March 2008

The International Elasticity Puzzle

Kim J. Ruhl*

University of Texas at Austin

ABSTRACT

In models of international trade, the elasticity of substitution between foreign and domestic goods—the Armington elasticity—determines the behavior of trade flows and international prices. International real business cycle models need low elasticities, in the range of 1 to 2, to match the quarterly fluctuations in trade balances and the terms of trade, but static applied general equilibrium models need high elasticities, between 10 and 15, to account for the growth in trade following trade liberalization. To reconcile these contradictory findings, we construct a model in which cyclical fluctuations are caused by temporary shocks, as in business cycle models, but tariff changes are permanent. In the model, firms do not change export status in response to temporary shocks, while tariff decreases induce some non-exporters to export. In a calibrated model, the entry of new exporters increases the measured elasticity with respect to a tariff change to 6.4, while the elasticity in response to temporary shocks is 1.2.

*I would like to thank Timothy Kehoe for his advice and encouragement. I also thank V.V. Chari, Patrick Kehoe, Sam Kortum, Erzo G.J. Luttmer, Edward Prescott, and Kei-Mu Yi helpful discussions and comments. I thank the National Science Foundation for financial support under grant number 0536970.

1. Introduction

A common feature of many international trade models is national product differentiation: countries produce and trade differentiated goods that are to some extent substitutable for each other. In these models, the elasticity of substitution between home and foreign produced goods—the *Armington elasticity*, after Armington (1969)—is the critical parameter for determining the behavior of trade flows and international prices. The importance of this parameter has manifested in two of the leading branches of international economics: the international business cycle literature that seeks to understand the high frequency fluctuations in macroeconomic aggregates, and the static applied general equilibrium literature that focuses on explaining the patterns of trade and the effects of trade policy. These two disciplines, however, have very different views on the value of the Armington elasticity.

International real business cycle (IRBC) models need small values of the elasticity to generate the volatility of the terms of trade and the negative correlation between the terms of trade and the trade balance that are found in the data. IRBC models commonly use Armington elasticities around 1.5, though sensitivity analysis suggests values even lower than this may be appropriate. (See for example, Backus, Kehoe and Kydland (1994), Zimmermann (1997), or Heathcote and Perri (2002).) Not surprisingly, when empirical researchers have estimated the Armington elasticity from high frequency data they find small estimates that range from about 0.2 to 3.5. In contrast, applied general equilibrium (GE) models need large Armington elasticities to explain the growth in trade volumes that result from a change in tariffs. As shown in Yi (2003), these models need elasticities of 12 or 13 to generate the large growth in trade found in the data. Empirical work on trade liberalizations, as well as cross country regressions relating trade patterns to tariff and non-tariff barriers, find Armington elasticities that range from about 4 to 15, similar to the ones needed in applied GE models.

The key to understanding these two different measurements of the Armington elasticity is to realize that the source of variation in the prices and quantities being measured is different. The high frequency variation in the time series data is likely caused by transitory shocks to supply or demand. These are exactly the transitory shocks used in the business cycle models that need small Armington elasticities. Trade liberalization, however, can be thought of as a permanent change. When agents react differently to temporary and permanent changes, the measured Armington elasticities will differ.

In this paper we show that a model combining elements of the real business cycle and the applied GE frameworks can reproduce both the low elasticities estimated from the time series data and the high elasticities implied by the growth in trade following a decrease in tariffs. The model features an entry cost of exporting and heterogeneous firms, as in Melitz (2003), but adds aggregate productivity shocks, as found in the IRBC models of Backus, Kehoe and Kydland (1994) and Stockman and Tesar (1995). As in Melitz (2003), entry costs interact with firm level heterogeneity to partition firms into exporters and non-exporters. A firm decides whether to become an exporter by comparing the expected future value of exporting to the cost of entry. If the expected gain from exporting is larger than the cost of entry, the firm becomes an exporter. The movement of firms into and out of exporting in response to temporary changes in productivity or permanent changes in tariffs drives the two very different elasticities measured in the model.

Since temporary changes in productivity change expected future profits from exporting little, few firms choose to change their exporting status. Thus, temporary changes in productivity tend to show up as price changes in the goods already being traded, and consumers substitute between goods at the low "true" elasticity specified as a parameter in the model. We call this change in the amount of goods that were already being traded a change on the *intensive margin.* When all change occurs on the intensive margin, as it does in most trade models, estimating the elasticity from aggregate trade flows will recover the true elasticity. In our model there is a small number of firms entering and exiting the export market in response to productivity shocks, but the trade from these firms is small compared to aggregate trade. Thus, when we estimate the Armington elasticity in the model in response to productivity shocks, we find small values that are close to the true elasticity specified by the parameters. The estimated elasticity is not exactly equal to the true elasticity because the price indices constructed for use in the estimation do not incorporate the changing set of goods being traded. As shown in Feenstra (1994), not including these newly traded goods imparts a bias to the indices. We discuss this mismeasurement and its implications in Section 6.

In contrast, a permanent change, such as lower tariffs, raises the profit from exporting in all states of nature and increases the expected future gain from exporting more than a temporary productivity shock. This induces more firms to begin exporting, resulting in large trade flows. We call the increase in trade flows from newly traded goods an increase on the *extensive margin*. The increase in the extensive margin is the key to understanding the large elasticities measured in response to trade liberalization. Following a decrease in tariffs, trade increases for two reasons: consumers buy more of the goods they already import, since the delivered price is now lower, and they buy new imports that were not previously available. The first kind of growth is intensive margin growth, while the second is extensive margin growth. If the change in aggregate trade is mistakenly assumed to be all intensive margin growth, the small changes in tariffs appear to induce large changes in imports on the intensive margin, which implies a large elasticity of substitution. When we make a similar measurement in our model after a decrease in tariffs, we find an Armington elasticity that is more than 3 times the true value. When we shut down the extensive margin in our model, however, our measured elasticity is the same as the true elasticity.

There is growing evidence that the extensive margin, which drives the central result of this paper, is an important facet of the data. Empirically, Hillberry and McDaniel (2002) find evidence of extensive margin growth for the U.S. following the implementation of the North American Free Trade Agreement (NAFTA). Kehoe and Ruhl (2002) document extensive margin growth in a study of trade liberalizations and lay out a simple Ricardian model to highlight the forces at work. In a test of several firm-level models of exporter behavior, Bernard, Jensen and Schott (2003) study data on U.S. manufacturing plants and find strong evidence that new plants choose to enter the export market as trade costs fall.[1] In a cross sectional study, Hummels and Klenow (2005) find that larger and richer countries have larger extensive margins. The larger, richer, countries also tend to be the countries with the most liberalized trade policy. Broda and

[1] Bernard, Jensen et al. (2003) test the predictions of the plant-level models of Bernard, Eaton, Jensen and Kortum (2003), Melitz (2003), and Yeaple (2005).

Weinstein (2006) document the increase in the number of imported goods available in the U.S. and compute an import price index, based on Feenstra (1994), which corrects for these newly imported goods. They find significant gains in welfare from these new imports.

Earlier theoretical work on export entry costs and uncertainty has focused on hysteresis in exporting. Baldwin (1988), Baldwin and Krugman (1989), and Dixit (1989) show how temporary increases in profitability (they considered exchange rates) could increase exports as firms enter the export market, but that high levels of exports would persist even as exchange rates appreciated. The firms that had incurred the sunk cost to export would continue to do so, even faced with a less attractive exchange rate, since they had already incurred the start-up cost.

In more recent work, Melitz (2003) incorporates export entry costs into a model of monopolistic competition, as in Dixit and Stiglitz (1977), in an environment with no aggregate uncertainty. His analysis focuses on the reaction of an industry to changes caused by trade liberalization. The model presented here has a structure of production similar to that in Melitz (2003), but focuses on the effects of industry structure on aggregate trade in the presence of aggregate uncertainty. In Alessandria and Choi (2007a) export entry costs are imbedded in a standard international real business cycle model to study the effects of exporter entry and exit on international correlations of consumption and output. They find, as do we, that the extensive margin is not important for aggregate quantities at business cycle frequencies. Ghironi and Melitz (2005) model fixed—but not sunk—export costs, in a real business cycle model to generate endogenous persistent deviations from purchasing power parity. Their models are conceptually similar to the one presented here, in that they involve aggregate uncertainty modeled as shocks to productivity. Their analysis, however, is focused on the characteristics of international business cycles, while we are concerned with the different behavior of trade flows in response to different sources of variation.

Empirical justification for the export entry costs that are central to this model, and the literature cited above, has come as the result of plant-level dynamic models of export entry decisions. The seminal work in this literature is Roberts and Tybout (1997), who develop an econometric model of a plant's decision to enter the export market. Using panel data on the Colombian manufacturing sector, they reject the null hypothesis that entry costs are unimportant. In a study of German plants, Bernard and Wagner (2001) find evidence of substantial sunk costs in export entry. Using a detailed data set on U.S. manufacturing plants, Bernard and Jensen (2004) find that export entry costs are significant and that plant heterogeneity is important in the export decision.

The next section reviews the evidence on the elasticity of substitution as estimated by different authors using different techniques. Section 3 lays out evidence that the extensive margin is active during times of policy changes, but not during business cycles. Section 4 presents a model of the extensive margin: firms choose whether or not to export in the presence of fixed costs and uncertainty about future productivity. Section 5 discusses computational issues and calibration of the model. Section 6 presents the model results, and Section 7 concludes.

2. Measuring Import Price Elasticities

In this section we review previous estimates of the Armington elasticity. Based on our hypothesis that transitory changes in profitability lead to different responses than permanent changes, we divide the literature into two subsections. We find that studies that use high frequency price variation to estimate the Armington elasticity find low values, while studies that use data on trade barriers or data from trade liberalizations, tend to find much higher values.

2.1. Incorporating Substitution into Models

In Armington (1969), it is posited that goods produced by different countries are intrinsically different goods. The utility that consumers derive from these nationally differentiated goods is represented by a constant elasticity of substitution utility function,

$$U = \left[\omega C_h^{\rho} + \left(1-\omega\right) C_f^{\rho} \right]^{\frac{1}{\rho}} \tag{1}$$

where C_k is the consumption of the good produced in country k. Maximizing this function with respect to the standard budget constraint and rearranging the first order conditions yields

$$\frac{C_h}{C_f} = \left(\frac{P_f}{P_h} \frac{\omega}{1-\omega} \right)^{\sigma}, \tag{2}$$

where P_h is the price of the good produced in the home country and P_f is the price of the good produced in the foreign country. It is easy to see from (2) that $\sigma = 1/(1-\rho)$ is the elasticity of substitution between the goods. When the two goods are differentiated by country of origin, as in this case, this elasticity is commonly referred to as the *Armington elasticity*.

An alternative way of incorporating national product differentiation is to assume that countries produce intermediate goods that are combined to produce an aggregate consumption-investment good. In these models, the feasibility condition is

$$C + X = \left[\omega\, q_h^{\rho} + (1-\omega)\, q_f^{\rho}\right]^{\frac{1}{\rho}}, \tag{3}$$

where C is consumption, X is investment, q_h is the intermediate good produced in the home country, and q_f is the intermediate good produced in the foreign country. When using this specification, the constant elasticity function on the right-hand side of (3) is commonly called the *Armington aggregator*. Minimizing the cost of producing one unit of the aggregate good implies the same first order condition as (2).

In trade models that feature imperfect competition and differentiated goods, such as those in Helpman and Krugman (1985), consumers are frequently modeled as having constant elasticity of substitution preferences over varieties of goods within an industry,

$$U_j = \left[\omega \int_{\iota \in \mathrm{I}_h} c_{h,j}(\iota)^{\rho}\, d\iota + (1-\omega) \int_{\iota \in \mathrm{I}_f} c_{f,j}(\iota)^{\rho}\, d\iota \right]^{\frac{1}{\rho}}, \tag{4}$$

where $c_{h,j}(\iota)$ is consumption of domestically produced variety ι, and $c_{f,j}(\iota)$ is consumption of the foreign produced variety in industry j. The set I_k is the set of varieties that the consumer has available for purchase. Maximizing (4) subject to a standard budget constraint yields the familiar condition,

$$\frac{c_{h,j}(\iota)}{c_{f,j}(\iota')} = \left(\frac{p_{f,j}(\iota')}{p_{h,j}(\iota)} \frac{\omega}{1-\omega}\right)^{\sigma}, \tag{5}$$

where $p_{f,j}(\iota')$ and $p_{h,j}(\iota)$ are, respectively, the price of the foreign and home goods. For example, $c_{h,j}(\iota)$ could be a shirt made domestically, while $c_{f,j}(\iota')$ is a shirt made abroad and imported. The empirical literature we survey below uses data collected at a level of aggregation higher than the "variety" level used in (5) so, it is common to write the first order condition in terms of industry level aggregate quantities and prices,

$$\frac{C_{h,j}}{C_{f,j}} = \left(\frac{P_{f,j}}{P_{h,j}} \frac{\omega}{1-\omega} \right)^{\sigma} \tag{6}$$

where $C_{h,j}$ is the aggregate amount of domestic consumption of goods in industry j and $P_{h,j}$ is an aggregate price index for the domestically produced goods. $C_{f,j}$ and $P_{f,j}$ are similarly defined for the goods imported into the home country.

Note that in contrast to the Armington model, goods in this specification are not different because they are made in different countries. Here goods are different by their very nature; these are the "differentiated goods" found in models of monopolistic competition such as Dixit and Stiglitz (1977). That the goods are made in different countries matters only to the consumer through the home bias parameter, ω. What is the same about the two models, however, is the implication of the first order conditions, (2) and (6). These equations are the basis for the estimation of Armington elasticities.

2.2. Estimates from Price Variation

We begin by reviewing the Armington elasticities used in the IRBC literature. The Armington elasticities used in the IRBC literature range from 0.5 to 2.0. In Backus, Kehoe, and Kydland (1994) each country produces a tradable intermediate good which is combined using an Armington aggregator, as in (3), to produce an aggregate consumption-investment good. The authors' baseline choice of the Armington elasticity is 1.5, but they perform sensitivity analysis to this parameter. They find that the model with a smaller elasticity (0.5) can better account for the volatility of the terms of trade and the negative correlation between the terms of trade and the trade balance, than can the model with a larger (2.5) elasticity. Heathcote and Perri (2002) use a similar two-intermediate-goods environment to study business cycles under different degrees of financial market completeness. Beginning with a baseline value of 1.0, they find that the volatility of the terms of trade, the cross-country correlation of investment, and the cross-country correlation of consumption and output are closer to those in the data when lower values of the elasticity are used.

In addition to the low elasticities needed for IRBC models to match the features of the high frequency data, low elasticities are also found when they are directly estimated from high frequency data on prices. The estimating equations are derived from the first order conditions, such as those in (2) or (6). Taking the logarithm of (6) yields the basic equation estimated by several authors,

$$\log\left(C_{f,jt}/C_{h,jt}\right)=\alpha_j+\sigma_j\log\left(P_{h,jt}/P_{f,jt}\right)+\varepsilon_{jt}\,, \tag{7}$$

where $C_{f,jt}$ is the real quantity of imports in industry j, $C_{h,jt}$ is the real consumption of domestically produced goods in industry j, and $P_{h,jt}$ and $P_{f,jt}$ are price indices for domestic sales and imports. To estimate this equation, quarterly data on imports and domestic consumption of the industry's good, as well as data on the relative prices is collected. Typically, the price data take the form of a unit price index for imports, and a producer price index for domestic goods. Reinert and Roland-Holst (1992) estimate an equation similar to (7) for 163 industries and find elasticities that range from 0.02 to 3.49, with an average value of 0.91. Blonigen and Wilson (1999) estimate elasticities for 146 sectors, and find an average elasticity of 0.81, with a maximum value of 3.52. Similar estimates are found in Reinert and Shiells (1993) and in the short run elasticities reported in Gallaway, McDaniel and Rivera (2003). The elasticities found by estimating (7) on high frequency time series data are fairly robust. Adjustments have been made for, among other things, serially correlated errors, differing levels of aggregation, and seasonal effects. The elasticities estimated using these expanded techniques still find low values.

The most complete study of trade elasticities to date is Broda and Weinstein (2006), in which the authors estimate tens of thousands of elasticities for the United States. Using a theoretical model with three tiers of goods: a composite imported good, imported goods, indexed by g, and varieties of a good, m_{gc}, where c indexes the country of origin, the authors extend the methodology developed in Feenstra (1994) and estimate a demand equation for varieties,

$$\Delta_{t-1,t}\log\left(s_{gct}\right)=\varphi_{gt}-\left(\sigma_g-1\right)\Delta_{t-1,t}\log\left(p_{gct}\right)+\varepsilon_{gct}\,, \tag{8}$$

where s_{gc} is the expenditure share of variety gc in good g, φ_{gt} is a random effect, and Δ is the difference operator. The demand equation is estimated with an export supply equation to allow for upward sloping supply curves,

$$\Delta_{t-1,t}\log\left(p_{gct}\right)=\psi_{gt}+\omega_g\Delta_{t-1,t}\log\left(s_{gct}\right)+\delta_{gct}. \quad (9)$$

These equations are differenced with respect to a reference county, eliminating the random effects, and the resulting system of equations is estimated using a general method of moments estimator. Identification comes from the cross country variation in prices (the between estimator).

Broda and Weinstein (2006) define a 10-digit Harmonized System code or 7-digit Tariff System of the U.S.A. code from a particular country as a variety. When defining a good as a HS code, they report good specific elasticities with a median value of 3.10. The median elasticity is higher than the others found in this section, but it is important to note that unlike the other studies in the section, the elasticities are restricted to be greater than one. Broda and Weinstein (2006) estimate some large elasticities as well—the maximum elasticity in the HS data is 4302.6—but the estimated values vary with the good type in ways one might expect. For example, goods that are commodities have higher elasticities than goods that are considered *a priori* differentiated.

What is clear, from both the calibration based literature and the empirical estimates, is that the cyclical fluctuations in prices and quantities seem to imply that the Armington elasticity is small. We now turn to the estimates of the Armington elasticity that use data from trade liberalizations to identify the elasticity of substitution.

2.3. Estimates from Trade Policy and Geography

In this subsection we consider estimates of the Armington elasticity that are derived from variation in geography and tariffs. In contrast to the high frequency data used in the studies above, the main sources of variation considered in these studies are permanent in nature. In the studies of specific trade liberalization episodes, the trade policy being analyzed is typically not a temporary policy to be reversed later. In the cross section regressions, cross-country variation in trade barriers, such as transportation costs, tends to be permanent. Though per-mile transportation costs may be falling, distances between countries are not; the relative transportation cost of importing a good from two different countries stays about the same. The heterogeneity in bilateral trade policy, which is also an important source of variation, may change, but again, the changes are likely permanent.

Before turning to the empirical estimates, we consider the model-based "estimates" of the Armington elasticity. Applied general equilibrium models have been used in the past to predict the consequences of trade policy, but these models can also be used *ex post* to study trade flows. This method uses the observed change in trade flows and tariff rates to back out values for the elasticity of substitution between foreign and domestic goods. Yi (2003) performs this exercise for three workhorse models: a Ricardian model, as in Dornbusch, Fischer, and Samuelson (1977), a model with differentiated goods as in Krugman (1980), and a model with an Armington aggregator, as in Backus, Kehoe, and Kydland (1994). He finds that the Armington elasticity needs to be at least 12 for any of the models to generate the observed trade flows in response to the observed tariffs rates. Similarly, in a "calibration-as-estimation" exercise, Anderson, Balistreri, Fox, and Hillberry (2005) find that the Armington elasticities needed to match the world bilateral trade pattern are, on average, 17.

When Armington elasticities are econometrically estimated from trade policy episodes, the results support the high elasticities found in the model-based exercises like Yi (2003). Elasticities derived from natural experiments, like policy changes, are typically computed from the changes in imports and domestic consumption given the observed changes in tariff rates. Comparisons are made over two points in time, typically several years, as liberalization is usually a gradual process. A simple version of the calculation is

$$\Delta_{t,t+T} \log\left(\frac{C_{hj}}{C_{fj}}\right) = \alpha_j + \sigma_j \Delta_{t,t+T} \log\left(1+\tau_j\right) + \varepsilon_{j,t+T}\ , \tag{10}$$

where τ is the *ad valorem* tariff rate and $\Delta_{t,t+T}$ is the $T-$year difference operator. The subscript j indexes the industry; the equation specified in (10) can be estimated as a panel. Calculating the elasticity this way assumes that the change in tariff is the only change in relative prices, although various adjustments are typically made to control for other sources of variation.

Using this methodology, Clausing (2001) finds a price elasticity of 9.6 in a study of the Canada-U.S. Free Trade Agreement.[2] Head and Ries (2001) find elasticities that range from 7.9 to 11.4 in a regression relating trade shares to both tariff and non-tariff barriers between Canada and the United States. In a detailed study of NAFTA and the Canada U.S. Free Trade Agreement that features data on thousands of goods, Romalis (2007) estimates elasticities that range between 4 and 13. These estimates are based on the substantial variation in tariff rates across partners and goods that is a feature of the disaggregated data.

Further evidence of high Armington elasticities can be found in cross sectional studies, where the identification comes from differences across countries, such as distance or tariff levels. In a model of economic geography with explicit consideration of transportation costs, Hummels (2001) uses data from Argentina, Brazil, Chile, New Zealand, Paraguay, and the United States to estimate the elasticity of substitution between varieties in a model with preferences similar to (4). The estimation produces elasticities ranging from 3 to 8. Baier and Bergstrand (2001) estimate a panel model over 16 OECD countries and find that trade flows are about 16 times more responsive to changes in tariffs than changes in relative prices.

In this section we have seen how estimates of the Armington elasticity differ considerably when the underlying source of variation differs. Elasticities measured by studying the high frequency changes in prices and quantities are likely capturing responses to transitory shocks, and range between 0.2 and 3.5. Estimates based on changes in trade policy, other trade costs, or cross country variation are likely capturing the response of trade flows to more permanent factors, and usually range between 4 and 15. The large differences in these estimates suggest that trade flows are more sensitive to permanent changes then to temporary shocks.

3. When is the Extensive Margin Important?

The estimates surveyed above imply that different kinds of variations in import prices drive different measured elasticities. In the model presented below, we use the extensive

[2] In models based on (4), the elasticity of substitution and the price elasticity of demand are equal. In models with a finite number of varieties the price elasticity converges to the elasticity of substitution as the number of varieties approaches infinity.

margin to generate the different response to temporary and permanent changes. In this section we review the evidence on when the extensive margin is most important. Kehoe and Ruhl (2002) use disaggregated trade flow data to determine when the extensive margin is important for trade growth. Using data on about 800 4-digit Standard International Trade Classification (SITC) codes, they construct a set of least traded goods—goods with no trade plus goods with very little trade—and study how this set of goods grows during trade liberalization episodes and over the business cycle.

To construct the set of least traded goods for a particular trade flow, they order the SITC codes by the value of trade in a base year. They cumulate the ordered codes to form 10 sets of codes, each representing one-tenth of total exports. The first set is constructed, starting with the smallest codes, by adding codes to the set until the sum of their values reaches one-tenth of total export value. The next set is formed by summing the smallest remaining codes until the value of the set reaches one-tenth of total export value. This procedure produces 10 sets of codes, each representing one-tenth of total trade in the base year. The first set consists of the least traded goods: the codes with the smallest export values, including all the SITC codes with zero trade value. If goods that were previously not traded, or traded very little, begin being traded this will show up in the data as growth in the set of least traded goods, that is, growth on the extensive margin.

3.1. Trade Liberalization and the Extensive Margin

Kehoe and Ruhl (2002) find substantial growth in the set of least traded goods following trade liberalization. As an example, consider the North American Free Trade Agreement's impact on exports from Mexico to Canada, the results of which are reproduced in Figure 1. Taking 1989 as the base year, the goods are partitioned into 10 sets: each bar in Figure 1 represents a set of goods that makes up 10 percent of total trade in 1989. The number above the bar is the number of SITC codes in the set. The first bar is the set of least traded goods: 736.6 of the least traded SITC codes that account for 10 percent of trade in 1989. The height of the bar represents the share of those same goods in 1989. The extensive margin growth is striking; the set of goods that accounted for 10 percent of trade in 1989 accounts for 23 percent of trade in 1999. Further evidence of the impact of the NAFTA can be seen in Figure 2. This figure focuses on the timing of the growth of the extensive margin. For each year, the least traded goods' share of total Mexican exports to Canada is plotted. The extensive margin decreases slightly prior to the NAFTA and grows as the agreement is implemented. Kehoe and Ruhl (2002) study various liberalization episodes and find extensive margin growth to be a robust feature of the data.

3.2. Business Cycles and the Extensive Margin

Kehoe and Ruhl (2002) show that while the extensive margin is active during trade liberalization, it is does not change much over the business cycle. Take the United States' exports to the United Kingdom from 1989-1999 as an example; the U.K. is one of the top trading partners of the U.S. during this time period. Trade policy between the two countries did not change significantly, but the period encompasses both a recession and an expansion. As in the previous example, trade between the U.S. and the U.K. is partitioned into 10 sets of goods. The results are plotted in Figure 3. The least traded goods, which make up 10 percent of trade in 1989, make up 11 percent of trade in 1999. In fact, the there is little change in the composition of goods being traded; no set of goods changes by more than four percentage points. In Figure 4 the evolution of the least traded goods is plotted. While there is some variation in the share of trade accounted for by the least traded goods, the variation is small compared to that in Figure 2.

Estimates of the elasticity of substitution between domestic and foreign goods show that trade flows are more sensitive to permanent changes, like trade liberalization than they are to cyclical fluctuations. The data from disaggregated trade flows reveals that the extensive margin is active during these trade policy episodes, but not during business cycles. In the next section we specify a model that ties these two ideas together. The model produces temporary changes through productivity shocks and permanent changes through trade policy. It is then possible to study the effects of temporary and permanent changes on trade flows and the extensive margin.

4. Model

The model is designed to incorporate the major elements of standard applied general equilibrium models into an environment with aggregate uncertainty as in IRBC models. Consumers derive utility from consuming differentiated goods sold in markets characterized by monopolistic competition. The economy departs from the standard framework by requiring firms to pay a one-time entry cost before exporting and by subjecting the economy to aggregate productivity shocks. This structure of production is similar to Melitz (2003), but here the economy is subject to uncertainty about future productivity. These elements create an economy in which firms respond differently to temporary shocks to productivity, than they do to permanent changes in policy.

The economy consists of two countries, denoted as home (h) and foreign (f). Each country is endowed with $\bar{L}_k$ units of labor, which is the only factor of production. Each country produces two types of goods: a non-traded, homogeneous good, q, which is sold in a competitive market, and a continuum of tradable differentiated goods, indexed by ι, which are produced by monopolistically competitive firms. A differentiated good produced in country h and consumed in country f is denoted by $c_h^f(\iota)$. For clarity, only the home country variables and maximization problems are described. The foreign country faces analogous problems.

At each date t, one of H possible events, η_t, occurs. Each event is associated with a vector of economy-wide productivity shocks, $z_t = \left(z_h(\eta_t), z_f(\eta_t)\right)$, and the initial event η_0 is given. We assume η follows a stationary first-order Markov chain with transition matrix Λ. An element of Λ, the probability of event η' happening tomorrow, given event η happened today, is $\lambda_{\eta\eta'} = \mathrm{pr}\left(\eta_{t+1} = \eta' \middle| \eta_t = \eta\right)$.

A key feature of this economy is that differentiated good firms are heterogeneous in their productivity and in the entry cost they must pay in order to export. A firm is indexed by its idiosyncratic productivity, $\phi \in \mathcal{F}$, its entry cost, $\kappa \in \mathcal{K}$, and whether or not it is an exporter. The firm's values of ϕ and κ are constant for the life of the firm. A firm is an exporter if and only if it has paid the entry cost κ. The distribution of firms that begin the period as exporters is represented by a measure over (ϕ,κ), $\mu_{hx}(\phi,\kappa)$, which has support $\mathcal{F} \times \mathcal{K}$. Plants that begin the period as non-exporters are tracked by a similar measure, $\mu_{hd}(\phi,\kappa)$. Since firms have market power, they will choose their prices taking into account the consumer's demand function. In order to compute the household's demand, the firm needs to know which firms are selling in a market, so the measures over exporters and non-exporters are state variables. Denoting the vector of firm distributions as $\mu = \left(\mu_{hd}, \mu_{hx}, \mu_{fd}, \mu_{fx}\right)$, the aggregate state variables for the economy can be represented by (η, μ). The agents in the economy take as given the laws of motion for μ,

$$\mu'_{hd}(\phi,\kappa) = \mathrm{M}_{hd}(\phi,\kappa,\eta,\mu) \tag{11}$$
$$\mu'_{hx}(\phi,\kappa) = \mathrm{M}_{hx}(\phi,\kappa,\eta,\mu),$$

with similar laws of motion for the distributions in the foreign country. For notational simplicity, define $\mathrm{M}(\;)$ as the law of motion over the distribution of all firms,

$$\mu' = \mathrm{M}(\eta,\mu). \tag{12}$$

4.1. Households

The economy is populated by a unit measure of identical households. It is assumed that each household owns an equal share in all domestic firms in operation. We do not allow countries to borrow and lend with each other, but within a country, households can trade a complete set of Arrow Securities. Given these assumptions, and the homotheticity of preferences, the households can be represented by a stand-in household. The household's period utility function is

$$u\left(c_h^h, c_f^h, q_h\right) = \frac{\gamma}{\rho}\log\left(\int_{\iota\in \mathrm{I}_h^h(\mu)} c_h^h(\iota)^\rho d\iota + \int_{\iota\in \mathrm{I}_f^h(\mu)} c_f^h(\iota)^\rho \, d\iota\right) + (1-\gamma)\log\left(q_h\right), \tag{13}$$

where $c_k^j(\iota)$ is the differentiated good ι made in country k and consumed in country j. We denote the pre-tariff price of differentiated good ι made in country k and consumed in country j as $p_k^j(\iota)$. Imports are subject to an ad valorem tariff, τ, which is modeled as an iceberg transportation cost for simplicity. Important objects in this model are the sets of good available for consumption in the period, $\mathrm{I}_h^h(\mu)$ and $\mathrm{I}_f^h(\mu)$. $\mathrm{I}_h^h(\mu)$ is the set of goods produced in the home country that are available to consume in the home country. Since firms do not pay an entry cost to sell to the domestic market, this set is the entire set of domestically produced goods, regardless of the state of the economy. The set of imported varieties that are available, $\mathrm{I}_f^h(\mu)$, will generally be a strict subset of the varieties produced in the foreign country. This set consists only of the goods whose firms have paid the entry cost to set up exporting operations. This set of varieties varies with the state of the economy.

As is common, we can define an artificial composite good,

$$C_h = \left[\int_{\iota \in \mathrm{I}_h^h(\mu)} c_h^h(\iota)^{\rho} d\iota + \int_{\iota \in \mathrm{I}_f^h(\mu)} c_f^h(\iota)^{\rho} d\iota \right]^{\frac{1}{\rho}}, \tag{14}$$

whose price can be found by minimizing the cost of producing one unit of C_h,

$$P_h(\eta,\mu) = \min_{c_h^h(\iota), c_f^h(\iota)} \int_{\iota \in \mathrm{I}_h^h(\mu)} p_h^h(\iota;\eta,\mu) c_h^h(\iota) d\iota + \int_{\iota \in \mathrm{I}_f^h(\mu)} p_f^h(\iota;\eta,\mu)(1+\tau) c_f^h(\iota) d\iota, \tag{15}$$

subject to $C_h = 1$. Solving this minimization problem yields the familiar expressions for the price of the composite good,

$$P_h(\eta,\mu) = \left[\int_{i \in \mathrm{I}_h^h(\mu)} p_h^h(\iota;\eta,\mu)^{\frac{\rho}{\rho-1}} d\iota + (1+\tau)^{\frac{\rho}{\rho-1}} \int_{\iota \in \mathrm{I}_f^h(\mu)} p_f^h(\iota;\eta,\mu)^{\frac{\rho}{\rho-1}} d\iota \right]^{\frac{\rho-1}{\rho}}, \tag{16}$$

demand functions for each of the differentiated varieties produced in the home country,

$$\tilde{c}_h^h\left(p_h^h(\iota);\iota,\eta,\mu\right) = \left(\frac{P_h(\eta,\mu)}{p_h^h(\iota;\eta,\mu)} \right)^{\frac{1}{1-\rho}} C_h(\eta,\mu), \tag{17}$$

and demand functions for each of the differentiated varieties produced in the foreign country,

$$\tilde{c}_f^h\left(p_f^h(\iota);\iota,\eta,\mu\right) = \left(\frac{P_h(\eta,\mu)}{p_f^h(\iota;\eta,\mu)(1+\tau)} \right)^{\frac{1}{1-\rho}} C_h(\eta,\mu). \tag{18}$$

The stand-in household in each country inelastically supplies labor to firms and chooses consumption of the domestically produced varieties, the available imported varieties, and the non traded good to maximize utility. The household's value function can be written as

$$W\left(\eta,\mu\right)=\max\left(\gamma\log\left(C_h\right)+\left(1-\gamma\right)\log\left(q_h\right)+\beta\sum_{\eta'}W\left(\eta',\mu'\right)\lambda_{\eta\eta'}\right), \qquad (19)$$

subject to the budget constraint,

$$P_h\left(\eta,\mu\right)C_h+p_{hq}\left(\eta,\mu\right)q_h+\sum_{\eta'}Q\left(\eta'|\eta,\mu\right)B\left(\eta'|\eta\right)=\bar{L}_h+\Pi_h\left(\eta,\mu\right)+B\left(\eta\right) \qquad (20)$$

and the laws of motion over the distribution of firms, (12).

The price of the non-traded good in the home country is p_{hq}. Aggregate profits, Π_h, are returned to the household, and the home country wage is normalized to 1. As the households within a country are identical, there will be no borrowing and lending within the country. The existence of these securities, however, allows us to compute

$$Q\left(\eta'|\eta,\mu\right)=\beta\frac{P_h\left(\eta,\mu\right)C_h\left(\eta,\mu\right)}{P_h\left(\eta',\mu'\right)C_h\left(\eta',\mu'\right)}\lambda_{\eta\eta'} \qquad (21)$$

so that firms value future profits in a consistent way.

4.2. Differentiated Good Producers

There is a continuum of differentiated good firms indexed by their idiosyncratic productivities, ϕ, and export entry costs, κ. A firm's marginal cost of production consists of two parts: the idiosyncratic, non-stochastic productivity, ϕ, and the economy wide, stochastic productivity, $z_h\left(\eta\right)$. The production function for a firm of type $\left(\phi,\kappa\right)$ in aggregate state $\left(\eta,\mu\right)$ is linear,

$$y\left(\phi,\kappa,\eta,\mu\right)=z_h\left(\eta\right)\phi l\,. \qquad (22)$$

When a firm chooses to begin exporting, it must pay the entry cost, κ. This cost must be paid before the realization of η and can not be recovered afterward. After paying this entry cost, the firm faces no further costs associated with exporting.

An incumbent firm enters the period as either an exporter or a non-exporter. After aggregate productivity is revealed, firms choose how much labor to hire and how much to produce. After production, a mass of entrants, ν, arrives who have not paid the fixed cost to export. The joint distribution of idiosyncratic productivity and entry costs over these entrants has p.d.f. Φ. At the end of the period, non-exporters decide whether to continue as non-exporters, or to pay κ and begin exporting. In addition, firms face an exogenous probability of death, δ. Exporting decisions, along with the exogenous death of firms, determine the next period's distributions of exporters and non-exporters, μ_{hd}' and μ_{hx}'. The timing of decisions and the evolution of the distributions over firms are displayed in Figure 5.

The firm's problem can be broken up into two sub-problems. The first problem is a static maximization of period profits. The second is the dynamic decision of exporter status. We turn to the static problem first. Plants are monopolistic competitors who choose prices to maximize profits, taken as given the aggregate price index and the wage. The firm realizes, however, that the household's demand is downward sloping, and thus the demand functions defined in (11) and (12) appear in the firm's problem. For clarity in the firm's dynamic problem, it is useful to define the value of maximized profits from selling domestically for a firm of type (ϕ,κ) in aggregate state (η,μ) as

$$\begin{aligned} \pi_{hd}(\phi,\kappa,\eta,\mu) &= \max_{p_h^h,\, l} \tilde{c}_h^h\left(p_h^h;\phi,\kappa,\eta,\mu\right) p_h^h - l \\ \text{s.t.} \quad z_h(\eta)\phi l &= \tilde{c}_h^h\left(p_h^h;\phi,\kappa,\eta,\mu\right) \end{aligned} \tag{23}$$

and maximized profits from exporting as

$$\begin{aligned} \pi_{hx}(\phi,\kappa,\eta,\mu) &= \max_{p_h^f,\, l} \tilde{c}_h^f\left(p_h^f;\phi,\kappa,\eta,\mu\right) p_h^f - l \\ \text{s.t.} \quad z_h(\eta)\phi l &= \tilde{c}_h^f\left(p_h^f;\phi,\kappa,\eta,\mu\right). \end{aligned} \tag{24}$$

Optimization implies that firms set prices as a constant markup over marginal costs,

$$p_h^h(\phi,\kappa,\eta,\mu) = p_h^f(\phi,\kappa,\eta,\mu) = \frac{1}{\rho\phi z_h(\eta)}, \tag{25}$$

and determines the labor demand functions, $l_h^h(\phi,\kappa,\eta,\mu)$ and $l_h^f(\phi,\kappa,\eta,\mu)$. Having defined the maximized values from the static problem, the firm's dynamic problem is reduced to choosing only exporting decisions. As there are no export continuation costs, an exporter will always choose to stay an exporter; monopolistic competition ensures that the firm always earns a positive profit. An exporter's value function is defined by

$$\begin{aligned} V_{hx}(\phi,\kappa,\eta,\mu) &= \pi_{hd}(\phi,\kappa,\eta,\mu) + \pi_{hx}(\phi,\kappa,\eta,\mu) + (1-\delta)\sum_{\eta'} Q\left(\eta' \middle| \eta,\mu\right) V_{hx}(\phi,\kappa,\eta',\mu') \\ \text{s.t.} \quad \mu' &= \mathrm{M}(\eta,\mu) \end{aligned} \tag{26}$$

where $\mathrm{M}(\ )$ is the law of motion for the aggregate state variable μ. The term $(1-\delta)$ reflects the probability of exogenous death that the firm faces. The summation represents an expected value calculation where the transition probabilities are included in Q, as defined in (21).

A non-exporter must decide whether to remain selling only to the domestic market, or to enter the export market. The non-exporter's problem can be written as

$$V_{hd}(\phi,\kappa,\eta,\mu)=\max\left\{\begin{array}{l}\pi_d(\phi,\kappa,\eta,\mu)+(1-\delta)\sum_{\eta'}Q(\eta'|\eta,\mu)V_{hd}(\phi,\kappa,\eta',\mu'),\\ \pi_d(\phi,\kappa,\eta,\mu)-\kappa+(1-\delta)\sum_{\eta'}Q(\eta'|\eta,\mu)V_{hx}(\phi,\kappa,\eta',\mu')\end{array}\right\}. \quad (27)$$
$$\text{s.t.}\qquad \mu'=\mathrm{M}(\eta,\mu)$$

The two terms in the maximization correspond, respectively, to the expected future profit from continuing as a non-exporter and the expected future profit from becoming an exporter. This choice is the crucial one for the results presented here. The small, temporary productivity shocks change the future expected profits from exporting little, and thus few firms are willing to sink the cost of becoming exporters. A permanent change, such as a tariff decrease, increases a firm's profit in every realization of η, and thus has a larger effect on the expected future profits from exporting. This larger effect induces more firms to enter the export market, and increases the amount of goods being traded. Thus, in this model, permanent changes in tariffs have larger impacts on trade than do temporary shocks to productivity. Solving the non-exporter's problem yields the decision rule over next period's export status, $d_h(\phi,\kappa,\eta,\mu)$, which is equal to 1 if the firm chooses to export next period and is equal to 0 if the firm chooses to continue as a non-exporter in the next period.

4.3. Nontraded Good Producers

The non-traded good, q, is produced by a constant returns to scale firm and is sold in a competitive market. The firm's problem is

$$\begin{array}{l}\max_{l}\ p_{hq}(\eta,\mu)q_h-l\\ \text{s.t.}\ \ q_h=z_h(\eta)l\end{array}, \qquad (28)$$

where the price of good q_h is p_{hq}.

4.4. Market Clearing

Goods market clearing conditions are standard. The labor market clearing condition is

$$
\begin{aligned}
L_h = & \iint_{\kappa,\phi} l_h^h(\phi,\kappa,\eta,\mu)\,\mu_{hd}(\phi,\kappa)\,d\phi d\kappa \\
& + \iint_{\kappa,\phi} \left[l_h^h(\phi,\kappa,\eta,\mu) + l_h^f(\phi,\kappa,\eta,\mu) \right] \mu_{hx}(\phi,\kappa)\,d\phi d\kappa \\
& + \iint_{k,\phi} \kappa\, d_h(\phi,\kappa,\eta,\mu)\big(\mu_{hd}(\phi,\kappa) + \nu\Phi(\phi,\kappa)\big)\,d\phi d\kappa \quad .
\end{aligned} \tag{29}
$$

The first term on the right hand side is the use of labor for production by non-exporting firms, the second is the use of labor for production by exporting firms and the third term is the use of labor in exporter entry, some of which is paid by newly created firms. Aggregate profits are the sum of gross profits earned by firms minus the costs of entry,

$$
\begin{aligned}
\Pi_h(\eta,\mu) = & \iint_{\kappa,\phi} \pi_{hd}(\phi,\kappa,\eta,\mu)\,\mu_{hd}(\phi,\kappa)\,d\phi d\kappa \\
& + \iint_{\kappa,\phi} \left[\pi_{hd}(\phi,\kappa,\eta,\mu) + \pi_{hx}(\phi,\kappa,\eta,\mu) \right] \mu_{hx}(\phi,\kappa)\,d\phi d\kappa \\
& - \iint_{k,\phi} \kappa\, d_h(\phi,\kappa,\eta,\mu)\big(\mu_{hd}(\phi,\kappa) + \nu\Phi(\phi,\kappa)\big)\,d\phi d\kappa \quad .
\end{aligned} \tag{30}
$$

The model is closed by assuming that trade is balanced each period,

$$
\int_{\iota\in \mathrm{I}_h^f(\mu)} p_h^f(\iota;\eta,\mu)\,c_h^f(\iota;\eta,\mu)\,d\iota \;-\; \int_{\iota\in \mathrm{I}_f^h(\mu)} p_f^h(\iota;\eta,\mu)\,c_f^h(\iota;\eta,\mu)\,d\iota \;=0\,. \tag{31}
$$

4.5. Equilibrium

Equilibrium is defined recursively. For simplicity, only the home country equilibrium objects are enumerated; the agents in the foreign country solve problems analogous to those in the home country, and have the corresponding decision rules.

A *recursive equilibrium* in this economy is value functions, $V_{hx}(\phi,\kappa,\eta,\mu)$ and $V_{hd}(\phi,\kappa,\eta,\mu)$, decision rules, $d_h(\phi,\kappa,\eta,\mu)$, $l_h^h(\phi,\kappa,\eta,\mu)$, $l_h^f(\phi,\kappa,\eta,\mu)$, $p_h^h(\phi,\kappa,\eta,\mu)$, and $p_h^f(\phi,\kappa,\eta,\mu)$, the sets of goods available for consumption

$\mathrm{I}_h^h(\mu)$ and $\mathrm{I}_f^h(\mu)$, decision rules $c_h^h(\iota,\eta,\mu)$ for each $\iota \in \mathrm{I}_h^h(\mu)$, $c_h^f(\iota,\eta,\mu)$ for each $\iota \in \mathrm{I}_f^h(\mu)$, and $q_h(\eta,\mu)$, the price function, $p_{hq}(\eta,\mu)$, and the laws of motion for the distribution of firms, $\mathrm{M}_{hd}(\eta,\mu)$ and $\mathrm{M}_{hx}(\eta,\mu)$, such that these functions satisfy:

1. the household's problem, (19),
2. the firm's problems for each type $(\phi,\kappa) \in \mathcal{F} \times \mathcal{K}$, (26) and (27),
3. the consistency of aggregate and individual decisions,

   $\mathrm{M}_{hx}(\phi,\kappa) = \left[\mu_{hx}(\phi,\kappa,\eta,\mu) + \left(\nu\,\Phi(\phi,\kappa) + \mu_{hd}(\phi,\kappa,\eta,\mu)\right) d_h(\phi,\kappa,\eta,\mu)\right](1-\delta)$

   and $\mathrm{M}_{hd}(\phi,\kappa) = \left(1 - d_h(\phi,\kappa,\eta,\mu)\right)\left[\mu_{hd}(\phi,\kappa,\eta,\mu) + \nu\,\Phi(\phi,\kappa)\right](1-\delta)$,

   for all $(\phi,\kappa) \in \mathcal{F} \times \mathcal{K}$,
4. the goods and labor market clearing conditions,
5. the balanced trade condition, (31).

4.6. Equilibrium Properties

The equilibrium in this economy is characterized by a cut-off rule, $\hat{\phi}(\eta,\mu,\kappa)$, that is a function of the aggregate state and the value of the entry cost. This is the productivity level at which the non-exporting firm with idiosyncratic productivity $\hat{\phi}(\eta,\mu,\kappa)$ and entry cost κ is indifferent between continuing as a non-exporter, and entering the export market when the aggregate state is (η,μ). The cutoff productivity satisfies

$$\kappa = (1-\delta)\sum_{\eta'}\left(Q(\eta'|\eta,\mu)\left[V_{hx}\left(\hat{\phi}(\eta,\mu,\kappa),\kappa,\eta',\mu'\right) - V_{hd}\left(\hat{\phi}(\eta,\mu,\kappa),\kappa,\eta',\mu'\right)\right]\right). \qquad (32)$$

The right-hand-side of (32) is the discounted expected future gain from exporting: it is the difference in future profits if the firm exports rather than only selling domestically. The left-hand-side of (32) is the cost of entering the export market. Plants with productivity below $\hat{\phi}(\eta,\mu,\kappa)$ sell only to the domestic market, while firms with productivity above $\hat{\phi}(\eta,\mu,\kappa)$ sell to both the domestic and the exports markets. As the aggregate state of the economy changes, the expected future profits of the firms change; this shifts the cutoff firms, generating the entry and exit of firms in the export market. It is easy to show that, for a given (η, μ), $\hat{\phi}$ is increasing in κ. That is, the larger is the entry cost, the more productive a firm needs to be in order to break even in the export market.

Plants moving into and out of exporting play the crucial role in explaining the different responses of aggregate exports to changes in trade policy and productivity. If the shocks to productivity are small, or not very persistent, there will be little movement of firms into or out of exporting. Thus, productivity shocks induce incumbent exporting firms to change prices, and households react to the change in prices by substituting according to the elasticity of substitution implied by ρ. This is growth on the intensive margin. With only a few firms entering the export market, extensive margin growth is small. If transitory shocks are the dominant source of the variation measured in time series regressions such as Reinert and Roland-Holst (1992) or Blonigen and Wilson (1999), they would mostly affect the intensive margin, and the low elasticities estimated in these studies would reflect a low true elasticity of substitution.

A permanent change, such as the lower tariffs that accompany free trade agreements, increases the profit from exporting for all realizations of future productivity, and thus has a larger impact on the expected future profits from exporting. The larger increase in expected future profits induces some firms that were only selling in the domestic market to enter the export market. These newly traded goods create growth on the extensive margin. These extensive margin effects increase trade flows by more than that implied by the fall in tariffs and the elasticity of substitution. This mechanism contributes to the seemingly large estimated elasticities found when regressing trade volumes on tariff changes or transportation costs as in Clausing (2001) or Hummels (2001).

In Figure 6 we plot the components of (32): the expected gain from exporting for different productivity types and the value of a particular entry cost. The intersection of these two lines defines $\hat{\phi}$: the cutoff productivity for firms with a particular entry cost κ. All the firms with productivity less than $\hat{\phi}$ do not export, while firms with productivity greater than $\hat{\phi}$ do. To see how the model responds to temporary shocks compared to permanent changes, consider two scenarios. In the first scenario, the economy is subject to a positive productivity shock. These shocks are persistent, but not permanent. In response to the shock, the expected future value of exporting shifts up; for any level of idiosyncratic productivity, the gain from exporting is larger when productivity is high. This shift lowers the cutoff productivity to $\hat{\phi}_{bc}$ and the firms between $\hat{\phi}_{bc}$ and $\hat{\phi}$ enter the export market. A typical conditional distribution over firm productivity types is shown in Figure 7. As the cutoff productivity moves to the left, the amount of new firms entering can be inferred from the slope of the distribution. If the slope is steep, the mass of firms entering from even a small change in the cutoff productivity can be large. The slope of this distribution near the marginal exporter is a function of the model's parameters, which will be determined in the calibration.

Now consider subjecting the economy to a *permanent* decrease in tariffs of the same magnitude as the productivity shock. The permanent change shifts the expected future value of exporting up by much more than the persistent, but not permanent, shock. The cutoff firm is now $\hat{\phi}_{tar}$, and the firms between $\hat{\phi}_{tar}$ and $\hat{\phi}$ begin exporting. The larger shift in expected future profits from exporting from the tariff change leads to a greater number of firms beginning to export. In Figure 7 we can see how the large shift in the cutoff productivity leads to a greater number of new entrants than in the case of the temporary shock.

It is the difference in the number of new exporters under the two scenarios that changes the implied Armington elasticity. If no new firms enter the export market in response to the productivity shock, households only substitute between the goods already being imported and domestically produced goods at the rate σ. This is trade growth on the intensive margin. In this case, the measured Armington elasticity is exactly $1/(1-\rho)$. If firms enter the export market in response to a change in future profits, trade flows increase from the trade of the new exporters as well as the increase in trade of the continuing exporters. If newly traded goods are not accounted for, this extensive margin growth shows up in the trade aggregates as intensive margin growth, and the response of imports appears very large. The large response in trade to the change in tariffs results in a large estimate of the Armington elasticity.

5. Calibration and Computation

The model has the ability to generate nonlinear dynamics; the crux of this paper is to show that these nonlinearities are large in response to changes in tariffs, but small in response to changes in productivity. To evaluate the size of the nonlinearities in response to productivity changes, we require a nonlinear solution method, in particular we will discretize the state space and solve the value functions in (26) and (27) using iterative methods.

In order to solve the firms' problems in (16) and (17), the firms need to know the aggregate price index, P_h, and the consumption of the composite good, C_h. These values depend on the state variables for this economy, which include the distributions over firms, μ_{hd} and μ_{hx}, and their counterparts for the firms in the foreign country. Due to the high dimensionality of these objects, standard computational techniques are not applicable. Rather than take the distributions as state variables, we proceed as in Krusell and Smith (1998) by replacing the distribution with a finite dimensional vector of the distribution's moments, m. We parameterize forecasting functions that predict the future aggregate price level, aggregate profits, the foreign wage, and moments of the distributions as a function of the current period's variables. These functions are parameterized by a vector of coefficients, α.

For a given parameter vector α we solve the model and use the policy functions to generate simulated data, where we solve for the equilibrium prices in each period of the simulation. If the forecasting functions are consistent with the simulated data we have arrived at a solution; if not, the vector α is adjusted and the process is repeated until the forecasting functions are consistent with the simulated data that they generate. The large number of state variables makes this process computationally intensive and many of the simplifications that we have made in the model have been to reduce this burden, which is still considerable. Appendix A provides more details about the computation.

5.1. Calibration

The model is calibrated to match the United States and a symmetric partner country in 1987, which is before both the Canada-U.S. Free Trade Agreement and the NAFTA. We choose parameters such that the model's deterministic steady state displays the key aggregate and plant level[3] patterns from the data. We set the ad valorem tariff rate to be 5 percent, and choose ρ to be 0.50, which implies a value of 2.0 for the true elasticity of substitution. This choice of elasticity is in the range found in the time series estimations, and is frequently used in the international real business cycle literature.

The model period is one year, so setting β equal to 0.96 implies and annual real interest rate of about 4.00 percent. We consider the traded goods sector in the model to be manufacturing, even though there is trade in primaries and services. We exclude primaries and services since the plant level data needed to calibrate the model is only available for manufacturing plants. This definition of the tradable goods sector implies that γ, the share of expenditures on manufactures, is 0.17. The parameter governing the exogenous death of plants is set so that 2.35 percent of jobs are lost to exiting plants. This is the average found in Davis, Haltiwanger and Schuh (1996) for the years 1973-1998. Given the value for δ, the parameter ν—the mass of entrants each period—is chosen so that the mass of firms in operation is normalized to one. The parameters are summarized in Table 1.

The idiosyncratic productivity can take values in the interval $[1, \bar{\phi}]$ and the entry cost can take values in the interval $\left[0, \bar{\kappa}\right]$. We discretize the interval over productivities into 200 different types, and the interval over entry costs into 200 different types, for a total of 40,000 firm types. The shape of the distribution over entrants, Φ, is chosen so that the shape of the equilibrium distribution over firm sizes matches that for U.S. manufacturing plants. This requires a joint distribution over $\left(\phi, \kappa\right)$. The unconditional probability density function over idiosyncratic productivity is

$$f\left(\phi\right) = \phi^{-\theta_\phi - 1}\theta_\phi. \tag{33}$$

In order to reproduce the size distribution of exporting plants, the size of a firm's export entry cost needs to be correlated with the firm's productivity. Without correlated entry costs, the economy would not have enough small exporting firms. The prevalence of

[3] The data in the Census of Manufacturing is collected at the establishment, or plant, level. While the model presented here has only one-plant firms, this is not true in the data.

small exporters—measured either by employment or sales—is explored in more detail in Arkolakis (2007). The distribution of export entry costs, conditional on ϕ, has the form

$$g\left(\kappa \middle| \phi\right) = \bar{\kappa}^{\theta_\kappa(\phi)} \kappa^{-\theta_\kappa(\phi)-1} \tag{34}$$

where we allow for correlation by specifying $\theta_\kappa\left(\phi\right) = \zeta + \xi\phi^2$. The joint probability distribution is built up from f and g in the usual way.

The remaining parameters, θ_ϕ, ζ, ξ, $\bar{\phi}$, and $\bar{\kappa}$, jointly determine the exporting and production structure of the traded goods sector. We choose these parameters to match the trade-output ratio in manufacturing, the employment size distribution of all plants in the economy, and the employment size distribution of exporting plants. To do so, we match the share of firms in seven (six independent observations) employee-size bins from the U.S. census of manufacturing. We match the share of firms in three (two independent observations) employee-size bins for exporting plants from the census of manufacturers. Fitting the two plant size distributions using so few parameters is difficult. Figure 8 and Figure 9 display the relative success of the calibration. Both figures indicate that the model fits the data fairly well considering the small number of free parameters. More details about the model's fit can be found in Table 2.

The final parameters to specify are the ones governing the exogenous productivity shocks. Aggregate productivity in each country is either $1-\varepsilon$, which we denote the low shock, or $1+\varepsilon$, which we denote the high shock. Each country has an identical symmetric transition matrix over its aggregate productivity,

$$\Lambda_i = \begin{bmatrix} \bar{\lambda} & 1-\bar{\lambda} \\ 1-\bar{\lambda} & \bar{\lambda} \end{bmatrix} \tag{35}$$

which requires only one parameter, $\bar{\lambda}$. Assuming that the shocks to productivity in the two countries are independent, the process requires calibrating only these two parameters. We choose $\bar{\lambda}$ and ε so that the volatility and persistence of the logged output in the model match those in the U.S. data. The persistence of logged output is measured as the one period autocorrelation, and we impose that one country's productivity does not have any "spillover" effects on the other country's productivity. We choose $\bar{\lambda}$ and ε so that the one period autocorrelation of output is 0.79, and the standard deviation is 0.036, the values for logged, linearly detrended, annual U.S. data from 1950-2000.

6. Model Results

In Section 4 we saw that qualitatively the model has different implications for temporary and permanent changes to exporting profitability, and that these different responses could lead to different estimates of the Armington elasticity. We now use the calibrated model to see if the exporting entry costs can quantitatively account for the different estimates of the Armington elasticity found in the literature.

6.1. Business Cycle Estimates

To see if the model can account for the low estimates derived from the high-frequency data, we use the model to generate simulated time series data on the prices and quantities of domestic goods and imports. To accurately test the model we must construct measures from the simulated data that are consistent with the methods used in the empirical works cited above. To do so, we construct a Laspeyres price index for imports, as in Reinert and Roland-Holst (1992), which weights period t prices by the quantities imported in a chosen base period,

$$\hat{P}_{f,t}^{h} = \frac{\sum_{i \in \mathrm{I}_{f,0}^{h}} p_{f,t}^{h}(i) c_{f,0}^{h}(i)}{\sum_{i \in \mathrm{I}_{f,0}^{h}} p_{f,0}^{h}(i) c_{f,0}^{h}(i)}. \tag{36}$$

where $\mathrm{I}_{f,0}^{h}$ is the set of goods traded in the base period, $t = 0$.[4] A similar Laspeyres index for domestic consumption is constructed, and these indices are used to deflate expenditures on imports and domestically produced goods. The data are used to estimate the Armington elasticity using a specification similar to that found in the empirical literature,

$$\Delta_{t-1,t} \log\left(C_{f,t}^{h} / C_{h,t}^{h}\right) = \alpha + \sigma \Delta_{t-1,t} \log\left(\hat{P}_{h,t}^{h} / \hat{P}_{f,t}^{h}\right) + \varepsilon_t. \tag{37}$$

---

[4] In both the computational model and the data, there are only a finite number of goods being traded, so the formulas expressed in this section are also written in terms of a finite number of goods. The counterparts to these formulas for the model presented in Section 4, which features a continuum of goods, are the same except the summations are replaced by integrals.

We simulate 250 economies, each for a 300 periods. After discarding the first 50 periods, we construct price indices as specified above, choosing period 51 to be the base year. We fit the equation in (37) to the data from each simulation. The mean value of the elasticity, σ, is 1.21 and the mean value of the constant, α, is 1.4×10^{-4}. A histogram of the elasticities from the simulations is plotted in Figure 10. The parameters are precisely estimated, and the average r-squared for the regressions is 0.56. This estimate fits well into the range of values estimated from time series data. As surveyed above, the time series estimates range from 0.2 to 3.5, with an average of about 1.0, a little lower than the elasticity estimated from the model. Our estimate is also in line with the values used in the international business cycle literature; these studies commonly use Armington elasticities between 1.0 and 2.0.

In interpreting the results, it is important to notice that the estimated elasticity is about 40 percent smaller than the true elasticity. In equilibrium, the household's first order conditions must hold with equality, which implies that the regression in (37) should fit exactly, and the elasticity of substitution should be measured as 2.0. The equation estimated with the Laspeyres price indices, however, does not fit exactly because the price index is not accounting for the changing set of goods being traded. Not taking into account the changing set of goods results in significant measurement error and a lower estimated value for the Armington elasticity.

Decomposing the price index exposes the source of the measurement error. Using the expression for prices in (18), the price index in (26) can be written as

$$\hat{P}_{f,t}^{h} = \frac{w_{f,t}}{z_{f,t}} \frac{\sum_{i\in \mathrm{I}_{f,0}^{h}} \left(\phi_i \rho\right)^{-1} c_{f,0}^{h}\left(i\right)}{\sum_{i\in \mathrm{I}_{f,0}^{h}} p_{f,0}^{h} c_{f,0}^{h}\left(i\right)}. \tag{38}$$

This splits the index into two parts. The first term on the right-hand side is the ratio of the wage and the aggregate productivity. This term is common to all firms, so we denote it the *aggregate component*. The second term depends on the set of goods being imported, so we denote it the *composition component.* Note that the set of goods over which the index is computed is held fixed through time. As a practical matter, goods that become traded subsequent to the base year must be ignored: the quantities imported of "newly traded" goods are by definition zero in the base year[5]. This formulation implies that any change in the composition of goods being imported will not be reflected in the price index.

In contrast to the Laspeyres index, the constant elasticity of substitution (CES) price index makes adjustments for the set of goods being traded. This index is exactly the cost of purchasing one unit of the composite import good,

$$P^h_{f,t} = \frac{\left[\sum_{i \in I^h_{f,t}} \left(p^h_{f,t}(i)\right)^{\frac{\rho}{\rho-1}}\right]^{\frac{\rho-1}{\rho}}}{\left[\sum_{i \in I^h_{f,0}} \left(p^h_{f,0}(i)\right)^{\frac{\rho}{\rho-1}}\right]^{\frac{\rho-1}{\rho}}}. \tag{39}$$

The summation in this index runs over all the goods being imported, incorporating the changing composition of goods. The domestic price index can similarly be defined. If we were to run the regression specified in (27) using the CES indices that account for the changing composition of imports, we would find the estimated elasticity is exactly 2.0.

As with the Laspeyres index, the CES price index can be split into an aggregate component and a composition component. The aggregate component is identical across the two indices. In contrast to the Laspeyres index, the composition component in the CES index changes as the set of goods being imported changes. As more goods are imported, this term decreases, lowering the price index. When goods exit the import market, the price index increases. The divergence of the composition component in these two indices is the source of the measurement error. This kind of measurement error has been noted by Helkie and Hooper (1988), Krugman (1989), Feenstra (1994) and Feenstra and Shiells (1997) in the context of income elasticity estimates. They argued that the growing number of goods being imported by the U.S. is not reflected in the import price index, so the new imports lead to an increase in the value of imports but not in the price index. The rising value of imports is then attributed to the growth in U.S. income, implying a large value for the income elasticity of imports. Broda and Weinstein (2006) extended the methodology in Feenstra (1994) to create a price index for U.S. imports that

[5] If the price index is "rebased," the set of goods over which the index is defined would change to reflect the mix of goods available in the new base period. If the new price index was ratio-spliced onto the old index, this would not change our findings, as the rebasing would only have an affect on the level of the index and not the volatility.

corrects for the appearance of new varieties. They calculate a 28 percent upward bias in the traditionally measured price of imports. In our model, goods both enter and exit the import market, and both situations lead to bias in the measurement of prices.

The difference between the two indices can be seen in Figure 11, where we plot the price indices for a particular simulation. The Laspeyres index, whose composition component is constant, tracks w_f / z_f. For most periods, the CES index moves similarly to the Laspeyres index, reflecting the small number of foreign firms entering and exiting the export market. The difference between the two is most apparent when both the home country and the foreign country have low productivity. Foreign firm export profits are lowest in this state: low foreign country productivity makes production more costly, and low home country productivity decreases home country income. When the marginal exporting firms die off, they are replaced by firms that do not enter the export market. When both countries have low aggregate productivity, the Laspeyres index increases, but the CES index increase more, as it factors in the decreasing number of imported varieties. The correlation coefficient between these two indices is 0.53. The imperfect correlation between the two price indices drives the difference in the estimated elasticities.

6.2. Trade Liberalization

To see if the model can generate the large response to changes in tariffs, we consider the complete removal of the 5 percent tariff in the baseline model. To capture this idea, we compute the equilibrium under the 5 percent tariff and then compute the equilibrium in the absence of tariffs. By proceeding in this way, we avoid having to compute the transition path from the high tariff equilibrium to the low tariff equilibrium. The transition path between tariff policies is certainly interesting from the point of view of welfare analysis, but is very difficult to compute in this model[6]. The results of the decrease in tariffs are dramatic. As can be seen in the first column of Table 3, the decline in tariffs leads to 23.39 percent increase in exports. The ratio of imports to domestic consumption grows by 30.01 percent as a result of the tariff decrease. Measuring the Armington elasticity with respect to the 5 percentage point tariff change yields an

[6] For an analysis of transition paths without aggregate uncertainty, but with idiosyncratic uncertainty, see Alessandria and Choi (2007b).

elasticity of 6.38, which is more than 5 times the elasticity estimated from the time series data. This value is in the middle of the range of values reported in Hummels (2001), whose estimates range from about 3 to 8, with an average of 5.6. Our estimate also falls into the range estimated in Romalis (2007), which finds elasticities typically between 4 and 13. The estimates in Clausing (2001) and Head and Ries (2001), who study the Canada-U.S. Free Trade Agreement, range from 8 to 11. Though the value of 6.38 found here is lower than their findings, our simple model has abstracted from other possibly important elements such as returns to scale, intermediate good trade, or capital accumulation.

The five-fold difference between the time series estimate of the Armington elasticity and the elasticity implied by trade liberalization is driven by the growth in trade on the extensive margin. Table 3 displays the results of trade liberalization in two different models: the baseline model specified in Sections 4 and 5, and a model identical to the baseline model except that the entry costs have been set to zero for all firms. The model without entry costs has the obvious counterfactual implication that all firms export. Given this characteristic of the model, modify the household's preferences to include a home bias parameter, ω, as in (4). We choose ω so that the country still only exports 9.1 percent of output when tariffs are 5 percent. The rest of the model is calibrated as in Section 5, although we no longer have to choose the maximum value of the entry cost or the shape parameters for the distribution over entry costs. The effects of trade liberalization in the two models are drastically different. Eliminating a 5 percent tariff in the baseline model increases exports by 23.39 percent, while exports increase by only 7.32 percent in the model without entry costs. The extra 16.07 percentage points of growth in exports are due to the new exporters that enter the market following the decrease in tariffs. The model without entry costs has no extensive margin; all the firms are already exporting. The increase in trade is driven by the lower delivered price that results from the elimination of the tariff. The baseline model, however, has substantial extensive margin growth; 2.23 percent more firms export following the elimination of the tariff. The extra trade generated by the new exporters, along with the intensive margin trade, drive up the ratio of imports to domestic consumption by 30.42 percent; a value 3.17 times larger than the change in the model without entry costs. This increase implies an Armington elasticity of 6.38 compared to the measured elasticity of 2.01 in the model without entry costs.

7. Conclusion

Models in which countries trade differentiated goods depend crucially on the Armington elasticity to determine both the short run and the long run behavior of trade flows and prices. Models that were built to explain high frequency fluctuations, such as those in the IRBC literature, need small values of the Armington elasticity, while applied GE models need large elasticities to match the increase in trade following trade liberalization. Econometric estimates support this dichotomy: estimates derived from high-frequency time series data are low, and range between 0.2 and 3.5, while estimates gleaned from trade liberalizations range from 4 to 15. This paper reconciles these two observations by recognizing that the sources of variation in the two approaches are different. The high-frequency variation in the time series studies is caused by small and persistent, but not permanent, shocks to productivity or demand. Trade liberalization, however, is typically a permanent change. When agents react differently to temporary and permanent changes, the measured elasticities will differ.

We build a model that incorporates elements from the international real business cycle literature and the applied GE models that are commonly used to study trade policy. The key feature of the model is that firms face a cost of entering the export market, as well as uncertainty about future profits from exporting. Temporary shocks will induce few firms to change export status and incumbent exporters respond to the shocks by changing prices. Measuring the elasticity with respect to these changes will recover the low elasticity implied by the model's parameters. A permanent tariff change increases the future value of exporting in all states, and induces a larger number of firms to begin exporting. When the trade from these newly imported goods is not properly accounted for, the response of trade to small changes in tariffs looks large, implying a high Armington elasticity. The calibrated model, with a "true" elasticity of substitution of 2.0, is capable of producing time series estimates of the Armington elasticity of about 1.21, while the elasticity in response to a reduction in tariffs is 6.38. The initial success of this model adds to the growing evidence that studying firm level choices can lead to a greater understanding of international trade.

References

Alessandria, G. and H. Choi (2007a), "Do Sunk Costs of Exporting Matter for Net Exporter Dynamics?" *Quarterly Journal of Economics*, 122(1), 289-336.

Alessandria, G. and H. Choi (2007b), "Establishment Heterogeneity, Exporter Dynamics, and the Effects of Trade Liberalization," *Federal Reserve Bank of Philadelphia Working Paper* 07-29.

Anderson, M. A., Balistreri, E. J., Fox, A. K., and R. H. Hillberry (2005), "Taste Parameters as Model Residuals: Assessing the 'Fit' of an Armington Trade Model," *Review of International Economics*, 13(5), 973-984.

Arkolakis, K. (2007), "Market Penetration Costs and the New Consumers Margin in International Trade," Yale University.

Armington, P. S. (1969), "A Theory of Demand for Products Distinguished by Place of Production," *International Monetary Fund Staff Papers*, 16(1), 159-178.

Backus, D. K., Kehoe, P. J., and F. E. Kydland (1994), "Dynamics of the Trade Balance and the Terms of Trade: The J-Curve?" *American Economic Review*, 84(1), 84-103.

Baier, S. L. and J. H. Bergstrand (2001), "The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity," *Journal of International Economics*, 53(1), 1-27.

Baldwin, R. E. and P. R. Krugman (1989), "Persistent Trade Effects of Large Exchange Rate Shocks," *Quarterly Journal of Economics*, 104(4), 635-654.

Baldwin, R. E. (1988), "Hysteresis in Import Prices: The Beachhead Effect," *American Economic Review*, 78(4), 773-785.

Bernard, A. B., Eaton, J., Jensen, J. B. and S. Kortum (2003), "Plants and Productivity in International Trade," *American Economic Review*, 93(4), 1268-1290.

Bernard, A. B. and J. B. Jensen (2004), "Why Some Firms Export," *Review of Economics and Statistics*, 86(2), 561-569.

Bernard, A. B., Jensen, J. B., and P. K. Schott (2003), "Falling Trade Costs, Heterogeneous Firms, and Industry Dynamics," *NBER Working Paper* 9639.

Bernard, A. B. and J. Wagner (2001), "Export Entry and Exit by German Firms," *Weltwirtschaftliches Archiv/Review of World Economics*, 137(1), 105-123.

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

Broda, C. and D. E. Weinstein (2006), "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121(2), 541-585.

Clausing, K. A. (2001), "Trade Creation and Trade Diversion in the Canada-United States Free Trade Agreement," *Canadian Journal of Economics*, 34(3), 677-696.

Davis, S. J., Haltiwanger, J. C., and S. Schuh (1996), *Job Creation and Destruction*. Cambridge, Mass., MIT Press.

Dixit, A. K. (1989), "Hysteresis, Import Penetration, and Exchange Rate Pass-Through," *Quarterly Journal of Economics*, 104(2), 205-228.

Dixit, A. K. and J. E. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3), 297-308.

Dornbusch, R., Fischer, S. and P. A. Samuelson, (1977), "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods," *American Economic Review*, 67(5), 823-839.

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

Feenstra, R. C. and C. R. Shiells (1997), "Bias in U.S. Import Prices and Demand," in T. F. Bresnahan and R. J. Gordon, eds., *The Economics of New Goods*, National Bureau of Economic Research Studies in Income and Wealth, vol. 58. Chicago and London: University of Chicago Press, 249-273.

Gallaway, M. P., McDaniel, C. A., and S. A. Rivera (2003), "Short-Run and Long-Run Industry-Level Estimates of U.S. Armington Elasticities," *North American Journal of Economics and Finance*, 14(1), 49-68.

Ghironi, F. and M. J. Melitz (2005), "International Trade and Macroeconomic Dynamics with Heterogeneous Firms," *Quarterly Journal of Economics*, 120(3) 865-915.

Head, K. and J. 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.

Heathcote, J. H. and F. Perri (2002), "Financial Autarky and International Business Cycles," *Journal of Monetary Economics*, 49(3), 601-627.

Helkie, W. L. and P. Hooper (1988), "An Empirical Analysis of the External Deficit, 1980-86," in R. C. e. Bryant, G. e. Holtham and P. e. Hooper, eds., *External Deficits and the Dollar: The Pit and the Pendulum*, Washington, D.C.: Brookings Institution, 10-56.

Helpman, E. and P. R. Krugman (1985), *Market Structure and Foreign Trade : Increasing Returns, Imperfect Competition, and the International Economy*. Cambridge, Mass., MIT Press.

Hillberry, R. H. and C. A. McDaniel (2002), "A Decomposition of North American Trade Growth since NAFTA," *International Economic Review*, 43(2), 1-6.

Hummels, D. (2001), "Toward a Geography of Trade Costs," Purdue University.

Hummels, D. and P. J. Klenow (2005), "The Variety and Quality of a Nation's Exports," *American Economic Review*, 95(3), 704-723.

Kehoe, T. J. and K. J. Ruhl (2002), "How Important is the New Goods Margin in International Trade?" *Federal Reserve Bank of Minneapolis Staff Report 324*.

Krugman, P. R. (1980), "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70(5), 950-959.

Krugman, P. R. (1989), "Differences in Income Elasticities and Trends in Real Exchange Rates," *European Economic Review*, 33(5), 1031-1046.

Krusell, P. and A. A. Smith, Jr. (1998), "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867-896.

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

Reinert, K. A. and D. W. Roland-Holst (1992), "Armington Elasticities for United States Manufacturing Sectors," *Journal of Policy Modeling*, 14(5), 631-639.

Reinert, K. A. and C. R. Shiells (1993), "Armington Models and Terms-of-Trade Effects: Some Econometric Evidence for North America," *Canadian Journal of Economics*, 26(2), 299-316.

Roberts, M. J. and J. R. Tybout (1997), "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, 87(4), 545-564.

Romalis, J. (2007), "NAFTA's and CUSFTA's Impact on International Trade," *Review of Economics and Statistics*, forthcoming.

Stockman, A. C. and L. L. Tesar (1995), "Tastes and Technology in a Two-Country Model of the Business Cycle: Explaining International Comovements," *American Economic Review*, 85(1), 168-185.

Yeaple, S. R. (2005), "A Simple Model of Firm Heterogeneity, International Trade and Wages," *Journal of International Economics*, 65(1), 1-20.

Yi, K.-M. (2003), "Can Vertical Specialization Explain the Growth of World Trade?" *Journal of Political Economy*, 111(1), 52-102.

Zimmermann, C. (1997), "International Real Business Cycles among Heterogeneous Countries," *European Economic Review*, 41(2), 319-356.

Appendix 1
Computational Algorithm

The state space of the model is made up of two continuous variables, ϕ and κ, two distributions, μ_{hx} and μ_{fx}, and one discrete variable, η. The two continuous variables are discretized as $\{\phi_i\}_{i=1}^{I}$ and $\{\kappa_j\}_{j=1}^{J}$ and the distribution over these variables is $f(\phi_i, \kappa_j)$.

The model is difficult to compute because the state variables μ_{hx} and μ_{fx}—the distributions over exporting firms—are infinite dimensional objects. To deal with this problem, we solve the model using the technique described in Krusell and Smith (1998), which replaces the distribution with a finite dimensional vector of the distribution's moments. The model's structure implies that the relevant moment of the distribution is

$$m = \sum_{j=1}^{J} \sum_{i=1}^{I} \phi_i^{\rho/(1-\rho)} d\left(\phi_i, \kappa_j, m, m^*, \eta_n\right) f\left(\phi_i, \kappa_j\right), \tag{40}$$

where the function d is equal to one if a firm of type (ϕ_i, κ_j) in aggregate state (m, m^*, η_n) is an exporter and equal to zero otherwise.

We need to forecast 4 objects: aggregate profits, the aggregate price index, the wage in the foreign country and the moment m:

$$\log\left(\Pi\left(m, m^*, \eta_n\right)\right) = \alpha_{1n}^{\Pi} + \alpha_{2n}^{\Pi} \log(m) + \alpha_{3n}^{\Pi} \log\left(m^*\right) \qquad n = 1, \ldots, 4$$

$$\log\left(P\left(m, m^*, \eta_n\right)\right) = \alpha_{1n}^{P} + \alpha_{2n}^{P} \log(m) + \alpha_{3n}^{P} \log\left(m^*\right) \qquad n = 1, \ldots, 4$$

$$\log\left(w^*\left(m, m^*, \eta_n\right)\right) = \alpha_{1n}^{w} + \alpha_{2n}^{w} \log(m) + \alpha_{3n}^{w} \log\left(m^*\right) \qquad n = 1, \ldots, 4$$

$$\log\left(m'\left(m, m^*, \eta_n\right)\right) = \alpha_{1n}^{P} + \alpha_{2n}^{P} \log(m) + \alpha_{3n}^{P} \log\left(m^*\right). \qquad n = 1, \ldots, 4$$

To characterize the forecasting function of a variable, we need to specify 3 parameters per value of the exogenous state variable η. Denote the vector containing all 48 of the parameters α.

The algorithm proceeds iteratively; let s denote the iteration. We begin by choosing initial values for α, and a set of grid points for m, ϕ, and κ. The algorithm for solving the model is:

1. Given α^s and the functional forms of the forecasting functions, solve for the exporter and non-exporter value functions and the export decision rules using value function iteration. When the forecasting function for m' yields a value that is not a grid point, use bilinear interpolation in the m and m^* dimensions.
2. Simulate the economy for 3000 periods using the export decision rules found in step 1, but computing *equilibrium prices* in each period. This produces a series of data: $\left\{\Pi_{ht}, P_{ht}, w_{ft}, m_{ht}\right\}_{t=0,\ldots,3000}$.
3. To ensure no dependence on initial conditions, discard the first 500 observations. Using the simulated data, find new values of α, α^{s+1}, by ordinary least squares.
4. If the maximum change between the elements of α^s and α^{s+1} is less than 10^{-5} stop. If not, repeat steps 2 through 4 until α converges.

Forecasting Function Values

Coefficient	P	Π	m'	w_f
α_1, $\eta=(H,H)$	-1.446	-2.314	-1.513	0.000
α_2, $\eta=(H,H)$	-0.006	0.037	0.031	-0.330
α_3, $\eta=(H,H)$	-0.011	-0.008	0.184	0.330
α_1, $\eta=(H,L)$	-1.445	-2.313	-1.562	-0.032
α_2, $\eta=(H,L)$	-0.006	0.036	0.049	-0.330
α_3, $\eta=(H,L)$	-0.011	-0.006	0.140	0.331
α_1, $\eta=(L,H)$	-1.396	-2.369	-0.228	0.032
α_2, $\eta=(L,H)$	-0.006	0.001	0.883	-0.331
α_3, $\eta=(L,H)$	-0.012	0.000	0.004	0.330
α_1, $\eta=(L,L)$	-1.395	-2.370	-0.190	0.000
α_2, $\eta=(L,L)$	-0.006	0.000	0.891	-0.331
α_3, $\eta=(L,L)$	-0.011	-0.001	0.016	0.331
SST	1.24E-4	3.61E-7	0.381	0.128
R^2	0.999	0.819	0.999	0.999

Table 1

Calibration: $\tau = 0.05$, $\rho = 0.50$

Parameter	Value	Data
β	0.961	Annual real interest rate (4.0%)
δ	0.094	Share of employment lost to plant closure (0.094)
γ	0.171	Share of manufacturing in total production (0.171)
ε	0.025	Standard deviation of log-output (0.036)
λ	0.803	Autocorrelation of log-output (0.793)

Table 2

Jointly Determined Parameters

Parameter	All Establishments		
	Employees	Share of Establishments	
		Data	Model
	1-20	0.657	0.645
	21-50	0.162	0.172
$\theta = 0.672$	51-100	0.080	0.078
$\bar{\phi} = 134.4$	101-250	0.085	0.062
	251-500	0.022	0.029
	501-1000	0.009	0.011
	>1001	0.005	0.003
	SSE		0.001
	Exporting Establishments		
	Employees	Share of Establishments	
		Data	Model
	1-100	0.771	0.771
$\zeta = -0.190$	101-500	0.182	0.182
$\xi = 0.0009$	>500	0.047	0.047
$\bar{\kappa} = 10.00$	SSE		0.000
	Trade-Output Ratio	0.091	0.091

Table 3
Response to Permanent Removal of 5% Tariff

Variable	**Entry Costs Model** (% change)	**No Entry Costs Model** (% change)
Exports	23.39	7.32
Imports/Dom. Consumption	30.42	9.57
Exporting Plants	2.23	0.00
Implied σ	6.38	2.01

Figure 1

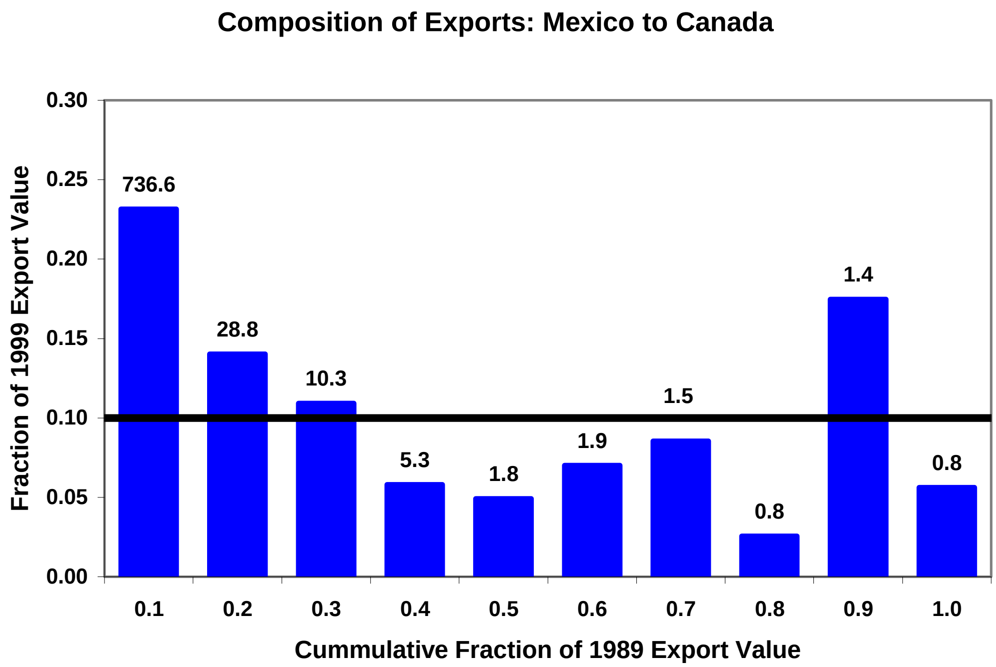


Figure 2

Figure 3

Composition of Exports: U.S. to U.K.

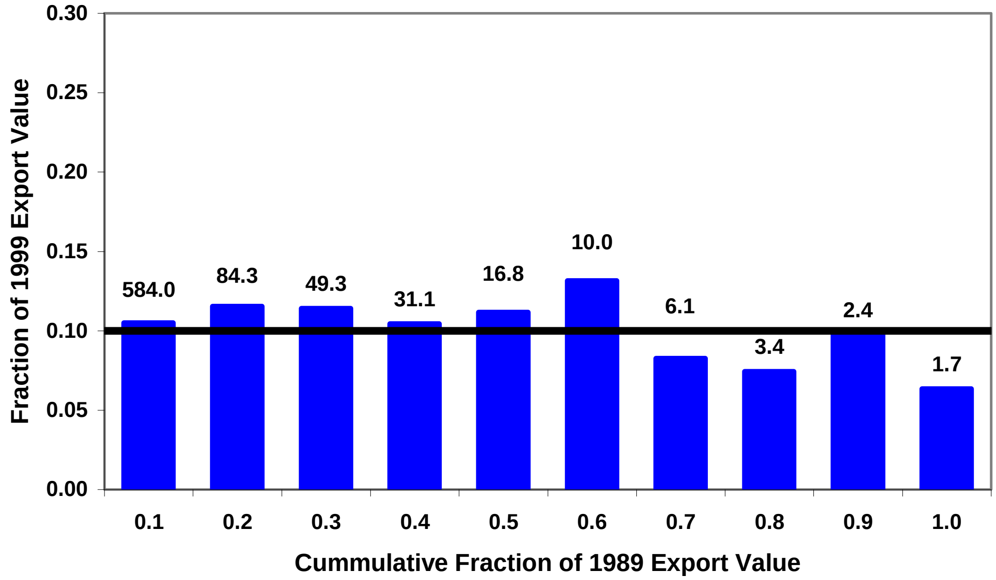


Figure 4

Least Traded Goods: U.S. to U.K.

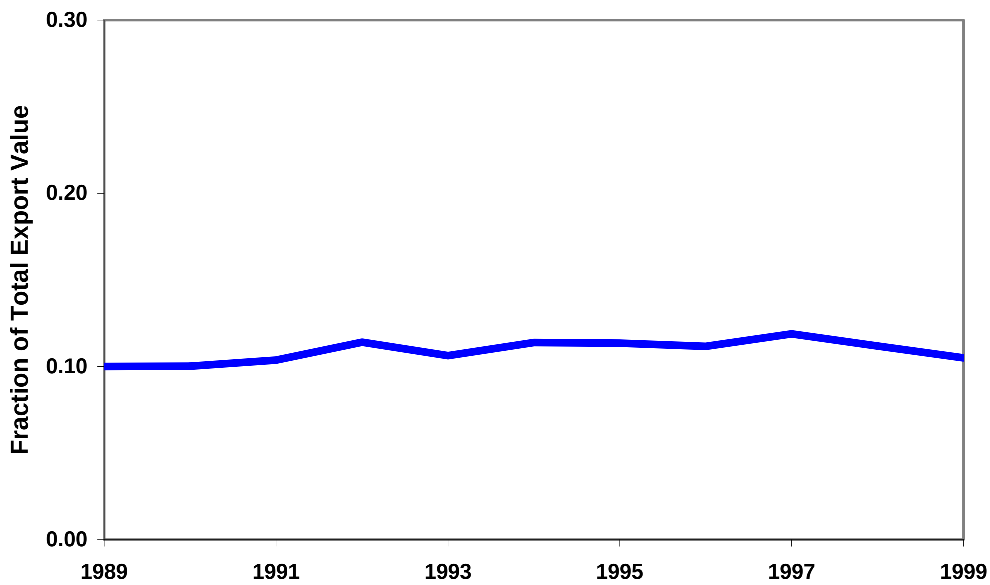
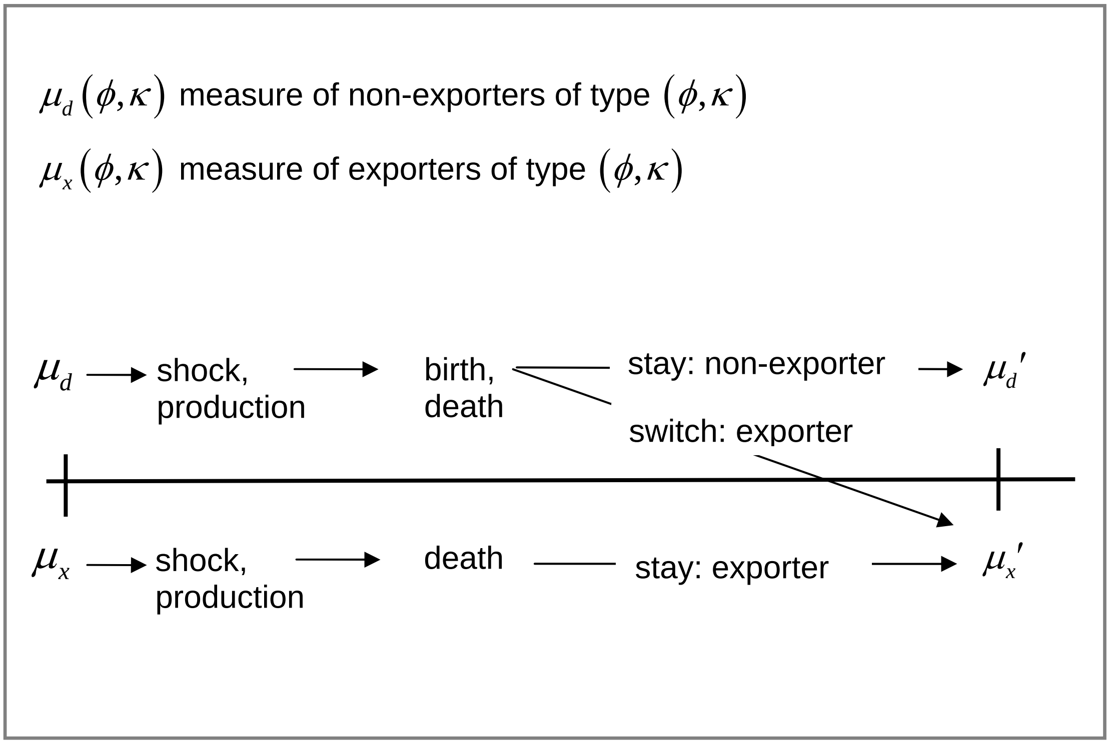


Figure 5



$\mu_d(\phi,\kappa)$ measure of non-exporters of type (ϕ,κ)
$\mu_x(\phi,\kappa)$ measure of exporters of type (ϕ,κ)
μ_d
shock, production
birth, death
stay: non-exporter
switch: exporter
μ_d'
μ_x
shock, production
death
stay: exporter
μ_x'

Figure 6

Persistent v. Temporary Shocks

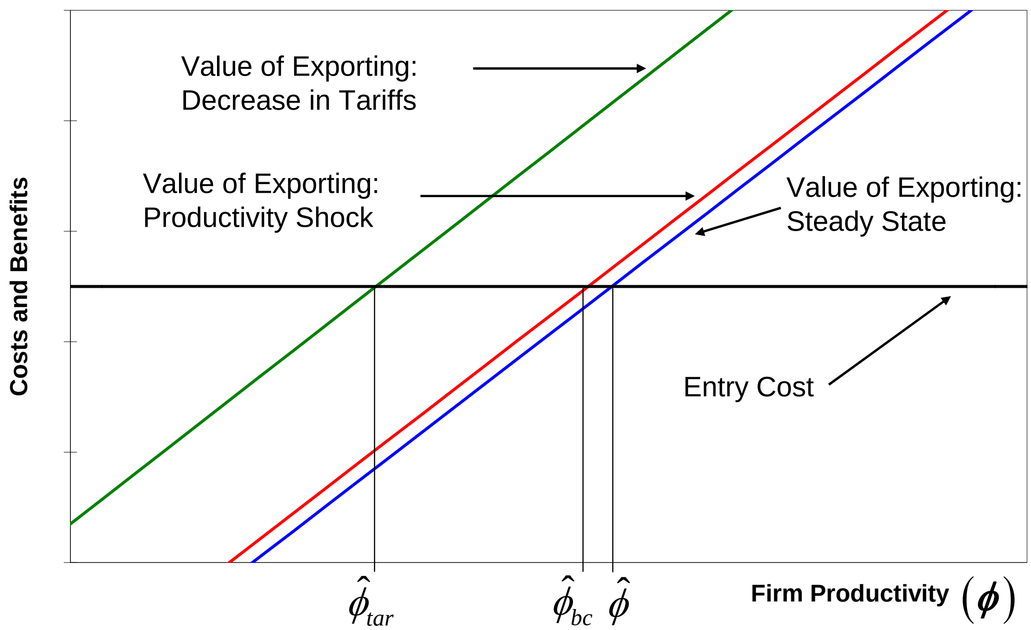


Figure 7

Productivity Distribution

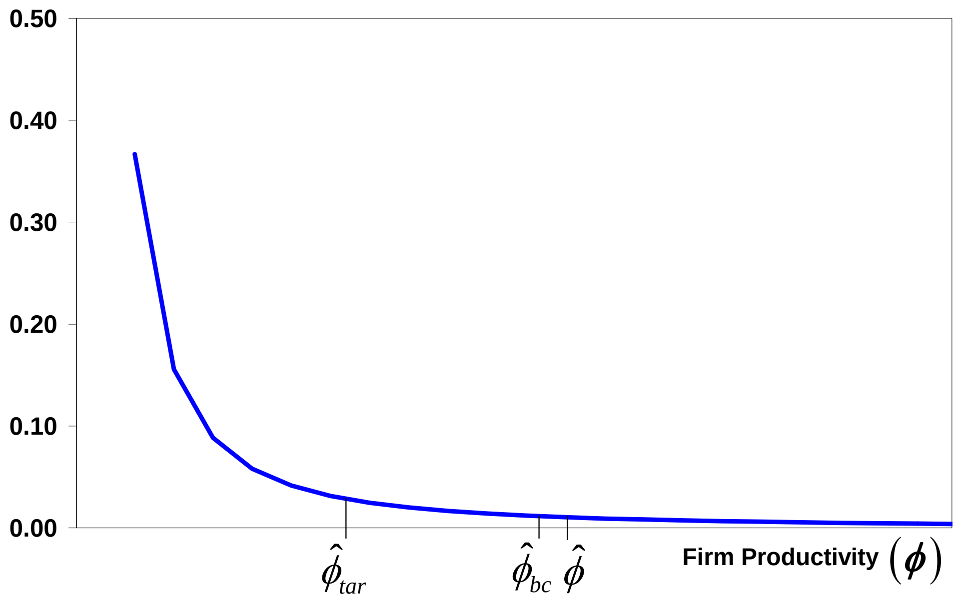


Figure 8

Establishment Size Distribution, All Establishments

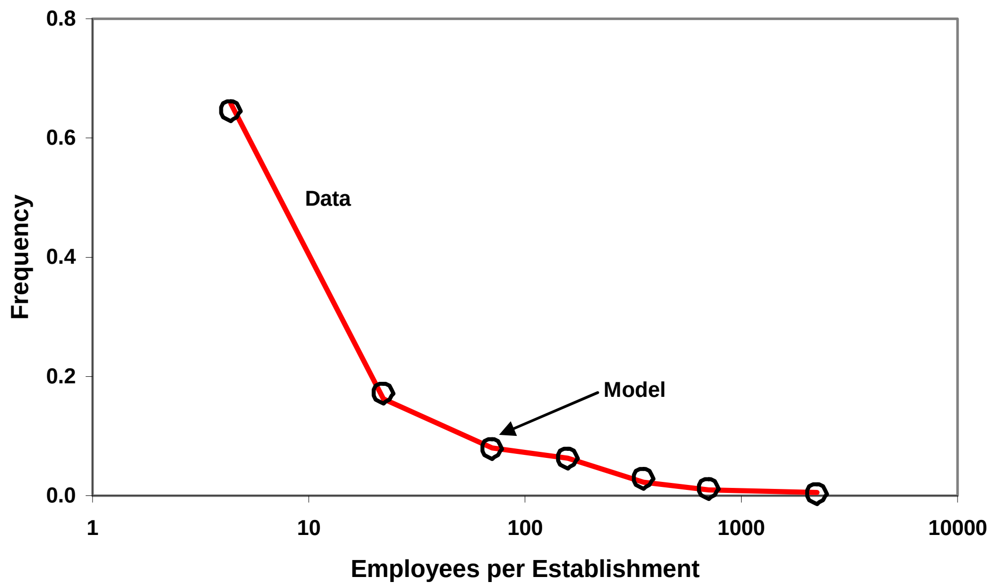


Figure 9

Establishment Size Distribution, Exporters

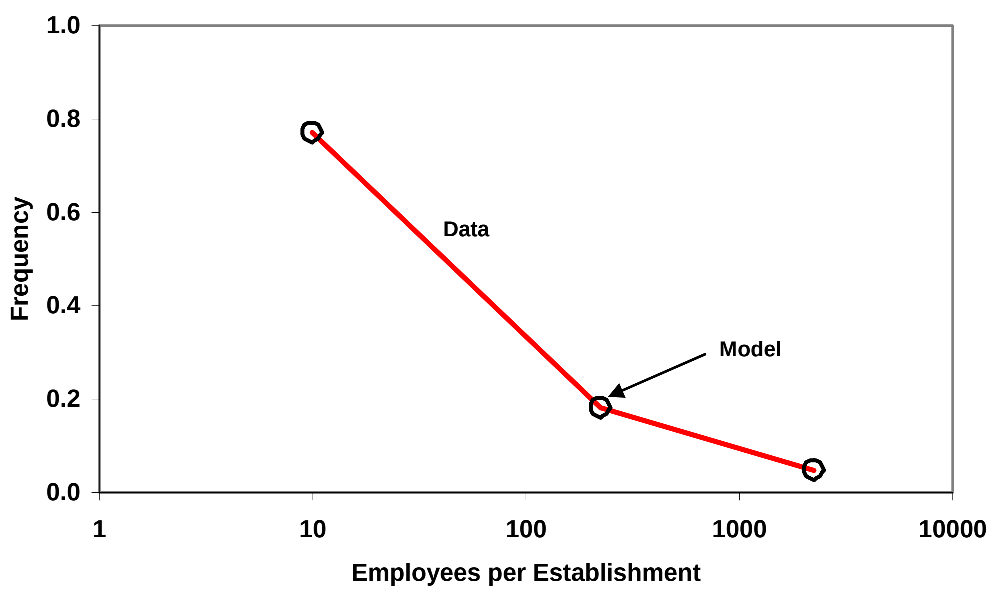


Figure 10

Measured Elasticity of Substitution

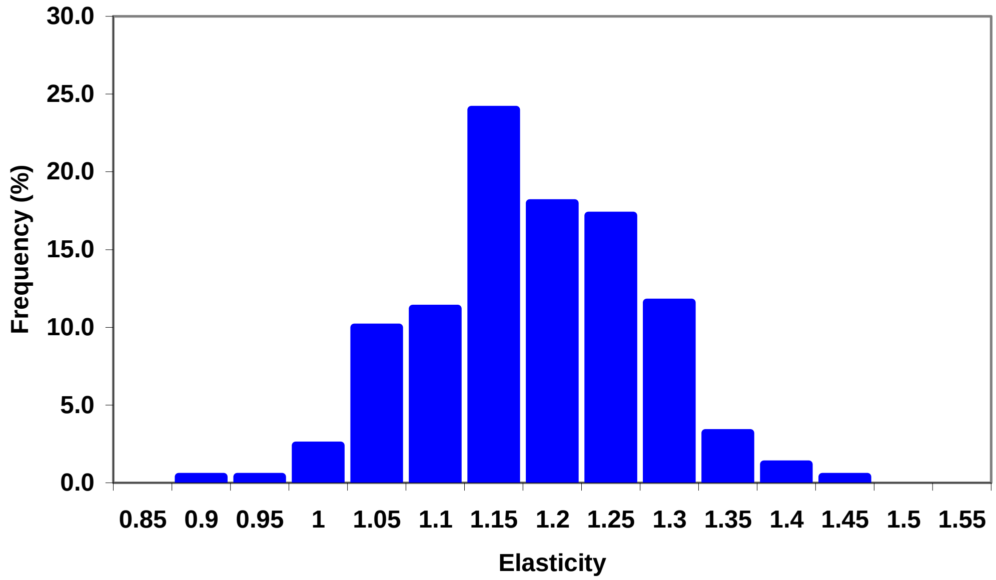


Figure 11

Laspeyres and CES Import Price Indices