prices," *American Economic Review* 84(1994), 157–177.

Feenstra, Robert C, *Product Variety and the Gains from International Trade* (Cambridge, MA: MIT Press, 2010).

Feenstra, Robert C., Philip A. Luck, Maurice Obstfeld, and Katheryn N. Russ, "In Search of the Armington Elasticity," NBER Working Paper 20063(2014)

Feenstra, Robert C, and John Romalis, "International Prices and Endogenous Product Quality," *Quarterly Journal of Economics* 129(2014), 477–528.

Gallaway, Michael P., Christine A. McDaniel, and Sandra A. Rivera, "Short and Long-Run Industry-Level Estimates of U.S. Armington Elasticities," *North American Journal of Economics and Finance* 14(2003), 49–68.

Goldstein, Morris and Mohsin S. Khan, "Income and Price Effects in Foreign Trade," In Ronald W. Jones and Peter B. Kenen, eds., *Handbook of International Economics* vol. 2, (Amsterdam: North-Holland, 1985).

Hansen, Lars Peter, John Heaton, and Amir Yaron, "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business and Economic Statistics* 14(1996), 262–280.

Harberger, Arnold C, "Some Evidence on the International Price Mechanism," *Journal of Political Economy* 65(1957), 506–521.

Harrison, Glenn W., Thomas F. Rutherford, and David G. Tarr, "Quantifying the Uruguay Round," *Economic Journal* 107(1997), 1405–1430.

Heathcote, Jonathan and Fabrizio Perri, "Financial Autarky and International Business Cycles," *Journal of Monetary Economics* 49(2002), 601–627.

Hillberry, Russell, and David Hummels, "Trade Elasticity Parameters for a Computable General Equilibrium Model," In Peter B. Dixon, and Dale W. Jorgenson, eds., *Handbook of Computable General Equilibrium Modeling* vol 1B, (Amsterdam: Elsevier, 2013).

Imbs, Jean, and Isabelle Méjean, "Elasticity Optimism," *American Economic Journal: Macroeconomics* 7(2015), 43–83.

Kehoe, Timothy J, and Kim J. Ruhl, "How Important is the New Goods Margin in International Trade?" Federal Reserve Bank of Minneapolis, Research Department Staff Report 324(2007).

Kucheryavyy, Konstantin, Gary Lyn, and Andres Rodríguez-Clare, "External Economies and Internatonial Trade: A Quantitative Framework," mimeo (2015), University of California-Berkeley.

Lai, Huiwen, and Daniel Trefler, "The Gains from Trade with Monopolistic Competition: Specification, Estimation, and Mis-specification," NBER Working Paper 9169 (2002).

Leamer, Edward E., and Robert M. Stern, *Quantitative International Economics*, (Boston: Allyn and Bacon, 1970).

Machlup, Fritz, "Elasticity Pessimism in International Trade," *Economia Internazionale* 3(1950), 118-141.

MacKinnon, James A., "Bootstrap Methods in Econometrics," *Economic Record* 82(2006), S2–S18.

Magee, Stephen P., "Prices, Incomes, and Foreign Trade," In Peter B. Kenen, ed., *International Trade and Finance: Frontiers for Research*, (New York: Cambridge University Press, 1975).

Marquez, Jaime. 2002. *Estimating Trade Elasticities*. Boston: Kluwer Academic Publishers.

Mattoo, Aaditya, Prachi Mishra, Arvind Subramanian, "Beggar-thy-Neighbor Effects of Exchange Rates? A Study of the Renminbi," *American Economic Journal: Economic Policy (forthcoming)*

McDaniel, Christine A., and Edward J. Balistreri, "A Review of Armington Trade Substitution Elasticities," *Économie Internationale* 23(2003): 301–313.

Melitz, Marc J.,"The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica* 71(2003), 1695–1725.

Orcutt, Guy H., "Measurement of Elasticities in International Trade," *Review of Economics and Statistics* 32(1950), 117-132.

Rauch, James E., and Vitor Trindade, "Ethnic Chinese Networks In International Trade," *The Review of Economics and Statistics* 84(2002), 116–130.

Rauch, James E., and Vitor Trindade, “Information, International Substitutability, and Globalization,” *American Economic Review* 93(2003), 775–791.

Reinert, Kenneth A., and David W. Roland-Holst, “Armington Elasticities for United States Manufacturing Sectors,” *Journal of Policy Modeling* 14(1992), 631–639.

Rigobon, Roberto, “Identification Through Heteroskedasticity,” *Review of Economics and Statistics* 85(2003), 777–792.

Romalis, John, “NAFTA’s and CUSFTA’s Impact on International Trade,” *Review of Economics and Statistics* 89(2007), 416–435.

Ruhl, Kim J., “The International Elasticity Puzzle,” New York University Stern School of Business (2008).

Sato, Kazuo, “The Ideal Log-Change Index Number,” *Review of Economics and Statistics* 58(1976), 223–228.

Shiells, Clinton R., Robert M. Stern, and Alan Deardorff, “Estimates of the Elasticities of Substitution between Imports and Home Goods for the United States,” *Weltwirtschaftliches Archiv* 122(1986), 497–519.

Soderbury Anson, “Investigating the Asymptotic Properties of Elasticity of Substitution Estimates,” *Economic Letters* 190(2010), 57–62.

Soderbury Anson, “Estimating Import Supply and Demand Elasticities: Analysis and Implications,” Purdue University (2012).

Stern, Robert M., Jonathan Francis, and Bruce Schumacher, *Price Elasticities in International Trade: An Annotated Bibliography*(London: Macmillan, 1976).

Tokarick, Stephen, “The Implications of China’s Pattern of Growth for the Rest of the World,” International Monetary Fund (2012).

Vartia, Y.O., “Ideal Log-Change Index Numbers,” *Scandinavian Journal of Statistics* 3(1976), 121–126.