In Search of the Armington Elasticity*

Robert C. Feenstra Philip Luck Maurice Obstfeld Katheryn N. Russ

April 16, 2014

Abstract

The elasticity of substitution between goods from different countries—the Armington elasticity—is important for many questions in international economics, but its magnitude is subject to debate: the "macro" elasticity between home and import goods is often found to be smaller than the "micro" elasticity between foreign sources of imports. We investigate these two elasticities in a model using a nested CES preference structure. We explore estimation techniques for the macro and micro elasticities using both simulated data from a Melitz-style model, and highly disaggregate U.S. production data matched to Harmonized System trade data. We find that in up to one-half of goods there is no significant difference between the macro and micro elasticities, but in the other half of goods the macro elasticity is significantly lower than the micro elasticity, even when they are estimated at the same level of disaggregation.

*UC Davis and NBER, UC Davis, 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 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 the role of international prices in trade balance adjustment, the optimal extent of international portfolio diversification, the effects of regional trade agreements, and the welfare benefits of expanding world trade.

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 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, 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

[1] Goldstein and Khan (1985) survey a large body of research on empirical aggregate import equations and endorse a much 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 (2013) 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).

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]

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 that can be significant in magnitude: there is not just bias in small samples, but inconsistency in large samples. We illustrate this bias/inconsistency in section 3 through OLS estimation using simulated as well as U.S. data. The simulated dataset incorporates shocks to tastes and technologies that are calibrated from the literature, and results in downward-biased estimates of both the micro and macro Armington elasticities. 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.

In section 4 we draw on Feenstra (1994) to propose a generalized method of moments (GMM) estimation strategy that – in large samples at least – corrects for the statistical 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. Soderbury (2010, 2012) has recently identified small-sample biases in this estimator using simulated data.[4] That concern is amplified by our extension of this estimator to the nested CES framework in this paper, where the estimating equation is considerably more complicated than in Feenstra (1994). We apply the GMM estimator to our simulated data and find that this small-sample bias persists even for sample sizes considerably larger than what we have in our U.S. data. We then suggest a possible solution: adding an additional moment condition to the estimation.

In general, GMM estimation is improved by adding additional moment conditions. But as noted by Wooldridge (2001, p.91), extra conditions can add noise unless they are conceptually well-founded. 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

[3] Ardelean and Lugovskyy (2010) also introduce a nested CES structure into a monopolistic competition and trade model, but do not allow for heteregeneous firms as in the Melitz-Chaney model. Ideally, the difference between the "micro" and "macro" elasticities should be derived from an underlying micro-structure, but we do not attempt to solve that problem here. See the recent paper by Cosar et al. (2011) for empirical evidence on home bias (which they describe as "border frictions") in the wind turbine industry in Europe. One relevant factor might be the perceived security of supply. Blonigen and Wilson (1999) identify some structural determinants of empirical Armington substitution elasticities between U.S. goods and imports.

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

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. As we discuss in section 4, this moment condition adds new information. We find that after using this extra condition, the sectoral macro elasticities estimates in our U.S. data rise in all sectors, becoming greater than unity in most cases. Still, the macro elasticity is found to below micro elasticities for about half of the nearly 100 goods.

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 HS level. In section 5 we briefly explore the theoretical implications of aggregation across goods. We find that there exists an aggregate import demand equation with the same macro elasticity that applies at more disaggregate levels. In other words, aggregation itself need not lead to a lower value of the macro elasticity.[5] This result is at first sight surprising: the nested CES utility function that we assume is not weakly separable across imports of all goods and domestic consumption of all goods, so it does not obviously allow one to treat imports as a separate aggregate with a well-defined price index. However, drawing an analogy to the “latent separability” concept of Blundell and Robin (2000), we show that our utility function allows for consistent aggregation across goods even when conventional weak separability does not hold. Section 6 applies out model and empirical estimates to the classic question of calculating the impact on imports of currency devaluation. Our discussion of devaluation effects highlights the key role of the macro Armington elasticity.

In summary, our paper gives a nuanced answer to the question of whether the micro and macro elasticites differ. For at least one-half of the goods in our sample, the elasticities are not significantly different. That result gives limited support to the new computable models initiated by Dekle, Eaton and Kortum (2007, 2008), which do not introduce allow differences between the micro and macro elasticities of substitution. For the remaining sectors, however, the macro elasticity is significantly lower, albeit not as low as the estimates around unity often found using macro time-series methods. For many policy questions, such as gauging the effect of a currency devaluation on aggregate imports, it therefore remains important to take account of potential differences between the micro and macro elasticities. Further conclusions are given in section 7, and various technical results are gathered in two 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.

[5] This theoretical result would be offset, however, if the macro elasticity differs across goods in a way that is correlated with the residual in the estimating equation. These circumstances would lead to "heterogeneity bias" or "aggregation bias" in the elasticity estimate, as analyzed by Dekle, Jeong, and Ryoo (2013) and Imbs and Méjean (2013). That source of bias is not the focus here. The classic treatment is Orcutt (1950).

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 varities of good g that country j imports from country i by N_g^{ij}. (It 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

$$\begin{aligned} P_g^j &= \left\{ \beta_g^j \left(P_g^{jj} \right)^{1-\omega_g} + \left(1 - \beta_g^j \right) \left(P_g^{Fj} \right)^{1-\omega_g} \right\}^{\frac{1}{1-\omega_g}} \\ &= \left\{ \beta_g^j \left(P_g^{jj} \right)^{1-\omega_g} + \left(1 - \beta_g^j \right) \left[\sum_{i=1, i \neq j}^{J} \kappa_g^{ij} \left(P_g^{ij} \right)^{1-\sigma_g} \right]^{\frac{1-\omega_g}{1-\sigma_g}} \right\}^{\frac{1}{1-\omega_g}} \\ &= \left\{ \beta_g^j \left[\int_{N_g^{jj}} p_g^j (\varphi)^{1-\sigma_g} \, \mathrm{d}\varphi \right]^{\frac{1-\omega_g}{1-\sigma_g}} \right. \\ &\quad \left. + \left(1 - \beta_g^j \right) \left[\sum_{i=1, i \neq j}^{J} \kappa_g^{ij} \int_{N_g^{ij}} \left(\tau_g^{ij} p_g^i (\varphi) \right)^{1-\sigma_g} \mathrm{d}\varphi \right]^{\frac{1-\omega_g}{1-\sigma_g}} \right\}^{\frac{1}{1-\omega_g}}. \end{aligned}$$

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 $\varphi \in [\hat{\varphi}_g^{ij}, \infty)$ export to j and the measure of varieties of g exported from i to j is given by $N_g^{ij} = 1 - H_g^i(\hat{\varphi}_g^{ij})$.

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:

$$\begin{aligned} \pi_g^{ij}(\hat{\varphi}_g^{ij}) &= \frac{\tau_g^{ij} p_g^i(\hat{\varphi}_g^{ij})\, \kappa_g^{ij}}{\sigma_g}\left[\frac{\tau_g^{ij} p_g^i(\hat{\varphi}_g^{ij})}{P_g^{Fj}}\right]^{-\sigma_g}(1-\beta_g^j)\left(\frac{P_g^{Fj}}{P_g^j}\right)^{-\omega_g}\alpha_g^j\left(\frac{P_g^j}{P^j}\right)^{-\eta} C_j \\ &= W^i f_g^{ij}. \end{aligned} \tag{5}$$

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

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

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:

$$H_g^i(\varphi) = 1 - \varphi^{-\gamma_g^i}. \tag{7}$$

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

$$P_g^{ij} = \left(\frac{\sigma_g}{\sigma_g - 1} \right) \left[\frac{\gamma_g^i}{\gamma_g^i - (\sigma_g - 1)} \right] \frac{\tau_g^{ij} W^i}{A^i A_g} (N_g^{ij})^{\frac{-[\gamma_g^i - (\sigma_g - 1)]}{\gamma_g^i(\sigma_g - 1)}}, \tag{8}$$

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 $t = 1, ..., T$.[6] 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 imports of good g from country $i \neq j$ (covering all varieties N_{gt}^{ij}) as:

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

Spending on good g from home supply is:

$$V_{gt}^{jj} = \alpha_{gt}^j \beta_{gt}^j \left(\frac{P_{gt}^{jj}}{P_{gt}^j} \right)^{1-\omega_g} \left(\frac{P_{gt}^j}{P_t^j} \right)^{1-\eta} P_t^j C_t^j. \tag{10}$$

[6] 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.

Dividing (9) and (10) we obtain imports from country i relative to home demand,

$$\frac{V_{gt}^{ij}}{V_{gt}^{jj}} = \kappa_{gt}^{ij} \left(\frac{1-\beta_{gt}^{j}}{\beta_{gt}^{j}} \right) \left(\frac{P_{gt}^{ij}}{P_{gt}^{Fj}} \right)^{1-\sigma_g} \left(\frac{P_{gt}^{Fj}}{P_{gt}^{jj}} \right)^{1-\omega_g} . \tag{11}$$

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}.[7]

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 (8), 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 average of prices:

$$UV_{gt}^{ij} = \int_{\hat{\varphi}_{gt}^{ij}}^{\infty} \tau_{gt}^{ij} p_{gt}^{i}(\varphi) \left[\frac{c_{gt}^{ij}(\varphi)}{\int_{\hat{\varphi}_{gt}^{ij}}^{\infty} c_{gt}^{ij}(\varphi)\, dH_g^i(\varphi)} \right] dH_g^i(\varphi) . \tag{12}$$

To simplify this expression, we make use of $c_{gt}^{ij}(\varphi_1) = c_{gt}^{ij}(\varphi_2)(\varphi_1/\varphi_2)^{\sigma_g}$ to evaluate the integral appearing in the denominator as

$$\begin{aligned} \int_{\hat{\varphi}_{gt}^{ij}}^{\infty} c_{gt}^{ij}(\varphi)\, dH_g^i(\varphi) &= \int_{\hat{\varphi}_{gt}^{ij}}^{\infty} c_{gt}^{ij}\left(\hat{\varphi}_{gt}^{ij}\right) \left(\frac{\varphi}{\hat{\varphi}_{gt}^{ij}} \right)^{\sigma_g} dH_g^i(\varphi) \\ &= c_{gt}^{ij}\left(\hat{\varphi}_{gt}^{ij}\right) \int_{\hat{\varphi}_{gt}^{ij}}^{\infty} \left(\frac{\varphi}{\hat{\varphi}_{gt}^{ij}} \right)^{\sigma_g} \gamma_g^i \varphi^{-\gamma_g^i - 1} d\varphi \\ &= \frac{\gamma_g^i}{(\gamma_g^i - \sigma_g)} c_{gt}^{ij}\left(\hat{\varphi}_{gt}^{ij}\right) \left[1 - H_g^i\left(\hat{\varphi}_{g}^{ij}\right)\right], \end{aligned}$$

where $\gamma_g^i > \sigma_g$ is assumed.

This expression illustrates a general property of integrating a power function of the form φ^n (where $n < \gamma_g^i$) using the Pareto distribution: the result is the *initial value* of the function, $\left(\hat{\varphi}_{gt}^{ij}\right)^n$, times the hazard rate $1 - H_g^i\left(\hat{\varphi}_{gt}^{ij}\right)$, times a factor of proportionality. Applying this rule to the rest

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

of the integral in (12), the initial values and hazard rates cancel and we readily obtain:

$$UV_{gt}^{ij} = \frac{\left(\gamma_g^i - \sigma_g\right)}{\left(\gamma_g^i - \sigma_g + 1\right)} \tau_{gt}^{ij} p_{gt}^i \left(\hat{\varphi}_{gt}^{ij}\right).$$

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

$$\frac{UV_{gt}^{ij}}{UV_{gt-1}^{ij}} = \left(\frac{\tau_{gt}^{ij} W_t^i / A_{gt} A_t^i}{\tau_{gt-1}^{ij} W_{t-1}^i / A_{gt-1} A_{t-1}^i}\right) \left(\frac{N_g^{ij}}{N_{gt-1}^{ij}}\right)^{1/\gamma_g^i} = \frac{P_{gt}^{ij}}{P_{gt-1}^{ij}} \left(\frac{N_{gt}^{ij}}{N_{gt-1}^{ij}}\right)^{1/(\sigma_g - 1)}, \tag{13}$$

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 (8). It is apparent from (13) that the unit value is *positively* associated with an increase in product variety N_{gt}^{ij}, in contrast to the CES price index in (8). 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 (13). 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:[8]

$$\frac{P_{gt}^{Fj}}{P_{gt-1}^{Fj}} = \prod_{i=1, i\neq j}^{J} \left[\left(\frac{\kappa_{gt}^{ij}}{\kappa_{gt-1}^{ij}}\right)^{\frac{1}{1-\sigma_g}} \frac{P_{gt}^{ij}}{P_{gt-1}^{ij}}\right]^{w_{gt}^{ij}}, \tag{14}$$

where s_{gt}^{ij} is the share of country j's imports of g from i in total imports from all foreign countries, (so that $\sum_{i\neq j} s_{gt}^{ij} = 1$) and:

$$w_{gt}^{ij} \equiv \frac{\left(\frac{s_{gt}^{ij} - s_{gt-1}^{ij}}{\ln s_{gt}^{ij} - \ln s_{gt-1}^{ij}}\right)}{\sum_{i=1, i\neq j}^{J} \left(\frac{s_{gt}^{ij} - s_{gt-1}^{ij}}{\ln s_{gt}^{ij} - \ln s_{gt-1}^{ij}}\right)}. \tag{15}$$

The numerator in (15) is the "logarithmic mean" of the import shares s_{gt}^{ij} and s_{gt-1}^{ij}, and lies in between these two shares, while the denominator ensures that the weights w_{gt}^{ij} sum to unity. The special formula for these weights in (15) is needed for the geometric mean in (14) precisely to

[8] To prove (14), note that $s_g^{ij} = \kappa_g^{ij}\left(P_g^{ij}/P_g^{Fj}\right)^{1-\sigma_g}$, so that $P_g^{Fj} = \left(\kappa_g^{ij}/s_g^{ij}\right)^{\frac{1}{1-\sigma_g}} P_g^{ij}$. 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 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 (14).

measure the ratio of the CES functions, $P_{gt}^{Fj}/P_{gt-1}^{Fj}$. But in practice, the Sato-Vartia formula will give very similar results to those obtained using other index number formulas.

If we use the unit values $UV_{gt}^{ij}/UV_{gt-1}^{ij}$ instead of the CES prices on the right of (14), 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,

$$\frac{UV_{gt}^{Fj}}{UV_{gt-1}^{Fj}} = \left(\frac{P_{gt}^{Fj}}{P_{gt-1}^{Fj}} \right) \left(\frac{\kappa_{gt}^{Fj} N_{gt}^{Fj}}{\kappa_{gt-1}^{Fj} N_{gt-1}^{Fj}} \right)^{\frac{1}{(\sigma_g - 1)}} \tag{16}$$

as is obtained by using (8) and (14), 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 (11), (13) and (16) we have,

$$\Delta \ln \left(\frac{V_{gt}^{ij}}{V_{gt}^{jj}} \right) = -(\sigma_g - 1) \Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) + (1 - \omega_g) \Delta \ln \left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}} \right) + \varepsilon_{gt}^{ij}, \; i = 1, ..., J, i \neq j, \tag{17}$$

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

$$\varepsilon_{gt}^{ij} \equiv \Delta \ln \left(\frac{\kappa_{gt}^{ij}}{\kappa_{gt}^{Fj}} \right) + \Delta \ln \left(\frac{N_{gt}^{ij}}{N_{gt}^{Fj}} \right) + \Delta \ln \frac{(1 - \beta_{gt}^{j})}{\beta_{gt}^{j}} - \frac{(1 - \omega_g)}{(\sigma_g - 1)} \Delta \ln \left(\frac{\kappa_{gt}^{Fj} N_{gt}^{Fj}}{N_{gt}^{jj}} \right), \tag{18}$$

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

3 Data and OLS estimation

We have every reason to expect that the error term is correlated with the relative prices that appear on the right of (17). 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. 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.

There is an interesting difference in the micro and macro biases, however. Notice that macro coefficient $(1-\omega_g)$ applies to the final variable in (17) and also applies – with a minus sign – to the final source of error in (18). That pattern is exactly the structure of classical measurement error, where the term βx in a regression equation becomes $\beta(x+u)-\beta u$ when x is measured with the error u. The term $-\beta u$ is of course added to the error in the regression model, and a familiar calculation show that this classical measurement error leads to a bias toward zero in β. We cannot claim that the measurement error in our model is classical (i.e. uncorrelated with other variables), because the final source of error in (18) is endogenous within our model. Neverthless, the similarity in structure to classical error – with $(1-\omega_g)$ applying to both the final variable in (17) and the final source of error in (18) – means that we can still expect a bias towards zero in the OLS estimate of $(1-\omega_g)$ so that the macro elasticity ω_g is *biased towards unity.*

That result does not apply to the micro elasticity σ_g, however, because the measurement error in (13) itself depends on the micro elasticity, appearing there as $1/(\sigma_g-1)$. As a result, the first two terms in (18) do not depend on any coefficients. We have argued above that σ_g is downward biased, but there is not necessarily a bias towards zero in (σ_g-1). Rather, it can be shown more formally that the downward bias in σ_g can easily lead to OLS estimates below unity (or even below zero). For both the micro and macro elasticity estimates, we should not expect their OLS biases to diminish in large samples: the simultaneity and measurement error biases also lead to inconsistency. We demonstrate this bias/inconsistency in the micro and macro elasticities in the next two subsections: first using simulated data, and then using a new U.S. dataset containing disaggregate data for both imports and domestic production.

3.1 Simulated Data

We simulate a small-scale example of our theoretical model to demonstrate that the shocks to productivity and preferences can generate the observed downward bias in OLS estimates. 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, as follows. Appendix A shows how we solve for the model's equilibrium.

We assume that all productivity and taste shocks are covarying and log-normally distributed. To this end, we draw random numbers for each from a multivariate normal distribution, then exponentiate them such that they form covarying log-normal random variables. We calibrate the standard deviation of the log of country-wide productivity shocks A^i and industry-level productivity shocks A_g using estimates by Basu, Fernald and Kimball (2006, hereafter BFK) as 0.015 and 0.02, respectively. These industries are disaggregated only at the 2-digit SIC level, but have the advantage of being perfectly matched in methodology and source to the aggregate productivity shock measure. While the value from BFK serves as the baseline standard deviation of country-specific shocks (applied to the largest country), we make the productivity and preference shocks heteroskedastic to be in accordance with identifying conditions for the GMM method that we will use later and that is discussed below. We scale the BFK values for the standard deviations (0.015 for $\log(A^i)$ and 0.02 for $\log(A_g)$) along the diagonal of the portion of the variance-covariance matrix corresponding to log productivity shocks across countries by a constant that is equally spaced along the interval [0.95, 1.05]. The same is done for the standard deviation of log productivity shocks across goods, using a constant equally spaced along the interval [0.4, 1]. Thus, these scaled values make up the diagonals of our variance-covariance matrix for the portion corresponding to the log of productivity shocks A^i and A_g.

As a guide to calibrate the standard deviations for our taste shocks, we use the standard deviation of demand shocks (in logs, demeaned by year and product and controlling for income) estimated by Foster, Haltiwanger, and Syverson (2008), for which industries are characterized by 7-digit SIC codes. Foster, Haltiwanger, and Syverson find that the standard deviation of log preference shocks is five times as large as that of log productivity shocks using plant-level data. We apply the same proportion to the relative size of the standard deviation of the log taste shocks here for a median good and the industry-level productivity shocks calibrated using Basu, Fernald, and Kimball (2006). That is, half of the good-country ($\log(\alpha_g^j)$) and good-country-pair ($\log(\kappa_g^{ij})$) taste shocks have standard deviations below the standard deviation of $\log(A_g)$ and half greater. We scale the diagonal values of the variance-covariance matrix corresponding to taste shocks α_g^j by a constant spaced over the interval [0.1, 20].

For the taste shocks κ_g^{ij} applying to bilateral country pairs, we include both a country-pair (κ^{ij}) and a country-pair-good ($\hat{\kappa}_g^{ij}$) component, setting $\kappa_g^{ij} = 0.5\kappa^{ij} + 0.5\hat{\kappa}_g^{ij}$. We do this because bilateral taste shocks that apply equally across goods are key to dampening OLS estimates of σ, but taste shocks that are heteroskedastic across goods are key to identifying ω_g in the GMM estimation.[9] Because there is little empirical guidance to calibrate the variances, we give the log of these shocks a mean of zero and a minimum variance of 0.25, scaling the diagonal values of the portion of the variance-covariance matrix corresponding to $\hat{\kappa}_g^{ij}$ by a constant equally spaced over the interval [0.5, 1] and the diagonal values of the portion of the variance-covariance matrix corresponding to $\hat{\kappa}_g^{ij}$ by a constant equally spaced over the interval [0.1, 10]. Again, the median good has log taste shocks α_g^j and κ_g^{ij} roughly five times the size of the variance of its corresponding productivity shocks, which is in line with the estimates by Foster, Haltiwanger, and Syverson (2008).[10] Because the κ_g^{ij} taste shocks are shares that must sum to one, we divide each one by their sum.

We set the mean of β_g^j equal to 0.5 for all countries and goods, allowing home bias to emerge from trade costs rather than a presumed preference weight. There are no obvious estimates of the variance of β_g^j, so we set the standard deviation of these shocks to be 1.5 percent of the size of the variance of the taste shocks κ_g^{ij} and α_g^j – making them about 20% of the size of aggregate productivity shocks – to give them macroeconomic significance without allowing them to exceed

[9]See Assumption 2 in our later discussion.

[10]The good with the least volatility in κ_g^{ij} is good one exported from country two to country one. The good with the most volatility is the good 10 exported from country four to country five.

the size of other shocks that have more precise estimates in the literature.[11]

There is little evidence in the literature measuring the correlation between productivity shocks and taste shocks. We set the mean covariances for all log variables (the off-diagonals of the multivariate normal distribution's variance-covariance matrix) equal to 0.75. Being agnostic as to the sign or relative size of these covariances, we shock them each period, multiplying the mean 0.75 by a normally distributed random variable with mean zero and standard deviation 0.25. So the off-diagonals of the variance-covariance matrix for the normal draws are given by 0.75 multiplied by these normally distributed shocks.

We vary country size (L_j) so that the largest country is five times as large as the smallest country and the size of the median country is normalized to equal 1.[12] The variable trade cost τ is 1.15, about halfway between the minimum and maximum values for trade costs estimated by Hummels (2007). The fixed cost of production f_g^{ij} is set to 0.05.[13] We set ω_g equal to 2, in accord with Ruhl (2008), and set σ_g equal to 3, within the range of estimates of the elasticity of substitution between imported varieties identified by Feenstra (1994) and Broda and Weinstein (2006). The Pareto shape parameter γ_g is set equal to 3.5 so that the distribution of firm size has a fat tail (di Giovanni, Levchenko, and Rancière 2010).[14]

Using this simulated dataset, we estimate (17) using panel OLS over all export sources i and goods g. We identify the median estimate from 1,000 simulations for both σ and ω separately and then use a bootstrap on the associated dataset to obtain confidence intervals.[15] The median estimate of σ for the 1,000 simulated datasets with $T = 10$ is 0.369, as shown in the first row of Table 1, with a 95% confidence interval of $(0.12, 1.95)$. The true value of σ, which is 3, does not lie within the confidence interval, demonstrating a very large downward bias in the OLS estimate. That finding is in line with our discussion above, where we expected a bias in the OLS estimate that is downward and not necessarily towards unity. The median estimate of 0.369 is obtained with only $T = 10$ time periods in the simulated data, but when we use more time periods such as 50 and 100 as reported in Table 1, the bias does not noticeably lessen. As we expected, the OLS bias becomes inconsistency in large samples.

[11] We also experimented with giving all shocks an autocorrelation coefficient of between 0 and 0.9, corresponding to the measured behavior of productivity shocks in the macroeconomics literature. Since GMM corrects for autocorrelation issues, our results are unchanged, and so shocks are not autocorrelated in our baseline analysis.

[12] Note that our assumption of a fixed labor supply corresponds to the common assumption in the macroeconomics literature that (dis)utility is quadratic in labor. Our results are robust to allowing small shocks to the labor supply, analogous to shocks in the elasticity of labor supply.

[13] We calibrate the fixed cost to roughly match the percentage of the U.S. working-age population that are business owners, as reported by the Global Entrepreneurship Monitor. The GEM reports that 15.5% of the working-age population owned a nascent, new, or established business in 2007 (Phinisee et al. 2008). It is difficult for us to target this level precisely due to the endogenous number of firms N_g^{ij}. The fraction $N_g^{ij} f_g^{ij} / L_i$ averages 15.1% for the largest country in our simulations. The results are robust to increases the size of the fixed costs.

[14] In addition to ensuring sufficient mass in the tail of the firm size distribution, choice of 3.5 for the Pareto shape parameter is broadly consistent with the literature. See Simonovska and Waugh (2011, 2012) for the most recent estimates.

[15] In the bootstrap we do not redraw observations for entire countries, as sometimes used in panel work, because the number of countries is too small. Instead, we hold the sample of countries fixed and redraw observations by year. This bootstrap method is used to obtain standard errors for all OLS and GMM estimation.

Table 1: Median OLS Estimates of the Micro and Macro Elasticities from Simulated and U.S. Data

	OLS estimates from Eq. (17)	
	Sigma	Omega
Data	(1)	(2)
Simulated	0.369	1.35
$T = 10$	(0.12, 1.95)	(0.61, 2.10)
Simulated	0.370	1.34
$T = 50$	(-0.90, 1.54)	(0.86, 1.85)
Simulated	0.373	1.40
$T = 100$	(-0.39, 1.13)	(1.01, 1.78)
U.S.	1.06	0.89
$T \leq 15$	(0.89, 1.23)	(0.54, 1.24)

Note: 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. The final row reports the median estimate over 109 U.S. goods. For both the simulated and U.S. data, the confidence intervals shown in parentheses are computed by bootstrapping the dataset corresponding to the median estimate.

The median estimate for ω is 1.35, with a 95% confidence interval of [0.607, 2.101] obtained through the same bootstrap process. Because the true value 2 for this macro elasticity is contained within the confidence interval, we are finding less downward bias than was the case for the micro elasticity. Once again, the median estimates of the macro elasticity do not vary much as we expand the number of time periods in the simulation, falling from 1.35 to 1.34 and then rising to 1.40 as T expand s from 10 to 50 to 100. These results suggest a bias/inconsistency towards unity in the OLS estimates of the macro elasticity, which is in line with our expectations.

3.2 U.S. Data

We run analagous OLS regressions to investigate U.S. import supply for multiple foreign countries at the most disaggregate level possible. 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,[16] and

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

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.

Limitations of the CIR data are that: (i) it is only a subset of U.S. manufacturing industries; (ii) the list of industries changes over time, especially with the shift from SIC (Standard Industrial Classification) to NAICS (North American Industry Classification System) in 1997; (iii) not all industries include import, export and U.S. domestic supply data for both values and quantities (as needed to compute unit values); (iv) while a concordance from HS to the "product codes" used to track industries in CIR is provided, a given HS is sometimes associated with more than one product code. In the latter case, we needed to aggregate U.S. shipments across multiple product codes to obtain a correspondence to the import and export data.

After this aggregation procedure, the resulting dataset has 191 goods, by which we mean an SIC-based product code (up to 7 digits) or a NAICS-based product code (up to 10 digits). Of these, 80 goods are based on a single 10-digit Harmonized System (HS) commodity, and another 42 goods are based on two or three 10-digit HS commodities. So the majority of the dataset is at a highly disaggregate level: this is the first time we are aware of that U.S. production data have been matched to imports and exports at such a disaggregate level (although there are earlier efforts at higher levels of aggregation, such as Reinert and Roland-Holst 1992 and Gallaway, McDaniel , and Rivera 2003). Since we have the matching exports and imports for these 191 goods, we can also compute U.S. apparent consumption, and consumption from U.S. supply, as appears in the denominator of the dependent variable in (17). When estimating this equation we pool across goods when they share some common HS commodities: this happens frequently for a SIC and NAICS-based product pair spanning 1992-1996 and 1997-2007, but sometimes a product code in one period will correspond to two codes in the other period, or there may be no correspondence over time. So after this pooling we end up with 109 slightly broader "goods" used in the estimation rather than 191. Each of these 109 goods is available for at most 16 years, giving $T = 15$ when the data are first differenced, but very often less than that.

Using this dataset, we estimate (17) using panel OLS over all export sources i for each good. At the bottom of Table 1 we report the median estimate of the micro elasticity σ_g obtained over the 109 goods, which is 1.06. Once again, its confidence interval is obtained by bootstrapping the dataset for the product corresponding to that median estimate. The kernel density of the estimates of σ_g is graphed in Figure 1, along with the kernel densities of the lower and upper 95% confidence bounds. Approximately 90% of the estimates of σ_g are below 1.5, which indicates rather low estimates for this micro Armington elasticity as compared with other estimates in the literature (e.g. Broda and Weinstein, 2006). So the median estimates of the micro elasticities from our U.S. data appear to be low, though not as low as those those from the simulated data, where the median estimates are 0.37. The fact that our estimates of the micro elasticity from U.S. data have a median so close to unity, and a sizable mass of the kernel density around that point, suggests to us that there is a good deal of classical measurement error in the U.S. unit values. That is not a surprising result when considering the erratic nature of import unit values at the Harmonized System level. The errors in our simulated model, by contrast, do not give rise to classical measurement error in the estimation of σ_g.

At the bottom of Table 1 we also report the median OLS estimate of the macro elasticity ω_g, which is 0.89. The kernel density of the estimates of ω_g is graphed in Figure 2. The bias in this OLS estimate of the macro elasticity is unknown because we do not have more reliable estimates from the literature – at the same level of disaggregation – to compare it to. But the fact that the median estimate of the macro elasticity is so close to unity, again with a sizable mass of the kernel density around this point, suggests to us again that classical measurement error applies to the OLS estimate of the macro elasticity, too.

4 Estimating with Moment Conditions

We turn now to an improved estimation technique that will allow us to estimate the micro and macro elasticities with less bias, and test whether these two elasticities differ. 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 moment condition that is used to obtain the micro elasticity σ_g and corrects for simultaneous equation bias and measurement error. This procedure is described as "step 1" below.

To also obtain the macro elasticity, we begin with the demand for imports *relative to home demand*. That 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. Assuming again that the error terms in the demand and supply equations are uncorrelated gives us a second moment condition that can be used to obtain the micro and macro elasticities, in "step 2." We will find from our simulated dataset that there is still some small-sample bias in the estimation of the micro and macro elasticities. We can offset this bias by increasing the number of years in our simulation, but cannot do so with our U.S. data, which has a quite limited time span. So instead, we investigate adding an additional equation to help offset the bias in the elasticities.

In particular, by aggregating over source countries, we can work with *total import demand relative to home demand,* in what we call "step 3." That demand equation only involves the macro Armington elasticity — not the micro elasticity from step 1. But we will find that the corresponding reduced-form supply equation involves the supply parameters estimated in steps 1 and 2. So this extra equation — at the level of total import demand — can help to identify the macro elasticity without introducing an extra parameters on the supply side. We will find that estimation of these three stages provides us with estimates of the Armington elasticities that are close to unbiased in our simulated data. This is the estimation method that we then use on the U.S. data.

4.1 Step 1: The Micro Elasticities

For the micro elasticity, we can start with a simplified demand equation that only relies on import data. Sum (17) across foreign countries $i \neq j$ using the Sato-Vartia weights w_{gt}^{ij}, and then take the difference between (17) and the resulting equation. The terms involved 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,

$$\varepsilon_{gt}^{iF} = \Delta \ln \left(\frac{\kappa_{gt}^{ij}}{\kappa_{gt}^{Fj}} \right) + \Delta \ln \left(\frac{N_{gt}^{ij}}{N_{gt}^{Fj}} \right). \tag{19}$$

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

$$\Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) = \frac{-\Delta \ln(V_{gt}^{ij}/V_{gt}^{Fj})}{(\sigma_g - 1)} + \frac{\varepsilon_{gt}^{iF}}{(\sigma_g - 1)}. \tag{20}$$

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}^{ijF}. 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,

$$\Delta \ln \left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}} \right) = \rho_{1g} \frac{\varepsilon_{gt}^{iF}}{(\sigma_g - 1)} + \delta_{gt}^{iF}, \tag{21}$$

for $i = 1, ..., J, i \neq j$, and $t = 1, ..., T$, 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 (21) as a reduced-form supply curve.

As it is stated, equation (21) is without loss of generality and the residual δ_{gt}^{iF} in this supply curve is uncorrelated with the variables on the right of (21) over all import sources i by construction. That is, conditional on the data on the right of (21), the OLS coefficient ρ_{1g} is chosen so that the following condition holds:

$$\sum_t \sum_{i \neq j} \varepsilon_{gt}^{iF} \delta_{gt}^{iF} = 0. \tag{22}$$

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.[17] 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 (21) 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 (20) and supply equation (21):

$$\varepsilon_{gt}^{iF} = \Delta \ln\left(\frac{V_{gt}^{ij}}{V_{gt}^{Fj}}\right) + (\sigma_g - 1)\Delta\ln\left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}}\right), \tag{23}$$

$$\begin{aligned} \delta_{gt}^{iF} &= \Delta\ln\left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}}\right) - \rho_{1g}\frac{\varepsilon_{gt}^{iF}}{(\sigma_g - 1)} \\ &= (1-\rho_{1g})\Delta\ln\left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}}\right) - \frac{\rho_{1g}}{(\sigma_{1g}-1)}\Delta\ln\left(\frac{V_{gt}^{ij}}{V_{gt}^{Fj}}\right). \end{aligned} \tag{24}$$

Here, the second line of (24) follows by substituting for ε_g^{iF} from (23). Multiplying these two equations together and dividing by $\left(1-\rho_{1g}\right)(\sigma_g - 1)$, we obtain the estimating equation:

$$Y_{gt}^{iF} = \theta_{1g}X_{1gt}^{iF} + \theta_{2g}X_{2gt}^{iF} + u_{gt}^{iF}, \tag{25}$$

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

$$\theta_{1g} = \frac{\rho_{1g}}{(\sigma_g - 1)^2(1-\rho_{1g})}, \theta_{2g} = \frac{(2\rho_{1g}-1)}{(\sigma_g-1)(1-\rho_{1g})} \tag{26}$$

and the error term

$$u_{gt}^{iF} = \frac{\varepsilon_{gt}^{iF}\delta_{gt}^{iF}}{(\sigma_g-1)(1-\rho_{1g})}. \tag{27}$$

Summing over time, the expectation of the error term in (27) is zero from Assumption 1, so that gives us $J-1$ moment conditions that we can use for estimation. Formally, we can proceed

[17] This logic does not go through, however, when an unmeasured factor influences *both* demand and supply. An example is unmeasured quality, which would affect both demand and supply costs, so that the demand and supply errors would be correlated with each other. 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 done by Feenstra and Romalis (2012). They show how the GMM estimates of σ_g are affected by introducing quality in this manner.

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 (27) 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 (25) 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 (25). If we regress the left and right-hand side variables on country indicators, we obtain the following equation:

$$\overline{Y}_g^{iF} = \theta_{1g}\overline{X}_{1g}^{iF} + \theta_{2g}\overline{X}_{2g}^{iF} + \overline{u}_g^{iF},\ i = 1, ..., J, i \neq j, \tag{28}$$

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 (25), then likewise that country appears for T_g^i periods in (28), i.e. just repeating the averaged obervation for that country T_g^i times. So the TSLS procedure it equivalent to estimating (28) over countries while weighting each country-observation by T_g^i (i.e. multiplying each country-observation in (28) 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.[18]

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.[19] That requirement is an example of "identification through heteroskedasticity," as labeled by Rigobon (2003), and applied here in a panel context.[20]

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 (26) (see Feenstra 1994). We can improve efficiency by using weighting estimates, obtained from the errors:

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

We weight the estimating equation (25) by the inverse of the variance of these errors over time, for each country i, and then re-estimate (25) to obtained 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 $u_{gt}^{iF} = Y_{gt}^{iF} - \theta_{1g}X_{1gt}^{iF} -$

[18] 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.

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

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

$\theta_{2g} X_{2gt}^{iF}$ 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(28) 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(28) 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.

Estimating the Micro Elasticities in Simulated and U.S. Data

The estimation of (25) 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 the OLS estimates, and another 5 goods had imaginary point estimates $\hat{\sigma}_g$ when solving the quadratic equation arising from (26), so in the GMM estimation we work with the remaining 98 goods.

In all estimators, we include a constant term in (25). Feenstra (1994) argued that this term can control for classical measurement error – that is, error uncorrelated with other variables – in the unit values. That result can be seen from the dependent variable $\overline{Y}_g^{iF}$ in (28), which is the second moment of the unit values and will therefore equal the second moment of prices plus the second moment of measurement error in the unit values, plus twice their cross-moment. If the cross-moment has expected value of zero and the variance of measurement error is constant across countries, then that variance appears as a constant term in (28), which we wish to capture.

In the simulated dataset with a true value of $\sigma = 3$, applying the above TSLS and GMM procedures to equation (25) gives the median results shown in Table 2 using 1,000 simulations. 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

Table 2: Median GMM Estimates of the Micro Elasticity Using Simulated and U.S. Data

	Sigma estimated from Eq. (25)			
	TSLS	2-step GMM	LIML	CUE
Data	(1)	(2)	(3)	(4)
Simulated	3.36	3.72	2.66	2.75
$T = 10$	(3.15, 3.97)	(3.54, 3.97)	(2.14, 2.97)	(2.73, 2.79)
Simulated	3.27	3.70	2.60	2.66
$T = 50$	(3.05, 3.57)	(3.54, 4.90)	(1.88, 3.34)	(2.52, 2.79)
Simulated	3.42	3.69	2.60	2.65
$T = 100$	(2.59, 5.96)	(2.73, 6.54)	(1.63, 3.58)	(2.52, 2.78)
U.S.	3.24	4.12	1.54	1.98
$T \leq 15$	(1.34, 20.75)	(2.16, 10.47)	(-17.20, 20.56)	(-16.19, 25.92)

Notes: Same as Table 1, except there are now 98 U.S. goods.

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 2. 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 consistent with the results from our U.S. data and those of Soderbery.

Turning to the U.S. data, the kernel density of the TSLS estimates for σ_g over the 98 goods is graphed in Figure 3. The median estimate is 3.24 from the TSLS estimates, as reported at the bottom of Table 2. 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. Comparing the density of estimates in Figure 3 and Figure 1 makes it clear that the OLS estimates for σ_g are strongly downward biased. To obtain confidence intervals on the GMM estimates we perform the same bootstrap used on the OLS estimates. That is, we randomly re-draw observations each year and re-estimate (25). The lower and upper 95%

confidence bounds are also graphed in Figure 3, and the median of the confidence intervals across goods for TSLS is (1.34, 20.75), as reported in Table 2.

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. On the other hand, both the median LIML and CUE estimates reported at the bottom of Table 2 are *lower* than found for TSLS. These medians are taken over slightly different sets of goods because LIML and especially CUE fail to converge in some cases.[21] 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 bias in LIML and CUE when estimating the micro Armington elasticity is a topic for further research.[22] 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.

4.2 Step 2: The Macro Elasticity

To estimate the macro elasticities, we now consider the complete demand equation (17), which we re-write slightly so that the unit values appear on the left and the import values on the right:

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

This specification unpacks (20), as it now shows the multilateral unit value UV_{gt}^{Fj} relative to the home unit value, UV_{gt}^{jj}, on the right. But, similar to our discussion of (20), 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 $UV_{gt}^{Fj}/UV_{gt}^{jj}$. In addition, 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 equation,

[21] Of the 103 potential goods TSLS provides real estimates in 100 cases, 2-step GMM 99, LIML 85 and CUE 100. There are 12 cases where LIML converges and CUE does not. Going forward we focus only on the 98 goods for which both TSLS and 2-step GMM both provide real estimates.

[22] 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 (28) is used on the left-hand side. It is surprising that a persistent downward bias is found despite this desirable property.

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

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

We regard (29) as the natural extension of the reduced-form supply curve (21), 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. We shall give a more formal justification for this reduced-form supply equation in the next sub-section. 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 2: $E\left(\sum_t \varepsilon_{gt}^{ij}\delta_{gt}^{ij}\right) = 0$ 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 2 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}. We now argue, however, that the macro demand and supply elasticities, ω_g and ρ_{2g}, cannot be identified separately for every good.

To see this, let us isolate the errors terms in (17) and (29) as:

$$\begin{aligned}
\varepsilon_{gt}^{ij} &= \Delta \ln\left(\frac{V_{gt}^{ij}}{V_{gt}^{jj}}\right) + (\sigma_g - 1)\Delta \ln\left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}}\right) + (\omega_g - 1)\Delta \ln\left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}}\right), \\
\delta_{gt}^{ij} &= \Delta \ln\left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}}\right) + \rho_{2g}\frac{(\omega_g - 1)}{(\sigma_g - 1)}\Delta \ln\left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}}\right) - \frac{\rho_{1g}\varepsilon_{gt}^{ij}}{(\sigma_g - 1)} \\
&= (1 - \rho_{1g})\Delta \ln\left(\frac{UV_{gt}^{ij}}{UV_{gt}^{Fj}}\right) + \frac{(\rho_{2g} - \rho_{1g})(\omega_g - 1)}{(\sigma_g - 1)}\Delta \ln\left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}}\right) - \frac{\rho_{1g}}{(\sigma_g - 1)}\Delta \ln\left(\frac{V_{gt}^{ij}}{V_{gt}^{jj}}\right).
\end{aligned}$$

Multiplying these two equations together, dividing by $(1 - \rho_{1g})(\sigma_g - 1)$, we obtain an estimating equation expressed in the convenient form:

$$Y_{gt}^{iF} = \sum_{n=1}^{2}\theta_{ng}X_{ngt}^{ij} + \sum_{n=3}^{4}(\omega_g - 1)\theta_{ng}X_{ngt}^{ij} + (\omega_g - 1)^2\theta_{5g}X_{ngt}^{j} + u_{gt}^{ij}, \tag{30}$$

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 (26), and also:

$$\theta_{3g} = \frac{-(1+\rho_{2g}-2\rho_{1g})}{(\sigma_g-1)(1-\rho_{1g})}, \theta_{4g} = \frac{-(\rho_{2g}-2\rho_{1g})}{(\sigma_g-1)^2(1-\rho_{1g})}, \theta_{5g} = \frac{-(\rho_{2g}-\rho_{1g})}{(\sigma_g-1)^2\,(1-\rho_{1g})}. \tag{31}$$

Notice that the dependent variable Y_{gt}^{iF} in (25) and (30) are the same but the X variables are different. Nevetheless, the coefficients θ_{1g} and θ_{2g} are the same in the two systems, and then the extended system (30) 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_g and supply elasticity ρ_{2g}.

As in the earlier estimation of (25), we can use source-country indicator variables as IV, in which case the dependent variables are averaged over time. A difficulty, however, is that the macro elasticity ω_g appears in (30) on variables that have little variation over the source countries i. This is most obvious for X_{5gt}^{j}, which after running on the IV becomes the second moment of $(UV_{gt}^{Fj}/UV_{gt}^{jj})$ and does not differ over the source countries at all. But this difficulty is also evident from X_{3gt}^{ij} and X_{4gt}^{ij}, which become cross-moments between a relative unit-value or value that *does* differ across i and the relative unit-value $(UV_{gt}^{Fj}/UV_{gt}^{jj})$ that *does not* depend on i. After averaging over time and allowing $T_g^i \to \infty$, we would not expect the probability limits of $\bar{X}_{3gt}^{ij}$ and $\bar{X}_{4gt}^{ij}$ to have meaningful variation across source countries i.

To address this difficulty, we shall pool the observations in (30) over a set of goods g for which we assume that the macro demand and supply elasticities are the same:

$$\omega_g = \omega \ \text{ and } \rho_{2g} = \rho_2 \ \text{ for } \ g = 1, ..., G. \tag{32}$$

When estimating (30), we shall rely on heteroscedasticity across goods rather than across countries to identify the coefficients of X_{3gt}^{ij} and X_{4gt}^{ij}, from which we obtain the macro elasticity. The IV used are indicator variables across source countries i interacted with indicator variables across goods g within broad sectors. For convenience, we shall estimate (30) in a sequential method by substituting the estimated values of the micro elasticities $\widehat{\sigma}_g$ and the supply elasticities $\widehat{\rho}_{1g}$ obtained from the import data into (30), and then estimating this equation using nonlinear TSLS to obtain the macro demand ω and supply ρ_2 elasticities.[23]

Because of the nonlinear structure of the coefficients in (30), in practice we perform IV estimation manually by averaging (30) over time for each source country and good. As in our discussion of the micro elasticities, 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 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

[23] Alternatively, we can estimate (25) and (30) simultaneously to obtain the micro and macro elasticities. It turns out that our estimates in simulated data – with a single micro elasticity σ and a single macro elasticity ω – do not differ much whether we adopt a sequential or a simultaneous strategy. But for U.S. data we allow for differing σ_g across goods as well as differing ρ_{1_g}, and in that case the simultaneous estimation of σ_g, ρ_{1_g} and ω is computationally difficult; the sequential procedure is therefore preferred.

Table 3: Median System GMM Estimates of the Macro Elasticity Using Simulated Data

T	Omega from Eq. (30)		Omega from Eqs. (30) and (40)	
	TSLS	2-step GMM	TSLS	2-step GMM
	(1)	(2)	(3)	(4)
10	1.38	1.40	1.44	1.52
	(1.30, 1.45)	(1.26, 1.63)	(1.22, 1.59)	(1.20, 1.90)
50	1.50	1.53	1.62	1.70
	(1.20, 1.71)	(1.28, 1.78)	(1.38, 1.83)	(1.37, 2.22)
100	1.76	1.73	1.78	1.93
	(1.44, 2.84)	(1.62, 2.02)	(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 equation (30) alone or (30) and (40) 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 (25). Reported in parentheses are the 95% confidence intervals obtained by bootstrapping the entire system.

(30) by the inverse of the variance of the residuals computed over time, for each country i and good g, and then re-estimating (30) 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 (30) to control for measurement error in the relative unit-value within Y_{gt}^{iF}.

Estimating the Macro Elasticities in Simulated and U.S. Data

We begin by reporting results from simulated data in Table 3. 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. Column (1) presents the median TSLS estimate of the macro elasticity $\hat{\omega}$ from (30), in step 2, that rely on the micro elasticity estimates $\hat{\sigma}_g$ from (25), in step 1. The true value of the macro elasticity ω is 2, and in comparison, we find that the median TSLS estimates in column (1) are biased downwards by $12-31\%$, depending on the sample size.

A small sample bias is a common feature of GMM estimates described by Cameron and Trivedi (2005, p.177).[24] The median estimate moves closer to the true value of ω as our time period increases from 10 to 50 to 100, however, and the confidence interval in the large sample includes the true value of 2. So in contrast to the OLS estimates in Table 1, we find that the TSLS bias

[24] Cameron and Trivedi (2005) describe this problem in the context of results from Altonji and Segal (1996), adding that "in the literature their results are interpreted as being relevant to GMM estimation with cross-section data or short panels," which is the case for our U.S. data where T is no greater than 15.

Table 4: System GMM Estimates of the Macro Elasticity Using U.S. Data

Sector	Number of Goods	Omega from Eq. (30) TSLS (1)	Omega from Eq. (30) 2-step GMM (2)	Omega from Eqs. (30) and (40) TSLS (3)	Omega from Eqs. (30) and (40) 2-step GMM (4)
Food Products	6	2.28	2.27	4.08	3.12
		(-17.37, 8.56)	(0.28, 4.32)	(-11.51, 6.41)	(0.64, 4.32)
Apparel Manufacturing	13	1.06	1.01	2.51	3.60
		(-2.83, 3.12)	(0.06, 1.60)	(0.55, 3.46)	(-1.82, 8.19)
Rubber, Stone, & Misc Metal	5	1.28	1.34	1.38	1.65
		(-1.94, 2.78)	(-4.53, 2.52)	(0.83, 1.96)	(0.75, 2.56)
Chemical Manufacturing	6	-3.65	1.16	2.10	1.46
		(-15.29, 16.28)	(-5.64, 3.38)	(-4.89, 3.56)	(-1.24, 8.58)
Primary Metals	20	1.03	1.821	2.064	1.16
		(-0.75, 3.43)	(0.572, 4.114)	(1.62, 2.80)	(-1.55, 2.06)
Metal Products	9	0.57	0.69	0.87	0.88
		(0.09, 1.44)	(-0.92, 2.44)	(0.73, 0.98)	(0.84, 0.97)
Machinery	15	1.39	1.41	2.01	2.36
		(-0.57, 2.81)	(-0.07, 2.43)	(0.72 2.82)	(0.87, 2.53)
Electronics	24	-0.29	0.75	2.40	3.48
		(-1.03, 3.26)	(-2.77, 0.43)	(1.69, 2.60)	(1.68, 3.71)

Notes: Same as Table 3, except that the Table 3 assumed value of $\omega = 2$ does not apply in this table which is based on estimates from U.S. data.

from the nested CES equation is in large part a small sample issue. Similar results hold for the 2-step GMM estimates reported in column (2), which again move closer to the true value of 2 as the sample size increases. Still, even with $T = 100$ the median estimate of 1.76 or 1.73 in columns (1) and (2), respectively, is still quite far from the true value of $\omega = 2$.

Turning to the U.S. data, We divide the 98 goods in our dataset into eight sectors with the number of goods in each sector as shown in Table 4.[25] In column (1) we report TSLS estimates of the macro elasticity, and in column (2) we report 2-step GMM estimates. For each sector there is a single estimate of ω, since we have constrained this macro elasticity to be the same for all goods within a sector. The TSLS and 2-step GMM estimators give similar results except for the chemical manufacturing and electronics sectors, for which TSLS results in negative estimates of ω whereas 2-step GMM gives positive estimates. For either estimator, the confidence intervals obtained by

[25] We chose the sectors so that they would have no fewer then 5 or more than 25 goods, and so that reasonable additions or substractions to the goods included in each did not strongly affect the estimated macro elasticity.

bootstrapping the micro and macro estimates from (25) and (30) are quite large: the value $\omega = 1$ is within the confidence intervals for all sectors, and the 2-step GMM point estimates are not that far from unity in most sectors. So if we were to view these results as reliable, we would not be able to reject the hypothesis that the macro Armington elasticity equals unity.[26]

We believe such a conclusion is premature, however, in view of the small-sample bias shown in columns (1) and (2) of Table 3 using simulated data. The U.S. data have at most $T = 15$ annual observations in first differences, so we could expect that our TSLS and 2-step GMM estimates using the U.S. data are subject to the same small sample bias as we see in the simulated data. We have no way of increasing the time span to offset this bias because information from *Current Industrial Reports* is not available in earlier years. So instead, we propose to offset the small-sample bias by adding an additional moment condition which comes from the macro structure of our model, as described in the next section.

4.3 Step 3: Aggregation over Countries

To reduce the bias in the macro elasticities, we consider a further moment condition obtained by aggregating across countries. To achieve this, we start from (9), which gives country j's spending on imports of good g from country i. 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}:

$$\begin{aligned} V_{gt}^{Fj} &= \sum_{i \neq j} V_{gt}^{ij} \\ &= \alpha_{gt}^{j}\left(1-\beta_{gt}^{j}\right)\left(\frac{P_{gt}^{Fj}}{P_{gt}^{j}}\right)^{1-\omega}\left(\frac{P_{gt}^{j}}{P_{t}^{j}}\right)^{1-\eta} P_{t}^{j} C_{t}^{j}\left[\sum_{i \neq j} \kappa_{gt}^{ij}\left(\frac{P_{gt}^{ij}}{P_{gt}^{Fj}}\right)^{1-\sigma_g}\right] \\ &= \alpha_{gt}^{j}\left(1-\beta_{gt}^{j}\right)\left(\frac{P_{gt}^{Fj}}{P_{gt}^{j}}\right)^{1-\omega}\left(\frac{P_{gt}^{j}}{P_{t}^{j}}\right)^{1-\eta} P_{t}^{j} C_{t}^{j}. \end{aligned} \tag{33}$$

The last line follows from definition (3) and we also impose $\omega_g = \omega$ from (32). Combining the foregoing expression with the demand V_{gt}^{jj} as computed from equation (10), we obtain (34):

$$\ln\left(\frac{V_{gt}^{Fj}}{V_{gt}^{jj}}\right) = (1-\omega)\ln\left(\frac{P_{gt}^{Fj}}{P_{gt}^{jj}}\right) + \ln\left(\frac{1-\beta_{gt}^{j}}{\beta_{gt}^{j}}\right), \tag{34}$$

so that the home-foreign Armington elasticity ω can be identified from a multilaterally aggregated equation for imports of good g.[27] We refer to (34) as the "macro" import demand equation.

[26] Indeed, in earlier seminar presentations we made that conclusion based on estimates like those shown in columns (1) and (2) of Table 4.

[27] 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. They match Unites States consumption to import data as we do, but at a higher aggregation level than in our data. At the same time, their estimation method aggregates across different foreign suppliers to the United States.

Of course, when estimating this equation we use unit value rather than the true price indexes, as shown in (13) (for $i = j$) and (16). It follows that the aggregate equation for estimation is:

$$\Delta \ln \left(\frac{V_g^{Fj}}{V_g^{jj}} \right) = (1 - \omega) \Delta \ln \left(\frac{UV_g^{Fj}}{UV_g^{jj}} \right) + \varepsilon_g^{Fj}, \tag{35}$$

with the error term,

$$\varepsilon_{gt}^{Fj} \equiv \Delta \ln \frac{(1 - \beta_t^j)}{\beta_t^j} + \frac{(\omega - 1)}{(\sigma_g - 1)} \left[\Delta \ln \kappa_{gt}^{Fj} + \Delta \ln \left(\frac{N_{gt}^{Fj}}{N_{gt}^{jj}} \right) \right]. \tag{36}$$

The properties of (35) are quite similar to those of the disaggregate equation (17), and we can similarly adapt the technique of Feenstra (1994) to estimate the macro elasticity. We begin by re-writing the demand equation (35) slightly as:

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

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:

$$\Delta \ln \left(\frac{UV_{gt}^{Fj}}{UV_{gt}^{jj}} \right) = \rho^F \frac{\varepsilon_{gt}^{Fj}}{(\omega - 1)} + \delta_{gt}^{Fj}, \tag{37}$$

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 (37) 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 (37) 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 3: $E \left(\sum_t \varepsilon_{gt}^{Fj} \delta_{gt}^{Fj} \right) = 0$ for $g = 1, ..., G$.

Notice that the error term ε_{gt}^{ij} in equation (18) is equal to $\varepsilon_{gt}^{iF} + \varepsilon_{gt}^{Fj}$, the sum of the micro and macro errors defined in (19) and (36), respectively. It follows that the reduced-form supply relation (21) can be re-written as:

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

where the second line makes use of (37) and we define

$$\rho_{2g} \equiv \frac{\rho_{1g}}{\rho^F} \quad \text{and} \quad \delta_{gt}^{ij} \equiv \rho_{2g}\left(\frac{\omega-1}{\sigma_g-1}\right)\delta_{gt}^{Fj} + \delta_{gt}^{iF}. \tag{38}$$

Like equation (29), the form of which we simply assumed, the preceding reduced-form supply equation is an aternative representation of the supply equation (21) – a representation that allows econometric identification of ω_g. The new equation above justifies our use of (29) in the last subsection, and it does so by making use of the macro supply equation (37), which allows us to deduce the cross-equation restrictions $\rho_{2g} = \rho_{1g}/\rho_F$. This additional structure is a potential advantage for estimation purposes. Imposition of the restrictions $\rho_{2g} = \rho_{1g}/\rho_F$ is one way that the macro equation adds information to our estimating system, and this extra information allows the parameter ρ_{2g} to differ across goods.

Why do equations (21) and (37) jointly lead to an equation isomorphic to (29)? For any good g, equation (21) shows how a relative demand shock ε_{gt}^{iF} in favor of imports from i, and at the expense of the aggregate of country-j imports of g, raises the relative (unit-value) price of those two goods, $UV_{gt}^{ij}/UV_{gt}^{Fj}$, along their relative supply curve. But equation (29), like our new equation above for $\Delta \ln\left(UV_{gt}^{ij}/UV_{gt}^{Fj}\right)$, shows how the same relative price is affected by the *different* shock ε_{gt}^{ij} governing relative demand between imports from country i and domestic goods. That shock adds to ε_{gt}^{iF} the additional shock ε_{gt}^{Fj} affecting the relative demand for aggregate imports of g as compared to domestic goods; compare equation (18) with (19) and (26). Thus, to make our alternative representation valid, we need to control for the extra additive shock component ε_{gt}^{Fj}. Equation (37) implies that we can do so by controlling for the relative price $\Delta \ln\left(UV_{gt}^{Fj}/UV_{gt}^{jj}\right)$.

We further note that Assumption 3 adds extra information to Assumption 2. 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 (38). For $\rho_{2g} \neq 0$ and $\omega \neq 1$,we can re-write the expection in Assumption 3 as:

$$\begin{aligned} 0 &= E\left(\sum_t \sum_{i\neq j} w_{gt}^{ij}\varepsilon_{gt}^{ij} \sum_{k\neq j} w_{gt}^{kj}\delta_{gt}^{kj}\right) \\ &= E\left(\sum_t \sum_{i\neq j} (w_{gt}^{ij})^2 \varepsilon_{gt}^{ij}\delta_{gt}^{ij}\right) + E\left(\sum_t \sum_{i\neq j} \sum_{k\neq i,j} w_{gt}^{ij} w_{gt}^{kj} \varepsilon_{gt}^{ij}\delta_{gt}^{kj}\right). \end{aligned} \tag{39}$$

The first summation in (39) equals zero in its unweighted form from Assumption 2, 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 3 imposes that the sum of these two complex summations equals zero. For that reason, the moment condition for the macro import demand equation is adding information to what we have used in Assumption 2 for the nested-CES demand equations. Accordingly, we will exploit both moment conditions in the estimation.

To make use of the macro demand and supply equations, we now proceed in the same manner as with the estimating equation (25) for the micro elasticities. Multiply the errors in the macro demand and supply equations, we can obtain:

$$Y_{gt}^{Fj} = \phi_1 X_{1gt}^{Fj} + \phi_2 X_{2gt}^{Fj} + u_{gt}^{Fj}, \tag{40}$$

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

$$\phi_1 = \frac{\rho^F}{(\omega-1)^2(1-\rho^F)}, \phi_2 = \frac{(2\rho^F-1)}{(\omega-1)(1-\rho^F)}$$

and the error term

$$u_{gt}^{Fj} = \frac{\varepsilon_{gt}^{Fj}\delta_{gt}^{Fj}}{(\omega-1)(1-\rho^F)}.$$

We jointly estimate the macro equation (40) together with equation (30) to obtain the macro elasticity, making use of the cross-equation restriction $\rho_{2g} = \rho_{1g}/\rho^F$ in (38) and using estimates of the micro elasticities and ρ_{1g} that come from (25).[28] As noted above in the estimation of the nested CES equation, we can use indicator variables for source countries interacted with indicator variables for goods as IV when estimating (30). Equivalently, we can perform the IV estimation manually by averaging (30) over time for each source country and good and then using weighted NLS to obtain the macro demand and supply elasticities. In this manual approach the averaged-over-time equation is weighted by T_g^i, the number of years that each country and good appears. Taking a similar approach to the macro equation (40), we first average this equation across goods within each broad sector. It would not make any sense to have the macro equation (30) appearing with a weight of unity, since in that case it would have little impact on the results. Instead, we weight each good-observation within the averaged-over-time macro equation by $\sum_i T_g^i$, so that it essentially receives the same weight in total as the nested CES equations. We refer to this approach as TSLS.

Using these initial estimates of the macro elasticities, we compute the variance of the residuals for each country and good in (30) and use T_g^i divided by the inverse of these variances as weights when re-estimating the averaged-over-time nested CES equation (30). Likewise, we average the variance of the residuals from the nested CES equation for each good, and divide $\sum_i T_g^i$ by the

[28] The estimates of the macro elasticity from (40) alone are available on request. Those estimates tend to be downward biased in the simulated data and rather high in the U.S. data, but with large confidence intervals that often include zero. We interpret those estimates as lacking precision on their own, but improving the power of the macro elasticity that comes from the nested CES equation (30) when the cross-equation restriction $\rho_{2g} = \rho_{1g}/\rho^F$ is also used.

average of these variances to obtain the weight that is applied to each good-observation in the averaged-over-time macro equation (40). Then estimating the nested CES and macro equations simultaneously using $\rho_{2g} = \rho_{1g}/\rho^F$, we obtain the 2-step GMM estimates.

Estimating the Macro Elasticities Once Again

Starting with the simulated data, the results from simultaneous estimation of the nested CES equation (30) and the macro equation (40) are shown in columns (3) and (4) of Table 3. 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 (25), or step 1, and substituting these into (30) and (40), or steps 2 and 3 estimated as a system. For the TSLS estimates in column (3), we see that the median point estimates of the macro elasticity all move slightly closer to their true value of 2 as compared to the TSLS estimates in column (1), that do not make use of the macro equation. Evidently, the extra moment condition obtained from the macro side of the model is adding useful information to the estimation.

The reduced bias that arises from using the macro moment condition shows through more strongly when applying 2-step GMM, shown in column (4) of Table 3. For values of $T = 100$, the median estimate of the macro elasticity is 1.93, only slightly lower than its true value of $\omega = 2$. For lower values of T the downard bias is more noticable, but it still an improvement over the 2-step GMM estimates obtained by using the nested CES equation alone, in column (2). We conclude that making use of the macro moment condition in conjunction with the nested CES moment condition always gives less bias than using the nested CES moment condition alone.

Turning to the U.S. data, our estimates of the macro elasticity obtained from simultaneous estimation of the nested CES equation (30) and the macro equation (40) are shown in columns (3) and (4) of Table 4. As we found in the simulations, the estimates of the macro elasticity are pulled up by adding the macro moment condition: TSLS estimates in column (3) exceed those in column (1), and the 2-step GMM estimates in column (4) generally exceed those in column (2). The point estimates of the macro elasticity are significantly less than unity in only one sector (Metals) 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 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 when the estimation uses all three moment conditions.

To investigate the size of the macro elasticity more formally, in Table 5 we report the results of tests for the null hypothesis that $\sigma_g \leq \omega$. To perform these one-sided tests we use the bootstrapped data from the three estimating equations: (25), (30) and (40).[29] 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

[29] The use of boostrapping to test hypotheses is discussed in MacKinnon (2006).

that sector. We then count the proportion of 500 bootstrap samples (as we use) where $\hat{\sigma}_g \leq \hat{\omega}$ is found for each good. With a 5% significance level, if there are fewer than 25 bootstrap samples where $\hat{\sigma}_g \leq \hat{\omega}$ then we reject the null hypothesis that $\sigma_g \leq \omega$. From this result we conclude that the macro elasticity ω is significantly less than the micro elasticity σ_g.

Within Table 5 we report three sets of hypothesis tests. The first test, in columns (1)-(3), uses the TSLS estimates of both the micro and macro elasticities. We found earlier in Table 2 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 3 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. That concern is particularly valid when using the 2-step GMM estimates of the micro elasticity, which had the greatest upward bias in Table 2. So we also consider a second version of the hypothesis test in columns (4)-(6) where we substitute the TSLS estimates of the micro elasticity into (30) and (40) to obtain new 2-step GMM estimates of the macro elasticity. The third version of the hypothesis test, in columns (7)-(9), uses the 2-step GMM estimates for both elasticities.

Looking first at the hypothesis test using TSLS estimates, in column (1) we repeat the sectoral estimates of the macro elasticities $\hat{\omega}$ from Table 4. In column (2) 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 (3) we report the number of goods for which we reject the null hypothesis that $\hat{\sigma}_g \leq \hat{\omega}$, meaning that ω is significantly less than σ_g. Looking at the total reported in the final row of column (3), there are only 22 out of the 73 goods where the lower point estimate of the macro elasticity means that it is significantly below the micro elasticity. In other words, less than one-quarter of the total number of goods lead to a significant difference in the elasticities. This inability to statistically distinguish the elasticities reflects somewhat large standard errors in the TSLS estimates, and we obtain sharper results using 2-step GMM estimates.

The increased precision of 2-step GMM is reinforced by the second set of tests we report, in columns (4)-(6), where now we mix the TSLS estimates of the micro elasticity with new 2-step GMM estimates of the macro elasticity from (30) and (40). The number of goods for which $\hat{\omega} < \hat{\sigma}_g$ in the point estimates falls to 68 (see column 5), but the number where ω is significantly less than σ_g increases to 29 (see column 6). In column (7) we again repeat the sectoral macro elasticities $\hat{\omega}$ from Table 4. In column (8) we show how many goods in each sector have estimates of the macro elasticity $\hat{\omega}$ less than that of the micro elasticities $\hat{\sigma}_g$, obtaining 73 goods, the same as the count with TSLS estimates. But now when we test the null hypothesis that $\hat{\sigma}_g \leq \hat{\omega}$, we are able to reject it for 50 of the goods. In other words, for one-half of the total number of goods, we find that the macro elasticity is significantly *below* the micro elasticity.[30]

[30] In lieu of reliable estimates of both the macro and micro elasticities 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 bootstrap methodology described previously, testing the null hypothesis $\sigma_g = 2\omega$ for each of our 98 goods.

Table 5: Testing that the Macro Elasticity is less than the Micro Elasticity

		Sigma estimated from Eq. (25) and Omega from Eqs. (30) and (40)								
	Number of Goods	σ_g from TSLS ω from TSLS			σ_g from TSLS ω from 2-step GMM			σ_g from 2-step GMM ω from 2-step GMM		
Sector			No. Goods with:			No. Goods with:			No. Goods with:	
		ω	$\omega < \sigma_g$	Significant	ω	$\omega < \sigma_g$	Significant	ω	$\omega < \sigma_g$	Significant
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Food Products	6	4.08	4	0	3.13	5	2	3.12	5	4
Apparel Manufacturing	13	2.51	12	6	2.28	12	7	3.60	11	9
Rubber, Stone & Misc Metal	5	1.38	5	2	1.32	5	2	1.65	3	2
Chemical Manufacturing	6	2.10	4	0	1.79	6	3	1.46	6	2
Primary Metals	20	2.06	16	5	6.57	6	6	1.16	20	14
Metal Products	9	0.87	9	8	0.90	9	8	0.88	9	9
Machinery	15	2.01	12	1	1.64	13	0	2.36	9	5
Electronics	24	2.40	11	0	2.25	12	1	3.48	10	5
Total	98		73	22		68	29		73	50

Notes: This table reports estimates of ω, the number of goods in each sector for which $\omega < \sigma_g$ in the point estimates, and the results of the one-sided test for $\sigma_g \leq \omega$ constructed from a bootstrap technique explained in the text. The column labeled "Significant" is the number of goods for which that test is rejected at the 5% level, so that many goods have ω significantly less than σ_g.

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 *between one-quarter and one-half* 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. In the remaining one-half (or more) of the goods there is no significant difference between the macro and micro elasticities.

5 Aggregation over Goods

Earlier sections have developed a model of international trade flows and estimated the implied import demand equations using United States data at a highly disaggregated level. Much previous literature focuses on estimation of aggregate import demand equations, in an attempt to ascertain directly the average relationship between aggregate measures of international competitiveness and aggregate national imports. Furthermore, many studies, while using data that are disaggregated to some degree, still combine potentially disparate goods into composite categories. In this section, we explore the performance of aggregate import demand equations within our framework in order to clarify the conditions under which estimates derived from aggregate data will be accurate.

At first glance, aggregation over goods would seem to be impossible with the nested CES functions specified in (1)-(2), because it is not possible to define CES aggregates of "total import goods" and "total domestic goods" that cut across these nests. But it turns out that, surprisingly, an aggregate of imports and domestic goods can be consistently achieved using the concept of *latent separability* (Blundell and Robin 2000).

For convenience we now drop the time subscript t, which is no longer needed. We start from the (33), which gives import demand aggregated across source countries. The next step is to sum these imports across all available goods g. Observe that the parameter σ_g — the possibly good-specific substitution elasticity between varieties of g purchased from the same source (domestic or foreign) — enters the preceding equations only through its role in constructing the price indexes. *Given* those indexes, the parameter σ_g does not appear in the aggregate demand for imports of good g. While we will *not* assume that σ_g is the same across goods, we will make the assumption that $\omega_g = \omega$ is invariant across different goods.

From (33), total country j expenditure on imports is:

$$V^{Fj} = \sum_g V_g^{Fj} = \sum_g \left[\alpha_g^j \left(1 - \beta_g^j\right) \left(\frac{P_g^j}{P^j}\right)^{\omega-\eta} \left(\frac{P_g^{Fj}}{P^j}\right)^{1-\omega} \right] P^j C^j. \tag{41}$$

Define,

$$\bar{P}^{Fj} \equiv \left[\sum_g \alpha_g^j \left(\frac{P_g^j}{P^j}\right)^{\omega-\eta} \left(P_g^{Fj}\right)^{1-\omega} \right]^{\frac{1}{1-\omega}}. \tag{42}$$

When estimating both σ_g and ω with 2-step GMM we can only reject the "Rule of Two" for 31 of our 98 goods. Additionally, of the 50 goods for which we reject $\sigma_g < \omega$ (reported in column 9 of Table 5), we also fail to reject the "Rule of Two" for 36.

For the case $\omega = \eta$, consumers substitute between domestic and foreign varieties just as readily as between different goods. In that case, therefore, the utility function can be written as a weakly separable function of import consumption and domestic-product consumption, and $\bar{P}^{Fj}$ is simply a standard CES price-index of the good-specific foreign price indexes P_g^{Fj} :

$$\bar{P}^{Fj} = \left[\sum_g \alpha_g^j \left(P_g^{Fj} \right)^{1-\eta} \right]^{\frac{1}{1-\eta}}.$$

Even in the case $\omega \neq \eta$, however, definition (42) provides us with a valid index where the weight on P_g^{Fj} in the overall index $\bar{P}^{Fj}$ depends on the variation in country j's comprehensive price index for good g relative to its overall CPI. The reason we define the index $\bar{P}^{Fj}$ is that, with one added assumption, the aggregate import share is simply a function of the latter index and the overall CPI. The added assumption is that $\beta_g^j = \beta^j$ for all g: in words, β^j is a uniform (across goods) country j demand shock in favor of domestic products. Then, whether or not $\omega = \eta$, it is immediate from (41) that $V^{Fj} = \left(1 - \beta^j\right) \left(\frac{\bar{P}^{Fj}}{P^j} \right)^{1-\omega} P^j C^j$.

Likewise, define V^{Hj} to be total country j spending on home-produced goods:

$$V^{Hj} \equiv \sum_g V_g^{jj} = \beta^j \sum_g \alpha_g^j \left(\frac{P_g^j}{P^j} \right)^{\omega-\eta} \left(\frac{P_g^{jj}}{P^j} \right)^{1-\omega} P^j C^j.$$

Furthermore, define the home index $\bar{P}^{Hj}$ (in analogy to) $\bar{P}^{Fj}$ as

$$\bar{P}^{Hj} = \left[\sum_g \alpha_g^j \left(\frac{P_g^j}{P^j} \right)^{\omega-\eta} \left(P_g^{jj} \right)^{1-\omega} \right]^{\frac{1}{1-\omega}},$$

which depends solely on domestic prices only in the weakly separable case $\omega = \eta$. Whether or not $\omega = \eta$, we can write,

$$V^{Hj} = \beta^j \left(\frac{\bar{P}^{Hj}}{P^j} \right)^{1-\omega} P^j C^j.$$

Dividing (41) by the above equation gives us the key result:

$$\frac{V^{Fj}}{V^{Hj}} = \frac{1 - \beta^j}{\beta^j} \left(\frac{\bar{P}^{Fj}}{\bar{P}^{Hj}} \right)^{1-\omega}, \tag{43}$$

which shows that aggregate imports relative to domestic demand are a simple log-linear function of their relative price, with elasticity ω.

To understand the properties of this aggregate demand equation, define the comprehensive CPI P^j as,

$$P^j \equiv \left[\sum_g \alpha_g^j \left(P_g^j \right)^{1-\eta} \right]^{\frac{1}{1-\eta}},$$

where $\left(P_g^j\right)^{1-\omega} = \beta^j \left(P_g^{jj}\right)^{1-\omega} + \left(1-\beta^j\right)\left(P_g^{Fj}\right)^{1-\omega}$. Then it turns out that we can express P^j in the form

$$P^j = \left[\beta^j \left(\bar{P}^{Hj}\right)^{1-\omega} + \left(1-\beta^j\right)\left(\bar{P}^{Fj}\right)^{1-\omega}\right]^{\frac{1}{1-\omega}}, \tag{44}$$

as can be shown by substituting above for $\bar{P}^{Hj}$ and $\bar{P}^{Fj}$, and then using the expression for P_g^j. This representation directly shows the CPI's functional relationship to the "domestic" and "foreign" price indexes $\bar{P}^{Hj}$ and $\bar{P}^{Fj}$. This description of import demand is precisely what would come out of the hypothetical consumer problem

$$\max_{D,M} \left[\left(\beta^j\right)^{\frac{1}{\omega}} D^{\frac{\omega-1}{\omega}} + \left(1-\beta^j\right)^{\frac{1}{\omega}} M^{\frac{\omega-1}{\omega}}\right]^{\frac{\omega}{\omega-1}}$$

subject to $\bar{P}^H D + \bar{P}^F M = PC$, where D stands for aggregate real domestic consumption and M stands for aggregate real imports.[31] In this sense, the model of import demand admits exact aggregation across goods, with ω as the substitution elasticity between aggregate imports and domestic consumption. Because $\bar{P}^H$ and $\bar{P}^F$ both depend on *all* prices, however, the aggregation is less straightforward than it would be in the case of weakly separable utility ($\omega = \eta$). Instead, aggregation is possible because of the property of *latent separability* analyzed by Blundell and Robin (2000). Whereas weak separability requires that the utility or expenditure function is partitioned into *mutually exclusive* sets of goods, the more general concept of latent separability allows the set of goods to be overlapping: some goods can appear in many of the sub-groups.[32]

6 Impact of a Devaluation

Imbs and Méjean (2013) argue that there are grounds for "elasticity optimism" regarding the responsiveness of imports to a change in the terms of trade. To make this argument, they contrast two approaches to the estimation of σ_g: first, estimating this elasticity separately for 56 sectors using a modification of the GMM method in Feenstra (1994); and second, pooling the data across all sectors and estimating a *single* elasticity. They show that if the sectoral estimates are weighted by their shares in expenditure and summed, a theoretically consistent aggregate elasticity results.[33] That aggregate elasticity of substitution is found to be significantly larger than the single estimate obtained by pooling the data. Therefore, they conclude, a pooled estimate that ignores heterogeneity across sectors is downward biased and gives too pessimistic a view of the impact of a devaluation on the value of imports.

[31] This formulation is the starting point for many empirical studies, for example, Reinert and Roland-Holst (1992), Blonigen and Wilson (1999), Gallaway, McDaniel, and Rivera (2003), and Broda and Weinstein (2006). Our analysis shows the exact form of the price indexes under which their approach would be valid.

[32] To see how this concept applies in our case, consider the aggregate $\bar{P}^{Fj}$ defined in (42). It is a summation over the import price indexes P_g^{Fj}, which depend on the import prices of good g from all source countries. But in addition, the "weights" $\alpha_g^j \left(\frac{P_g^j}{P^j}\right)^{\omega-\eta}$ appear in the formula. When $\omega \neq \eta$ these weights depend on the prices for all imported and domestic goods.

[33] More general versions of this formula were derived and discussed by Barker (1970) and Magee (1975).

We have repeated the exercise of Imbs and Méjean on our own data by pooling across the industries and estimating a single value for σ. Estimating equation (25) with TSLS, we obtain $\hat{\sigma} = 2.67$, while using 2-step GMM, we obtain $\hat{\sigma} = 4.98$. The first of these is 18% below the median TSLS estimate across all goods, 3.24, whereas the second is 21% above the median 2-step GMM estimate across all goods, 4.12. So in contrast to Imbs and Méjean, we do not find that pooling across goods in our sample necessarily gives a substantial downward bias to the estimate of σ.[34]

However, the more important message of our paper is that the aggregate elasticity they compute by taking a weighted average of the sectoral estimates *does not indicate the impact of a devaluation on aggregate imports.* The reason for this is that the data Imbs and Méjean (2013) use in their estimation is for imports only, without any matching domestic production data. Therefore, they are estimating the *micro* Armington elasticity. But in order to understand the impact of exchange rate changes on imports, as shown in (43), we need to use the *macro* Armington elasticity, about which they have no information. Therefore, their results cannot be interpreted as supporting either "elasticity optimism" or "elasticity pessimism," at least in regard to the impact of a devaluation on imports.[35]

When the macro Armington elasticity ω_g differs across goods g, then the impact of a devaluation cannot be obtained from the simple aggregate demand equation (43). Rather, we should instead compute the total derivatives of imports while adding up across sectors. This yields a weighted average formula that is broadly similar to that found in Imbs and Méjean (2013), but now using the macro Armington elasticities ω_g rather than the micro elasticities σ_g. For simplicity, we omit the home country j superscript in this calculation, as well as the time subscript t.

Total imports of the home country are given by

$$V^F = \sum_{g=1}^{G} V_g^F = \sum_{g=1}^{G} P_g^F C_g^F = \sum_{g=1}^{G} \left[P_g^F \alpha_g (1-\beta_g) \left(\frac{P_g^F}{P_g} \right)^{-\omega_g} \left(\frac{P_g}{P} \right)^{-\eta} C \right].$$

Assume that $P_g^F = E P_g^{F*}$, with P_g^{F*} fixed, implying full immediate pass-through from the exchange rate E to import prices. Assume also that domestic producers' prices do not respond to changes in E. For a given level of consumption C, the aggregate import demand elasticity with respect to

[34] Pooling across goods appears to result in a downward bias in the estimation of ω. Estimating ω using only eq. (30) and constraining ω to be constant across goods yield estimates of 0.62 and 0.67 for TSLS and 2-step GMM respectively. Estimating ω using eq. (30) and eq. (40) yields 0.04 and 2.61 for TSLS and 2-step GMM respectively. Comparing these constrained estimates to the unconstrained estimates reported in Table (4) we see that 75% of the time the unconstrained ω is greater than the constrained ω.

[35] As we note in the concluding section, the impact of a devluation on *exports* will depend on both the "micro" and "macro" elasticities found in foreign countries. So Imbs and Méjean (2013) are providing some optimism regarding the elasticity of exports with respect to terms of trade.

E is:

$$\begin{aligned}\frac{d\ln V^F}{d\ln E} &= \frac{E}{V^F}\frac{dV^F}{dE} \\ &= 1+\left(\frac{E^2}{V^F}\right)\frac{d}{dE}\sum_{g=1}^{G}\left[P_g^{F*}\alpha_g(1-\beta_g)\left(\frac{P_g^F}{P_g}\right)^{-\omega_g}\left(\frac{P_g}{P}\right)^{-\eta}C\right] \\ &= 1+\sum_{g=1}^{G}\left(\frac{E^2}{V^F}\right)\frac{d}{dE}\left[P_g^{F*}\alpha_g(1-\beta_g)\left(\frac{P_g^F}{P_g}\right)^{-\omega_g}\left(\frac{P_g}{P}\right)^{-\eta}C\right].\end{aligned}$$

In Appendix B we simplify this equation to obtain

$$\frac{d\ln V^F}{d\ln E} = 1-\sum_{g=1}^{G}(1-m_g)w_g^F\omega_g-\eta\sum_{g=1}^{G}w_g^F\left(m_g-m\right), \tag{45}$$

where $w_g^F \equiv V_g^F/V^F$ is the share of good g in total imports, $m_g \equiv V_g^F/V_g$ is the import share of good g, and $m \equiv V^F/V$ is the share of imports (of all goods) in total consumption spending.

The intuition for (45) is as follows. The first term of unity is the *valuation* effect on import spending, which the other effects must offset for a devaluation to reduce the value of imports. The second term reflects the impact of the rise in E on $\left(P_g^F/P_g\right)^{-\omega_g}$: this negative effect is smaller when a bigger share of good g is imported, because the percent rise in P_g will then be closer to that in P_g^F. The third and last term reflects the impact on $(P_g/P)^{-\eta}$: the negative influence of lower demand for good g is larger when good g has a higher than average import share (P_g will then rise relative to P). An alternative way to group the preceding terms would be as

$$\frac{d\ln V^F}{d\ln E} = 1-\sum_{g=1}^{G}w_g^F\left[\omega_g-m_g\left(\omega_g-\eta\right)\right]+m\eta.$$

Given estimates of η and data on import shares, it is straightforward to calculate the preceding devaluation elasticity. We see from this formula that if goods with higher macro Armington elasticities – or more precisely a higher value of $[\omega_g - m_g(\omega_g-\eta)]$ – also have a higher share of imports w_g^F, then they will contribute more towards obtaining a negative value for this devaluation elasticity. In this respect we agree with Imbs and Méjean (2013); but contrary to them, the Armington elasticities appearing in the formula are the macro and not the micro elasticities, which in general seem likely to be relatively smaller in general.

7 Conclusions

In this paper, we distinguish between the substitution elasticity among alternative foreign import sources and the substitution elasticity between domestic and foreign import sources. These two elasticities are conceptually quite distinct, except within the two-country models that predominate in macroeconomic discussion. They are in some cases empirically quite distinct, as we demonstrate using a new data set of highly disaggregated and concorded domestic production and import data for the United States. We find evidence in our data 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 significantly lower in up to one-half of the goods we analyze, as in the approach to calibration in traditional CGE policy analysis. The fact that the micro and macro elasticities are not significantly different from each other in the other half of cases offers some limited support for the newer generation of computable structural models, which do not allow for any difference between them.

Our results also have important implications for quantitative welfare assessments of trade policy. In their comprehensive recent survey, Costinot and Rodríguez-Clare (2013) argue within a simple gravity model that the "trade elasticity" ε relevant for welfare analysis is ω, the elasticity of substitution between foreign and domestic goods. However, they note that the formula for ε is more complicated in other models, including ours. In our model, if σ is equal to ω, then the trade elasticity ε is simply the Pareto shape parameter $\gamma > \sigma$, as in the Melitz-Chaney model. (There is a corresponding result for the Fréchet parameter θ in the Ricardian model of Eaton and Kortum 2002.) However, if $\sigma \neq \omega$, as we found in many cases, then the trade elasticity ε is a function of σ, ω, and γ, and is increasing in ω. Moreover, one can show that $\varepsilon < \gamma$ if $\omega < \sigma$. Since our results are often consistent with $\omega < \sigma$ and trade gains are inversely proportional to ε, our results also imply quantitatively bigger trade gains than one might have surmised based on the typical parameter assumption that $\sigma = \omega$.[36]

Regardless of whether the macro elasticity emerges as lower than its micro counterpart, we 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. Like the earlier studies, we also find macro elasticity values near unity in U.S. and simulated data, until we add an additional moment condition to overcome small-sample bias. We frame the emprical 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

[36] In our model the elasticity of trade (elasticity of the aggregate import/domestic consumption ratio with respect to relative trade costs for imports) is:

$$\varepsilon = \frac{\gamma(\omega - 1)}{\sigma - 1 + \left(\frac{\gamma}{\sigma - 1} - 1\right)(\sigma - \omega)}$$

We are grateful to Andrés Rodríguez-Clare for supplying this calculation.

estimation approaches.

Our empirical findings raise the question of why substitution between home goods and imports are 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. Anderson, de Palma, and Thisse (1992), shows that a CES indirect utility function for the aggregate consumer can be derived from certain discrete choice models with random utility. In that framework, a relatively smaller elasticity for the macro Armington elasticity is 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 varieties.[37]

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 and 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.[38]

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 one moves 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.

[37] For a recent model of search among alternative foreign suppliers, see Cadot, Carrère, and Strauss-Kahn (2014).

[38] 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. Theirs is one mechanism generating differences in the elasticity of trade with respect to temporary versus permanent price relative changes.

References

[1] Altonji, Joseph G. and Lewis M. Segal. 1996. "Small-Sample Bias in GMM Estimation of Covariance Structures." *Journal of Business and Economic Statistics* 14(3): 353–366.

[2] Anderson, James E. and Eric van Wincoop. 2004. "Trade Costs." *Journal of Economic Literature* 42(3): 691-751.

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

[4] Ardelean, Adina and Volodymyr Lugovskyy. 2010. "Domestic Productivity and Variety Gains from Trade." *Journal of International Economics*, 80(2): 280-291.

[5] Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. 2012. "New Trade Models, Same Old Gains?" *American Economic Review* 102(1): 94–130.

[6] Armington, Paul S. 1969. "A Theory of Demand for Products Distinguished by Place of Production." *IMF Staff Papers* 16(1): 159-178.

[7] Balistreri, Edward J. and Thomas F. Rutherford. 2013. "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.

[8] Barker, Terry S. 1970. "Aggregation Error and Estimates of the U.K. Import Demand Function." In Kennth Hilton and David F. Heathfield, eds., *The Econometric Study of the United Kingdom.* London: Macmillan.

[9] Basu, Susanto, John G. Fernald, and Miles S. Kimball, 2006. "Are Technology Improvements Contractionary?" *American Economic Review* 96(5): 1418-1448.

[10] Baum, Christopher F., Mark E. Schaffer, and Steven Stillman. 2007. "Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing" *Stata Journal* 7(4): 465–506.

[11] Bergin, Paul. 2003. "Putting the 'New Open Economy Macroeconomics' to a Test." *Journal of International Economics* 60(1): 3-34.

[12] Bergin, Paul. 2006. "How Well Can the New Open Economy Macroeconomics Explain the Exchange Rate and Current Account?" *Journal of International Money and Finance* 25(5): 675-701.

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

[14] Blundell, Richard and Jean-Marc Robin. 2000. "Latent Separability: Grouping Goods without Weak Separability." *Econometrica* 68(1): 53-84.

[15] Broda, Christian and David E. Weinstein. 2006. "Globalization and the Gains from Variety." *Quarterly Journal of Economics* 121(2): 541-585.

[16] Cadot, Olivier, Céline Carrère, and Vanessa Strauss-Kahn. 2014. "OECD Imports: Diversification of Suppliers and Quality Search." *Review of World Economics* 150(1): 1-24.

[17] Cameron, A. Colin and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* New York: Cambridge University Press.

[18] Chaney, Thomas. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." *American Economic Review* 98(4): 1707-1721.

[19] Cheng, H. S. 1959. "Statistical Estimates of Elasticities and Propensities in International Trade.". *IMF Staff Papers* 7(1): 107-158.

[20] Cosar, A. Kerem, Paul L. E. Grieco, and Felix Tintelnot. 2011. "Borders, Geography, and Oligopoly: Evidence from the Wind Turbine Industry." University of Chicago and Pennsylvania State University.

[21] Costinot, Arnaud and Andrés Rodríguez-Clare. 2014. "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.

[22] Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2007. "Unbalanced Trade." *American Economic Review* 97(2): 351-355.

[23] Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2008. "Global Rebalancing with Gravity: Measuring the Burden of Adjustment." *IMF Staff Papers* 55(3): 511-540.

[24] Dekle, Robert, Hyeok Jeong, and Heajin H. Ryoo. 2013. "Firm-Level Heterogeneity and the Aggregate Exchange Rate Effect on Exports." University of Southern California.

[25] di Giovanni, Julian, Andrei Levchenko, and Roman Rancière. 2011. "Power Laws in Firm Size and Openness to Trade: Measurement and Implications." *Journal of International Economics* 85(1): 42-52.

[26] di Giovanni, Julian, Andrei Levchenko, and Jing Zhang. 2012. "The Global Welfare Impact of China: Trade Integration and Technological Change." International Monetary Fund Working Paper 12/79 (February).

[27] Eaton, Jonathan and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica* 70(5):1741-1779.

[28] Feenstra, Robert C. 1994. "New Product Varieties and the Measurement of International Prices." *American Economic Review* 84(1): 157-177.

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

[30] Feenstra, Robert C. and John Romalis. 2012. "International Prices and Endogenous Product Quality." National Bureau of Economic Research Working Paper No. 18314.

[31] Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98(1): 394-425.

[32] Gallaway, Michael P., Christine A. McDaniel, and Sandra A. Rivera. 2003. "Short and Long-Run Industry-Level Estimates of U.S. Armington Elasticities." *North American Journal of Economics and Finance* 14(1): 49-68.

[33] Goldstein, Morris and Mohsin S. Khan. 1985. "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.

[34] Hansen, Lars Peter, John Heaton and Amir Yaron. 1996. "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business and Economic Statistics* 14(3), 262-280.

[35] Harrison, Glenn W., Thomas F. Rutherford, and David G. Tarr. 1997. "Quantifying the Uruguay Round." *Economic Journal* 107(444): 1405-1430.

[36] Harberger, Arnold C. 1957. "Some Evidence on the International Price Mechanism." *Journal of Political Economy* 65(6): 506-521.

[37] Head, Keith and John Ries. 2001. "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade." *American Economic Review* 91(4): 858-876.

[38] Heathcote, Jonathan and Fabrizio Perri. 2002. "Financial Autarky and International Business Cycles." *Journal of Monetary Economics* 49(3): 601-627.

[39] Hillberry, Russell and David Hummels. 2013. "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.

[40] Hummels, David. 2007. "Transportation Costs and International Trade in the Second Era of Globalization." *Journal of Economic Perspectives* 21(3): 131-154.

[41] Imbs, Jean and Isabelle Méjean. 2013. "Elasticity Optimism." Manuscript, HEC Lausanne.

[42] Kemp, Murray C. 1962. "Errors of Measurement and Bias in the Estimates of Import Demand Parameters." *Economic Record*. 38(83): 369–372.

[43] Kehoe, Timothy J. and Kim J. Ruhl. 2009. "How Important is the New Goods Margin in International Trade?" Federal Reserve Bank of Minneapolis, Research Department Staff Report 324.

[44] Lai, Huiwen and Daniel Trefler. 2002. "The Gains from Trade with Monopolistic Competition: Specification, Estimation, and Mis-specification." National Bureau of Economic Research Working Paper No. 9169 (September).

[45] Leamer, Edward E. and Robert M. Stern. 1970. *Quantitative International Economics*. Boston: Allyn and Bacon.

[46] MacKinnon, James A. 2006. "Bootstrap Methods in Econometrics." *Economic Record* 82(s1): S2-S18

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

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

[49] McDaniel, Christine A. and Edward J. Balistreri. 2003. "A Review of Armington Trade Substitution Elasticities." *Économie Internationale* 23(3): 301-313.

[50] Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71(6): 1695-1725.

[51] Orcutt, Guy H. 1950. "Measurement of Elasticities in International Trade." *Review of Economics and Statistics* 32(2): 117-132.

[52] Phinisee, Ivory, I. Elaine Allen, Edward Rogoff, Joseph Onochie, and Monica Dean. 2008 "Global Entrepreneurship Monitor National Entrepreneurial Assessment for the United States of America: 2006-2007 Executive Report." Global Entrepreneurship Monitor.

[53] Reinert, Kenneth A. and David W. Roland-Holst. 1992. "Armington Elasticities for United States Manufacturing Sectors." *Journal of Policy Modeling* 14(5): 631-639.

[54] Rigobon, Roberto. 2003. "Identification Through Heteroskedasticity." *Review of Economics and Statistics* 85(4): 777-792.

[55] Romalis, John. 2007. "NAFTA's and CUSFTA's Impact on International Trade." *Review of Economics and Statistics* 89(3): 416-435.

[56] Ruhl, Kim J. 2008. "The International Elasticity Puzzle." New York University Stern School of Business.

[57] Shiells, Clinton R., Robert M. Stern, and Alan Deardorff. 1986. "Estimates of the Elasticities of Substitution between Imports and Home Goods for the United States." *Weltwirtschaftliches Archiv* 122: 497-519.

[58] Simonovska, Ina and Michael Waugh. 2011. The Elasticity of Trade: Estimates and Evidence. National Bureau of Economic Research Working Paper No. 16796.

[59] Simonovska, Ina and Michael Waugh. 2012. Different Trade Models, Different Trade Elasticities? UC Davis and NewYork University Stern School of Business.

[60] Soderbury Anson. 2010. "Investigating the Asymptotic Properties of Elasticity of Substitution Estimates." *Economic Letters* 190(2): 57-62.

[61] Soderbury Anson. 2012. "Estimating Import Supply and Demand Elasticities: Analysis and Implications." Purdue University.

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

[63] Tokarick, Stephen. 2012. "The Implications of China's Pattern of Growth for the Rest of the World." International Monetary Fund (September).

[64] Wooldridge, Jeffrey M. 2001. "Applications of Generalized Method of Moments Estimation." *Journal of Economic Perspectives*, 15(4): 87–100.

A Equilibrium of the Model

This appendix shows how to solve for the model's general equilibrium. To review the model: There are J countries and G tradable goods. The elasticity of substitution between any two goods g and g' is η, a constant. Each country j can produce multiple varieties of every good g. The constant elasticity of substitution between any two varieties of the same tradable good g is σ_g – regardless of the foreign country $i \neq j$ producing the variety of good g. However, the elasticity of substitution between the domestic basket of good g varieties, C_g^{jj}, and the composite foreign basket, C_g^{Fj} is ω_g; see equation (2).

Recall that N_g^{ij} is the measure of varieties of good g that country i produces for export to country j. The extensive margin is limited by fixed costs and a zero-profit condition, such that the firm-specific productivity of the marginal exporting firm, $\hat{\varphi}_g^{ij}$, entails zero profits, as shown by equation (5) in the main text. If $H_g^i(\varphi)$ is the cumulative distribution function for productivity of varieties of good g in country i, then the law of large numbers implies that $N_g^{ij} = 1 - H_g^i(\hat{\varphi}_g^{ij})$, as was noted earlier in the paper. We now show how the cutoff φ for each producer/good/destination is determined endogenously, in a global equilibrium, under the assumption that firm-specific productivity follows a Pareto distribution, as in equation (7). Under that distribution, $N_g^{ij} = \left(\hat{\varphi}_g^{ij}\right)^{-\gamma_g^i}$.

For importing country j, exporting country $i \neq j$, and good g, the cutoff productivity level for exports, $\hat{\varphi}_g^{ij}$, is defined by the equality of the marginal firm's profits and fixed costs, as in equation (5) in the main text:

$$
\begin{aligned}
W^i f_g^{ij} &= \pi_g^{ij}\left(\hat{\varphi}_g^{ij}\right) = \frac{p_g^i(\hat{\varphi}_g^{ij}) y_g^{ij}(\hat{\varphi}_g^{ij})}{\sigma} \Leftrightarrow \\
\sigma W_i f_g^{ij} &= p_g^i(\hat{\varphi}_g^{ij}) \tau_g^{ij} \kappa_g^{ij} \left[\frac{\tau_g^{ij} p_g^i(\hat{\varphi}_g^{ij})}{P_g^{Fj}}\right]^{-\sigma_g} (1-\beta_g^j) \left(\frac{P^{Fj}}{P_g^j}\right)^{-\omega_g} \alpha_g^j \left(\frac{P_g^j}{P^j}\right)^{-\eta} C^j \Leftrightarrow \\
\left(\hat{\varphi}_g^{ij}\right)^{\sigma_g - 1} &= \frac{\sigma W^i f_g^{ij}}{\kappa_g^{ij}(1-\beta_g^j)\alpha_g^j P^j C^j} \left(\frac{\sigma}{\sigma-1} \frac{\tau_g^{ij} W^i}{A_g A^i P^j}\right)^{\sigma_g - 1} \left(\frac{P_g^{Fj}}{P_g^j}\right)^{\omega_g - \sigma_g} \left(\frac{P_g^j}{P^j}\right)^{\eta - \sigma_g}. \qquad (46)
\end{aligned}
$$

For $i = j$ (that is, in the case of a domestic firm's home sales), preceding formula holds with i set equal to j (and $\tau_g^{jj} = 1$), β_g^j in place of the product $\kappa_g^{ij}(1-\beta_g^j)$, and P_g^{jj} in place of P_g^{Fj}.

Using equation (8), the overall price level for imports of good g in country j is

$$
P_g^{Fj} = \frac{\sigma_g}{(\sigma_g - 1) A_g} \left[\sum_{i=1, i \neq j}^{J} \left\{\frac{\gamma_g^i}{\gamma_g^i - (\sigma_g - 1)} \left(\frac{\tau_g^{ij} W^i}{A^i}\right)^{1-\sigma_g} \left(\hat{\varphi}_g^{ij}\right)^{(\sigma_g - 1) - \gamma_g^i}\right\}\right]^{\frac{1}{1-\sigma}}, \qquad (47)
$$

while the index for country j's own production of good g for domestic consumption, P_g^{jj}, is given by equation (8) with i set equal to j. As above, the relationship

$$
P_g^j = \left\{\beta_g^j \left(P_g^{jj}\right)^{1-\omega_g} + \left(1-\beta_g^j\right) \left(P_g^{Fj}\right)^{1-\omega_g}\right\}^{\frac{1}{1-\omega_g}} \qquad (48)
$$

then shows the dependence of the general price level for good g on the endogenous country-specific wages and country-pair-specific productivity cutoffs.[39]

To compute the model's general equilibrium, we have to determine the $(GJ+2)\times J$ unknowns $\left\{\hat{\varphi}_g^{ij}, W_i, C^i\right\}$. We begin with the equilibrium conditions for the national labor markets. Assume labor supplies are *exogenously fixed*, with L_i denoting the total labor supply in a country i (which will be divided between fixed and variable production costs). As explained above, the labor requirement for a good g exported to country j by a producer from country i is

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

with output demand given by

$$y_g^{ij}(\varphi) = \tau_g^{ij}\kappa_g^{ij}\left[\frac{\tau_g^{ij} p_g^i(\varphi)}{P_g^{Fj}}\right]^{-\sigma_g}(1-\beta_g^j)\left(\frac{P^{Fj}}{P_g^j}\right)^{-\omega_g}\alpha_g^j\left(\frac{P_g^j}{P^j}\right)^{-\eta}C^j$$

for an import and by

$$y_g^{jj}(\varphi) = \left[\frac{p_g^j(\varphi)}{P_g^{jj}}\right]^{-\sigma_g}\beta_g^j\left(\frac{P^{jj}}{P_g^j}\right)^{-\omega_g}\alpha_g^j\left(\frac{P_g^j}{P^j}\right)^{-\eta}C^j$$

for varieties of good g produced domestically, in country j itself.

Thus the total $foreign$ demand for a country i's labor is given by summing over all goods g and all importing countries $j \neq i$:

$$\begin{aligned}
&\sum_{g=1}^{G}\sum_{j\neq i}^{J}\left\{\int_{\hat{\varphi}_g^{ij}}^{\infty}\left[\frac{y_g^{ij}(\varphi)}{A_g A^i\varphi}+f_g^{ij}\right]\mathrm{d}H_g^i(\varphi)\right\}\\
&=\sum_{g=1}^{G}\sum_{j\neq i}^{J}\left[N_g^{ij}f_g^{ij}+\frac{\kappa_g^{ij}(1-\beta_g^j)\alpha_g^j(\sigma_g-1)^{\sigma_g}}{(\sigma_g W^i)^{\sigma_g}(A_gA^i)^{1-\sigma_g}}\cdot\frac{\left(P^j\right)^{\eta}C^j}{\left(\tau_g^{ij}\right)^{\sigma_g-1}\left(P_g^{Fj}\right)^{\omega_g-\sigma_g}\left(P_g^j\right)^{\eta-\omega_g}}\int_{\hat{\varphi}_g^{ij}}^{\infty}\varphi^{\sigma_g-1}\mathrm{d}H_g^i(\varphi)\right]\\
&=\sum_{g=1}^{G}\sum_{j\neq i}^{J}\left[N_g^{ij}f_g^{ij}+\frac{\gamma_g^i\kappa_g^{ij}(1-\beta_g^j)\alpha_g^j(\sigma_g-1)^{\sigma_g}}{(\sigma_g W^i)^{\sigma_g}(A_gA^i)^{1-\sigma_g}}\cdot\frac{\left(P^j\right)^{\eta}C^j}{\left(\tau_g^{ij}\right)^{\sigma_g-1}\left(P_g^{Fj}\right)^{\omega_g-\sigma_g}\left(P_g^j\right)^{\eta-\omega_g}}\int_{\hat{\varphi}_g^{ij}}^{\infty}\varphi^{\sigma-\gamma_g^i-2}\mathrm{d}\varphi\right]\\
&=\sum_{g=1}^{G}\sum_{j\neq i}^{J}\left[\left(\hat{\varphi}_g^{ij}\right)^{-\gamma_g^i}f_g^{ij}+\frac{\gamma_g^i\kappa_g^{ij}(1-\beta_g^j)\alpha_g^j(\sigma_g-1)^{\sigma_g}}{\left[\gamma_g^i-(\sigma_g-1)\right](\sigma_g W^i)^{\sigma_g}(A_gA^i)^{1-\sigma_g}}\cdot\frac{\left(P^j\right)^{\eta}C^j\left(\hat{\varphi}_g^{ij}\right)^{\sigma-1-\gamma_g^i}}{\left(\tau_g^{ij}\right)^{\sigma_g-1}\left(P_g^{Fj}\right)^{\omega_g-\sigma_g}\left(P_g^j\right)^{\eta-\omega_g}}\right]\\
&=\sum_{g=1}^{G}\sum_{j\neq i}^{J}\left(\hat{\varphi}_g^{ij}\right)^{-\gamma_g^i}\left[f_g^{ij}+\frac{\gamma_g^i\kappa_g^{ij}(1-\beta_g^j)\alpha_g^j\tau_g^{ij}}{\left[\gamma_g^i-(\sigma_g-1)\right]}\left(\frac{\sigma_g}{\sigma_g-1}\frac{\tau_g^{ij}W^i}{A_gA^i\hat{\varphi}_g^{ij}P_g^{Fj}}\right)^{-\sigma_g}\left(\frac{P_g^{Fj}}{P_g^j}\right)^{-\omega_g}\left(\frac{P_g^j}{P^j}\right)^{-\eta}\frac{C^j}{A_gA^i\hat{\varphi}_g^{ij}}\right].
\end{aligned}$$

[39] Recall that $N_g^{jj} = \left(\hat{\varphi}_g^{jj}\right)^{-\gamma_g^j}$ in equation (8).

Similarly, the *domestic* demand for country i's labor is

$$\sum_{g=1}^{G}\left\{\int_{\hat{\varphi}_g^{ii}}^{\infty}\left[\frac{y_g^{ii}(\varphi)}{A_gA^i\varphi}+f_g^{ii}\right]\mathrm{d}H_g^i(\varphi)\right\}$$
$$=\sum_{g=1}^{G}\left(\hat{\varphi}_g^{ii}\right)^{-\gamma_g^i}\left[f_g^{ii}+\frac{\gamma_g^i\beta_g^i\alpha_g^i}{\left[\gamma_g^i-(\sigma_g-1)\right]}\left(\frac{\sigma_g}{\sigma_g-1}\frac{W^i}{A_gA^i\hat{\varphi}_g^{ii}P_g^{ii}}\right)^{-\sigma_g}\left(\frac{P_g^{ii}}{P_g^i}\right)^{-\omega_g}\left(\frac{P_g^i}{P^i}\right)^{-\eta}\frac{C^i}{A_gA^i\hat{\varphi}_g^{ii}}\right].$$

As a result, we get the J labor-market equilibrium conditions (one for each country $i = 1,2,...,J$):

$$L_i=\sum_{g=1}^{G}\left(\hat{\varphi}_g^{ii}\right)^{-\gamma_g^i}\left[f_g^{ii}+\frac{\gamma_g^i\beta_g^i\alpha_g^i}{\left[\gamma_g^i-(\sigma_g-1)\right]}\left(\frac{\sigma_g}{\sigma_g-1}\frac{W^i}{A_gA^i\hat{\varphi}_g^{ii}P_g^{ii}}\right)^{-\sigma_g}\left(\frac{P_g^{ii}}{P_g^i}\right)^{-\omega_g}\left(\frac{P_g^i}{P^i}\right)^{-\eta}\frac{C^i}{A_gA^i\hat{\varphi}_g^{ii}}\right]$$
$$+\sum_{g=1}^{G}\sum_{j\neq i}^{J}\left(\hat{\varphi}_g^{ij}\right)^{-\gamma_g^i}\left[f_g^{ij}+\frac{\gamma_g^i\kappa_g^{ij}(1-\beta_g^j)\alpha_g^j\tau_g^{ij}}{\left[\gamma_g^i-(\sigma_g-1)\right]}\left(\frac{\sigma_g}{\sigma_g-1}\frac{\tau_g^{ij}W^i}{A_gA^i\hat{\varphi}_g^{ij}P_g^{Fj}}\right)^{-\sigma_g}\left(\frac{P_g^{Fj}}{P_g^j}\right)^{-\omega_g}\left(\frac{P_g^j}{P^j}\right)^{-\eta}\frac{C^j}{A_gA^i\hat{\varphi}_g^{ij}}\right]. \tag{49}$$

The outline for determining the $(GJ+2)\times J$ unknowns $\left\{\hat{\varphi}_g^{ij}, W_i, C^i\right\}$ is now clear. Substituting all country/good versions of the various price-index equations into (46), we obtain the first $G\times J^2$ equations that we need. We obtain another J equations from the conditions listed in (49), for a itotal of $(GJ+1)\times J$ conditions. To find the final J conditions giving us the $(GJ+2)\times J$ equations determining all of the $(GJ+2)\times J$ unknowns $\left\{\hat{\varphi}_g^{ij}, W_i, C^i\right\}$, we need to explain the J country-specific consumption levels.

To do so, notice that under balanced trade, the budget constraint of any country i is:

$$P^iC^i=\left\{\sum_{g=1}^{G}\sum_{j=1}^{J}\int_{\hat{\varphi}_g^{ij}}^{\infty}p_g^i(\varphi)y_g^{ij}(\varphi)\mathrm{d}H_g^i(\varphi)\right\}.$$

Using the production function and markup equation, however,

$$py=\frac{\sigma}{\sigma-1}\frac{W}{A_gA\varphi}A_gA\varphi(\ell-f)=\frac{\sigma}{\sigma-1}W(\ell-f).$$

Restoring the appropriate super/subscripts and substituting into the equation for P^iC^i above, we thus get the J final equations that we need:

$$P^iC^i=W^i\sum_{g=1}^{G}\left\{\frac{\sigma_g}{\sigma_g-1}\sum_{j=1}^{J}\int_{\hat{\varphi}_g^{ij}}^{\infty}\left[\ell_g^{ij}(\varphi)-f_g^{ij}\right]\mathrm{d}H_g^i(\varphi)\right\}.$$

These J consumption equations can be simplified by use of the cutoff conditions (5); they are equivalent to:

$$P^iC^i=W^i\sum_{g=1}^{G}\left\{\left[\frac{\sigma_g\gamma_g^i}{\gamma_g^i-(\sigma_g-1)}\right]\sum_{j=1}^{J}\left(\hat{\varphi}_g^{ij}\right)^{-\gamma_g^i}f_g^{ij}\right\}. \tag{50}$$

Using the cutoff conditions (5) a second time, one can likewise simplify the J equations in (49) dramatically, so that they become:

$$\sum_{g=1}^{G}\left\{\left[\frac{\sigma_g\gamma_g^i-(\sigma_g-1)}{\gamma_g^i-(\sigma_g-1)}\right]\sum_{j=1}^{J}\left(\hat{\varphi}_g^{ij}\right)^{-\gamma_g^i} f_g^{\,ij}\right\}=L_i. \tag{51}$$

A special case offers some intuition. If $\sigma_g=\sigma$ and $\gamma_g^i=\gamma^i$, for all goods g, then

$$\left[\frac{\sigma\gamma^i-(\sigma-1)}{\gamma^i-(\sigma-1)}\right]\sum_{g=1}^{G}\sum_{j=1}^{J}\left(\hat{\varphi}_g^{ij}\right)^{-\gamma^i} f_g^{\,ij}=L_i;$$

and using this to eliminate $\sum_{g=1}^{G}\sum_{j=1}^{J}\left(\hat{\varphi}_g^{ij}\right)^{-\gamma^i} f_g^{\,ij}$ from the corresponding version of equation (50), we find that

$$P^iC^i=\left[\frac{\sigma}{\sigma-(\sigma-1)/\gamma^i}\right]W_iL_i<\left(\frac{\sigma}{\sigma-1}\right)W_iL_i.$$

(The last inequality follows from the earlier assumption that $\gamma^i>\sigma-1$.) Equilibrium consumption is below markup-adjusted labor costs because some labor is used to cover fixed production costs.

Without assuming that $\gamma_g^i=\gamma^i$ or $\sigma_g=\sigma$ for all goods g, we can compute the equilibrium as follows. Use equation (50) to eliminate C^i in equation (46); then use the GJ^2 resulting productivity cutoff equations plus the J equations of (51), together with the price-level equations, to solve for the $(GJ+1)\times J$ unknowns $\left\{\hat{\varphi}_g^{ij},W_i\right\}$; and finally, solve for the J real consumption levels $\{C^i\}$ using the equation for consumption, (50).

B Impact of a Devaluation

As derived in the text,

$$\frac{d\ln V^F}{d\ln E}=1+\sum_{g=1}^{G}\left(\frac{E^2}{V^F}\right)\frac{d}{dE}\left[P_g^{F*}\alpha_g(1-\beta_g)\left(\frac{P_g^F}{P_g}\right)^{-\omega_g}\left(\frac{P_g}{P}\right)^{-\eta}C\right].$$

We shall analyze the summation term by term. Observe that we can write a generic term in the summation as

$$\begin{aligned}&\left(\frac{EP_g^FC_g^F}{V^FP_g^{F*}C_g^F}\right)\frac{d}{dE}\left[P_g^{F*}\alpha_g(1-\beta_g)\left(\frac{P_g^F}{P_g}\right)^{-\omega_g}\left(\frac{P_g}{P}\right)^{-\eta}C\right]\\ =\;& w_g^F\frac{d}{d\ln E}\ln\left[P_g^{F*}\alpha_g(1-\beta_g)\left(\frac{P_g^F}{P_g}\right)^{-\omega_g}\left(\frac{P_g}{P}\right)^{-\eta}C\right],\end{aligned}$$

where, $w_g^F\equiv V_g^F/V^F$. So we wish to compute

$$\frac{d}{d\ln E}\left[-\omega_g\ln P_g^F+(\omega_g-\eta)\ln P_g+\eta\ln P+\text{ constants}\right].$$

The result of taking this derivative is

$$-\omega_g + (\omega_g - \eta)m_g + \eta \sum_g m_g w_g,$$

where, $m_g \equiv V_g^F/V_g,\ w_g \equiv V_g/V$. Thus each term in the summation above is given by

$$\begin{aligned} & \left(\frac{E^2}{V^F}\right) \frac{d}{dE} \left[P_g^{F*} \alpha_g (1-\beta_g) \left(\frac{P_g^F}{P_g}\right)^{-\omega_g} \left(\frac{P_g}{P}\right)^{-\eta} C\right] \\ = \quad & w_g^F \left[-\omega_g + (\omega_g - \eta)m_g + \eta \sum_{g=1}^{G} m_g w_g\right] \end{aligned}$$

and so,

$$\frac{d \ln V^F}{d \ln E} = 1 + \sum_{g=1}^{G} w_g^F \left[-\omega_g + (\omega_g - \eta)m_g + \eta \sum_{g=1}^{G} m_g w_g\right].$$

To simplify, notice that because $\sum_{g=1}^{G} w_g^F = 1$, the last equation becomes:

$$\begin{aligned} \frac{d \ln V^F}{d \ln E} &= 1 - \sum_{g=1}^{G} w_g^F \omega_g + \sum_{g=1}^{G} w_g^F m_g(\omega_g - \eta) + \eta \sum_{g=1}^{G} m_g w_g \\ &= 1 - \sum_{g=1}^{G} (1 - m_g) w_g^F \omega_g + \eta \sum_{g=1}^{G} m_g \left(w_g - w_g^F\right). \end{aligned}$$

Note further that

$$m_g w_g = \frac{V_g^F}{V_g} \frac{V_g}{V} = \frac{V_g^F}{V^F} \frac{V^F}{V} = w_g^F m,$$

where, $m \equiv V^F/V$ is the share of imports (of all goods) in total consumption spending. Thus, we can rewrite the derivative above in the final form:

$$\frac{d \ln V^F}{d \ln E} = 1 - \sum_{g=1}^{G} (1 - m_g) w_g^F \omega_g - \eta \sum_{g=1}^{G} w_g^F (m_g - m).$$

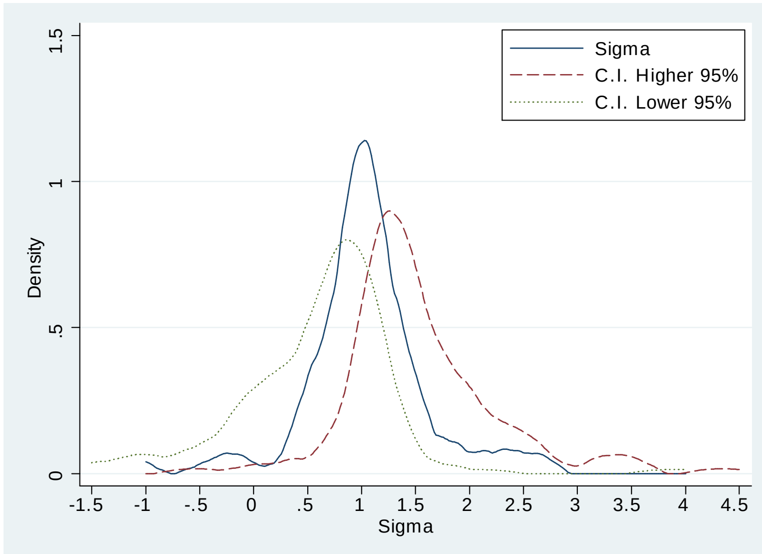

Figure 1: OLS results for sigma

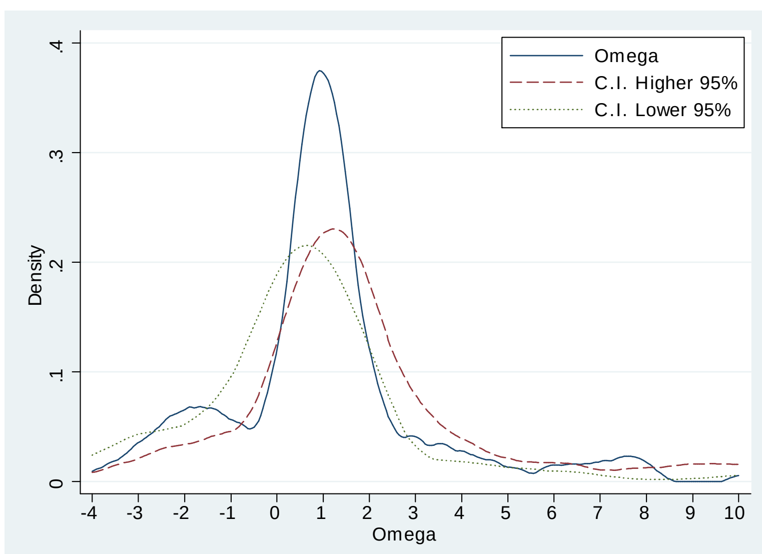

Figure 2: OLS results for omega

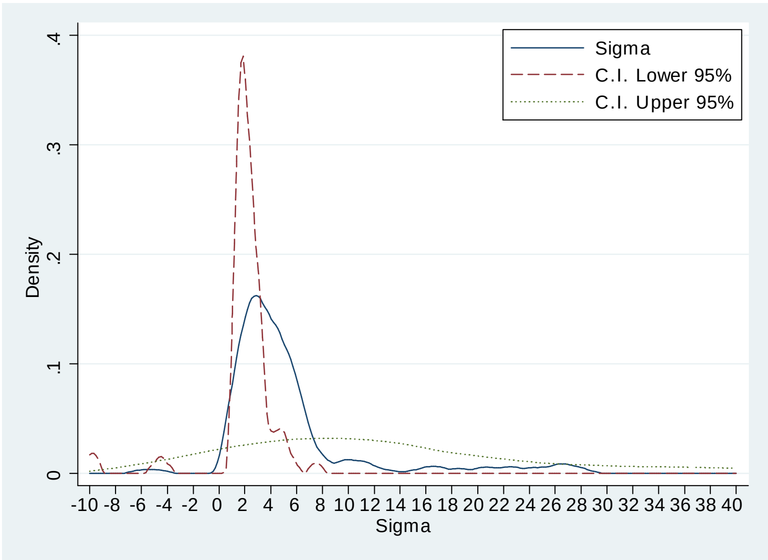

Figure 3: 2-step GMM results for sigma