

1-1-2000

# Alternative Hypotheses and the Volume of Trade: the Gravity Equation and the Extent of Specialization

Jon Haveman  
*Federal Trade Commission*

David Hummels  
*Purdue University*

Follow this and additional works at: <http://docs.lib.purdue.edu/ciberwp>

---

Haveman, Jon and Hummels, David, "Alternative Hypotheses and the Volume of Trade: the Gravity Equation and the Extent of Specialization" (2000). *Purdue CIBER Working Papers*. Paper 4.  
<http://docs.lib.purdue.edu/ciberwp/4>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact [epubs@purdue.edu](mailto:epubs@purdue.edu) for additional information.

# Alternative Hypotheses and the Volume of Trade: the Gravity Equation and the Extent of Specialization

Jon Haveman  
*Federal Trade Commission*

David Hummels  
*Purdue University*

February 2001

**Abstract:** We examine why the gravity equation works and the implications for its use. There are three contributions. First, we demonstrate that the gravity equation as a statistical relationship (and not as a specific prediction about bilateral trade) can be generated from a model with incomplete specialization. Second, we analyze two facts or “puzzles” that are broadly inconsistent with the complete specialization models typically used to derive the gravity equation. In both cases, the data are consistent with the alternative hypothesis of incomplete specialization. “Home bias”, for example, is only puzzling when viewed through the lens of the complete specialization model. Third, we demonstrate that the explanation for why the gravity equation works has considerable relevance for how the gravity equation is interpreted and used, and how we view bilateral trade.

**Acknowledgements:** The authors thank Jeff Bergstrand, Don Davis, Alan Deardorff, Simon Evenett, James Harrigan, Robert Heiny, and Wolfgang Keller for very helpful comments. Financial support from the CIBER center at Purdue University was also beneficial. The views expressed in this paper are those of the authors and do not necessarily represent the views of the Federal Trade Commission or any individual Commissioner.

## I. Introduction

The gravity model is one of the most commonly used devices in empirical trade research. Literally scores of studies employing gravity equations have appeared in the literature, with considerably increased use of late. The reason for this use is summed up in leading surveys of empirical evidence on international trade theory. Deardorff (1984) writes that gravity models are “extremely successful empirically” judging by their ability to explain variance in bilateral trade volumes. Leamer and Levinsohn (1997) write that gravity models “have produced some of the clearest and most robust empirical findings in economics.”

In its most common use, the gravity model provides a baseline against which to measure deviations from “normal” or frictionless bilateral trade. Three recent and prominent examples include Frankel, Stein and Wei (1995), who measure the effects of regional trading blocs, McCallum (1995) who measures the effect of international borders and the existence of “home bias”, and Wei (1998), who measures the effect of exchange rate variability on trade patterns.<sup>1</sup>

The theoretical motivation for the framework is found in a set of general equilibrium models that derive specific predictions for bilateral trade. Examples include Anderson (1979), Bergstrand (1985), Deardorff (1998), and Helpman and Krugman (1985). The common elements in each of these papers are complete specialization and identical preferences; each good is produced in only one country and consumers value and therefore purchase every good.

More recently, the fit of the gravity equation has been used as a test of monopolistic competition trade theory. Harrigan (1993) estimates a unitary elasticity of bilateral trade with respect to exporter output, as predicted by the theory. Helpman (1987) notes that the model predicts greater trade when countries are more similar in size, and suggests that trade growth within the OECD may be caused by size convergence over time. Hummels and Levinsohn (1995) show that this property also holds for non-OECD countries with little intra-industry trade. They interpret this as evidence that the fit of the simple gravity equation may be due to something other than trade in differentiated goods. A possible answer to this puzzle may be found in the work of Anderson (1979) and Deardorff (1998).<sup>2</sup> Both employ models characterized by complete specialization in production and derive the gravity equation without assuming monopolistic competition. As Grossman (1998) notes, the key to generating gravity predictions is complete specialization itself, and not the particular details on the supply-side that give rise to it.

A consensus appears to have formed around this position – the fit of the gravity equation suggests that production is extremely specialized, even if it does not inform us as to the source of that

---

<sup>1</sup> This is a very abbreviated list of correlates; others include shipping costs, tariff rates, product standards, information costs, exchange rate regimes, migration, political uncertainty, insecurity, and internet penetration.

specialization. If correct, this is a powerful insight into the nature of trade, as well as a highly useful insight for empirical work. Complete specialization models provide precise structural predictions for bilateral trade that can be taken directly to the data. In the context of the model the baseline value for bilateral trade is very well defined, and deviations from that baseline have direct and specific meaning in both a positive and normative context. For example, the elasticity of bilateral trade with respect to a tariff increase maps directly into a key parameter in the underlying utility function, and the resulting quantity change has a direct interpretation in terms of lost utility.<sup>3</sup>

The contrast with models that do not assume complete specialization is stark. The canon of neoclassical models is generally characterized by incomplete specialization (each homogeneous good is produced by more than one country). With more than two countries such models say little about the pattern of bilateral trade other than predicting the set of partners with which a country may trade. The reason for this neglect is that the bilateral pattern of trade is difficult to resolve. If there are multiple producers of an identical good willing to sell at the same price, importers will be indifferent between them and bilateral trade is indeterminate. In this case, there is no baseline value for bilateral trade, and no sense in which the bilateral pattern is tied to any question of normative importance. With smallish trade barriers importers will not be indifferent – they will purchase goods from whoever is the low cost supplier into their market. Rough intuition thus suggests that whatever these models might have to say about bilateral trade there is no reason to suppose it at all resembles the gravity equation.

This paper asks three questions. One, can we derive the gravity equation under incomplete rather than complete specialization models? Two, can we find corollary evidence suggesting that incomplete rather than complete specialization might explain the success of the gravity equation? Three, does it matter why the gravity equation works?

After reviewing the trade volume predictions of complete specialization models in Section II, the contribution of this paper is contained in three main sections. Section III analyzes a simple neoclassical trade model with incomplete specialization in order to identify trade volume predictions that are similar to or different from the complete specialization model. We provide multilateral trade volume predictions in a general setting and use world endowment data to simulate a multi-country Heckscher-Ohlin (HO) version of the model. The model yields insights regarding the income elasticity of trade that are very similar to the complete specialization model, but predicts a much lower level of trade. If trade is costless, then bilateral trade is indeterminate in this model. However, trade is not costless in the world, and we resolve the indeterminacy by letting importers choose partners in such a way as to minimize transportation

---

<sup>2</sup> See also Evenett and Keller (1998) and Jensen (2000).

<sup>3</sup> See Lai and Trefler (1999).

costs. This resolution yields (as a statistical relationship) bilateral trade in proportion to partner incomes, also known as the gravity equation.

Having demonstrated where the models are similar, we turn to the question of where they are different. We ask: if both complete and incomplete specialization models yield the gravity equation, can we identify corollary evidence that is consistent with only one of these models? Section IV analyzes evidence on trade in varieties and the level of trade. The complete specialization model used to derive the gravity equation implies that consumers highly value variety, importing all goods that are produced. In contrast, the incomplete specialization model suggests that importers will buy from only a small fraction of available sources, strongly consistent with the data. Second, the complete specialization model predicts trade levels much higher than incomplete specialization models, and much higher than we observe or can explain with observable barriers. In each instance, the data suggest that the complete specialization model used to explicitly derive the gravity equation considerably overstates either the extent of specialization (the degree to which goods are differentiated) or the degree to which consumers value that differentiation.

Previous authors have derived the gravity equation under a variety of assumptions, so simply adding another model to the list is of limited consequence. A case must be made that the explanation for why the gravity equation works has considerable relevance for how the gravity equation is interpreted and used, and how we view bilateral trade. In Section V, we examine the positive interpretation and normative content of gravity regressions under complete and incomplete specialization. Section VI concludes.

## II. The Null Hypothesis: Complete Specialization

We begin by examining predictions about the volume of trade in what we term complete specialization (CS) models. These models have at their core two common assumptions: production is completely specialized (only one exporter produces each good), and importers purchase all goods. The familiar monopolistic competition framework with CES utility is one such model, though its virtue is that complete specialization is derived rather than assumed.<sup>4</sup>

To show the central role these assumptions play, we provide the following simple model that is closely related to Anderson (1979). We assume all goods are final goods, that trade is balanced, and there is no international lending. With identical and homothetic preferences, and assuming no trade barriers, each country  $i$  will consume an income share,  $b^k$ , of good  $k$ , or

---

<sup>4</sup> With the CES utility structure, consumers demand all varieties and producers always produce unique varieties in order to have monopoly power over that variety.

$$(1) \quad C_i^k = b^k Y_i$$

World output of good  $k$ ,  $X_w^k$ , must equal world consumption, so summing purchases of good  $k$  over all countries,  $X_w^k = b^k \sum_l Y_l$ . Solving for  $b^k$  and substituting into equation (1) we find that  $i$  consumes its world income share ( $s_i = Y_i / Y_w$ ) of world production of good  $k$ .

$$(2) \quad C_i^k = s_i X_w^k$$

By assumption, production is completely specialized. That is, every country is the sole supplier of the goods that it produces, or  $X_w^k = X_j^k$ . As a result, the consumption vector directly pins down the pattern of bilateral imports of good  $k$  as

$$(3) \quad M_{ij}^k = s_i X_w^k = s_i X_j^k$$

Country  $i$  will demand a similar fraction of all the goods produced in country  $j$ . Summing over all sectors, and noting that the sum of sectoral output equals national income in  $j$ , we arrive at a prediction for bilateral trade:

$$(4) \quad M_{ij} = \sum_k s_i Y_j^k = \frac{Y_i Y_j}{Y_w}$$

This equation is known as the simple or frictionless gravity model.

Complete specialization models also provide straightforward predictions about the multilateral volume of trade. Sum equation (4) over all partner countries for a particular importer,

$$(5) \quad M_i = Y_i(1 - s_i)$$

This multilateral model provides a baseline prediction for the level of trade, and an implication that the trade share of income ( $M/Y$ ) is decreasing in income.

For several exercises, we wish to evaluate the predictions of the model in the presence of trade barriers. This requires a specific form of the model and a specific assumption about utility. We employ the most common form of the complete specialization framework, a one-sector model in which firms are monopolistically competitive and produce differentiated varieties that enter a CES utility function as follows.

$$U_i = \left( \sum_j (c_{ij})^\theta \right)^{1/\theta}$$

where  $C_{ij}$  denotes consumption in importer  $i$  of a variety  $j$ ,  $\theta = 1 - 1/\sigma$ , and  $\sigma$  is the elasticity of substitution between goods. An important benefit of this model is that it provides a flexible way to evaluate implied trade barriers over a range of substitutability between goods.

As the trade volume predictions in this model are well known, we do not repeat the derivation here. Country  $i$ 's purchases from a source  $j$  are given by

$$(6) \quad M_{ij} = Y_i Y_j \frac{(p_j t_{ij})^{-\sigma}}{P_i}$$

where  $Y_j$  is output from exporter  $j$ ,  $t_{ij} - 1$  is the ad-valorem trade barrier,  $p_j t_{ij}$  is the price of the good inclusive of all transactions costs, and  $P_i$  is a CES price index over all available goods.

This model contains four predictions that we wish to highlight and then compare to an alternative hypothesis of incomplete specialization. The first two predictions concern the income elasticity of trade. One, the gravity equation captured in equations (4) and (6) predicts a unitary elasticity of bilateral trade with respect to the product of partner incomes. Two, equation (5) predicts that the elasticity of multilateral trade with respect to income is less than one -- that small countries trade a larger fraction of income than do large countries.

The third prediction concerns the level of trade. Absent trade costs, equation (5) predicts multilateral trade equal to 90 percent of income for a country with 10 percent of world GDP. With trade barriers, equation (6), trade may be considerably lower depending on the size of the barrier and the elasticity of substitution. The fourth prediction concerns which countries trade. Equation (3) indicates that if a country produces a good, all other countries purchase it. Note that this is true even in the presence of (finite) iceberg transport costs.<sup>5</sup>

### III. The Alternative Hypothesis: Incomplete Specialization

The models used to derive the gravity equation as a specific and determinate prediction for bilateral trade have at their core the assumption of complete specialization -- each good is produced in only one country. In this section we examine a model with incomplete specialization (multiple countries may produce each homogeneous good). This is not an esoteric choice of models to examine -- the canonical models of neoclassical trade theory including Heckscher-Ohlin, Ricardian, and Specific Factors models commonly exhibit incomplete specialization.

Our goal is to examine the income elasticity of trade, the level of trade, and the number of partners from which a country imports a particular good using the most general possible setting. In some cases we will require a more specific formulation to make specific predictions and here we employ a simple 2 good, 2 factor, N country Heckscher-Ohlin world to illustrate model properties. The point here

---

<sup>5</sup> Equation (6) implies that, with finite trade barriers and finite substitution elasticity, all varieties are purchased.

is not to set up a horse race, but rather to illuminate which predictions are unique to complete or to incomplete specialization and which predictions are shared.

### III.1 Multilateral Trade

We begin by using a simple accounting exercise to derive the multilateral volume of trade in the presence of incomplete specialization. As above we assume away intermediate goods, and note that national output  $X$  and national income  $Y$  must be equal. Define  $\gamma_j^k$  as the output share of good  $k$  in country  $j$ , so that country  $j$ 's output of good  $k$  is

$$X_j^k = \gamma_j^k X_j = \gamma_j^k Y_j.$$

As in the previous model, identical and homothetic preferences ensure that every country consumes its world income share of each consumption good.

$$C_j^k = s_j X_w^k$$

An accounting identity expresses the volume of exports in a sector as production less consumption, or

$$EX_j^k = \gamma_j^k Y_j - s_j X_w^k = Y_j \left( \gamma_j^k - \frac{X_w^k}{X_j} \right) = Y_j (\gamma_j^k - b^k)$$

where  $b^k$  is the share of good  $k$  in world consumption (production). Summing over those sectors with positive exports ( $k \in EXP$ ) gives the total volume of exports.

$$(7) \quad EX_j = Y_j \sum_{k \in \text{exp}} (\gamma_j^k - b^k)$$

The volume of exports depends on the income of the country and the trade share of income (the summation term). The trade share of income can also be thought of as a measure of the extent of specialization - the difference between what a country produces and the average of world production.

For example, assume goods production is completely specialized (one exporter per good) as in the previous section. First rewrite equation (7) as

$$(8) \quad X_j = Y_j (1 - s_j) \sum_{k \in \text{exp}} (\gamma_j^k - \gamma_r^k)$$

where subscript  $r$  denotes rest-of-world. Since each country is the sole supplier of its export good,

$\gamma_r^k = 0$ , and since all goods produced in  $j$  are exported,  $\sum_{k \in \text{exp}} (\gamma_j^k) = 1$ . This simplifies to yield the

predicted trade level from the complete specialization model, equation (5). It also makes clear that trade

levels are maximized under complete specialization.<sup>6</sup> In the more general case of incomplete specialization (when  $g^k \neq 0$  in (8)), we expect much lower trade volumes. This is the first major result of this section: observed trade levels look low relative to a model that predicts the maximum possible volume of trade.

Equation (7) also provides insights into the elasticity of multilateral trade with respect to income. Suppose that in a cross-country sample the extent of specialization is uncorrelated with incomes.

$$\text{cov}\{Y_j, (\sum_{k \text{ exp}} (\gamma_j^k - b^k))\} = 0$$

That is, output shares do not differ from world production averages to a greater degree in large countries. In this case, the measured elasticity of multilateral trade with respect to income is one. This is true regardless of the details of why countries specialize. In a statistical sense, the supply-side details of the model are a disturbance around an accounting baseline. This baseline says simply: other things equal, large countries trade more.

### III.2 The Textbook Case: 2x2 Heckscher-Ohlin

Since (7) was derived using only an accounting identity and a common assumption on preferences it is independent of issues such as the number of goods, factors, or countries, whether production is characterized by increasing or constant returns to scale, or whether cross-country heterogeneity is due to factor endowments, technology, or goods differentiation. However, since it is only an accounting identity it has limited economic content. Next we examine equation (7) in the context of a familiar model.

We employ the textbook 2 goods, 2 factors Heckscher-Ohlin model, which we analyze in a many country context. In this model, the level of trade and the elasticity of trade with respect to income depend on the distribution of the capital-labor ratio ( $k$ ). Larger endowment differences (i.e., increasing  $kdif_i = |k_i - k_w|$ ) yield greater specialization in country  $i$  and more trade.<sup>7</sup> The estimated elasticity of trade with respect to income depends on the sample covariance of  $kdif$  and  $Y$ . If large and small countries have endowments that are equally far from the world endowment, then the extent of specialization given by endowments is uncorrelated with incomes. In this case, the elasticity of trade with respect to income is one. However, if large countries have endowments  $k_i = k_w$  and only small countries differ from the world endowment, in our cross-country sample we will observe that trade is decreasing in incomes.

---

<sup>6</sup> Of course, one can specify preferences to exceed trade volumes predicted by a model of complete specialization and identical homothetic preferences. In the limiting case, country A produces only good 1, and consumes only good 2. Country B produces only good 2, and consumes only good 1, yielding trade/GDP = 1.

In short, the HO model contains no theoretical presumptions for the extent of specialization, the level of trade, or the income elasticity of trade that are independent of the country distribution of  $k$  and  $Y$ . A sensible exercise in this case is to check world endowments and examine predictions about trade given the observed distributions. We assume taste and technology parameters, and using data on  $k$  and  $Y$ , we generate predicted values for trade in a  $2 \times 2 \times N$  country HO model. Details on the model solution are provided in an appendix, and summarized here.

Preferences are Cobb-Douglas with equal expenditure shares on both goods. We employ the standard assumption of infinitesimal trade barriers operating at the border. This ensures that countries will first satisfy domestic consumption demands with domestic production before turning to foreign sources of supply. Production is Cobb-Douglas for the two goods.

$$Z_1 = K^\alpha L^{1-\alpha} \quad Z_2 = K^{1-\alpha} L^\alpha$$

We experiment with the parameters, varying  $\alpha$  from 0.55 to 0.95. Data on  $k$  and  $Y$  come from World Bank sources and cover 96 countries in 1990. As there are large differences across countries in the fraction of output devoted to traded goods, we employ non-service sector GDP in place of total GDP.<sup>8</sup> The numerical model solves for the output of both goods in each country, and the vectors of multilateral imports and exports. Bilateral patterns are left to the following section.

The numerical model generates trade levels between 20 and 40 percent of output.<sup>9</sup> This is in the range of trade shares observed for most countries. The exercise provides a useful illustration that trade levels resembling actual data can be generated by the most common textbook model with no trade barriers. Section IV draws a contrast between this result and the size of trade barriers necessary to reconcile complete specialization models with actual trade levels.

Next we estimate the elasticity of multilateral trade with respect to income generated by the numerical HO model, both with and without endowment data as a control.

$$(9) \quad \ln M_i = a + b_1 \ln Y_i + e_{ij}$$

$$(10) \quad \ln M_i = a + b_1 \ln Y_i + b_2 kldif_i + e_{ij}$$

---

<sup>7</sup> However, once all countries are outside of the diversification cone and producing only their export goods, increasing factor dissimilarity does not affect specialization or trade volumes.

<sup>8</sup> This requires the following additional assumptions. One, the income share spent on services is exogenous. Two, factor usage in the service sector is identical to the national K/L ratio in each country. While these are strong assumptions, the results we present are very similar when using total GDP in the numerical model.

<sup>9</sup> Trade levels vary inversely with  $\alpha$ . Small values of  $\alpha$  imply that countries heavily specialize in their export good and export more.

In Table 1, we report estimates using data from the  $\alpha = .75$  case.<sup>10</sup> For contrast, we also report estimates employing actual multilateral trade data from the Statistics Canada World Trade Database.

Cross-section estimates on the HO model, omitting endowments, show elasticities close to, but slightly less than, one. Thus, the numerical HO model provides the same insight as the complete specialization model – while the *level* of trade is greater in large countries, small countries trade a larger *share* of income. In the HO model this is explained by the actual distribution of  $k$  and  $Y$ . Endowments of small countries deviate from the world endowment to a slightly greater extent than large countries. As a consequence, small countries are more extensively specialized, but the effect is not so important that their level of trade is greater. Note that, as equation (7) predicts, when we include the endowment data in the regression we cannot reject a unitary income elasticity.

### *III.3 Bilateral Trade and the Gravity Model*

Neoclassical models with incomplete specialization and more than two countries say little about bilateral trade, other than predicting the set of partners with which a country may trade. The first reason for this neglect is that the bilateral pattern of trade is difficult to resolve. If there are multiple producers of an identical good willing to sell at the same price, importers will be indifferent between them and bilateral trade is indeterminate. This is the second reason for the neglect of bilateral patterns – if importers are truly indifferent to partner selection, researchers should be indifferent as well. That is to say, if bilateral volumes resist positive description and their resolution is a matter of no consequence from a normative standpoint, why bother?

However, pure indeterminacy is a problem more relevant to theoretical exercises than to empirical applications. Even with multiple suppliers of homogeneous goods, in the real world trade will not be literally indeterminate, as variation in trade costs will quickly sort out bilateral pairings. To address this more empirically relevant case, we experiment with a model in which trade costs resolve the bilateral trade pattern.<sup>11</sup> We examine whether the trade patterns that arise from this model mimic the gravity equation, in the sense that bilateral trade is proportional to the product of partner incomes.

Suppose trade barriers are rising in distance so that importers of homogeneous goods buy only from the closest, and therefore cheapest, source of supply. This resolves the indeterminacy, but at the cost of an exceptionally complicated equilibrium (location-specific prices for each good). To avoid this problem, we allow bilateral trade costs to become arbitrarily small while retaining the cost ranking of partners. Then equilibrium prices are unaffected, but bilateral indeterminacy is resolved. This is fundamentally the same method used in neoclassical models to resolve indeterminacy between home and

---

<sup>10</sup> Results are qualitatively similar for other parameter values.

<sup>11</sup> We thank Jim Harrigan and Don Davis for this suggestion.

foreign sources of supply. Rather than assuming an arbitrarily small barrier operating at national borders, we assume an arbitrarily small bilateral barrier that is rising in distance.

Of course, this model provides an immediate departure from an important prediction of the complete specialization models used to derive the gravity equation. Importers do not value multiple sources of homogeneous goods, and so will not buy goods from every available supplier. Instead, importers shift purchases to the low-cost supplier(s).<sup>12</sup>

This model differs substantially from an earlier examination of bilateral trade under incomplete specialization. Deardorff (1998) analyzes a model in which trade is frictionless and all producers sell at the same price.<sup>13</sup> To resolve indeterminacy, Deardorff employs a particular probability model. Let all production be placed into a common world pool from which bilateral trade is allocated, one dollar at a time. To allocate this trade unit we randomly select an exporter and an importer, with selection probabilities given by their shares of the world pool, i.e., exporter  $j$  and importer  $i$  are chosen with probabilities  $P_j = Y_j / Y_w$  and  $P_i = Y_i / Y_w$  so that the expected probability of an  $ij$  match for any one draw is given by  $P_{ij} = Y_i Y_j / Y_w^2$ . With  $Y_w$  draws, the expected number of times a particular pair is chosen (and the volume of trade) is the simple, or frictionless gravity equation.  $E(M_{ij}) = Y_i Y_j / Y_w$

This is an interesting theoretical point but the approach has limited application. First, it is equivalent to showing the gravity equation holds by assuming the gravity equation holds – the resolution is not unique, and different random matching models generate different outcomes that are equally “likely”. Indeed, given indeterminacy *any* resolution of bilateral trade is acceptable so long as adding up constraints are met. Second, the model is extremely sensitive to the inclusion of trade barriers – even infinitesimally small barriers invalidate the prediction. As such it is uniquely unqualified to motivate the use of the gravity equation for measuring trade barriers, and after all, this is one of the principle uses to which gravity equations are put. Finally, we note that the Deardorff model predictions regarding the level of trade and numbers of exporters from whom an exporter buys are identical to the complete specialization model without barriers. We show in the next section that the data strongly reject these predictions.

To complete the analysis, we employ the 2x2xN numerical HO model from the previous section with data on  $k$  and  $Y$  drawn from world endowments. The model generates vectors of output and multilateral trade for each country, and determines the set of partners with which a country may trade.

---

<sup>12</sup> There will not necessarily be a single low cost supplier. Limao and Venables (1999) develop an analytical model with the properties of our current setup. Importers buy from multiple sellers, all of whom sell at the same (cif) price, but they do not buy from all available sellers.

<sup>13</sup> Evenett and Keller (1998) also examine incomplete specialization, but restrict themselves to a two-country case. As a result, bilateral trade is identical to multilateral trade and is therefore well determined in the model.

We then resolve indeterminacy in bilateral trade using the “minimum distance” rule. This yields a bilateral trade matrix for the numerical HO model that can be compared both to real bilateral data and to the prediction of the complete specialization model.

We estimate the simple gravity equation in cross-section using data generated from the numerical HO model and real bilateral trade data. For the HO model data we also include the variable *kldif* for both the importer and exporter. When examining real data, equation (6) implies that trade barriers influence trade volumes, and their omission may affect the estimated elasticity. To address this problem, we follow Hummels and Levinsohn (1995) in estimating the relationship in a panel with yearly data from 1970-1992 and including country pair fixed effects and year dummies.<sup>14</sup> This sweeps out time-invariant, pair-specific trade barriers, as well as pair-invariant, time specific trade barriers. This is our preferred specification as it controls for a long list of plausibly important omitted variables, enabling a cleaner examination of the income elasticity. The three estimating equations are:

$$(11) \quad \ln M_{ij} = a + \mathbf{b}_1 \ln Y_i Y_j + \mathbf{e}_{ij}$$

$$(12) \quad \ln M_{ij} = a + \mathbf{b}_1 \ln Y_i Y_j + \mathbf{b}_2 kldif_i + \mathbf{b}_3 kldif_j + \mathbf{e}_{ij}$$

$$(13) \quad \ln M_{ijt} = a + \mathbf{b}_1 \ln Y_{it} Y_{jt} + c_{ij} + yr_t + e_{ij}$$

We report the results in Table 2. A strict test that  $\mathbf{b}_1 = 1$  can be rejected in all regressions, including those employing real data, but a looser null that  $\mathbf{b}_1 > 0$  cannot be rejected. Note that the coefficients generated from the HO model are not significantly different from those estimated using panel techniques on real data.

In our HO model two things resolve the bilateral pattern of trade. Relative endowments determine the set of partners with which a country may trade, and relative distance determines which of that set will be chosen. Partner income plays no role in the theoretical model. Why then do partner incomes enter significantly in the empirical model?

The simple intuition is that the bilateral trade matrix is governed by a set of adding up constraints. When summing country *i*'s imports over all sources, one must get back *i*'s multilateral volume of imports, and similarly for country *j*'s exports. For any particular importer, the trade pattern may be unrelated to the

---

<sup>14</sup> These results are very similar when using either five year or ten-year intervals instead of annual data, and results are very similar when we exclude year dummies altogether.

size of exporters' shipments. But some country must buy the goods of large exporters. On average, bilateral shipments involving countries that export a great deal must exceed the bilateral shipments involving countries that export very little.

The evidence in Table 1 then allows us to connect this intuition back to partner incomes, as opposed to multilateral trade volumes. There we showed that the HO model (given world data on  $k$  and  $Y$ ) predicts a tight relationship between country size and the level of trade. This suggests that incomes operate as a good proxy for the multilateral import and export terms. That is, incomes predict multilateral trade, and multilateral trade is related to bilateral trade mechanically through adding up constraints. In this sense the model generates a statistical or "on-average" relationship between bilateral trade and incomes. This is very different from a model prediction linking the two variables, as we detail in Section V.

A few notes on robustness are in order. First, we have not proven that *every* allocation of bilateral trade must look like the gravity equation.<sup>15</sup> However, we have experimented with using many different barrier types (tariffs, NTB's, shipping costs) to resolve bilateral indeterminacy in the manner of the "minimum distance" rule used above, and all deliver quite similar results. Second, our argument clearly turns on the elasticity of multilateral trade with respect to income. Under the HO model version of our alternative hypothesis this elasticity depends in turn on the distribution of endowments and incomes. As a result it is easy to generate examples in which multilateral trade (and therefore bilateral trade) is unrelated to, or negatively related to, incomes. Thus it is not true that, as Deardorff (1998) conjectures, "almost any model" will fit the gravity equation. It just happens that endowments and income in the world are distributed in such a way that the HO model does generate the gravity relationship.

It should be emphasized that we have focused entirely on patterns that are revealed through bilateral trade volume regressions. This ignores *which* countries a nation trades with and how the models stack up in that regard. It is well known that 2x2 HO models cannot explain north-north trade, though recent work by Davis (1997) shows that higher dimensional models can. Our argument goes through for higher dimensions so long as the extent of specialization is not strongly negatively correlated with incomes. In this case we would match the determinants and level of trade volumes, while also resolving the north-north problem.<sup>16</sup>

---

<sup>15</sup> In previous drafts we have experimented with a wide variety of rules for randomly determining the bilateral trade pattern under indeterminacy and all yield something like the Table 2 results. This is clearly not a proof that *every* allocation under indeterminacy will work. However, since trade costs undoubtedly resolve indeterminacy in the real world, worrying about possible outcomes under indeterminacy seems a sterile pursuit.

<sup>16</sup> We have not attempted this calibration because we lack data on the finely differentiated factor supplies that motivate Davis' argument for higher dimensions.

## IV. Corollary Evidence

The previous sections suggest that the income elasticities of multilateral and bilateral trade look quite similar when derived from models with complete specialization and those with incomplete specialization. In this section we examine model predictions that are dissimilar. These two predictions concern “corners” or zero values in the bilateral trade matrix, and the overall level of trade. The complete specialization models used to derive the gravity regression have very particular predictions that differ substantially from the incomplete specialization model we have described. We do not intend this look at the data as a “test” of one model over another, but rather a broad check of the consistency of the complete specialization explanation for gravity regressions. We show that data patterns that appear as “puzzles” to be explained in the context of complete specialization models may not be puzzles at all if the world is incompletely specialized.

### *III.1 Zero Values in the Bilateral Trade Matrix*

Can we find direct evidence in trade data that production is completely specialized (only one country produces each good)? One possibility is that goods production is completely specialized at the precise level of aggregation at which we chose to examine trade data. In this case, we should see only one producer for each product code. A second possibility is that goods production is completely specialized at much more disaggregated levels than we are able to measure. (This includes “varieties” of goods, as in the monopolistic competition model.) In this case, we might observe many producers within a single product code. Aggregation issues thus preclude a direct investigation as to whether production is completely specialized.

However, by restricting our attention to the complete specialization models used to derive the gravity equation, a fundamental implication is observable. These models assume that production is completely specialized *and* that all products (varieties) are valued and purchased by all consumers.<sup>17</sup> We examine the equation (3) implication that if a country exports a good to at least one country, it will export that good to all countries.<sup>18</sup>

Figure 1 provides evidence on this point. We define a “good” as a 4-digit SITC category. Using 1990 bilateral trade data from the Statistics Canada World Trade Database (WTDB) we calculate the number of exporters in each good, with the count frequency displayed at the top of Figure 1. The median

---

<sup>17</sup> We acknowledge the contribution of referees who have pointed out that this implication is not a perfectly general implication of *all* complete specialization models with identical preferences. This is addressed below.

<sup>18</sup> Previous authors have noted the existence of zero values in the trade matrix, but this is generally acknowledged in the context of motivating proper econometric technique to handle zero values. We are unaware of work that uses the existence of zero values as evidence on the nature of specialization.

number of suppliers in each category is 81. For each importer and good we calculate the number of exporters from whom the importer purchases that good. We then express this as a ratio over the total number of exporters in that good. If an importer does not purchase a good from any exporter, the ratio is zero. Complete specialization models used to derive the gravity equation predict that the ratio should be one for all goods and countries. Using four-digit SITC level we have 173 countries and 438 4-digit sectors, for 75774 total data points.

The message of Figure 1 is clear. Countries do not buy all available goods, as zero values (no imports from any source of that good) represent fully 27 percent of the distribution. Conditional on importing the good from at least one exporter, countries buy from very few exporters. In 58 percent of these cases, importers buy from fewer than 10 percent of available exporters, and in 99.4 percent of the cases fewer than half of the available varieties purchased. Considering only rich (OECD) countries or only manufactured goods, the ratio increases but the distribution remains bunched much closer to zero than to one.<sup>19</sup> This data pattern is not unique to this year or this dataset. A very similar pattern can be found in the World Trade Database for every year from 1970-1997. UN COMTRADE data from 1962-1983 evaluated at the 4 and 5 digit SITC level exhibits the same pattern but to an even greater degree – in *eighty* percent of observations importers buy from fewer than 10 percent of available suppliers.

Of course, the model captured in equation (3) may not apply to all goods or all countries. It is therefore unsurprising to find some values less than one. The surprising thing is that in *none* of the 75,774 cases does an importer avail itself of all varieties. While the median number of suppliers in a good is 81, the modal value for number of partners with which a country trades a particular good is zero; conditional on buying from someone, the modal value is one.

This rejection is made more striking by the fact that our data examination is biased in favor of complete specialization models used to generate the gravity equation. Strictly speaking, equation (3) implies that if a country *produces* a good, all other countries import it. These data cannot identify cases in which countries produce (but fail to export) certain goods.<sup>20</sup> Also, we cannot identify the full breadth of varieties available from an exporter. There are many car models produced in the US, but we can only determine if at least one model is purchased. While we have no direct evidence on the full set of varieties produced, we suspect that the fraction of available varieties that are actually imported is even lower than the figure suggests.<sup>21</sup>

---

<sup>19</sup> Over the whole sample, importers buy from a median of 4 percent of exporters; for the OECD importers only it is 25 percent; for manufactures over all importers, 7.4 percent; for manufactured imports into the OECD, 29 percent.

<sup>20</sup> This would require a comprehensive and detailed set of production data and existing sources, such as the OECD COMTAP or STAN databases, are too aggregated and too limited in geographic scope to be of use.

<sup>21</sup> This conjecture is supported by calculations performed at the one-, two-, and three-digit level. Higher levels of aggregation pull the distribution further to the right, though even for very broad categories it is clear that importers buy only a fraction of available varieties.

What models would deliver Figure 1? In the most common formulation of the section II model, preferences are CES over varieties of goods, and monopolistically competitive firms produce unique, differentiated, varieties in order to charge markups. Introducing non-infinite iceberg transport costs to this model will diminish trade, but never force it to zero.<sup>22</sup>

To generate the data pattern in the figure requires a model in which importers do not consume every variety. Consider several examples. First, if there are multiple producers of a good that consumers regard as homogeneous, they will buy only from the low cost exporter. Thus, our alternative hypothesis of incomplete specialization can be used to reconcile a decidedly inconvenient fact from the complete specialization framework used to derive gravity equations.

Figure 1 can be reconciled with a complete specialization model coupled with some reason that consumers value and purchase only a subset of available varieties. For example, with preferences as in Lancaster (1979), importers only value and therefore will purchase a small number of “ideal” varieties. Even if the utility function exhibits some “love of variety”, the presence of fixed costs of trade (see Romer (1994) and Klenow and Rodriguez (1998) for two examples) or declining marginal utility of varieties (see Ottaviano, et al. (1999)) will cause importers to buy only a handful of varieties.

Clearly then, we cannot reject these other hypotheses as explanations for the Figure 1 result, and indeed they seem like sensible directions to pursue.<sup>23</sup> However, their implications for trade volumes and for the pattern of bilateral trade differ markedly from the simple world captured in the Section II model. For example, in a model with fixed costs of trade, larger countries are willing to pay the fixed costs for a greater set of varieties and therefore trade more as a fraction of income than do small countries. This exactly reverses the implication of equation (5). And with zeros, the bilateral pattern of trade becomes very difficult to resolve, and in many cases indeterminate. Absent some heterogeneity, the model implies that importers buy a subset of available varieties, but not *which* varieties.<sup>24</sup> Put another way, while it may be possible to reconcile Figure 1 with a modified “love of variety” preference formulation, it is much more difficult to reconcile the zero values with the model used to generate the gravity equation.

### *III.2 The Level of Trade*

Beginning with McCallum (1995), a number of authors have used gravity regressions to reveal a significant puzzle in international economics regarding the level of trade. The key idea is to estimate (6)

---

<sup>22</sup> One could argue that finite costs might diminish trade to the point that is observably zero, that is, below levels captured by trade statistics. If so, the zeros should be predictable in the sense that estimating the model with OLS on positive values should predict trade values extremely close to zero for all observed zeros. We have experimented with this and found that, for most observed zeros, the model predicts positive values of trade greater than the threshold value at which trade statistics are reported.

<sup>23</sup> We regard an investigation into partner selection, largely ignored to this point, as a promising method for learning about the nature of differentiation and trade barriers. See Haveman, et al. (1999) for an example.

treating domestic purchases as one of many bilateral “trade” options, and to include a dummy variable to indicate domestic purchases. These studies find large border effects – domestic purchases are many times larger than foreign purchases. The implication usually drawn is that international borders appear remarkably closed to trade. We argue here that trade levels are only puzzlingly low when evaluated in the context of a particular model of complete specialization. We briefly compare these model predictions to the data, and then turn to incomplete specialization as an alternative hypothesis for the “puzzle”.

Equations (4) and (6) provide predictions for the level of trade with and without trade barriers. We provide calculations showing how large barriers must be in order for the model with trade barriers to generate actual trade levels.<sup>25</sup> Suppose that the price of imported goods inclusive of ad-valorem barriers is constant over all external sources.<sup>26</sup> Using equation (6), we can then express the ratio of internal over external purchases (all imports,  $r$ ) as

$$(14) \quad \frac{M_{ii}}{M_{ir}} = \frac{s_i}{1-s_i} \frac{P_r t_{ir}}{P_i t_{ii}}^\sigma$$

$S = s_i/(1-s_i)$  is the predicted value of internal to external purchases provided that internal and external prices are equal. We observe  $S$  and  $M = M_{ii}/M_{ir}$ . Given values for  $\sigma$ , we can solve for the ratio of internal to external transactions costs implied by purchases. We provide calculations using aggregate trade data and use non-service sector gross output to construct the income share terms,  $s_i$ .<sup>27</sup> Multilateral imports are constructed from the WTDB. Internal purchases are non-service sector gross output less multilateral exports.

The first column of Table 3 reports  $M$ , the observed ratio of internal to external purchases, for 19 countries. The second column reports  $S$ , the predicted value of internal to external purchases under the assumption that internal and external prices are equal. The third column is  $M/S$ , and shows that the observed ratio of internal to external purchases is anywhere from 13 to 2000 times larger than what is predicted under the equal transactions cost case.

<sup>24</sup> See Klenow and Rodriguez (1998).

<sup>25</sup> We eschew the typical approach of estimating a gravity regression with border dummies for several reasons. One, the model implies particular elasticities for the income terms and we impose rather than estimate these elasticities. Two, estimation runs into a host of measurement problems that we avoid. For example, differences in measures of internal distance (how far is a country from itself?), the treatment of proxy variables like language and adjacency, and the manner in which the CES price index is handled all seem to dramatically affect estimates.

<sup>26</sup> This approach assumes there is little bilateral variation in transactions costs over external sources. This is precisely what gravity model-based studies imply. Border effects are many times larger than effects associated with exporter-varying barriers such as distance, language, or adjacency.

<sup>27</sup> To get non-service sector gross output, we use World Bank data to construct non-service sector GDP, then multiply by the gross output to value-added ratios found in Wei (1996). The difference between value added and gross-output is not a problem in the construction of income shares so long as the the ratio is similar for the importer and the world.

Several authors (Wei, 1996; Evans 1999) have noted that, in the context of this model, trade levels may be low because transactions costs are high or because the elasticity of substitution is high. In short, if  $\sigma$  is a free parameter it can be chosen to match any trade level given any transactions cost. How then can we use the level of trade to generate any inferences about the validity of the complete specialization model? Simply,  $\sigma$  is not a free parameter and can be identified by how bilateral import quantities respond to relative price variation.

Two recent papers that assume a complete specialization framework (in particular, monopolistic competition) and use aggregate bilateral trade data provide helpful estimates. Hummels (1999) uses the coefficient on the distance term from an aggregate gravity regression along with the technological relationship between freight and distance to identify  $\sigma$  and finds elasticities in the range of 2.2 to 4.6. Baier and Bergstrand (1999) relate time series variation in aggregate freight and tariff barriers to trade growth in a gravity framework and identify  $\sigma = 6$ .<sup>28</sup> Based on these estimates we use a value of  $\sigma = 5$  to interpret the trade data.

The third column of Table 3 can also be interpreted as the ratio of external to internal prices implied by trade flows in the  $\sigma = 1$  case. Exponentiating values in this column by  $(1/\sigma)$  yields the price ratio for higher substitution elasticities. The US is, according to this model, the most open to trade. External transactions costs raise the price of imported goods by only (!) 68 percent relative to domestic goods. Elsewhere, imported goods are from 2 to 4 times more expensive than domestic goods. A comparison to measurable trade barriers is useful here. US import data for 1994 indicate that aggregate expenditures on tariffs and transportation costs for imported goods are approximately 6 percent ad-valorem.

What conclusion do we draw from this exercise? Given evidence on the elasticity of substitution, complete specialization models predict extremely high trade levels. When filtering evidence on actual trade levels through these models, one must conclude that national borders are extremely costly barriers to surmount. While introspection and the size of measurable barriers suggest this interpretation is in error, we cannot formally reject it without knowing precisely what trade costs borders pose. Perhaps some enormous yet difficult to measure barrier looms at national borders.

Another interpretation is that borders pose a moderate barrier to trade and that complete specialization models dramatically over-predict trade volumes. The home bias literature assumes that the frictionless gravity equation (4) accurately predicts the trade volume that would prevail were barriers non-existent. And as the previous section shows, that model provides the limiting case for maximal trade flows given identical preferences. Models with incomplete specialization predict much lower trade

---

<sup>28</sup> Estimates on sector level data yield higher elasticity estimates, on the order of 7 to 8. See Hummels (1999).

volumes, and our 2x2 HO model employing world data on endowments and incomes generates trade volumes that are on the order of those observed in the world.<sup>29</sup> In that model domestic demands are first met by domestic supplies when both home and foreign producers manufacture the same good. This can be justified by very small border costs.

## V. Implications: Does it Matter Why the Gravity Equation Works?

Previous authors have shown that the gravity equation can be derived from any number of underlying models, so simply adding another model to the list is of limited consequence. A case must be made that the explanation for why the gravity equation works has some relevance for how the gravity equation is interpreted and used, and how we view bilateral trade.

The most immediate implication is seen in the previous section. Viewed through the lens of complete specialization models used to derive the gravity equation, both the level of trade and the existence of zeros in the bilateral trade matrix are significant puzzles. Much ink has been spilt trying to understand why home bias appears to so strongly characterize trade flows. Under the alternative hypothesis, *there is no puzzle*.

### V.1 Normative Implications of Bilateral Trade

The normative consequences of deviations from the gravity equation prediction for trade differ considerably under the complete and incomplete specialization models. Under complete specialization, the gravity equation is not merely a positive description of bilateral trade; it has normative content as well. Consider the most common version of the model that proceeds from CES preferences and is captured in equation (6). It is a straightforward exercise to evaluate the utility function using observed trade quantities. In this model deviations in bilateral trade from the predicted baseline result in direct and quantifiable welfare losses.<sup>30</sup>

The normative implications under incomplete specialization are quite different. As we argue in the previous section, indeterminacy may result from a world of costless trade precisely because importers are indifferent to multiple sources of supply. However this trade is arranged, whether it looks country pair by country pair exactly like the frictionless gravity equation as Deardorff suggests it may, or whether it deviates wildly from this benchmark, it simply does not affect welfare. In the case where indeterminacy is resolved via trade barriers, Table 2 indicates that something like the gravity equation appears to hold as a kind of “on-average” statistical relationship. Yet many country pairs will exhibit

---

<sup>29</sup> Another possibility is that fixed costs of trade prevent some firms from exporting. See Evans (2000).

<sup>30</sup> See Lai and Trefler (1999).

trade looking nothing at all like equation (6), and many of these values will be zero, a prediction confirmed for the real world by the data in Figure 1. Still, bilateral deviations that leave net supplies and terms of trade unchanged are of no normative consequence because under the alternative hypothesis, importers do not attach any value to multiple sources of supply. They care only for purchasing the goods from the low cost supplier.

A particular example of this logic can be seen in the debate over regional trade agreements. Krugman (1991) models the welfare gains from regional versus multilateral integration in a monopolistic competition framework, and Frankel, Stein and Wei (1995) implement the model empirically using a gravity regression. One basic insight is that regional agreements can be costly relative to multilateral liberalization because consumers value distinct varieties produced outside the bloc. Indeed, as Deardorff and Stern (1994) note, trade diversion must exist in this model because all outside varieties are distinct and valued. They show that the welfare consequences are very different in a neoclassical model in which the goods produced inside the bloc are identical to those produced outside the bloc. Here regional integration may not be costly precisely because there is no value to purchasing the homogeneous good from multiple sources.

## *V.2 Evidence on the Nature and Extent of Specialization*

Several authors have examined the elasticity of bilateral trade with respect to partner incomes hoping to generate inferences regarding the nature of specialization. Our findings on incomplete specialization provide some insight into this literature.

Hummels and Levinsohn (1995) find that a variant of the gravity model works well even for non-OECD countries whose trade does not appear well described by monopolistic competition. One response is to suggest, in the vein of Anderson (1979) and Deardorff (1998), that these countries trade goods whose production is completely specialized (though this inference is problematic given the importance of commonly produced bulk commodities in the trade of the non-OECD countries that Hummels and Levinsohn examine). An alternative response inspired by Table 2 is that it matters little whether these countries trade completely or incompletely specialized goods. Something like the gravity equation would have worked in any case.

DeBaere (1999) suggests that the broad success of the gravity equation, and a resolution of the Hummels and Levinsohn puzzle can be found in the particular specification employed. Denoting the sum of partner's income as  $(Y_{ij} = Y_i + Y_j)$ , we can rewrite the simple gravity equation as

$$M_{ij} = (Y_{ij})[1 - (Y_i / Y_{ij})^2 - (Y_j / Y_{ij})^2].$$

The income sum is a scaling term and the term in the square bracket is a measure of size dispersion that takes on maximal value when the countries are of equal size. DeBaere claims (but does not derive formally) that bilateral trade volumes will be correlated with the scaling term in any trade model whereas the size dispersion term will affect bilateral trade only under complete specialization (monopolistic competition).<sup>31</sup>

Employing DeBaere's formulation and using the data from the HO numerical model in which bilateral trade is resolved with the minimum distance rule, we estimate

$$\ln(M_{ij} / Y_{ij}) = -6.87 + 7.8 \cdot \ln(1 - (Y_i / Y_{ij})^2 - (Y_j / Y_{ij})^2) + e_{ij}$$

$$R^2 = .49$$

With a t-stat of 14, we find that size dispersion remains an important determinant of bilateral trade even under the HO model, thereby rejecting DeBaere's inference.

Feenstra, Markusen, and Rose (1999) estimate a gravity relationship on disaggregated trade data as a test of a reciprocal dumping model. They find income coefficients significantly less than one for estimates involving reference priced commodity groups, and take this as evidence for the dumping model. An alternative explanation for these results is that the reference priced commodities are best described by the HO model (that is, homogeneous goods produced under perfect competition). Our results suggest income coefficients for trade involving these goods will be positive, but less than one.

Finally, our results provide some insights into the *extent* of specialization, in that gravity equations cannot be regarded as arising only from complete specialization and that corollary evidence (zeros, trade levels) is highly problematic for complete specialization models. We liken the insights to research provided by Don Davis (1995, 1997) into the existence of intra-industry trade, and the large volume of bilateral trade between similar countries. A narrow reading indicates that both facts are inconsistent with the simplest neoclassical models, and suggests the need for models (such as monopolistic competition) that can explain these facts. However, Davis shows that extensions of neoclassical models can explain both intra-industry and north-north trade. The point is not that simple neoclassical models are wrong, or that new trade theory models are unnecessary. Instead, the exercise helps us to identify which features of models are important for explaining trade patterns.

In this vein, we recognize that models with complete specialization are a critical component of our empirical and theoretical toolkits. Yet, our facts on trade levels and variety trade indicate something

---

<sup>31</sup> Hummels and Levinsohn estimate this equation as a product of these two terms, whereas Debaere moves the scaling term to the left hand side. His result is that size dispersion is significant for trade within OECD countries but not for trade outside the OECD.

is amiss with these models. In each case the data suggest that the complete specialization model considerably overstates either the extent of specialization (the degree to which goods are differentiated) or the degree to which consumers value that differentiation. Even if we are unwilling to go as far as the incomplete specialization framework suggested as an alternative here, these data should provoke significant amendments so that the basic models conform to the included facts.

### *V.3 Coefficient estimates – consistency under the alternative hypothesis*

The gravity equation is most commonly employed as a simple device for describing trade, and identifying correlates. What do our results say about this basic enterprise? Recall that, under the alternative hypothesis, multilateral trade and bilateral trade are mechanically linked through an adding up constraint while variables such as national incomes merely operate as good proxies for multilateral trade. That is, national incomes and capital/labor differences explain multilateral trade, not bilateral trade. Because these variables may be noisy proxies for the true variable of interest, they are subject to the usual errors in variables problem. Attenuation biases their coefficients toward zero, with the bias growing worse as the proxies get noisier.

To show this property we alter the endowment data from our HO model so that the variable  $kdif_i = |k_i - k_w|$ , varies over countries  $i$  in a way that is uncorrelated with incomes  $Y_i$  (i.e., on average, large and small countries have endowments that are equally far from the world endowment). We re-solve the model for output and multilateral trade levels in each of our 96 countries, and estimate the income elasticity of multilateral trade in the manner of equation (9). With this setup, the estimated elasticity is one, but the relationship between trade and income is noisy. The amount of noise is described by the regression  $R^2$  from estimating equation (9). We repeat this exercise 50 times, varying  $kdif$  but holding  $cov(kdif, Y) = 0$  so that the estimated multilateral income elasticity remains one. The  $R^2$ 's for the 50 draws are plotted on the horizontal axis in Figure 2.

Using the data generated from the HO model, we then distribute bilateral trade using the minimum distance algorithm and estimate a bilateral gravity equation in the manner of equation (11). The income coefficient from that regression is plotted on the vertical axis of Figure 2. This is done separately for each of the 50 draws. With multiple draws we generate a relationship between how noisy income is as a proxy for multilateral trade and the resulting bilateral income coefficients. Figure 2 shows that a noisier relationship between multilateral trade and income result in a stronger downward bias on gravity equation estimates, even though the mechanism for allocating bilateral trade remains constant.

We have focused on incomes for simplicity, but under the alternative hypothesis, the insights apply more generally. It is easy to generate examples where variables such as per capita incomes, capital/labor ratios, and MFN tariff levels only affect multilateral trade directly. However, they enter into

bilateral regressions as proxies. As with the income example just shown, the noisier proxies will enter with downward biased coefficients.

#### IV.4 Interpreting Regression Fit

Finally, much ado is made about the high  $R^2$ s of gravity regressions.<sup>32</sup> This can be thought of in two ways. One, the included variables explain trade “well”. Two, the model is a useful baseline in the sense that it controls for many things that matter so we need worry less about what is omitted.<sup>33</sup> In this section we examine the assertion that the gravity model “fits the data well” under the null and alternative hypotheses.

The key to the analysis is a model-based approach to measuring regression fit in which we calculate the ad-valorem equivalent of the residuals.<sup>34</sup> Suppose that the null hypothesis of equation (6) is exactly correct so that if we have a complete characterization of trade barriers,  $t$ , we will estimate the model with no error. However, in practice we do not know  $t$  exactly, and variation in trade volumes due to this omission shows up in the residuals. Interpreting the residuals as missing components of  $t$ , we calculate from them the unexplained variation in trade barriers associated with unexplained variation in trade volumes.<sup>35</sup>

More formally, represent  $t_{ij}$  in (6) as  $\ln t_{ij} = \mathbf{d}_0 \ln DIST_{ij} + \mathbf{d}_1 LANG_{ij} + \mathbf{d}_2 ADJ_{ij} + \ln \mathbf{t}_{ij}$  where distance, language, and adjacency are common correlates used in gravity regressions, and  $\mathbf{t}_{ij}$  is the ad-valorem equivalent of the unobserved barrier. This is estimated by

$$(15) \quad \ln M_{ij} = a + \mathbf{b}_1 \ln Y_i Y_j + D_0 \ln DIST_{ij} + D_1 LANG_{ij} + D_2 ADJ_{ij} + e_{ij}$$

Using equation (6) we note that the size of the barrier implied by the residual is  $\mathbf{t}_{ij} = \exp(-e_{ij} / \mathbf{s})$  while the ad-valorem equivalent of the included variables can be similarly discerned ( $\mathbf{d}_n = D_n / \mathbf{s}$ ). This

---

<sup>32</sup> This sentiment is clear in several leading surveys of empirical evidence on international trade theory. Deardorff (1984) writes that gravity models are “extremely successful empirically” as witnessed by the variance explained, while Leamer and Levinsohn (1997) write that gravity models “have produced some of the clearest and most robust empirical findings in economics.” McCallum’s (1995) famous piece on national borders justifies the use of the gravity equation entirely by citing Deardorff’s quote.

<sup>33</sup> Wei (1998), a well-known application that uses the gravity model to assess the effect of exchange rate variability on trade, argues for using the gravity model on the grounds that it outperforms factor endowment models in an  $R^2$  sense, and that adding factor endowment terms to a gravity specification does little to improve fit.

<sup>34</sup> This technique is closely related to a calculation done by Hummels (1999). See Leamer (1988) for a calculation involving a factor endowments model that is similar in spirit.

<sup>35</sup> Of course, for this to be exactly correct, the unobserved barriers must be uncorrelated with the included regressors.

interpretation of the residual underscores the limited content of traditional measures of regression fit focused on unexplained variation in trade volumes. Suppose there are two states of the world, a high  $\sigma$  state and a low  $\sigma$  state, and that trade patterns are identical in both cases. With identical trade patterns, a regression  $R^2$  will yield the same measure of model fit in both states. However, in the high  $\sigma$  world, the residuals correspond to relatively small unobserved trade barriers, i.e., whatever is left does not much affect prices.

We estimate (15) using all (aggregate) bilateral trade flows reported in StatCan World Trade Database in 1990. We estimate the constant and the income terms rather than directly imposing values consistent with equation (6). This abstracts from concerns about the level of trade and allows us to focus entirely on the bilateral distribution of trade.<sup>36</sup> We also omit country pairs with zero trade values, as implied trade barriers must be infinitely large to explain a zero trade in this model. Note that both of these corrections make the residuals look much smaller than would be the case in an estimate strictly guided by the underlying theory.

$$\ln M_{ij} = -13.9 + .9 \ln Y_i Y_j - 1.03 \ln DIST_{ij} + .56 LANG_{ij} + .63 ADJ_{ij} + e_{ij}$$

$$R^2 = .62 \quad obs = 9083 \quad RMSE = 2.2$$

All included variables are significant at the 1% level. For each  $i,j$  pair there is a unique residual so we characterize the average unmeasured barrier as the root mean squared error of the regression. At  $\sigma=5$  (the preferred value from Section III.2) the RMSE is equivalent to a 55 percent tariff, while speaking a different language raises costs by 11 percent, sharing a border lowers costs by 12 percent, and increasing distance by one standard deviation above the mean (a 9 percent increase) increases costs by  $(1.03/5) * 0.09 = 1.8$  percent ad-valorem. The ad-valorem equivalent of the regression error is large both absolutely and relative to the other barriers (more than twice as large as the combined effect of traditionally included variables). Suppose we reject the literature's estimate of  $\sigma$  and claim a much higher value. This would lower the ad-valorem equivalent of the residual in absolute terms. However, it would not change the importance of the residual relative to the included barriers, as all are proportional to  $\sigma$ .

What do we conclude from this exercise? Under the null hypothesis of complete specialization captured in equation (6), the interpretation is that we have identified a few variables that are robustly correlated with trade volumes, but we simply have not done an adequate job of explaining barriers. Of course, our barrier list is far from comprehensive and the model might fare much better with explicit

measures of tariffs and transport costs (not to mention the relative price terms). But the simple form estimated here is very common in the literature, and generally hailed as providing a good explanation of the data. This interpretation is not sustainable, even under the null.

Under the alternative hypothesis of incomplete specialization, the interpretation is different. As goods become perfect substitutes, importers concern themselves only with relative prices. In this case, minute changes in trade barriers may generate enormous changes in bilateral trade volumes as importers shift to the low cost supplier. The implication is that the “missing barriers” captured by the residual, as well as the included barriers (distance, language, adjacency) may be quite small. To put this in the language of the null model, let  $\sigma$  go to infinity and evaluate the ad-valorem equivalents of

$$t_{ij} = \exp(-e_{ij}/s) \text{ and } d_n = D_n/s.$$

Even if we are only concerned with a positive description of bilateral trade flows, this implication is troubling. The interpretation under the alternative hypothesis suggests a considerable role for “noise” that may not yield to economic analysis, as small and difficult to measure price changes may swamp other effects.

## VI. Conclusions

There are three main contributions. First, we demonstrate that the gravity equation as a statistical relationship (and not as a specific prediction about bilateral trade) can be generated from a model with incomplete specialization. Second, we analyze two facts or “puzzles” that are broadly inconsistent with the complete specialization models typically used to derive the gravity equation. In both cases, the data are consistent with the alternative hypothesis of incomplete specialization. “Home bias”, for example, is only puzzling when viewed through the lens of the complete specialization model. Third, we demonstrate that the explanation for why the gravity equation works has considerable relevance for how the gravity equation is interpreted and used, and how we view bilateral trade.

---

<sup>36</sup> We have already shown in Section III.2 that the level of actual trade is much lower than is predicted by the complete specialization model.

## References

- Anderson, James (1979). "A Theoretical Foundation for the Gravity Equation." *American Economic Review* Vol. 69, 106-116.
- Bergstrand, J (1985). "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence." *Review of Economics and Statistics*, Vol. 69, 474--481.
- Davis, Donald (1997) "Critical Evidence on Comparative Advantage? North-North Trade in a Multilateral World" *Journal of Political Economy* Vol. 105, pp. 1051-60.
- Davis, Donald (1995) "Intra-industry trade: A Heckscher--Ohlin--Ricardo approach" , *Journal of International Economics* Vol. 39, pp. 201-226
- Deardorff, Alan, "Testing Trade Theories and Predicting Trade Flows". in Jones and Kenen (eds) *Handbook of International Economics*, vol 1. North Holland.
- Deardorff, Alan and Stern, Robert (1994) "Multilateral Trade Negotiations and Preferential Trading Arrangements," in Alan V. Deardorff and Robert M. Stern (eds.) Analytical and Negotiating Issues in the Global Trading System, Ann Arbor, MI: The University of Michigan Press.
- Deardorff, Alan (1998). "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" in Jeffrey Frankel (ed.), The Regionalization of the World Economy University of Chicago Press: Chicago.
- DeBaere, Peter (1999), Testing "New" Trade Theory without Testing for Gravity: Re-interpreting the Evidence" mimeo, University of Texas.
- Evans, Carolyn (2000), "The Economic Significance of National Border Effects", mimeo.
- Evans, Carolyn (2000), "National Border Effects, Heterogeneous Fixed Costs of International Trade, and Variety Availability", mimeo.
- Evenett, Simon and Wolfgang Keller (1997). "On Theories Explaining the Success of the Gravity Equation." NBER Working Paper 6529.
- Feenstra, Robert C., James A. Markusen and Andrew K. Rose (1998). "Understanding the Home Market Effect and the Gravity Equation: The Role of Differentiating Goods," NBER Working Paper 6804.
- Frankel, J., Stein, E. and Wei, S. (1995). "Trading Blocs and the Americas: The Natural, the Unnatural, and the Super-Natural." *Journal of Development Economics*, Vol. 47, 61--95.
- Grossman, Gene. {1998} "Comment on `Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?'" in Jeffrey Frankel (ed.), The Regionalization of the World Economy University of Chicago Press: Chicago.
- Harrigan, James (1993). "Scale Economies and the Volume of Trade," *Review of Economics and Statistics*, Vol. 76, 321--28.
- Haveman, Jon, Usha Nair-Riechert and Jerry Thursby (1999). "Trade Reduction, Compression, and Diversion: Empirical Regularities in the Effects of Protective Measures," mimeo, Purdue University.
- Helpman, Elhanan. (1987). "Imperfect Competition and International Trade: Evidence from 14 Countries," *Journal of the Japanese and International Economies*, Vol. 1, 62--81.

- Helpman, Elhanan and Krugman, Paul (1985). Market Structure and Foreign Trade. MIT Press.
- Hummels, David and Levinsohn, James (1995). "Monopolistic Competition and International Trade: Reconsidering the Evidence." *Quarterly Journal of Economics* Vol. 110, 799--836.
- Hummels, David (1999), "Toward a Geography of Trade Costs", mimeo.
- Jensen, Paul (2000), "Analysis of Bilateral Trade Patterns with Panel Data", *Review of International Economics*, Vol. 8.
- Klenow, Peter and Andres Rodriguez-Clare (1996). "Quantifying Variety Gains from Trade Liberalization," mimeo, University of Chicago.
- Krugman, P. R. (1991), "Is Bilateralism Bad?" in E. Helpman and A. Razin (eds.), International Trade and Trade Policy. Cambridge: MIT Press.
- Lai, Huiwen and Trefler, Daniel (1999), "The Gains from Trade: Standard Errors with the CES Monopolistic Competition Model", mimeo, University of Toronto.
- Lancaster, Kelvin (1979), Variety, Equity and Efficiency. New York: Columbia University Press.
- Leamer, Edward and Levinsohn James, "International Trade Theory, the Evidence" in Grossman and Rogoff (eds) *Handbook of International Economics*, Vol 3. North-Holland
- Limao, Nuno and Venables, Anthony (1999), "Geographical Disadvantage: a Heckscher-Ohlin-VonThunen model of International Specialization" CEPR Discussion Paper 2305.
- McCallum, John (1995). "National Borders Matter: Canada-U.S. Regional Trade Patterns." *American Economic Review*, Vol. 85, June, 615--623.
- Ottaviano Gianmarco, Takatoshi Tabuchi, Jacques-Francois Thisse, (1999), "Agglomeration and Trade Revisited" mimeo.
- Romer, Paul M. (1994). "New Goods, Old Theory, and the Welfare Costs of Trade Restrictions," *Journal of Development Economics*, Vol. 43, 4--38.
- Wei, Shang-Jin (1996). "Intra-National versus International Trade: How Stubborn are Nations in Global Integration?" NBER Working Paper 5531.
- Wei, Shang-Jin (1998). "Currency Hedging and Goods Trade" NBER Working Paper 6742.

## Appendix A – Simulation Methodology

In the text, multilateral trade flows are generated in a 2x2xn numerical Heckscher-Ohlin model. The solution technique for this model is discussed in this appendix.

We assume a Cobb-Douglas production technology for the two goods,

$$Z_1 = K^a L^{1-a}$$

$$Z_2 = K^{1-a} L^a$$

and experiment with the parameters, varying  $\alpha$  from 0.55 to 0.95. Preferences are also Cobb-Douglas, with equal expenditure shares on each good. Using our assumptions on taste and technology, and using data on endowments, we solve for the output and multilateral volume of trade in both goods for each country.

In order to determine the output of each good in each country, we need to determine the relative factor use for each good. For countries outside of the cone of diversification, this is trivial. For countries inside the cone of diversification, factor intensities are determined from the following two equations:

$$P_{Z1} MPL_{Z1} = P_{Z2} MPL_{Z2}$$

$$P_{Z1} MPK_{Z1} = P_{Z2} MPK_{Z2}$$

or more precisely, denoting  $f(k/l)$ ,  $g(k/l)$  as the production functions for goods one and two,

$$\left( \frac{P_{Z1}}{P_{Z2}} \right) f' \left( \frac{L_{Z1}}{K_{Z1}} \right) - g' \left( \frac{L_{Z2}}{K_{Z2}} \right) = 0$$

$$\left( \frac{P_{Z1}}{P_{Z2}} \right) \left[ f \left( \frac{L_{Z1}}{K_{Z1}} \right) - \left( \frac{L_{Z1}}{K_{Z1}} \right) f' \left( \frac{L_{Z1}}{K_{Z1}} \right) \right] - \left[ g \left( \frac{L_{Z2}}{K_{Z2}} \right) - \left( \frac{L_{Z2}}{K_{Z2}} \right) g' \left( \frac{L_{Z2}}{K_{Z2}} \right) \right] = 0$$

These are two equations in three variables,  $\frac{L_{Z1}}{K_{Z1}}$ ,  $\frac{L_{Z2}}{K_{Z2}}$ ,  $\frac{P_{Z1}}{P_{Z2}}$ .

Discovering the equilibrium values of all three variables is then a matter of searching over prices for that corresponding to factor intensities resulting in balanced world trade. The algorithm that determines equilibrium prices is one that searches over prices for those that equalize world imports and exports of both goods. The algorithm proceeds as follows. First, pick some relative prices of the final goods, the ratio of the world capital supply to the world labor supply is generally used here. For a given price ratio, the equations above uniquely determine the borders of the cone of diversification, the labor-capital ratios used in the production of the two goods by countries that do not specialize. Within the cone of diversification, the supply of labor and capital to each industry can be determined from full employment conditions on the factor endowments, and noting that the factor usage will be identical to the factor usage that define the cone. If trade does not balance for the chosen price vector, the relative price of the good in excess demand rises, and the algorithm is repeated until trade is balanced.

Figure 1 – “Variety” in Trade

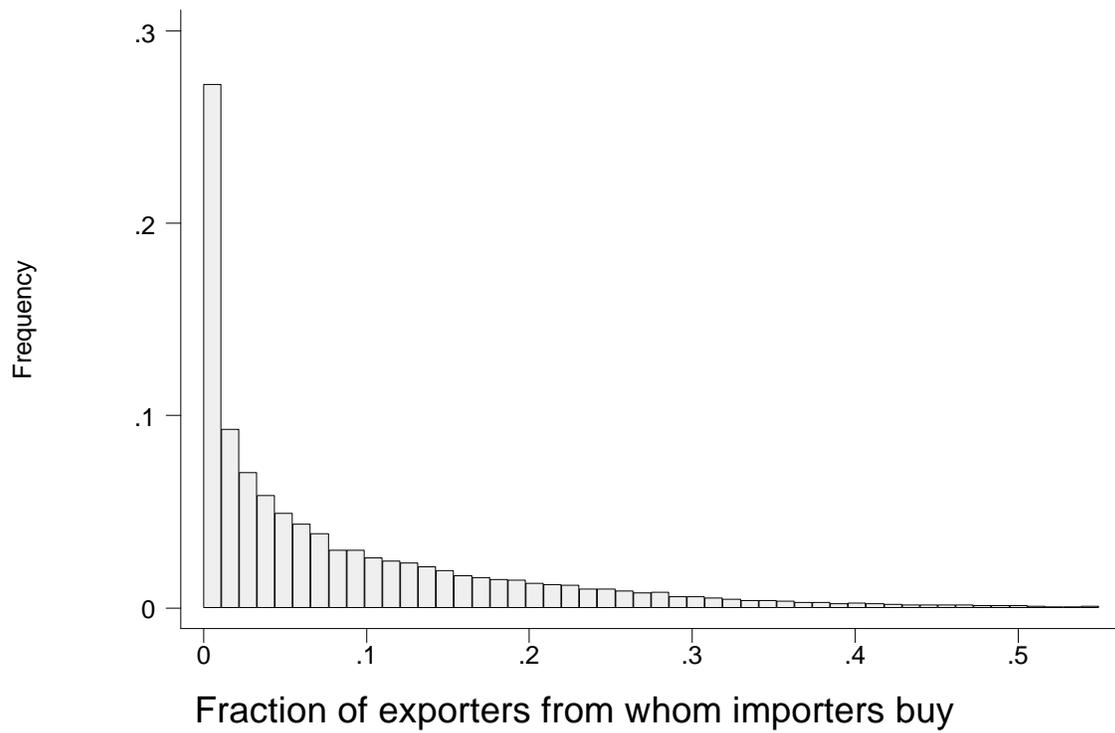
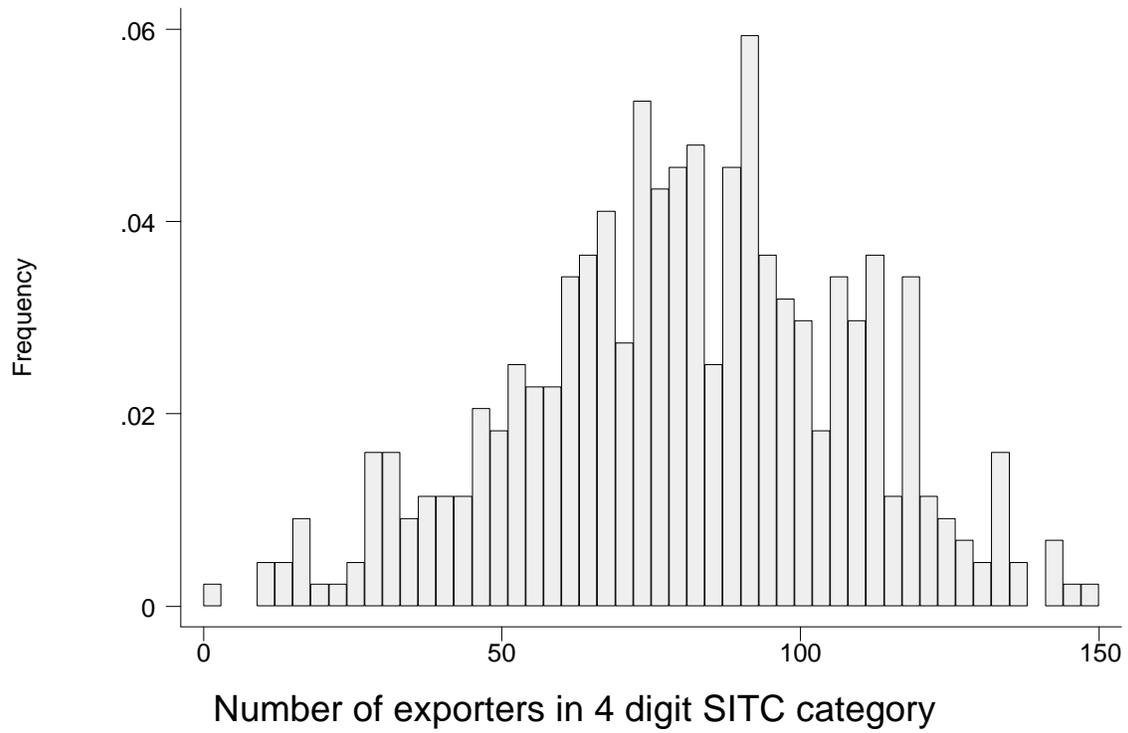


Figure 2: Attenuation Bias under Alternative Hypothesis

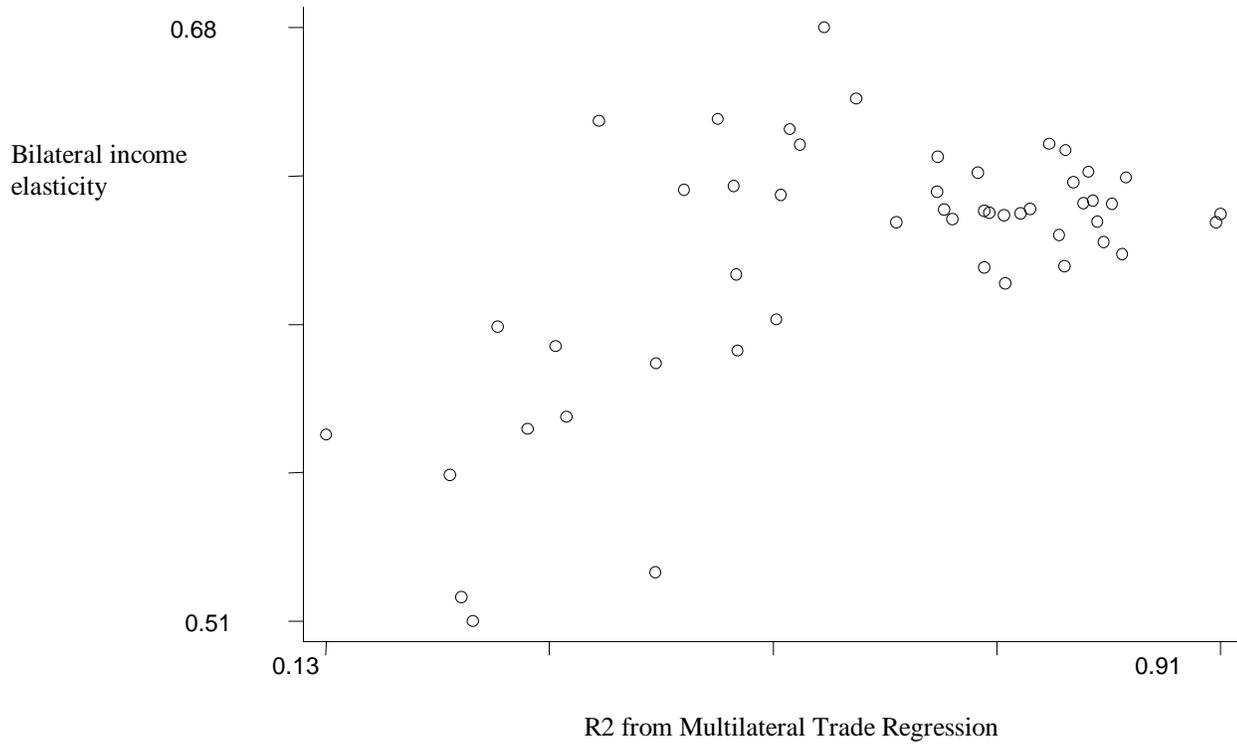


Table 1  
Elasticity of Multilateral Trade with Respect to Income

	HO Simulation		Real data
Income	0.81 (.04)	0.96 (.03)	0.92 (.03)
KLDIF		1.05 (.09)	
R2	0.83	0.92	0.93
N	96	96	96

Table 2  
Gravity Equation Estimates

	HO Simulation		Real data	
ln Yi Yj	0.57 (.03)	0.66 (.03)	0.9 (.01)	0.61 (.03)
kldif(i)		0.59 (.11)		
kldif(j)		0.59 (.11)		
R2	0.62	0.69	0.51	0.48
N	190	190	8359	189105

Table 3  
The Level of Trade

Country	M= Relative purchases	S= Relative size	Ratio of external to internal prices	
			$\sigma=1$	$\sigma=5$
Australia	4.37	0.013	334.0	3.20
Austria	2.77	0.008	328.8	3.19
Canada	2.30	0.025	90.7	2.46
Denmark	1.60	0.006	255.3	3.03
Finland	3.54	0.005	771.4	3.78
France	3.27	0.063	51.8	2.20
Germany	2.36	0.089	26.6	1.93
Italy	6.47	0.058	111.9	2.57
Japan	16.60	0.202	82.1	2.41
Mexico	4.43	0.015	298.5	3.13
Netherland	1.37	0.015	93.0	2.48
New Zeala	3.67	0.002	2037.1	4.59
Norway	1.49	0.005	298.9	3.13
Portugal	3.27	0.004	847.2	3.85
Spain	5.14	0.026	194.3	2.87
Sweden	3.22	0.011	293.0	3.11
Switzerlanc	1.94	0.012	168.7	2.79
UK	2.79	0.050	56.2	2.24
US	5.01	0.370	13.6	1.68