A REVIEW OF ARMINGTON TRADE SUBSTITUTION
ELASTICITIES

Christine A. McDaniel & Edward J. Balistreri

Article received on February 4, 2002
Accepted on July 4, 2003

ABSTRACT. Results from applied partial and general equilibrium models used to examine
trade policy are almost universally sensitive to Armington elasticities. While econometric esti-
mates have been offered for 30 years, many economists view these estimates as small. A
few robust findings emerge from the econometric literature: (1) long-run estimates are
higher than short-run estimates, (2) more disaggregate analyses find higher elasticities, and
(3) reduced-form time-series analyses find lower elasticities relative to cross-sectional studies.
We review the literature and offer simulation results to illustrate the sensitivity of general
equilibrium models to Armington elasticities.

JEL Classification: C20; C68; F10.
Keywords: Armington elasticities; Applied general equilibrium.

International Trade Commission (ebalistreri@usitc.gov).
Christine A. McDANIEL, Senior Economist, International Trade Administration, US Department of Commerce.
Economists have been called upon to assess the economic effects of proposed trade liberalization schemes, such as the Canada-US Free Trade Agreement, the North American Free Trade Agreement, and several bilateral agreements (for example, see Cox and Harris, 1985; Brown, 1992; Brown, Deardorff and Stern, 1992). Applied general equilibrium models are commonly used to analyze economy wide impacts of proposed policy changes. These models will continue to be employed as WTO members near the next round of multilateral talks and other regional trading agreements are pondered.

Using economic models to evaluate changes in international trade policy generally requires the conversion of policy changes into price effects. Model analyses use these price shifts to determine how policy is expected to affect output, employment, trade flows, economic welfare, and other variables of interest. The direction and magnitude of a trade policy change on individual variables depends on the size of the shock as well as the behavioral relationships present in the economy. When evaluating policy shifts within an economic model, these behavioral relationships largely take the form of elasticities reflecting the responsiveness of one set of variables to a change in a second set. For example, trade policy can affect the price of traded goods relative to domestically produced goods. As a result, a key relationship for model analysis is the degree of substitution between imported and domestic goods. This key relationship is commonly identified as the Armington elasticity.

In general, knowledge of elasticities is important for policy considerations. Changes in tariffs and taxes will affect a country’s trade opportunities, level of income, and employment. The size of these impacts will largely depend on the magnitude of elasticities. The Armington elasticity is an essential component of trade policy analysis. Applied partial and general equilibrium models that rely on the Armington structure are universally sensitive to these elasticities. Indeed, a modeler’s central Armington choice will drive key quantitative, and sometimes qualitative, results that policymakers use.

In the following section we review literature estimates of Armington elasticities. We point out some of the key differences and robust findings in these studies. In section II we exemplify the importance of the elasticity estimates by offering results from a stylized policy analysis. We make concluding remarks in section III.

**Econometric estimates**

Comprehensive industry-level estimates of Armington elasticities have appeared intermittently over the last few decades. The five main studies available for US imports include Stern, Francis, and Schumacher (1976), Shiells, Stern, and Deardorff (1986), Reinert and Roland-Holst (1992), and Shiells and Reinert (1993), and Gallaway, McDaniel and Rivera (2000). These studies employ standard, transparent approaches to Armington estimation.

---

2. The constant elasticity of substitution (CES) specification for the trade substitution elasticity is derived from Armington (1969).
Many trade economists, however, view these elasticity estimates with skepticism and believe that domestic and imported goods are much more substitutable than most estimates suggest. We highlight specification and identification issues with many of the techniques employed.

One of the first systematic studies to provide import-demand elasticities for the US was carried out by Stern, Francis, and Schumacher (1976). This study offers “best estimates” of US import-demand elasticities for 28 industries at the 3-digit ISIC level. Interestingly, rubber products, wearing apparel, metal products excluding machinery and transport equipment were among the sectors found to be “extremely import sensitive,” while food, beverages, textiles, tobacco, machinery including electrical machinery, and iron and steel were classified as “moderately import sensitive.” The wood and paper products industries were considered “import inelastic.”

Shiells, Stern, and Deardorff (1986) estimated trade substitution elasticities using a simple stock-adjustment model with annual data from 1962-1978 for 163 disaggregated industries. The authors obtained statistically significant Armington elasticities for 122 of 163 sectors estimated. Their estimates compared adequately with previous estimates from Stern et al. (1976).

Shiells and Reinert (1993) disaggregated US imports into those from the NAFTA members and those from the rest of the world (ROW). Using quarterly data over 1980-1988, they obtained estimates for 128 mining and manufacturing sectors. Elasticities were estimated using three specifications: (i) generalized least squares estimation technique, based on a Cobb-Douglas price aggregator; (ii) maximum likelihood estimation using a CES price aggregator; and, (iii) a simultaneous equation estimator that uses a Cobb-Douglas price aggregator and employs a distributed lag model. Shiells and Reinert found the estimates to be relatively insensitive across the three alternative estimation procedures.

Reinert and Roland-Holst (1992) estimated Armington elasticities for 163 US mining and manufacturing sectors. They obtained significant estimates for approximately two-thirds of the 3-digit SIC industries estimated using quarterly data from 1980-1988. Their statistically significant estimates range from 0.14 to 3.49. It is worth noting that a comparison of the higher-tier estimates from Reinert and Roland-Holst (1992) to the lower-tier nested estimates from Shiells and Reinert (1993) varies across sectors. In some instances, lower-tier estimates are greater than the upper-tier estimates, but that is not always the case. Shiells and Reinert attribute the failure to obtain consistently large, lower-tier elasticities across sectors to large terms-of-trade effects.

Gallaway, McDaniel and Rivera (2000) offer some of the most comprehensive and disaggregated set of Armington elasticity estimates. The authors consider explicitly the long-run aspect that is applicable to applied partial and general equilibrium modeling. They provide estimates for 309 industries at the 4-digit SIC level over the period 1989 to 1995. Significant long-run estimates range from 0.52 to 4.83. Long-run estimates are up to five times as large
as short-run estimates, and on average twice as large as the short-run estimates. The authors also find significant differences across 3-digit SIC series within 4-digit SIC groupings. These findings are important since long-run estimates are more appropriate for most trade policy analysis than short-run estimates and much of the applied trade policy analysis is conducted at the detailed commodity level. A comparison of the Gallaway, McDaniel and Rivera 4-digit estimates to Reinert's and Roland-Holst's 3-digit provides further insight to the well-known aggregation bias: the more detailed commodity level, the greater the measure of substitutability.

The magnitude of the elasticity estimates is notably modified when instrumental variables regression analysis is used, and also sensitive to the set of instruments used. As demonstrated by Orcutt (1950) and discussed in Goldstein and Khan (1985), the simultaneity between quantities and prices can bias price elasticities in trade relationships. Single-equation estimation techniques will commonly generate price elasticities that are biased downward because they are a weighted average of the actual demand and supply elasticity. Using rich panel data Erkel-Rousse and Mirza (2002) exploit the supply side considerations using instrumental variables and report a range from 1 to 13 – a wide range, but one with an upper bound much higher than other estimates.

Hummels (1999) uses a multi-sector model of trade to isolate channels through which trade cost or resistance affects trade volume across countries or regions he is able to solve for the implied substitution elasticity. Assuming that all distance related trade resistance is a freight charge he computes a range for the substitution elasticity of 2 to 5.3. More compelling are Hummels' direct substitution elasticity estimates in a framework that includes a more general interpretation of trade resistance. His average estimates are 4.8, 5.6, and 6.9 for aggregation at the 1-digit, 2-digit, and 3-digit levels respectively.

The cross-sectional estimates presented by Hummels, and the panel estimates by Erkel-Rousse and Mirza, are much higher than the central values obtained in the time-series studies of US data. The average of approximately 7 reported by Hummels at the 3-digit level is well above the average long-run estimate of nearly 2 reported by Gallaway et al. at the 4-digit level. The large divergence in these estimates might indicate misspecification in the single-equation time-series analyses.

The substitution elasticities estimated by Hummels (1999) and Erkel-Rousse and Mirza (2002), depart from the standard Armington formulation because the underlying structural model is one of firm-level product differentiation. Firm-level differentiation implies firm-level mark ups over marginal cost and is consistent with new trade theory. Although some computational models have adopted this richer theory, most rely on the standard constant returns Armington formulation, which only differentiates between regional varieties and does not accommodate mark ups. If one interprets the standard model as a general (constant returns) approximation to richer industrial-organization models then application of the Hummels or Erkel-Rousse and Mirza estimates is appropriate. Expanding the support for these higher
elasticities, empirical work in industrial economics generally suggests low mark-up ratios: low estimated mark-up ratios imply high elasticities (a point made by Erkel-Rousse and Mirza).

Another example of evidence for higher trade elasticities is found in Riedel (1988). Although Riedel does not directly estimate Armington elasticities, his estimation of Hong Kong’s export-demand elasticity has been cited as evidence that Armington elasticities are likely higher than suggested in most empirical work. Riedel argues that direct estimation of export demand for a small economy performs poorly because observed quantities are determined by price and supply conditions (export demand is perfectly elastic). He identified the system by inverting the demand function and considering the export supply function. Under this specification he presents compelling evidence that Hong Kong is a price taker. The indication is that the world market does not distinguish between Hong Kong goods and those of other countries and that the Armington elasticity is very high. This is important evidence that researchers need to consider when disaggregating small economies, because the optimal tariff is inversely related to the Armington elasticity.

One of the most useful aspects of econometric estimates is that they offer guidance on relative ease of substitutability across sectors. For example, in reviewing the econometric estimates reported in the papers mentioned above, plastic materials and resins, photographic equipment, paperboard boxes, malt beverages, and softwood lumber are sectors with some of the highest Armington estimates. On the other side, brooms and brushes, house furnishings, and food containers are sectors with some of the lowest Armington estimates.

In an attempt to explain the wide variation in these substitution elasticities across sectors, Blonigen and Wilson (1999) examine whether product, industry, and political characteristics between domestic and import goods are related to systematic differences in the substitution elasticity across US industries. The authors find that increased multinational presence in the downstream industries increases the elasticity of substitution unless importing behavior in the downstream industries is unusually high as well. The level of product differentiation may be another possible determinant of the degree of substitutability across sectors (Feenstra, Markusen, and Rose (1998), Eaton and Kortum (1999), and Erkel-Rousse and Mirza (2002)).

The reviewed papers that offer econometric estimates are useful to modelers in that they provide a starting point for specifying key behavioral parameters. However, the lack of consensus on point estimates reflects the sensitivity of the estimation results to the technique employed. There are, however, a few robust findings. First, long-run estimates are much

---

3. Specifically Harrison, Rutherford and Tarr (2001) make this argument. In their “low elasticities” case an Armington of 4 and a lower-level elasticity of substitution between imports from different regions of 8 is assumed. Although high this assumption is generally consistent the direct econometric estimates. In their central case, however, they assume an Armington of 15 and a lower-level elasticity of 30. They contend that these higher elasticities are similar to the point estimates obtained by Riedel (1988).

4. Athukorala and Riedel (1991) use similar methodology and show that Korea is a price taker in the market for machinery.
higher than short-run. Researchers using short-run estimates should be concerned about understating trade response.

The second general finding is that the level of aggregation is important: the more disaggregate the sample the higher the estimated substitution elasticity. In light of this, it is important to question common practices in flexible aggregation models. These include: applying aggregate estimates to individual sectors that are the focus of study, and applying the average elasticity from disaggregate estimates to an aggregated commodity. We feel that it is important for modelers to consider these known biases and not to adopt econometric point estimates indiscriminately.

Third, recent cross-sectional studies find higher substitution rates than the central values obtained in the time-series studies. The source of this difference is attributed to a fundamental difference in methodology. In general, cross-sectional studies exploit observed resistance in bilateral trade (controlling for country-specific fixed effects) and the size of trade-and-transport costs to infer the substitution elasticity (Hummels (1999)). The key identifying assumption is that the source of the uncontrolled bilateral resistance is due solely to substitution away from trade costs. In contrast, time-series studies are generally agnostic about the level of trade flows; only relating changes in prices to changes in flows. Therefore, the very long-run differences in flows observed in the cross-section, and attributed to persistent distortions that affect long-run supply, are muted in the time-series analysis. This potential downward bias in time-series analyses is supported by the data that shows larger long-run responses.

Finally we note that model misspecification is particularly problematic when estimated parameters are exogenous inputs to an entirely independent structural trade model. The specifications of Stern, Francis, and Schumacher (1976), Shiells, Stern, and Deardorff (1986), Reinert and Roland-Holst (1992), and Shiells and Reinert (1993), and Gallaway, McDaniel and Rivera (2000) all suffer from the general critique that they are structurally inconsistent with the general equilibrium because they do not give adequate consideration to the supply side of the market. The negative estimates commonly found in these studies may, at least in part, be attributed to model misspecification. Even the work by Hummels (1999) and Erkel-Rousse and Mirza (2002), which is consistent with a general equilibrium, is not necessarily consistent with the structure of most applied computational models. In sum, sensitivity to estimation technique and misspecification are reasons that prudence is in order when adopting estimates from the literature.

Challenges remain in econometric estimations of these important parameters. The literature is scarce on a number of issues that would affect applied modeling exercises, such as whether country-specific characteristics or the composition of trade affect the degree of substitutability. More research on the determinants of substitutability across sectors is also needed.

---

5. Large-scale applied models often assume constant returns and perfect competition. Hummels (1999) motivates his empirical trade model with a richer general equilibrium theory that includes monopolistic competition.
The sensitivity of policy analysis to Armington estimates

The assumption of national product differentiation is a defining feature of the Armington structure. Brown (1987) argues that this often yields results that are dominated by changes in terms of trade, rather than changes in efficiency. In order to illustrate the importance of the choice of the Armington elasticity in applied general equilibrium models, we compute examples in which the quantitative and qualitative results vary across assumed elasticities, which are within the range of the cited literature. The model is a direct extension of the widely available GTAP in GAMS system.6 The multi-region social accounts are based on an aggregation of GTAP version 4, which uses 1995 as a base year. The accounts are projected onto a steady-state expansion from 1997 to 2050. The dynamic structure is one of neo-classical capital accumulation driven by exogenous population and productivity growth. Agents are assumed to be forward-looking and accumulate capital to maximize intertemporal utility.7

For these illustrative experiments we simply use the steady-state equilibrium as the baseline.8 The model is also greatly simplified by aggregating up to include only three goods (agricultural products, manufactured products, and services), and four regions (Colombia, NAFTA members, other Latin American countries, and the rest of the world). We use the standard nested-Armington structure, in which the lower-tier substitution elasticity (between imported varieties) is twice the substitution elasticity between imported and domestic varieties.9

We focus on Colombia to illustrate an important implication of the Armington trade structure. It is typical to find applied models with a regional aggregation that explicitly breaks out focus countries of interest but most of the rest of the world is captured in residual aggregates. An implication of this is that relatively small countries (Colombia in our example) are assigned a unique variety in the demand system of their trade partners. The Armington structure results in a great deal of market power, and relatively high optimal tariffs, even for small countries (each country is a monopoly suppliers of its variety). This is the point originally highlighted by Brown (1987).

Using the simulation model we compute welfare impacts in terms of equivalent variation for the representative Colombian agent under various levels of protection holding the response

---

6. The GTAP in GAMS system is maintained by Tom Rutherford, University of Colorado. Documentation is available in Rutherford and Paltsev (2000).
7. We consider the dynamic equilibrium because we are interested in exploring the dependency of trade responses on capital flows. In the central case we assume perfect international intertemporal capital markets over the computed horizon (2005 to 2050). Lau, Fahlke, and Rutherford (1997) present a practical method for representing the neo-classical dynamic equilibrium for computation.
8. In actual policy analysis a more credible baseline might be considered that includes, at least, medium-run differential growth rates across countries.
9. In the simulations, in which we change the substitution elasticities, the ratio of 2 between the lower-tier and upper-tier elasticities is maintained. Thus, an Armington (upper-tier) elasticity of 5 implies an import-variety (lower-tier) substitution elasticity of 10.
of the other regions constant. Figure 1 presents these results over a range of Armington elasticities. With low Armington elasticities, 1 or 3, (which are roughly consistent with the long-run time-series econometric estimates of Gallaway, McDaniel, and Rivera, 2000) unilateral liberalization from the benchmark rate of protection is harmful to Colombia. Adopting a higher estimate, 5, (roughly consistent with Hummels, 1999) indicates that marginal liberalization is beneficial to Colombia. Only when we adopt an elasticity of 8 is total liberalization superior to the benchmark, and even then the optimal for Colombia is at a rate of protection that is about 40% of the current level. In this example, both the qualitative and quantitative effects of liberalization are sensitive to our choice of elasticity.

The welfare impacts from unilateral Colombian liberalization are conditional on the stylized dynamic structure that we have adopted. However, we note that past literature, generally

Figure 1 - Unilateral Colombian trade policy: welfare sensitivity to Armington specification
relying on static models, have indicated that simulations of bilateral free trade areas are sen-
sitive to Armington elasticities, but that the qualitative welfare results from simulations of
multilateral liberalization are less sensitive. Static models also generally support our central
result that welfare effects increase in the substitution elasticity.

Figure 1 illustrates that the Armington parameter not only determines the sign and magni-
tude of liberalization, but that it also might be used to mitigate some of the undesirable fea-
tures of the Armington structure. If we believe that Colombia, the country of focus, is
generally a price taker on world markets, but the Armington structure is chosen for its tracta-
bility, then assuming a high elasticity compensates for the implicit market power granted
when Colombia produces a unique variety. Obviously, a more satisfying solution would be to
get rigorous proof that Colombia is truly a price taker (i.e., estimate the trade elasticities in
an econometric model that includes the Colombian variety). Unfortunately, compromises
and simplifying assumptions must be made in applied work. Given the infeasibility of a full
structural estimation of the applied computational model, a researcher might identify the
sensitivity and offer an argument for the chosen elasticities. The professional audience can
then make an informed judgment about the validity of the results.

Another important point to recognize is that, although many simulation models adopt the
Armington structure, other structural assumptions play an important role in determining the
trade response independent of the chosen elasticity value. For example, an important struc-
tural assumption in dynamic models is the extent that international capital markets are avai-
able. In general, changes in trade policy will generate incentives for countries to adjust their
capital position to take advantage of the best rates of return.

To illustrate the impact of structural assumptions about capital flows on trade response, we
again simulate liberalization by Colombia. The Armington elasticity is held fixed at 5 across
scenarios, and we model a complete liberalization of tariffs. In the central case we make an
assumption that within the model horizon there can be no change in net indebtedness.
Within horizon international capital flows are unrestricted, but any net accumulations of debt
have to be paid back before the terminal period (2050). In the most restrictive scenario we
make an assumption that there is no change in Colombia’s capital flows (period-by-period
balance of payments). In the least restrictive case we approximate the infinite horizon solu-
tion with perfect capital markets. In this case net accumulated debt need only be serviced,
but never retired. Figure 2 presents the effect of these alternative assumptions on the
change in Colombian imports.

When the model is simply closed by maintaining no change in net indebtedness the trade
response is very similar to the perfect capital markets case. In the period-by-period balance
of payments case, however, the initial response is notably less. This is because, when capital
markets are available, Colombia can borrow against future exports and take advantage of
lower import prices in the short-run. All of the models converge to approximately the same
response in the long run. So, although the Armington elasticity largely determines the ulti-
mate response, the structure of international capital markets interacts with the Armington elasticity to determine the adjustment dynamics. We highlight this because dynamic models are gaining influence in policy arenas, and adjustment dynamics are often of central importance.

**CONCLUDING REMARKS**

There is no question that measurement of Armington elasticities is of fundamental importance in determining the response of trade models to policy experiments. There is equally little doubt that measurement of these elasticities is very difficult. The estimates from the
literature provide a wide range of point estimates, and little guidance on the correct estimate to apply to a given commodity in a given model for a given regional aggregation. Most of the controversy surrounding Armington elasticities reduces to a general structural inconsistency between the econometric models used to measure the response and the simulation models used to evaluate policy.

Three robust findings emerge from the econometric literature. First, long-run estimates of Armington elasticities are larger than short-run estimates. Second, more disaggregated analyses find higher elasticities, which indicates that aggregation matters and interacts with the Armington specification. Finally, single equation time-series approaches identify smaller responses relative to cross-sectional estimation that includes a consideration of supply conditions and a broader interpretation of industrial organization.

Given the sensitivity of econometric estimates many modelers turn to sensitivity analysis of simulation results, which is quite useful to economists but has had limited impact in policy arenas. Quantitative impacts from simulation models are often used in a way that is at odds with the notion of comparative welfare analysis that acknowledges simplifying assumptions and uncertainty over key parameter values. Policymakers often utilize single figures to support policy positions. Unfortunately, acknowledging the existence of second moments in policy forums may weaken the same argument that it strengthens in academic review.

The consumers of simulation results, thus, demand a definitive literature point estimate of key parameters to further their cause. However, all estimated parameters are conditional upon the structure under which they were estimated. Structurally different simulation models need to maintain the flexibility to adjust and reject estimated parameters that produce unrealistic responses. The responsibility, of course, falls on the researcher to define and identify a range of responses. Once sensitivity analysis reveals the problem, adjustments away from econometric estimates are warranted. Ironically, when dealing with models that do not maintain the same structure, rejection of econometric estimates might be necessary to produce simulation responses consistent with the data used to estimate the econometric model.

Structural inconsistency across simulation models is also important to consider when adopting a given set of elasticities. Obviously, the same parameter set will produce different responses across different models, and even across different aggregations of the same model. Thus, the adoption of parameter sets used in previous studies or standard models does not abdicate the researcher’s responsibility to think carefully about and defend the estimates. Outside the reviewed traditional econometric measurement approach to parameter identification, structurally consistent estimation/calibration of computational models is an area of promising research. Following the lead of the real business cycle literature and a philosophi-

---

10. Harrison, Jones and Kimbell (1993) urge that applied general equilibrium models should be subject to sensitivity analysis.
cal acceptance of calibration as a method of estimation (Dawkins, Srinivasan, and Whalley (2001)), creative researchers have combined aspects of stochastic estimation in structural general equilibrium models (Liu, Arndt, and Hertel (2001), and Francois (2001)). The key advantage this research offers is a methodological link between the structural policy model and parameter estimation. These ideas are in their infancy, but appear encouraging in their useful approach towards solving a very difficult measurement problem11.

C.A. McD. & E.J. B.

REFERENCES


11. We acknowledge helpful comments from two anonymous referees. Any errors remain our own.


