ON THEORIES EXPLAINING THE SUCCESS
OF THE GRAVITY EQUATION

Simon J. Evenett
Wolfgang Keller

Working Paper **6529**

ON THEORIES EXPLAINING THE SUCCESS
OF THE GRAVITY EQUATION

Simon J. Evenett
Wolfgang Keller

ABSTRACT

We analyze two main theories of international trade, the Heckscher-Ohlin theory and the Increasing Returns trade theory, by examining whether they can account for the empirical success of the so-called Gravity Equation. Since versions of both models can generate this prediction, we tackle the *model identification problem* by conditional bilateral trade relations on factor endowment differences and the share of intra-industry trade, because only for large factor endowment differences does the Heckscher-Ohlin model generate specialization of production and the Gravity Equation, and it predicts inter-, not intra-industry trade. There are three major findings: First, little production is perfectly specialized due to factor endowment differences, making the perfect specialization version of the Heckscher-Ohlin model an unlikely candidate to explain the empirical success of the Gravity Equation. Second, increasing returns are important causes for perfect product specialization and the Gravity Equation, especially among industrialized countries. Third, to the extent that production is not perfectly specialized across countries, we find support for *both* Heckscher-Ohlin and Increasing Returns models. Based on these findings, we argue that both models *explain difference components* of the international variation of production patterns and trade volumes, with important implications for productivity growth, labor and macroeconomics.

Simon J. Evenett
Department of Economics
Rutgers University
New Brunswick, NJ 08901-1248
evenett@fas-econ.rutgers.edu

Wolfgang Keller
Department of Economics
University of Wisconsin
Madison, WI 53706
and NBER
wkeller@ssc.wisc.edu

1. Introduction

The so-called Gravity Equation of trade predicts that the volume of bilateral trade is positively related to the product of the countries' gross domestic products (GDPs) and negatively related to trade barriers between trade partners. Empirical research has found that various versions of the Gravity Equation well describe the variation in the volume of trade across country-pairs as well as over time (see Leamer and Levinsohn 1995).

Since Anderson (1979) it has been increasingly recognized that the gravity equation prediction can be derived from very different structural models, including Ricardian models, Heckscher-Ohlin (H-O) models, and increasing returns to scale (IRS) models.[1] When consumers have both identical homothetic preferences and access to the same goods prices, a sufficient condition for obtaining a gravity equation is perfect product specialization, in the sense that each commodity is produced in only one country. The three types of trade models differ in the way product specialization is obtained in equilibrium: Technology differences across countries in the Ricardian model, factor proportions outside the vector space-of-diversification in the H-O model, and increasing returns at the firm level in the IRS model. Hence, as Deardorff (1998) has recently emphasized, the gravity prediction per se cannot be used to test any of these trade theories. Yet, the gravity prediction constitutes, perhaps along with the H-O-Vanek factor service trade prediction (see recent work by Trefler 1995, Davis et al. 1997), the most important result regarding the volume of international trade. Therefore, major insights are to be gained if it could be determined which model generates gravity-like trade volumes in a given sample of data, a problem we refer to as a model iden-

---

[1] Bergstrand (1990), Deardorff (1998), Eaton and Kortum (1997), Helpman and Krugman (1985), Leamer (1992), and Markusen and Wigle (1990).

tification issue. In addition, it would allow us to assess the suggestion by Hummels and Levinsohn (1995, 828) that something other than IRS is responsible for the empirical success of the gravity equation.

In this paper, first we address this identification problem by noting that, on the one hand, in a constant returns (CRS) H-O world, bilateral factor proportions differences must be very large in order to ensure that the economies lie outside a common space of diversification and to generate product specialization (see Figure 1). On the other hand, when product specialization is the result of IRS, the gravity prediction can be obtained even when there are no factor proportions differences. This difference has the following implication for the type of trade in the two models: In the H-O model, trade is exclusively inter-industry trade, defined as trade in goods with different factor intensities. However for the IRS model at least some, and potentially all, trade is intra-industry trade. Consequently, in the analysis below, we will take samples with low shares of intra-industry trade in total trade together with large factor proportions differences as those where a model of H-O-based product specialization might be behind the Gravity Equation, and samples with high shares of bilateral intra-industry trade (irrespective of factor proportions differences) as those where IRS-based product specialization might drive the Gravity Equation.

This might suggest a country-by-country reconciliation of the perfect specialization models of the H-O and IRS-based trade theories. Along these lines, the perfect specialization H-O theory would be responsible for the gravity prediction's success in explaining the variation in bilateral trade flows among countries with large factor proportions differences and low shares of intra-industry trade (the so-called 'North-South' trade). At the same time, the IRS theory would account for the success of the gravity prediction in trade between

2

countries where product differentiation and intra-industry trade is prevalent (the so-called 'North-North' trade.) However, no such reconciliation emerges from our analysis of a large and heterogeneous set of bilateral trade relations in the year 1985. We present evidence that perfect product specialization due to factor proportions differences is not a major part of explaining the success of the gravity prediction. In contrast, there is robust evidence that IRS-based theories of trade are an important reason why the gravity equation fits trade flows among industrialized countries well.

Secondly, the paper assesses the empirical relevance of the IRS-based trade models in general, with their far-reaching implications for productivity growth, labor, and open macroeconomics. This is more important now than ever before, as it has been demonstrated that there is no need to resort to IRS models to explain intra-industry trade (Davis 1995), or large trade volumes between countries with similar factor endowments (Chipman 1992). Even at an empirical level, Helpman's (1987) work showing that key implications of the IRS-based model are consistent with OECD countries' trade data was followed up by Hummels and Levinsohn (1995), who repeated Helpman's analysis with a set of non-OECD countries, the trade between which was not expected to contain much IRS-based trade. They showed that many correlations found by Helpman continue to hold in their non-OECD sample. To take Hummels and Levinsohn's result as evidence against the IRS-based trade models would be incorrect,[2] however, because their finding is just an expression of the model identification problem. In this paper, we account for both Helpman's (1987) and Hummels and Levinsohn's (1995) findings by showing that the former obtained his results because IRS-based

---

[2]Hummels and Levinsohn (1995, abstract) in particular are led to "question the apparent empirical success" of the IRS-based trade models.

trade is prevalent among OECD countries, whereas the latter found a similar correlation among non-OECD countries because there is some (but not perfect) product specialization driven by factor proportions differences.

More generally, we present strong evidence that the volume of international trade is determined by the extent of product specialization, which in turn is due to IRS and factor proportions differences. Where there is little or no two-way trade between countries, we find that a H-O model that predicts imperfect specialization better explains trade flows than a H-O model with perfect specialization. Also where there is two-way trade among nations, international trade flows are better explained by a model that incorporates both factor proportions differences and IRS than by a model where IRS alone generates perfect product specialization. These findings highlight the significance of both factor proportions differences and IRS as determinants of the extent of specialization and international trade flows.[3]

There are several important caveats to our approach: First of all, in our analysis of the gravity equation we focus on the proportionality of the volume of trade to the countries' GDPs rather than its relationship to trade resistance. This corresponds to the fact that to date there is relatively little theory on what form the gravity equation takes in the presence of transport costs.[4] Although we take up the influence of trade resistance in section 4.2.1 below and conclude that our major findings are not sensitive to that omission, we plan to

---

[3]This parallels recent results by Antweiler and Trefler (1997), who nest IRS and CRS models in the H-O-Vanek factor service trade expression to estimate scale parameters from these input demand equations. The authors find that although a majority of industries seems to be well-characterized by the assumption of CRS, there is evidence for IRS in a number of sectors. Also see Davis and Weinstein (1996, 1997) who test for the presence of IRS by exploiting the 'home-market' effect prediction (strong idiosyncratic demand for a certain good means a country will be exporting the good in the presence of IRS and transport costs); their results so far are mixed.

[4]It includes, however, Anderson (1979), Bergstrand (1990), and Deardorff (1998).

address the issue more comprehensively in future work.

Our approach relies on identifying intra-industry trade with IRS-based trade and large factor proportions differences with H-O based trade. Therefore, the second caveat is that instead of IRS we cannot rule out that Ricardian technology differences are what is really behind intra-industry trade (see Figure 1). To make progress on this issue requires product-level estimates of production technologies across countries, which are unavailable. From more aggregated evidence, however, we do not expect that technology differences are as important in generating product specialization as is IRS.[5] Thirdly, we note that also in the H-O model it is possible to generate trade in goods of identical factor intensities if one follows Armington (1969) and assumes that products are differentiated by country of production. However, for all practical purposes, the Armington hypothesis is essentially non-falsifiable, so we feel we do not give up much by not being able to rule it out. More generally, by relying on the assumption of identical homothetic preferences across countries throughout, we do not give any role to demand as, for instance, emphasized by Markusen (1986). We also ignore unbalanced trade in the analysis below. Those effects play, as in Hummels and Levinsohn (1995), a quantitatively very minor role in what follows. Lastly, our model identification relies on factor input requirement differences between IRS industries being smaller than those between IRS and CRS industries, because if this were not the case, then large factor proportions differences cease to indicate H-O based trade. Studies estimating scale elasticities by industry (such as Antweiler and Trefler 1997, Table 7) suggest that this

---

[5]Estimates of total factor productivity (TFP) levels (proxying for levels of technology) for 2-3 digit SIC manufacturing industries tend to find that there are big differences across OECD countries, which is inconsistent with production in several countries according to the technology differences hypothesis. Moreover, there is ample evidence that the industries in which the share of intra-industry trade in total trade is high tend to be those where marketing, R&D, and other fixed costs lead to IRS in conjunction with product differentiation.

assumption is likely to hold.

The remainder is as follows. In the following section, we derive the gravity equation predictions of four different trade models, and describe in more detail how we attempt to identify a particular model in our empirical analysis. Section 3 discusses the data set employed in this study, whereas section 4 presents the key empirical results. Section 5 concludes.

2. Derivation of the Estimating Equations

2.1. The Gravity Equation with Perfect Product Specialization

Throughout the paper, we assume that consumers in all countries have identical homothetic preferences, balanced trade, and zero trade and transport costs. First, we consider a typical model as laid out in Helpman and Krugman (1985, Ch. 8.1), where there are two countries, i and j, and two goods, X and Z (see Appendix A for more details). Both X and Z come in many symmetric differentiated varieties which are identically produced with increasing returns to scale. With preferences valuing product variety, both countries will demand all foreign varieties according to the countries' GDP as a share of world GDP, denoted $s^c, c = i, j$. Given that IRS leads to perfect product specialization for every variety, country i's imports from j, denoted M^{ij}, will be

$$M^{ij} = s^i \left[p_x n_x^j x^j + n_z^j z^j \right].$$

(2.1)

Here, n_g^c is the number of varieties of good g, $g = X, Z$, in country c, and p_x is the relative price of good x. Country j's imports from i, M^{ji}, will be

$$M^{ji} = s^j \left[p_x n_x^i x^i + n_z^i z^i \right].$$ (2.2)

The terms in the hard brackets of (2.1) and (2.2) are equal to the GDP of country j (denoted Y_j) and country i (denoted Y_i) respectively. Substitution of Y_i and Y_j yields

$$M^{ij} = s^i Y^j = \frac{Y^i Y^j}{Y^w} = s^j Y^i = M^{ji}.$$ (2.3)

Equation (2.3) is the gravity prediction referred to above,[6] and hence, this IRS trade model is potentially candidate to explain the success of the gravity equation.

The derived equation (2.3) is very general (see Helpman and Krugman 1985); for instance, nothing has been said about factor price equalization, factor endowment differences across countries, or factor intensities in the production of goods X and Z. Equation (2.3) holds whenever there is perfect product specialization in equilibrium, all consumers face the same goods prices and have identical homothetic preferences, and trade is balanced. Therefore, if instead one assumes that goods X and Z are homogeneous and produced with CRS, one can still derive (2.3) when perfect product specialization is obtained in equilibrium. For the following exposition, we assume without loss of generality that country i is capital-abundant and good X capital-intensive. Consider a model where factor endowments between the

---

[6]This gravity prediction, as well as those below, is readily generalized for the case that tradable goods make up a share λ, $0 \leq \lambda \leq 1$ of GDP which is constant across all countries. In that case, one obtains $M^{ij} = M^{ji} = \lambda \frac{Y^i Y^j}{Y^w}$. The parameter λ varies from country to country (λ should be in the order of $0.3 < \lambda < 0.8$), affected by the size of the government and services sectors which are largely non-tradables.

countries differ by so much that each country specializes in the production of one good.

Let k^x and k^z denote the equilibrium capital-labor ratios which are employed in industry X and Z, respectively, in the integrated equilibrium (see Helpman and Krugman 1985, Ch.1). Then, if

$$\frac{K^i}{L^i} \geq k^x > k^z \geq \frac{K^j}{L^j}, \tag{2.4}$$

the capital-abundant country i will specialize in producing the capital-intensive good X, and the labor-abundant country j will produce only the labor-intensive good Z. If X^c and Z^c denote the production of good X and Z, respectively, then $X^i = X^w$ and $Z^j = Z^w$; further, it is clear that the value of good $X-$production in country i is equal to its GDP, i.e., $p_x X^i = Y^i$, and similarly, $Z^j = Y^j$. This means that

$$M^{ij} = s^i Z^j = s^i Y^j = \frac{Y^i Y^j}{Y^w}, \qquad M^{ji} = s^j p_x X^i = s^j Y^i = \frac{Y^i Y^j}{Y^w}, \tag{2.5}$$

which is identical to the gravity equation (2.3) above.

As discussed by Helpman and Krugman (1985), this analysis generalizes to the multi-sector, multi-factor, and multi-country settings as long as specialization is the equilibrium outcome in all sectors of all countries. In the IRS-based model from equation (2.1), if there more countries and/or more goods, country i will import the same share s^i of total production of each, in the same way country i imports that share of country-j's varieties X and Z. Product specialization for all goods due to factor proportions differences will not occur when there are more than two countries (or goods), but only two factors of production. With more than two factors, factor-specificity can generate product specialization. In the extreme case where the production of any good in any country requires inputs that are specific to that good

8

in that country, equation (2.3) will reappear. Therefore, for an appropriately defined metric of factor endowment differences, also the H-O rationale for the gravity equation generalizes to a multi-sector, multi-country, and multi-factor setting.

So far our discussion has left open the question of whether one can derive gravity-type import volume predictions for the case where at least some goods are produced in more than one country. While product specialization has recently emphasized to be an important phenomenon of the patterns of world production (e.g., Haveman and Hummels 1997), no doubt there are many commodities which are produced in several countries. We now turn to two of those trade models with imperfect product specialization: the simplest, textbook Heckscher-Ohlin model, as well as its generalization to include a IRS sector due to Helpman (1981).

2.2. Generalizations: Gravity Equations with Imperfect Product Specialization

First, we derive a gravity equation for the generalization of the H-O model where one sector (Z) produces a homogeneous good under CRS, whereas a second sector (X) produces a differentiated good under IRS (Helpman 1981; see also Helpman and Krugman 1985, Ch.7,8). There are two countries $(i$ and $j)$ and two factors $(K$ and $L)$. For endowments inside the factor price equalization set, the volume of bilateral trade, defined as the sum of a country's exports and imports, is given by

$$VT^{ij} = s^j p_x X^i + s^i p_x X^j + \left(Z^j - s^j Z^w \right), \qquad (2.6)$$

9

where the first term on the right hand side gives country i's exports (M^{ji}), and the remaining

two terms are country i's imports (M^{ij}).[7]

Let γ^c be the share of good Z in country c's GDP, $\gamma^c = \frac{Z^c}{p_x X^c + Z^c}$. It is easy to show that

the assumption that country i is capital-abundant implies that $0 \leq \gamma^i < 0.5$. The following

reproduces a result from Keller (1998).

Proposition 1. *If good X is capital-intensive and produced under IRS, good Z labor-*

intensive and produced under CRS, and country i is relatively capital-abundant, then country

i's imports from country j are given by the following gravity equation:

$$M^{ij} = \left(1 - \gamma^i\right) \frac{Y^i Y^j}{Y^w}. \tag{2.7}$$

Proof. With balanced trade, $M^{ij} = M^{ji}$; $M^{ji} = s^j p_x X^i$, and using the definition of γ^i

gives $M^{ji} = s^j \left(1 - \gamma^i\right) Y^i = \left(1 - \gamma^i\right) \frac{Y^i Y^j}{Y^w}$ ∎

This proposition states that for any value $\gamma^i > 0$, the level of bilateral imports is lower

than in the case where both goods are differentiated (compare 2.7 to the gravity equation 2.3).

Furthermore, as the share of homogeneous good production in GDP declines, the predicted

level of imports rises, and in the limit, as $\gamma^i \to 0$, the generalized gravity equation (2.7)

reverts back to the simple gravity equation (2.3) above. In a sense, therefore, the volume of

trade is higher, the lower is the share of homogeneous goods in GDP.[8] It is also interesting

---

[7]With zero trade costs and Z being homogeneous and produced under CRS, consumers are indifferent with regard to its country of origin. Therefore, underlying equation (2.6) is a "minimum trade-volume" assumption, that only a country's excess demand for a good is imported from abroad. It is well-known that the H-O model determines only net, but not gross, trade flows.

[8]Note that this finding is in part due to H-O reasons, because γ^i is inversely related to country i's capital-labor ratio. A decline in γ^i therefore implies an increase in the volume of imports due to an increase in the countries' factor proportions differences (triggering a production response a la Rybczynski). What we emphasize is that for given factor proportions differences, the more product specialization there is the higher is the level of imports; see below.

10

to note that except for country j's GDP, only the characteristics of the capital-abundant country i enter the bilateral imports prediction (2.7).

Second, we show the particular form of the gravity equation in the case of a simple $2 \times 2 \times 2$ H-O model; that is, now both good Z and good X are homogeneous and produced under CRS. The volume of trade between countries i and j is given by

$$VT^{ij} = p_x \left(X^i - s^i X^w \right) + \left(Z^j - s^j Z^w \right),$$ (2.8)

where, according to the H-O theorem, the capital-abundant country i exports the capital-intensive good X and imports the labor-intensive good Z.[9] For this model, the following proposition, reproduced from Keller (1998), states the corresponding gravity-type relationship. The proof is in Keller (1998).

Proposition 2. *If both goods are homogeneous and produced under CRS, with country i being relatively capital-abundant and good X being relatively capital-intensive, then country i's imports from country j are given by the following gravity equation:*

$$M^{ij} = \left(\gamma^j - \gamma^i \right) \frac{Y^i Y^j}{Y^w}.$$ (2.9)

The gravity equation in the imperfect specialization H-O case, equation (2.9), depends not only on the product of the GDP's in the familiar way, but also on γ^j and γ^i, which are characteristics of both countries. Note that as the capital-labor ratios in the two countries converge, so do γ^j and γ^i. In the limit, when the factor proportions in i and j are equal,

---

[9]The "minimal trade-volume" rule has again been used to derive (2.8).

we have that $\gamma^j = \gamma^i$, in which case equation (2.9) gives the familiar result that there is no trade in a H-O model when factor proportions are identical across countries. Equation (2.9) includes the volume of imports prediction of the multi-cone H-O model given in (2.5) as a special case, because as factor proportions differences between i and j increase, the share of GDP derived from good Z in country j, γ^j, approaches one, whereas the share of good Z in the GDP of country i, γ^i, tends to zero. Indeed, when $\gamma^j = 1$ and $\gamma^i = 0$, equation (2.9) reverts to $M^{ij} = \frac{Y^i Y^j}{Y^w}$, the gravity equation for the perfect specialization model (see 2.5). Also, we can rewrite equation (2.9) in the following way:

$$M^{ij} = \left(\gamma^j - \gamma^i\right) \frac{Y^i Y^j}{Y^w} = \left[\left(1 - \gamma^i\right) - \left(1 - \gamma^j\right)\right] \frac{Y^i Y^j}{Y^w}.$$

Therefore, if we denote the volume of imports prediction for the case where the production of both goods is specialized across countries, equations (2.3) and (2.5), by M_S; the generalized H-O case with specialization for one good, but not the other good (equation 2.7) with M_G; and the H-O case in which both countries produce all goods, equation (2.9), with M_H, the following inequalities hold, ceteris paribus:

$$M_S > M_G > M_H. \tag{2.10}$$

This confirms that the bilateral volume of imports is higher, the more product specialization there is.

2.3. Model Identification

In the preceding section, we have derived the specific form of the gravity equation of trade for four models: (1) Multi-cone H-O: factor endowment differences lead to perfect product specialization; (2) Pure IRS: IRS leads to perfect product specialization; (3) Generalized H-O: with one good produced using IRS technology and perfectly specialized, the other good being produced with CRS technology and not specialized, and (4) Uni-cone H-O: where both goods are produced using CRS technology and are not specialized. If we restrict ourselves to $2 \times 2 \times 2$ models, then M_S is the import prediction for models (1) and (2), M_G is the prediction for model (3), and M_H is the prediction for model (4).

It is unlikely, however, that any of the observed trade flows are solely determined by any one of these four archetypal models. First of all, the data comes from a world with more sectors, countries, and factors than our $2 \times 2 \times 2$ models. Secondly, there should be positive amounts of IRS-based trade even between countries with the lowest recorded shares of IRS-based trade. Similarly, we expect there to be some factor-proportions-based trade even among countries where the share of IRS-based trade is highest.[10] Observed bilateral trade among countries is likely to be the result of the combination of the determinants of trade flows formalized in the four models considered here, and perhaps by other determinants which we have not addressed. However, in different circumstances (such as different distributions of factor endowments across trading partners) we expect different trade models to account for large proportions of the observed trade flows. Our inferences are based on whether each of the four trade models actually performs better in the very sample(s) (which are identified using explicitly stated criteria) where one expects the theory to perform better.

---

[10]Note that this point does not rule out the generalized H-O model (3).

Consider a cross-section of country-pairs where there is little (or no) product specialization due to IRS, but the absolute difference between the two countries' factor proportions, denoted $FDIF$, differs from one pair to another.[11] From equation (2.4), we expect ceteris paribus more product specialization in a pair where $FDIF$ is larger relative to another country-pair where factor proportions differences are smaller. Consequently, we expect the gravity prediction (2.3), i.e., prediction M_S, to be more accurate for the pair where $FDIF$ is larger, compared to the other with the lower value of $FDIF$. This observation allows to identify the H-O motivation for product specialization and the gravity prediction: If factor proportions differences are important in explaining the success of the gravity equation, then the prediction M_S should fit better in samples where the observed factor proportions differences $FDIF$ are higher.

Moreover, we can test the multi-cone H-O model against the uni-cone H-O model for different values of factor proportions differences: If product specialization is principally the result of factor proportions differences, then the uni-cone H-O model should be preferred to the multi-cone H-O model in samples where the observed factor proportions differences $FDIF$ are lower.

We will employ the index proposed by Grubel and Lloyd (1975) to control for the extent of IRS-based trade. Trade due to IRS and product differentiation can result in a country simultaneously importing and exporting varieties of a particular product (intra-industry

---

[11]We treat from now on a country-pair observation drawn from the actual multilateral trading world as resembling the two-country world observations of models (1)-(4). This introduces a bias which is related to the fact that bilateral trade between any two countries is not necessarily balanced. Generally, the trade imbalances are smaller than the volumes of imports by a full magnitude, so we expect that our analysis will be a reasonable first-order approximation to a multilateral trading world. We plan to address this issue in future work, though.

14

trade).[12] The index, denoted GL^{ij}, measures the share of intra-industry trade in the total trade

$$GL^{ij} = 1 - \left(\frac{\sum_g \left| M_g^{ij} - M_g^{ji} \right|}{\sum_g \left(M_g^{ij} + M_g^{ji} \right)} \right), \quad 0 \le GL^{ij} \le 1. \tag{2.11}$$

In the extreme case where a given good g is either exported or imported (no intra-industry trade), the Grubel-Lloyd index will be equal to zero. With positive amounts of intra-industry trade, it will be between zero and one;[13] it is generally higher when the share of intra-industry trade in total trade is higher. We employ the Grubel-Lloyd index to identify the samples where the pure IRS model is likely to determine trade flows. Specifically, the IRS model is likely to account for the performance of the gravity equation in those samples where the bilateral Grubel-Lloyd indices are larger, indicating that a larger proportion of bilateral trade is two-way trade in perfectly specialized differentiated products. In section 4.2 we examine whether in fact the prediction M_S (of the IRS model) is more accurate in samples with higher Grubel-Lloyd indices.

We emphasize that prediction M_S is common to both the multi-cone H-O and the pure IRS models. However, if we were to find that the prediction M_S is less at odds with the data in samples with higher Grubel-Lloyd indices, it would be incorrect to interpret this finding as evidence in favor of the multi-cone H-O theory (unless there was a perfect positive correlation across country pairs between the observed Grubel-Lloyd indices and factor

---

[12]We define intra-industry trade as trade in goods with similar factor input requirements; for our empirical analysis, though, intra-industry trade is taken as two-way trade of goods in the same four-digit SITC class. The two concepts need not be the same in a deterministic sense, but our analysis remains valid as long as they are the same on average; see below.

[13]In the pure IRS model (1), the maximum value of $GL^{ij} = 1$ is obtained if the share of GDP derived from producing X-varieties is the same in both countries; in the $2 \times 2 \times 2$ model, this will happen if factor proportions are identical in the two countries.

15

endowment differences–which there is not[14]). Furthermore, we can test the generalized H-O model against the pure IRS model. We expect the latter model to be preferred to the former in samples where there is a greater share of intra-industry trade, as measured by the Grubel-Lloyd index.

In the following section, we briefly discuss the data which will be used.

3. Data

We have developed a cross-sectional data set for fifty-eight countries for the year of 1985. The data set includes all countries with both GDPs above 1 billion US dollars and where internationally comparable capital-per-worker estimates (measured in US dollars) are available. These fifty-eight countries account for 67% of world imports and 79% of world GDP in 1985. The countries are listed in Table 1A. The data set includes nearly all industrialized countries, but relatively few of the less developed countries. This reflects the paucity of capital stock estimates available from the latter.

The source for data on GDP and capital-per-worker is the current Summers and Heston (Version 5.6) database (see Summers and Heston 1991 for a description of this database). Both variables are in internationally comparable (purchasing power adjusted), real US dollars for the year 1985.[15] Each country's GDP and capital-per-worker data are reported in Table 1A. The trade data comes from the NBER World Bilateral Trade Database, see Feenstra et al. (1997). The overall imports of country i from country j (M^{ij}) are immediately available

---

[14]Helpman (1981) shows that the share of intra-industry trade and factor proportions differences are negatively related in a $2 \times 2 \times 2$ model.

[15]One reason of using data for 1985 is that this is the benchmark year for inflation and PPP adjustments in the Summers and Heston dataset.

16

from this source. We have used the values reported by the importing country, as these are known to be more accurate than trade reported by the exporting country. With fifty-eight countries, we have $58 \times 57 = 3306$ bilateral import relations.

The bilateral Grubel-Lloyd indices GL^{ij} are also derived from data in Feenstra et al. (1997). These indices are computed using all goods at the four-digit SITC classification; that is, we do not confine ourselves to the manufacturing sector as is frequently done when Grubel-Lloyd indices are calculated. The Grubel-Lloyd index (2.11) can only be computed for country-pairs where there are positive amounts of trade. This is the case for 87% of all bilateral relations, so that our sample consists of 2870 observations.[16] The number of bilateral trade relations with partner countries varies by country from a low of 27 for Nepal to the maximum of 57 for most industrialized countries. Furthermore, the number of different goods classes traded with a partner country varies from country to country. At one end, Mauritius trades on average 36 types of goods with its partners, whereas at the other end, the U.S. trades on average 332 types of goods with its trade partners.

Keeping these facts in mind, the average Grubel-Lloyd indices we compute are presented in Figure 2. Across each of their respective trading partners, Bolivia has the lowest average Grubel-Lloyd index in this sample, with a value of 0.0006; and the U.K. has the highest value of 0.1495. Each nation's average Grubel-Lloyd index is shown in Table 1A, along with national GDP per capita, the number of bilateral relations the average is computed from, and the average number of industry classes where there is positive trade. In Table 1B we report the correlation among the variables of Table 1A. All correlations are positive;

---

[16]In addition to the 2870 bilateral observations for which we can compute the Grubel-Lloyd index, the NBER database reports positive levels of trade at an aggregate level for three more pairs: Kenya/Sierra Leone, Madagascar/Ireland, and Zambia/Guatemala. We have not been able to track down the reason for this inconsistency, but do not expect that it will affect the results significantly.

in particular, the correlation of the average Grubel-Lloyd index with GDP per capita is 0.78, and the correlation between the average Grubel-Lloyd index and the average number of industry classes traded is 0.93.

In Figure 3, we plot the cumulative distribution of bilateral Grubel-Lloyd indices (there are 1435 distinct indices, corresponding to 1435 country-pairs in our sample). The distribution is very skewed: 44% of all indices are equal to 0, and 78% of all (1120 country-pairs) have a value of 0.05 or less. In the empirical analysis below, we will treat these 1120 country-pairs as trade relations where there is little or no IRS-based trade, whereas the remaining 315 country-pairs with Grubel-Lloyd indices above 0.05 are taken as those where IRS-based trade is present .[17] It is important to note that this separation of the data set into 1120 and 315 country-pairs does not mean that IRS-based trade must account for a small share of international trade; in fact, the 315 country-pairs with Grubel-Lloyd indices above 0.05 account for 87.1 % of all of the imports by the 1435 country-pairs.

We are concerned about heteroskedasticity in our sample. To that end, let the total imports of country i and total exports by country j be denoted by $\bar{M}^i$ and $\bar{X}^j$, respectively. Then, we scale all the import values M^{ij} and the GDP term $\frac{Y^i Y^j}{Y^w}$ by $\psi^{ij} = (\bar{M}^i \times \bar{X}^j)^{0.5}$ and assume that this creates a scalar variance in regressions based on the Gravity Equation prediction. Finally, the gravity prediction gives a relationship between two endogenous variables, so the results might be affected by simultaneity bias (as discussed by Saxonhouse 1989 and Harrigan 1996). In a world with positive levels of production of all goods in all countries[18] factor endowments might be valid instruments for countries' GDPs. Here,

---

[17]The threshold level of the Grubel-Lloyd index of 0.05 is based on our priors. In Appendix B we show that our main findings do not depend on this choice of threshold level.

[18]This and other conditions need to be satisfied for factor endowments to be linearly related to GDPs with

however, we are especially concerned with perfect product specialization across countries. Because this makes factor endowments much less desirable instruments, we have decided to leave the gravity prediction un-instrumented.

4. Empirical Results

4.1. Heckscher-Ohlin-Based Product Specialization (Low-GL Sample)

We first consider the bilateral pairs where the computed Grubel-Lloyd (GL^{ij}) indices are 0.05 or less. As discussed above, those are the country-pairs in which we expect only negligible amounts of IRS-based trade to be present. We assume that the calculated value of factor proportions differences, $FDIF^{ij}$, determines the share of bilateral trade which is based on product specialization due to factor proportions differences–as in the multi-cone H-O model– in a stochastic sense: $FDIF^{ij}$ is equal to the factor proportions difference which indicates the true share of perfectly specialized goods trade, denoted $^*FDIF^{ij}$, plus a confounding variable. Hence,

$$FDIF^{ij} = {}^*FDIF^{ij} + \theta^{ij}, \forall ij, \qquad (4.1)$$

where we assume that $\theta^{ij} \sim N(0, \sigma_\theta^2)$, $\sigma_\theta^2 > 0$. The confounding variable θ^{ij} accounts for two effects: First, it captures measurement error in the underlying capital-labor ratio data. Second, it picks up changes in the share of product-specialized trade which are due to effects other than changes in factor proportions differences (due to, for example, Ricardian technology differences or differences in demand across countries.) The assumption that $E[\theta^{ij}] = 0$ means that we assume that the latter effects are not systematically related to

---

equal coefficients across countries; see, e.g., Harrigan (1996).

19

factor proportions differences.

Our resampling approach proceeds as follows. Using (4.1), we create artificial distributions for $FDIF$ which are centered on the data we have collected from the Summers and Heston dataset. Second, for a specified number K of equally sized $FDIF$ classes, we collect all the observations which fall into the kth class, $k = 1, ..., K$. Third, from these K sets of observations, we compute K sets of import parameters corresponding to the gravity predictions of the models under consideration.

To implement this, we rank the 2240 country-pair observations under consideration (those with $GL^{ij} < 0.05$) by the log of the factor proportions difference we obtained from the Penn World Tables. Then, for a given Monte-Carlo replication $r, r = 1, ..., R$, we draw a (2240×1) vector of θ^{ij} and add this to the factor proportions differences column. This is our factor proportions differences vector for that replication, $FDIF^{ij}(r)$. In the following step, we sort the data on the column $FDIF^{ij}(r)$. The kth class, $k = 1, .., K$, class of observations of this replication is denoted $M_k^{ij}(r)$ and $\frac{(Y^i Y^j)_k}{Y^w}(r)$.[19] We repeat this for the following replications $r + 1, r + 2, ..., R$, collecting all bilateral relations falling into a given group k across all R replications. Finally, with these sets of data on imports and GDP's, we run the following least squares regression

$$y_k^{ij}(R) = \alpha_k x_k^{ij}(R) + \epsilon_k^{ij}, \forall k = 1, ..., K,$$

where $y_k^{ij}(R)$ is the set of all $M_k^{ij}(r)$, and $x_k^{ij}(R)$ is the set of all $\frac{(Y^i Y^j)_k}{Y^w}(r)$ across all repli-

---

[19]For the uni-cone H-O model, this is independent variable is replaced by $\left[(\gamma^j - \gamma^i) \frac{(Y_k^i Y_k^j)}{Y^w} \right] (r)$.

cations which fall into percentile class k, where ϵ_k^{ij} has mean zero.[20]

Recall that there are two models under consideration in this section, the multi-cone H-O model where all trade is in (homogeneous) specialized products, and the uni-cone H-O model, where all trade is in imperfectly specialized products. The two corresponding imports predictions are $M^{ij} = \frac{Y^i Y^j}{Y^w}$ and $M^{ij} = (\gamma^j - \gamma^i)\frac{Y^i Y^j}{Y^w}$ (if country i is capital abundant and X is capital-intensive, whereas $M^{ij} = (\gamma^i - \gamma^j)\frac{Y^i Y^j}{Y^w}$ if country j is capital-abundant). Therefore, in the multi-cone H-O model, α_k is only one parameter, predicted to equal 1, whereas for the uni-cone H-O model, we estimate (up to) fifty-eight parameters γ, the share of the labor-intensive good Z in the countries GDP's.

Figure 4 shows the result of this approach when there are $K = 5$ classes, with average values of $FDIF$ rising from class $FDIF = 1$ to $FDIF = 5$.[21] The lower line indicates for the multi-cone H-O model how the estimated value of α_k varies as $FDIF$ changes. First, note that the estimated values are around 0.02 to 0.03, which are much smaller than the predicted value of $\alpha = 1$. Also, the value of α_k falls as k is increasing; as the average factor proportions differences in the sample is increasing, the estimated import parameter falls. This implies that as there is more product specialization due to factor proportions differences in the sample, the estimated parameter moves further away from its theoretically predicted value. This is inconsistent with product specialization due to factor proportions

---

[20]Our approach is related to the resampling literature (see Efron 1982) as well as to non-parametric smoothing techniques (see, e.g., Haerdle 1990). In particular, note that ϵ_k^{ij} is not independent, because the resampling procedure leads by design to observations being repeated in a given class k, as in standard bootstrapping techniques. The variance on θ^{ij} which we assume determines the degree of smoothing, related to the choice of the bandwidth in non-parametric regression analysis. What is different here is that we apply the smoothing to the identifying variable $FDIF$, not to the variables entering the regression equation.

[21]The graphs are based on simulations with $R = 5$, $K = 5$, $\sigma_\theta^2 = 1.2$. We have experimented with different choices on these parameters, but that did not qualitatively change the results noted in the text. The log of $FDIF$ has a mean of 9.31 and a standard error of 1.21 in this sample of low-GL country-pairs.

differences being an important component in explaining the success of the gravity equation.

The upper curve in Figure 4 plots the results for the uni-cone H-O model. The figure shows the average of the estimated differences $\gamma^j - \gamma^i$, that is the difference in the share of the labor-intensive good in the labor-abundant country from that in the capital-abundant country. This average difference is estimated to be in order of 0.03 to 0.08. Furthermore, the average difference tends to increase as the average value of factor proportions differences rises in the sample. This is what the theory (in the $2 \times 2 \times 2$ case) predicts: Larger differences in factor endowment ratios are mapping into larger differences in the shares of GDP derived from a given good.[22]

We now employ a widely-used model selection criteria to compare the performance of the multi-cone H-O and uni-cone H-O models for this sample of low-GL^{ij} country-pairs. The Amemiya's prediction criterion (denoted as APC) is defined as

$$ APC = \frac{e'e}{v - w} \left(1 + \frac{w}{v} \right), $$

where $e'e$ is the residual sum of squares, v is the number of observations, and w is the number of estimated parameters.[23] The APC criterion penalizes more heavily losses of degrees of freedom than the adjusted R^2 criterion; lower values indicate better model performance. Figure 5 shows that the uni-cone H-O model is preferred to the multi-cone H-O model across all $FDIF$ classes. This figure also shows the relative performance of the multi-cone to the

[22]The average standard errors of the estimates α_k for the multi-cone (uni-cone) H-O model are as follows; FDIF=1: 0.002 (0.033), FDIF=2: 0.002 (0.036), FDIF=3: 0.001 (0.022), FDIF=4: 0.001 (0.022), FDIF=5: 0.001 (0.020). The finding that $(\gamma^j - \gamma^i)$ is less precisely estimated than α_k for the multi-cone model is largely due to the fact that α_k is estimated with more degrees of freedom than any of the differences $(\gamma^j - \gamma^i), \forall ij$.

[23]We have also considered the Akaike Information Criterion; using that gives very similar results as with the Amemiya criterion, so we report only the latter.

uni-cone H-O model across FDIF classes, computed as the ratio of the *APC* criteria of the two models for a given FDIF class (the vertical axis on the right). There is not much change of the relative performance of the multi-cone model across FDIF classes; if anything, the multi-cone H-O model performs relatively worst in the very sample where factor proportions differences are largest (with a ratio larger than 1.4 for FDIF=5.)

Overall, there is little evidence to support the notion that the multi-cone H-O model plays a role in explaining why the gravity model 'works well' empirically. As far as the success of the gravity equation has something to do with the H-O model, it appears to be related to increased specialization within the cone of diversification, not outside of it. We now turn to the subsample of country-pairs where we suspect a larger than zero share of trade based on IRS, which we take to be the country-pairs with Grubel-Lloyd indices greater than 0.05.

4.2. IRS-Based Product Specialization (High-GL Sample)

Our approach here is analogous to the one used above. We assume that the actual share of IRS-based trade in total trade (denoted $^*GL^{ij}$) in a given country-pair equals the calculated Grubel-Lloyd index (denoted GL^{ij}) plus a confounding variable:

$$GL^{ij} = \ ^*GL^{ij} + \chi^{ij}, \forall ij,$$

where $\chi^{ij} \sim N(0, \sigma_\chi^2)$, with $\sigma_\chi^2 > 0$. We interpret χ as not only capturing measurement error in the Grubel-Lloyd index, but also the two-way trade that is caused by factors other than IRS. On average, though, we assume that the calculated Grubel-Lloyd index determines the share of IRS-based trade in total trade.

In this subsample, there are 630 country-pair observations accounting for 87% of all imports. We classify the 630 observations into $K = 5$ classes, with the Grubel-Lloyd indices rising from class $K = 1$ to class $K = 5$. The two models considered here are the pure IRS model (where both X and Z are produced under economies of scale) and the generalized H-O model (where the capital-intensive X sector is IRS, whereas the labor-intensive Z sector is assumed to be labor-intensive.) The former implies a volume of imports prediction of $M^{ij} = \frac{Y^i Y^j}{Y^w}$, and the latter $M^{ij} = (1 - \gamma^i)\frac{Y^i Y^j}{Y^w}$ if country i is capital-abundant relative to j ($M^{ij} = (1 - \gamma^j)\frac{Y^i Y^j}{Y^w}$ otherwise). The regression is again

$$y_k^{ij}(R) = \alpha_k \, x_k^{ij}(R) + \epsilon_k^{ij}, \forall k = 1, ..., K, \tag{4.2}$$

with $x_k^{ij}(R)$ is the relevant independent variable for each model. With the regression (4.2), we estimate one parameter α_k for the pure IRS model, and fifty-eight γ's for the generalized H-O model.

Figure 6 reports the results of this resampling experiment.[24] The lower line shows for the pure IRS model how the estimated value of α_k varies as the Grubel-Lloyd index increases. The parameter is estimated to be between 0.04 to 0.1, which is somewhat higher than the comparable parameter estimates for the multi-cone H-O model. Second, the estimate rises in samples with higher values of the Grubel-Lloyd index; the higher the share of IRS-based trade, the closer is the estimated parameter to the theoretical value of $\alpha_k = 1$. Therefore, this finding is consistent with IRS-based trade theory being the reason why the gravity prediction 'works' empirically. Note, however, that the parameter is never anywhere close

---

[24]This is based on choosing $R = 30$, $K = 5$, and $\sigma_X^2 = 0.3$; the noted features are not sensitive to this choice. The average (standard deviation) of GL is equal to 0.199 (0.133) for these high$-GL$ country pairs.

to 1, the theoretically predicted value when all goods are tradable.

The other line in Figure 6 shows the average estimate of the share of the IRS sector in the capital-abundant country, $(1 - \gamma^i)$, across the five different classes.[25] This share is estimated to be between 0.47 and 0.72, and rises as the Grubel-Lloyd index rises. Note that from country-pair to country-pair there are changes as to which of the two, exporter or importer, is the relatively capital-abundant country. Therefore, a non-structural interpretation of the average of the estimated $(1 - \gamma^i)$ is that it gives the average size of the share of the differentiated goods sector. According to this interpretation, we estimate a higher GDP share for the IRS-sector in samples with higher values of the Grubel-Lloyd index. This is consistent with the generalized H-O theory, where a higher share of intra-industry trade is driven by IRS and product differentiation.

The Amemiya Prediction Criterion is used to compare the performance of the two models across the five classes, with the results presented in Figure 7: the generalized H-O is preferred to the pure IRS model for all five classes. As in the previous subsection, the model which incorporates imperfect specialization performs better. Still, there is an important difference relative to the samples with predominantly H-O-based trade: the relative performance of the generalized H-O model is better in samples with lower Grubel-Lloyd indices. That means that although the model in which all goods are produced with IRS is never as successful as the generalized H-O model, the pure IRS model puts in its best relative performance where it should (namely, in the sample with the highest value of the Grubel-Lloyd index).

---

[25]The average standard errors of the estimates α_k for the Pure IRS (Generalized H-O) model are as follows; GL=1: 0.001 (0.119), GL=2: 0.001 (0.146), GL=3: 0.002 (0.206), GL=4: 0.002 (0.294), GL=5: 0.003 (0.512).

4.2.1. The Effect of Distance

One potential problem with the above analyses is that they do not control for differences across $FDIF$- or GL-classes in the average distance between countries. It is well-established empirically that bilateral trade volumes fall as bilateral distance increases (the second pillar of the gravity equation), raising the possibility that the different estimated parameters are driven by differences in the average distance between trade partners across the $FDIF$- or GL-classes.

In fact among the bilateral trade relations with Grubel-Lloyd indices less than 0.05, the average bilateral distance between trade partners varies little between $FDIF$-classes.[26] The analysis of bilateral trade relations with Grubel-Lloyd indices greater than 0.05 might be strongly affected by differences in the average distance among country-pairs in a given GL-class, because the average distance falls from about 6000 miles in GL=1 to about 2600 miles in GL=5. Within the set of our zero-trade cost theories, therefore, it is important to determine whether there remains a monotonically increasing relationship between the estimated parameter and the Grubel-Lloyd index once the effect of distance is controlled for. We conduct a two-dimensional resampling analysis analogous to the earlier ones, smoothing both the Grubel-Lloyd and distance variables:

$$GL^{ij} = {}^{*}GL^{ij} + \chi^{ij},$$

$$DIST^{ij} = {}^{*}DIST^{ij} + \pi^{ij}, \forall ij,$$

---

[26]It hovers around 8750 miles, with 8800 miles for $FDIF$=1 and 8700 miles for FDIF=5; we would like to thank Jon Haveman at Purdue University for making the distance data available at his webpage (http://intrepid.mgmt.purdue.edu/Jon/Data/TradeData.html#Gravity); the data gives the Great Circle distance between capital cities, as the crow flies.

where $\pi \sim N(0, \sigma_\pi^2), \sigma_\pi^2 > 0$, and $DIST^{ij}$ denotes the bilateral distance between i and j.[27]

In Figure 8 we have plotted the average estimated $(1 - \gamma^i)$, denoted α_{km}, for $K \times M$ classes of GL ($K = 5$ classes) and $DIST$ ($M = 5$ classes), where the latter denotes the average bilateral distance in a sample. It is clear from Figure 8 that we estimate a lower value of α_{km}, the higher is the average bilateral distance. The correlation between GL and α_{km} remains positive. However, this still does not determine whether there is an independent effect associated with GL, because the average distance for $(GL = 1, DIST = 1)$ is much higher than for $(GL = 5, DIST = 1)$, i.e., GL and $DIST$ co-vary. Therefore, we have taken the 25 values of α_{km}, GL_{km}, and $DIST_{km}$, and run the following least-squares regression:

$$\alpha_{km} = \beta_0 + \beta_1 GL_{km} + \beta_2 \ln(DIST_{km}) + \mu_{km}, \tag{4.3}$$

where μ_{km} is assumed to be mean zero and constant variance σ_μ^2. We find the following parameters (standard errors) for β_1 and β_2 : 0.174 (0.064), and -0.084 (0.011), respectively. In sum, higher values of the Grubel-Lloyd index are associated with higher import parameter estimates even when differences in the average distance between trade partners across subsamples are controlled for.

Figure 9 shows the analogous picture for the pure IRS model. The import parameter estimate rises both with the Grubel-Lloyd index rising and the average bilateral distance falling; for maximum values of GL and minimum values of $DIST$ (the combination $GL = 5$, $DIST = 1$), a value of $\alpha_{51} = 0.26$ is estimated, whereas for the minimum values of GL and maximum values of $DIST$, we estimate $\alpha_{15} = 0.03$. Because the resampling experiments are

---

[27]Here, $R = 50$, $\sigma_\chi^2 = 0.3$, and $\sigma_\pi^2 = 1.5$.

identically designed, the same co-variance of distance and the Grubel-Lloyd index is present as above for the generalized H-O model analysis. We run again the least squares regression (4.3), this time with the data generated by the pure IRS model underlying Figure 9. The estimated coefficients (standard errors) are: $\beta_1 = 0.08$ (0.02) and $\beta_2 = -0.02$ (0.003). Also here, the estimated import parameter is higher as the Grubel-Lloyd index rises, even when the effects of distance are controlled for. This suggests that our earlier interpretation of the results is robust.

5. Conclusions

We have compared the accuracy of the perfect specialization versions of the Heckscher-Ohlin model and the Increasing Returns-based model, both of which predict the gravity equation. In addition, we have demonstrated that, as far as their predictions for bilateral trade volumes are concerned, each of these perfect specialization models is a limiting case of a model with imperfect specialization. Our empirical strategy has exploited two factors to identify which models might explain trade flows in a given sample. First, in the multi-cone H-O model large factor proportions differences are required to induce product specialization, whereas product specialization in the pure IRS model occurs for arbitrary differences in factor proportions. Secondly, there is no intra-industry trade in the multi-cone H-O model, whereas there is intra-industry trade, and possibly only intra-industry trade, in the pure IRS model.

There are three major findings: First, little production is perfectly specialized due to factor proportions differences, making the perfect specialization version of the H-O model an unlikely candidate to explain the empirical success of the Gravity Equation. Secondly, increasing returns are important causes for perfect product specialization and the Gravity

28

Equation, especially among industrialized countries. Thirdly, models of imperfect specialization better explain the variation of bilateral trade flows than perfect specialization models. Factor proportion differences are important determinants of trade flows within the context of imperfect specialization models only, whereas there is evidence that increasing returns is a cause of product specialization along the lines of both trade models with imperfect as well as with perfect specialization of production.

In the light of this paper, there is no reason to believe that the results by Hummels and Levinsohn (1995) throw doubt on the empirical relevance of IRS trade theory. Our results suggest that the H-O model with imperfect specialization is likely to account for the regression results that these authors found in their sample where little intra-industry trade was expected. More generally, the data is supportive of several predictions of IRS-based trade theory, and we take our results as indicating that IRS-based trade plays, in conjunction with trade based on factor proportions differences, an important role in determining world trade flows.

Several important determinants of trade flows are missing from our analysis. First of all, we do not allow for preferences to differ across countries, even though other studies have found that countries tend to overproportionately demand home produced goods (or, exhibit a 'home bias'). This appears to be one reason of why standard trade models predict trade volumes which are far larger than what is found in the data (Trefler 1995). Secondly, a related point is that we do not consider explicitly models with transport costs and trade barriers. We know that the latter are empirically relevant and can also lead to a home bias. However, CRS production of the same good in two or more countries in the presence of transport costs is inconsistent with factor price equalization; moreover, as emphasized by

the recent economic geography literature (e.g. Davis and Weinstein 1996), CRS and IRS models might behave differently in the presence of transport cost and differences in demand across countries.

Thirdly, we have already noted that part of the product specialization which we attribute to factor proportions differences or IRS might actually be due to technology differences across countries. A fourth point is that we estimate two-country models with data from a multi-country world. The problem of developing multi-lateral trade predictions arises from the well-known indeterminacy of bilateral trade flows when there are imperfectly specialized homogeneous goods produced with CRS. Traditionally, researchers have instead considered the H-O-Vanek equation which predicts net factor service trade of a country with the rest of the world. However, as Trefler (1996) has shown, generally, the H-O-Vanek equation holds only in cases when also the simplest gravity equation with perfect product specialization holds. It is clear, though, and our results confirm that imperfect product specialization is empirically important. Therefore, we hope to make further progress by extending this work in the context of the gravity equation to a multilateral world with trade costs.

Table 1A

COUNTRY	GDP per Capita ($ US)	GDP (billion $ US)	Capital per Worker ($ US)	Number of Bilateral Relations	Average Number of Industry Classes Traded	Average Grubel-Lloyd Index
U.S.A.	16570	3964.9	29925	57	332	0.1371
CANADA	15589	392.3	34535	57	190	0.055
SWITZERLAND	14864	96.2	62769	57	209	0.1227
NORWAY	14144	58.7	44866	57	142	0.0537
AUSTRALIA	13583	214	33875	56	151	0.0306
SWEDEN	13451	112.3	31326	57	198	0.1163
DENMARK	12969	66.3	29286	57	181	0.0923
W. GERMANY	12535	765.4	47695	57	323	0.1473
BELGIUM-LUX.	12230	116.1	33494.5	57	221	0.1347
ICELAND	12209	2.9	15991	44	84	0.0064
FRANCE	12206	673.4	31796	57	282	0.1377
FINLAND	12051	59.1	38591	57	136	0.0625
JAPAN	11771	1421.4	28106	57	244	0.064
NETHERLANDS	11539	167.2	29325	57	255	0.1283
NEW ZEALAND	11443	37.4	30511	53	106	0.0189
U.K.	11237	636.2	17636	57	307	0.1495
AUSTRIA	11131	84.1	29708	57	171	0.1031
ITALY	10808	617.6	28117	57	265	0.1111
HONG KONG	10599	57.8	11220	57	166	0.0805
ISRAEL	8310	35.2	22307	52	119	0.0721
SPAIN	7536	290.7	21831	57	209	0.0807
IRELAND	7275	25.8	20700	56	117	0.0746
VENEZUELA	6225	106.3	20419	48	91	0.0078
GREECE	6224	61.8	22151	55	112	0.0304
MEXICO	5621	420.3	14016	52	100	0.0239
TAIWAN	5449	104.9	19194	57	167	0.0569
ARGENTINA	5324	161.5	12084	52	92	0.0201
PORTUGAL	5070	51.5	9503	56	111	0.0473
SYRIA	4240	43.9	15681	41	70	0.0011
MAURITIUS	4226	4.3	2393	35	36	0.0106
S. KOREA	4217	172.1	12036	55	139	0.0644
IRAN	4043	187.5	10686	41	81	0.0033
PANAMA	3499	7.6	16417	45	73	0.0074
CHILE	3467	42	7060	48	100	0.0081
TURKEY	3077	154.8	6989	49	103	0.0195
COLOMBIA	2968	87.5	12790	48	85	0.0092
ECUADOR	2913	27.1	15536	44	67	0.0031
PERU	2565	49.7	9480	48	90	0.0077
THAILAND	2463	127.3	4051	53	114	0.0286
JAMAICA	2215	5	3725	41	47	0.005
DOMINICAN REP.	2111	13.5	5282	43	56	0.012
GUATEMALA	2090	16.6	3985	45	57	0.0079
PARAGUAY	2072	7.7	849	37	42	0.0021
SRI LANKA	2045	32.4	8238	54	66	0.0186
MOROCCO	1956	43.2	2803	48	76	0.0067
BOLIVIA	1754	11.1	6987	37	49	0.0006
IVORY COAST	1545	15.1	1171	51	57	0.0076
PHILIPPINES	1542	84.4	4087	55	82	0.024
HONDURAS	1387	6.1	4507	41	51	0.0028
ZIMBABWE	1216	10.2	5750	48	52	0.0033
NIGERIA	1062	88.4	1103	44	83	0.0024
INDIA	1050	803.4	1712	57	115	0.0305
NEPAL	936	15.6	758	27	41	0.0031
SIERRA LEONE	905	3.3	215	34	38	0.0029
ZAMBIA	808	5.5	1785	37	53	0.0053
KENYA	794	16.1	1187	43	75	0.0052
MADAGASCAR	769	7.7	1714	34	40	0.0059
MALAWI	518	3.7	469	37	38	0.0028
Average	6248.55	222.31	16214.2	49.48	123.4	0.0427
Median	4233	58.9	12437	52	100	0.0198
Standard Deviation	4934.91	563.33	14136.79	8.05	78.57	0.0469

Table 1B

Correlation Matrix

	GDP per Capita	GDP	Capital per Worker	No. of Bilateral Pairs	Av. No. of Industries	Av. Grubel-Lloyd Index
GDP per Capita	1					
GDP	0.41	1				
Capital per Worker	0.89	0.26	1			
No. of Bilat. Pairs	0.68	0.31	0.66	1		
Av. No. of Industr.	0.81	0.6	0.73	0.74	1	
Av. Grubel-L. Index	0.78	0.45	0.73	0.71	0.93	1

Figure 1

The Model Identification Issue in the Gravity Equation Context

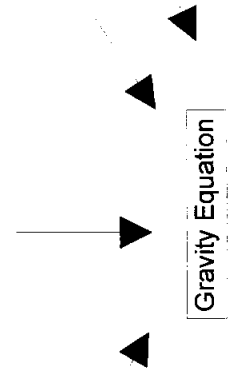|	Technology Differences (Ricardian)	Increasing Returns	Heckscher-Ohlin	Other Models
Type of Model				
1. Structural Assumptions: Identical homothetic demand, free trade, and…	Technology Differences within industry classes across countries	Increasing Returns at the firm level, monopolistic competition, product differentiation	Homogeneous goods and Multi-coneworld (Large Factor Proportions Differences)	
Consistent with absence of factor proportions differences?	yes	yes	no	
2. Implication for Nature of Trade	Intra-industry trade in goods with alternating technological superiority	Intra-industry trade in product varieties with potentially identical technologies across countries	Inter-industry trade	
consistent with trade in goods with identical factor requirements?	yes	yes	no	
3. Import Volume Prediction				

Gravity Equation

Figure 2

Average Grubel-Lloyd Index by Country, four-digit SITC Classification, all Sectors
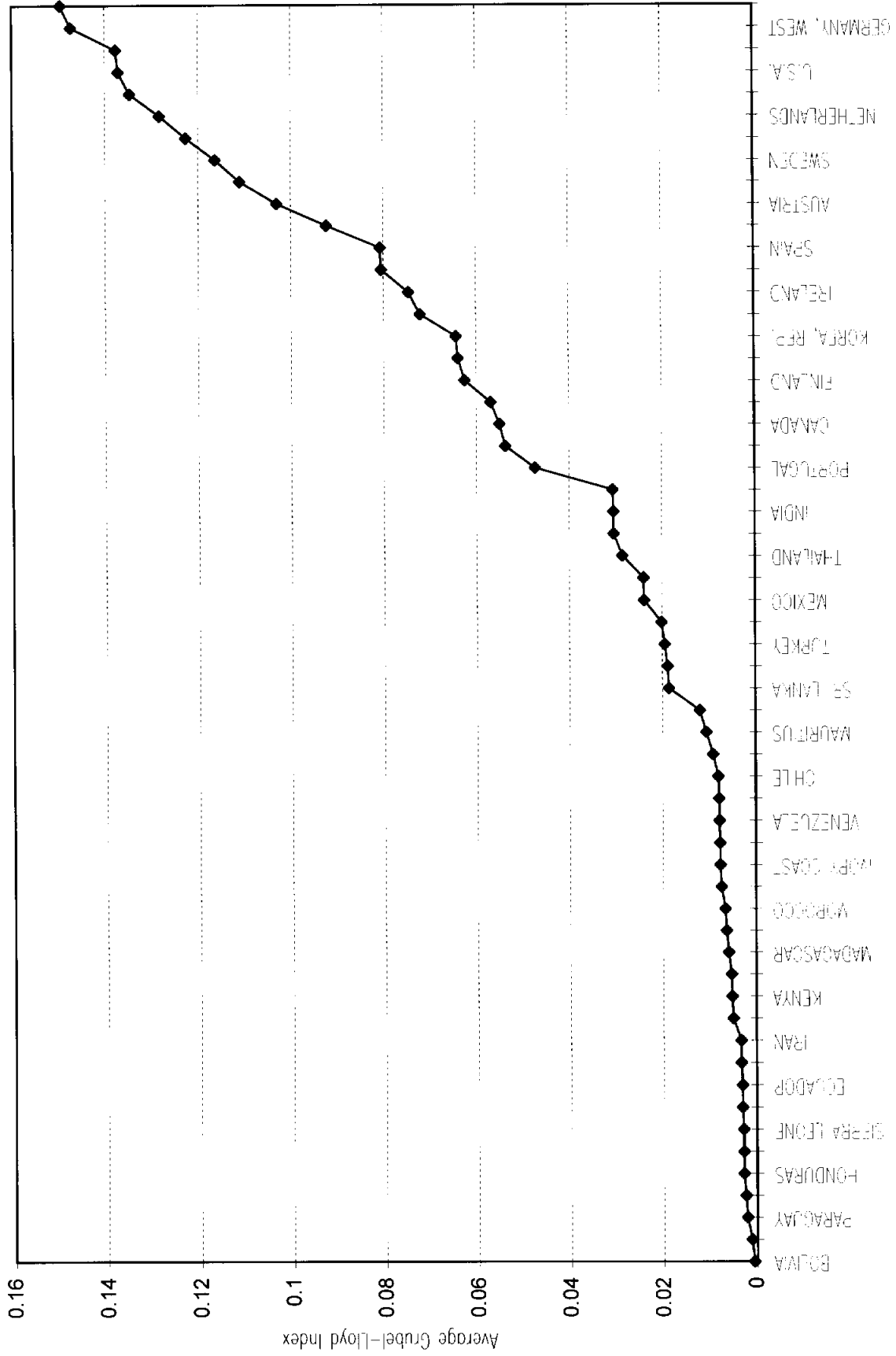
Figure 3
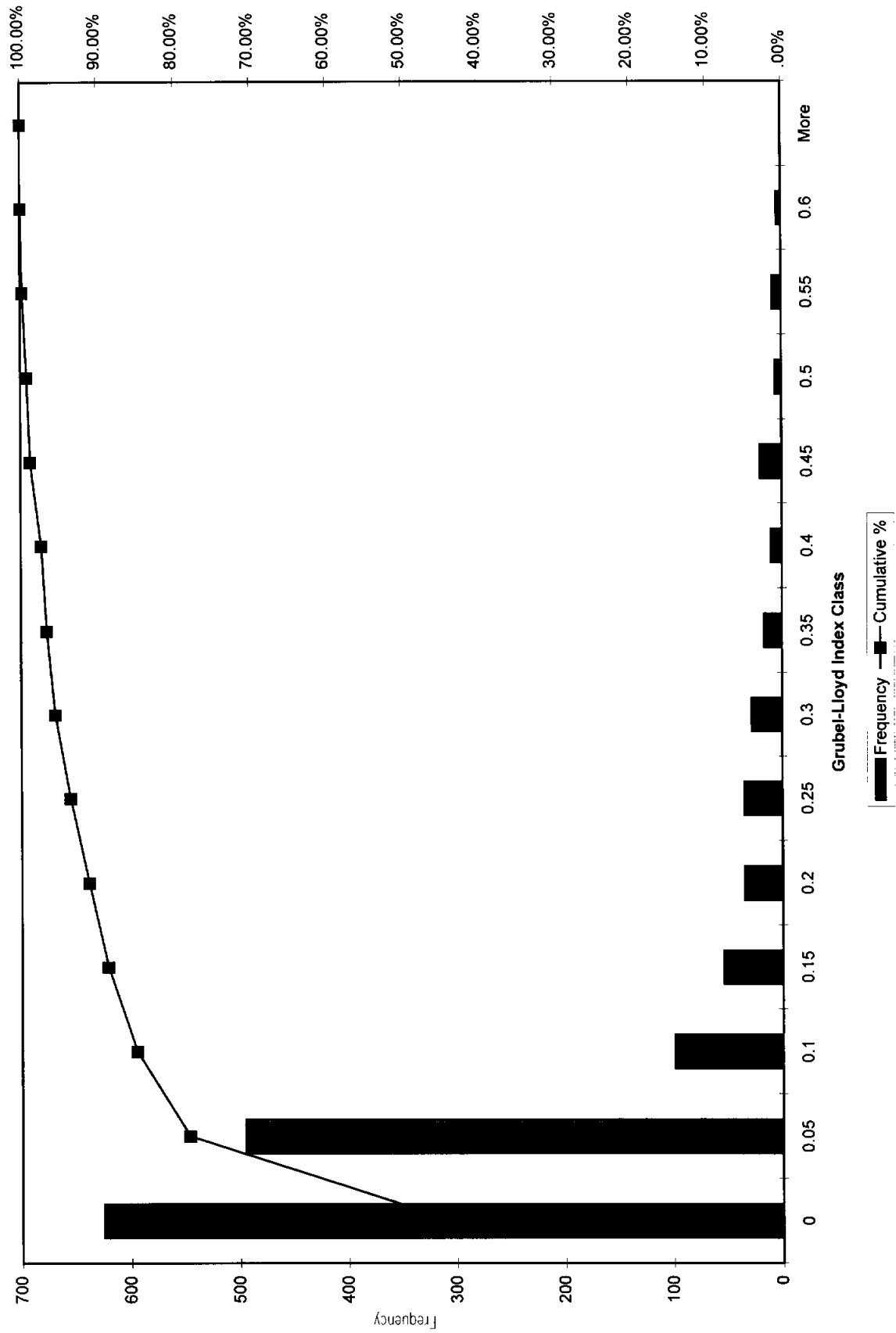
Distribution of Grubel-Lloyd Indices (1435 Country-Pairs)

Figure 4

Import Parameter Estimates for Samples with Heckscher-Ohlin based Trade for varying Factor Proportions Differences
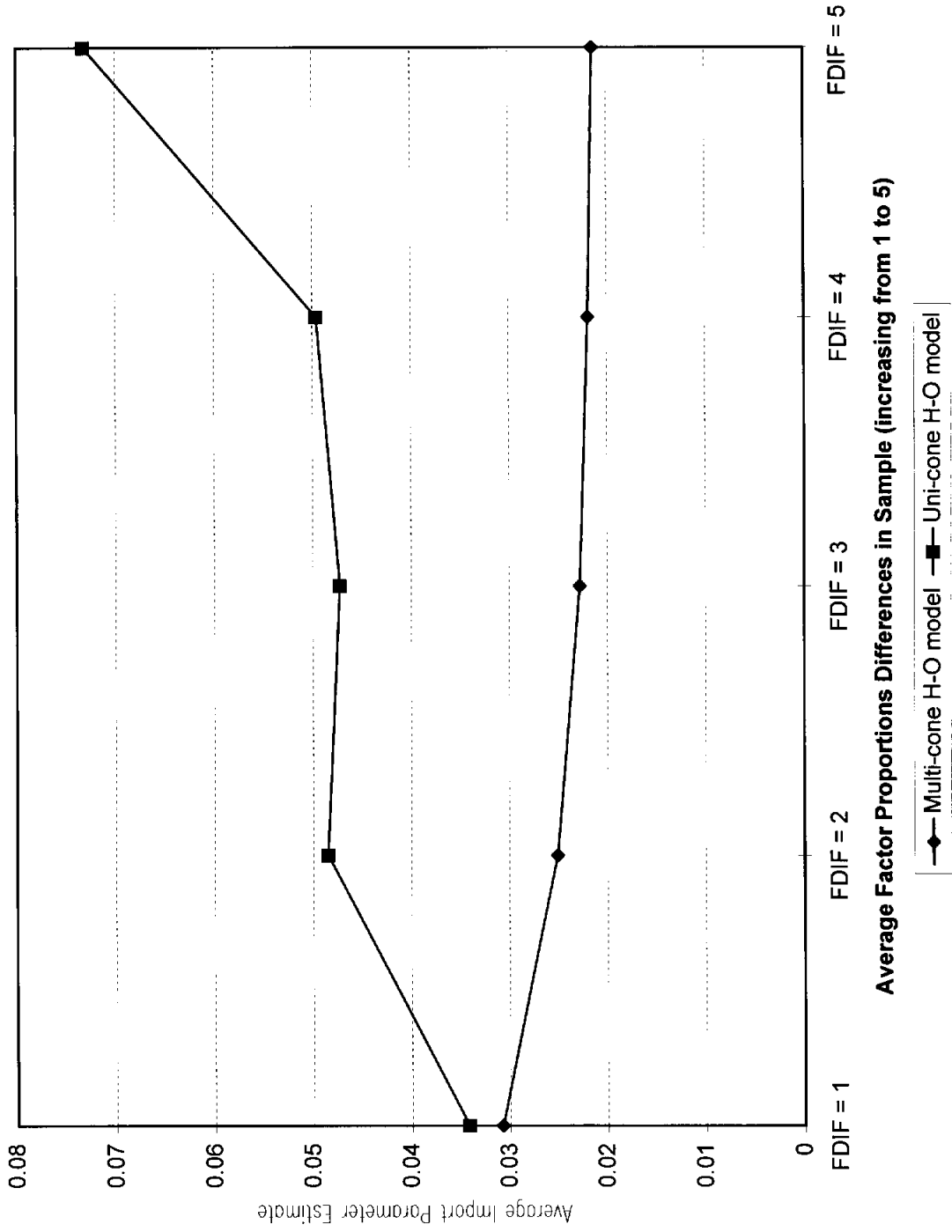

Average Factor Proportions Differences in Sample (increasing from 1 to 5)

Average Import Parameter Estimate

FDIF = 1   FDIF = 2   FDIF = 3   FDIF = 4   FDIF = 5

Multi-cone H-O model   Uni-cone H-O model

Figure 5

Absolute and Relative Model Performance for Samples with Heckscher-Ohlin based Trade for varying Factor Proportions Differences

Average Factor Proportions Differences (increasing from FDIF=1 to FDIF=5)

→ APC for Multi-cone H-O model   ■ APC for Uni-cone H-O model   ▲ Ratio of APC(Multi-cone H-O) to APC (Uni-cone H-O)

Figure 6

Import Parameter Estimates for Samples with Increasing Returns-based Trade, varying by Grubel-Lloyd Index

Average Grubel-Lloyd Index (increasing from GL=1 to GL=5)

— Pure IRS model   — Generalized H-O model

Figure 7

Absolute and Relative Model Performance for Samples with Increasing Returns based Trade for varying levels of Grubel-Lloyd Indices

Average Grubel-Lloyd index (increasing from GL=1 to GL=5)

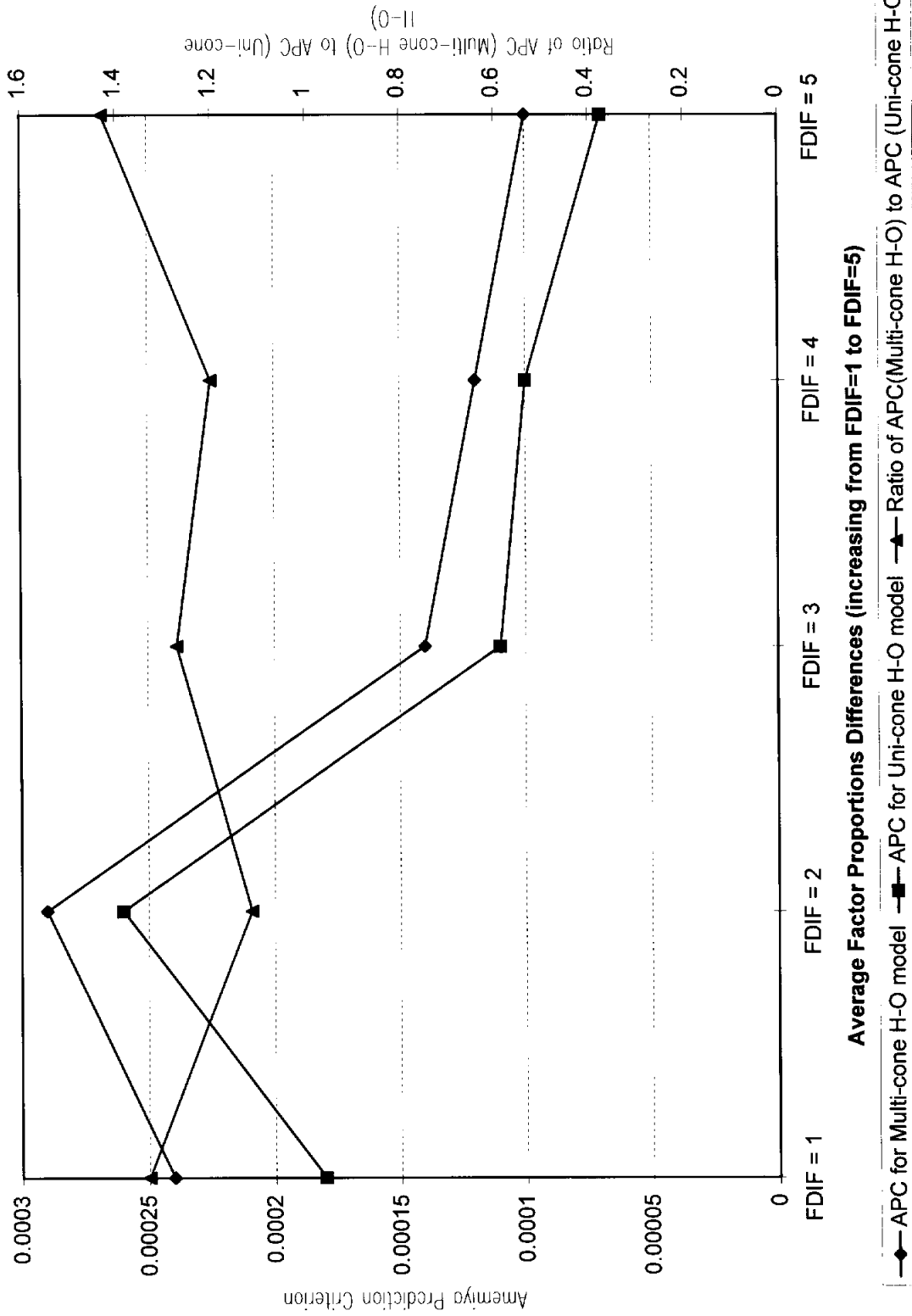—■— APC for Pure IRS model —▲— APC for Generalized H-O model —◆— Ratio of APC(IRS) to APC(Generalized H-O)
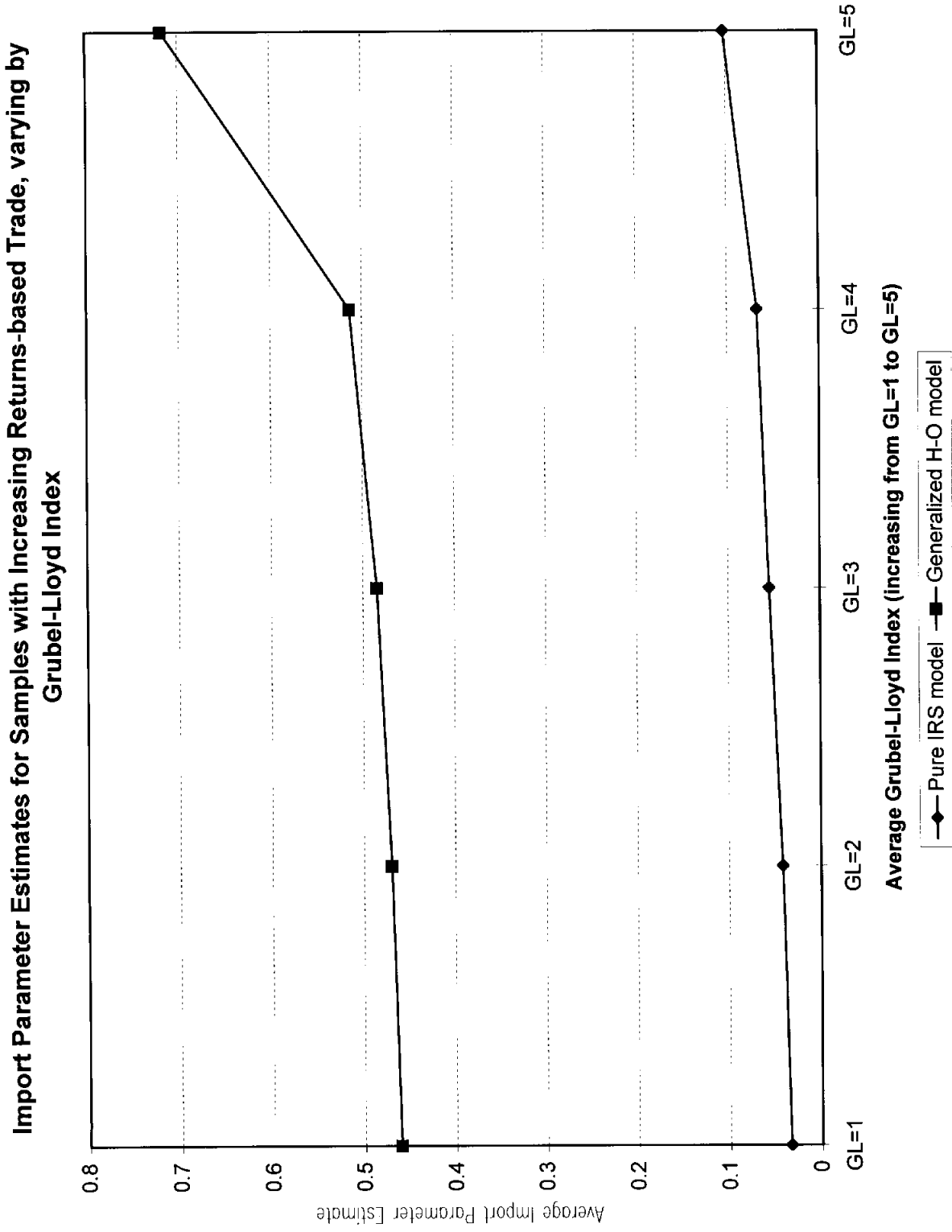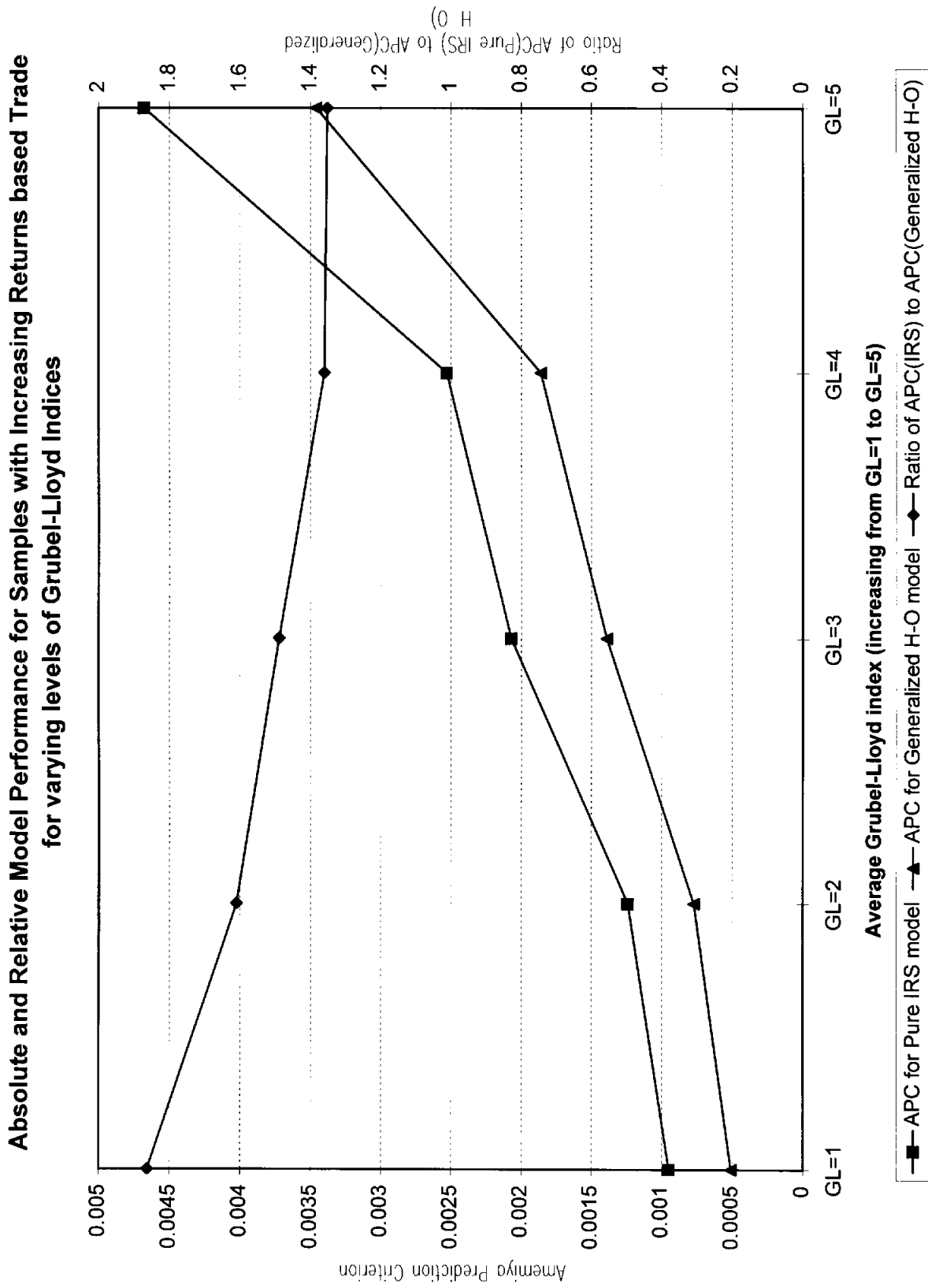
Figure 8

Import Parameter Estimates in Relation to Bilateral Grubel-Lloyd Index and Distance: the
Generalized Heckscher-Ohlin Model

■ 0.9-1.1
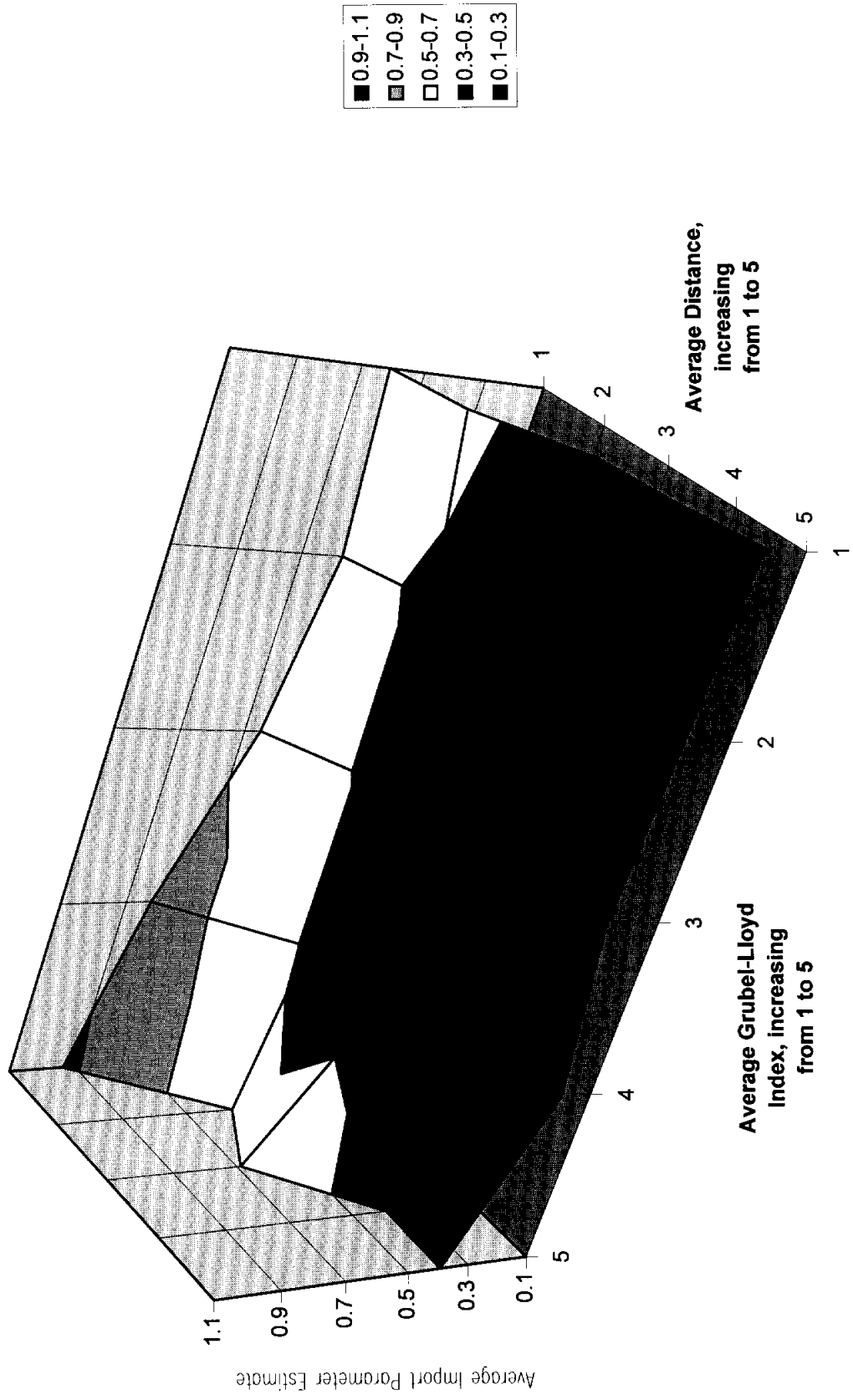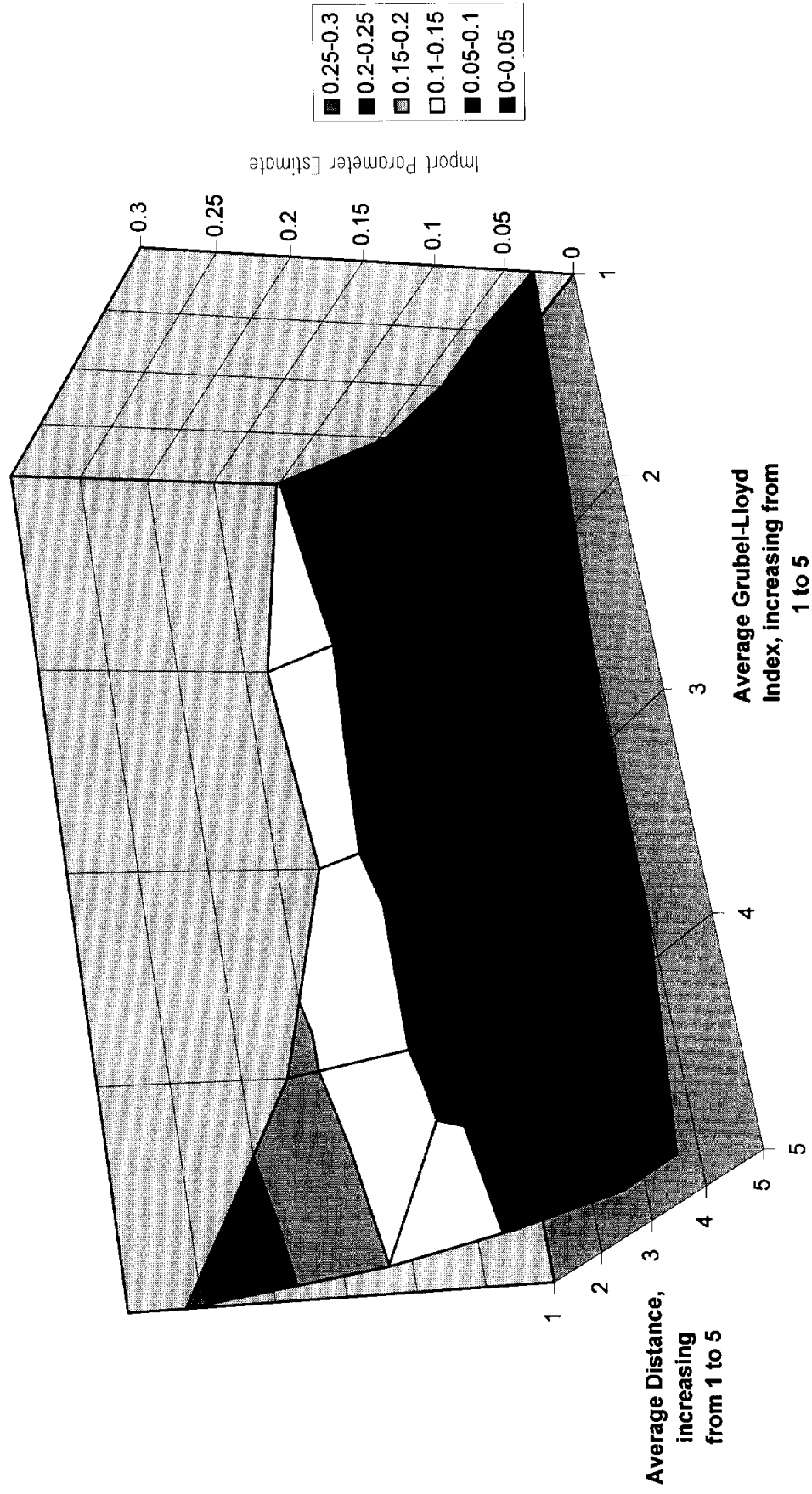▨ 0.7-0.9
□ 0.5-0.7
■ 0.3-0.5
■ 0.1-0.3

Average Distance,
increasing
from 1 to 5

Average Grubel-Lloyd
Index, increasing
from 1 to 5

Average Import Parameter Estimate

Figure 9

Import Parameter Estimates in Relation to Average Grubel-Lloyd Index and Distance: The Pure Increasing Returns Model

Legend:
- 0.25-0.3
- 0.2-0.25
- 0.15-0.2
- 0.1-0.15
- 0.05-0.1
- 0-0.05

Import Parameter Estimate

Average Grubel-Lloyd Index, increasing from 1 to 5

Average Distance, increasing from 1 to 5

References

[1] Armington, P. S. (1969), "A theory of demand for products distinguished by place of production", *IMF Staff Papers* 16: 159-177.

[2] Anderson, J. (1979) "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, 69(1): 106-116.

[3] Antweiler, W., and D. Trefler (1997), "Increasing Returns and All That: A View from Trade", mimeo, University of Toronto, August.

[4] Bergstrand, J. (1990), "The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis and the Determinants of Bilateral Intra-Industry Trade", *Economic Journal* 100: 1216-1229.

[5] Chipman, J. S. (1992), "Intra-Industry Trade, Factor Proportions, and Aggregation", in *Economic Theory and International Trade. Essays in memoriam J. Trout Rader*, W. Neuefeind and R. Riezman (eds.), pp.67-92. Berlin: Springer.

[6] Davis, D., (1995), "Intra-industry Trade: A Ricardo-Heckscher Ohlin Approach", *Journal of International Economics*, 39: 201-226.

[7] Davis, D., and D. Weinstein (1997), "Increasing Returns and International Trade: An Empirical Confirmation", mimeo, Harvard University, October.

[8] Davis, D., and D. Weinstein (1996), "Does Economic Geography Matter for International Specialization?", NBER Working Paper # 5706, Cambridge, MA.

[9] Davis, D., D. Weinstein, S. Bradford, and K. Shimpo (1997), "Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works", *American Economic Review* 87: 421-446.

[10] Deardorff, A. (1998) "Determinants of Bilateral Trade: Does Gravity Work in a Neo-classical World?", chapter 1 in J. A. Frankel (ed.),*The Regionalization of the World Economy*, Chicago: The University of Chicago Press.

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

[12] Eaton, J., and S. Kortum (1997), "Technology and Bilateral Trade", mimeo, Boston University, April.

[13] Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, Philadelphia: Society for Industrial and Applied Mathematics.

[14] Feenstra, R., R. E. Lipsey, and H. P. Bowen (1997), "World Trade Flows, 1970-1992, with Production and Tariff Data", NBER Working Paper # 5910 and accompanying CD-ROM, Cambridge, MA.

[15] Grubel, H., and P. Lloyd (1975), *Intra-Industry Trade: The Theory and Measurement of International Trade in Differentiated Products*, London: Macmillan.

[16] Härdle, W. (1990), *Applied Nonparametric Regression*, Econometric Society Monograph, Cambridge, U.K.: Cambridge University Press.

[17] Harrigan, J. (1996), "Openness to trade in manufactures in the OECD", *Journal of International Economics* 40: 23-39.

[18] Haveman, J., and D. Hummels (1997), "What can we learn from bilateral trade? Gravity and beyond", mimeo, University of Chicago Business School, June.

[19] Helpman, E. (1981), "International Trade in the Presence of Product Differentiation, Economies of Scale, and Monopolistic Competition: A Chamberlin-Heckscher-Ohlin Approach", *Journal of International Economics* 11: 305-340.

[20] Helpman, E. (1987) "Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries," *Journal of Japanese and International Economics*, 1: 62-81.

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

[22] Hummels, D. and J. Levinsohn (1995) "Monopolistic Competition and International Trade: Reconsidering The Evidence," *Quarterly Journal of Economics*, 110(3): 799-836.

[23] Keller, W. (1998), "Gravity Models of Trade without Specialization of Production", mimeo, University of Wisconsin, March.

[24] Leamer, E. (1992) "The Interplay of Theory and Data in the Study of International Trade", chapter 9 of M. Nerlove (ed.) *Issues in Contemporary Economics*, Vol. 2, New York University Press, New York, NY.

[25] Leamer, E. and J. Levinsohn (1995) "International Trade Theory: The Evidence", in G.M. Grossman and K. Rogoff (eds.), *Handbook of International Economics*, Vol. 3, Elsevier, Amsterdam.

[26] Markusen, J. (1986), "Explaining the Volume of Trade: An Eclectic Approach", *American Economic Review* 76: 1002-11.

[27] Markusen, J., and R. Wigle (1990), "Explaining the Volume of North-South Trade", *Economic Journal* 100: 1206-1215.

[28] Saxonhouse, G. (1989), "Differentiated Products, Economies of Scale, and Access to the Japanese Market", chapter 5 in R. C. Feenstra (ed.), *Trade Policies for International Competitiveness*, Chicago: University of Chicago Press.

[29] Summers, R., and A. Heston (1991), "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988", *Quarterly Journal of Economics* 106: 327-368.

[30] Trefler, D. (1996), "The Structure of Factor Content Predictions", mimeo, University of Toronto, March 1996.

[31] Trefler, D. (1995), "The Case of the Missing Trade and other Mysteries", *American Economic Review* 85: 1029-1046.

A. The Pure IRS Model

(See Helpman and Krugman 1985). Consider two countries, i and j, two goods, X and Z, and two factors, K (capital) and L (labor), the latter being immobile across countries. The goods X and Z come in many differentiated varieties which are identically produced with increasing returns to scale. Assume that countries have identical homothetic preferences, and that consumers value all varieties symmetrically in the Dixit-Stiglitz (1977) CES fashion. In equilibrium, any product variety will only be manufactured by one producer (perfect product specialization). Producers of differentiated products behave monopolistically competitively, and free entry determines the number of varieties (and firms) in equilibrium. Let n_g^c be the number of good f varieties produced in country c, $g = X, Z$, and $c = i, j$, s^c the share of country c in world spending and $x^c(z^c)$ the equilibrium quantity of a type X (type Z) variety. Let Y^c denote a country's GDP, and world GDP is given by $Y^w = Y^i + Y^j$. Finally, choose good Z as numeraire and let p_x denote the relative price of a good X variety.

With balanced trade–so that $s^c = \frac{Y^c}{Y^w}, \forall c$–and zero trade and transport costs, both countries will demand all foreign varieties according to the countries' GDP as a share of World GDP. Given that the varieties produced in country j are available in country i only through imports, country i's imports from j, denoted M^{ij}, will be

$$M^{ij} = s^i \left[p_x n_x^j x^j + n_z^j z^j \right],$$

whereas country j's imports from i, M^{ji}, will be

$$M^{ji} = s^j \left[p_x n_x^i x^i + n_z^i z^i \right].$$

The last two equations are shown as (2.1) and (2.2) in the text.

B. Alternative Prior Beliefs on the Presence of IRS-based Trade in Relation to the Grubel-Lloyd Index

Here we briefly review the sensitivity of our results with respect to the assumption that there is little to no IRS-based trade for Grubel-Lloyd values of $GL^{ij} < 0.05$. First, we compare the multi-cone H-O (M1) to the uni-cone H-O model (M4) for two alternative assumptions: (a) No IRS-based trade for $GL^{ij} < 0.075$, and (b) No IRS-based trade for $GL^{ij} < 0.033$.

Here, $\emptyset(M1)$ means, e.g., the average import parameter for model 1.[28]

		FDIF=1	FDIF=2	FDIF=3	FDIF=4	FDIF=5
	$\emptyset$(M1)	0.033	0.021	0.019	0.019	0.017
Case (a)	$\emptyset$(M4)	0.062	0.049	0.038	0.038	0.044
	$\frac{APC(M1)}{APC(M4)}$	1.28	1.2	1.13	1.2	1.21
	$\emptyset$(M1)	0.03	0.022	0.022	0.02	0.02
Case (b)	$\emptyset$(M4)	0.045	0.055	0.042	0.048	0.033
	$\frac{APC(M1)}{APC(M4)}$	1.16	1.08	1.17	1.25	1.33

According to the table, one finds irrespective of the GL cut-off point that the import parameter falls as FDIF rises in the multi-cone model M1. This parallels the findings in the main text. Furthermore, the ratio of the Amemiya Prediction Criterion (APC) does not vary systematically for either case (a) or case (b). Again this is identical to what we find with a $GL = 0.05$ cut-off, and it is inconsistent with perfect product specialization driven by factor proportions differences being a major element in explaining the success of the gravity equation. The average import parameter for the uni-cone H-O model, however, increases monotonically with FDIF for the $GL = 0.05$ cut-off point discussed in the text, whereas now, the relation between the average parameter estimate and FDIF is non-monotonic. In case (a), for instance, it first falls and then rises with FDIF. Note that case (a) includes

---

[28]For these experiments, $R = 10$ and $\sigma_\theta^2 = 2$.

country-pairs with a higher GL-value than that of 0.05 as discussed in the text. The finding of high values of $\emptyset(M4)$ for FDIF=1 and FDIF=2 is consistent with country-pairs with relatively high GL values (which have relatively low FDIF values) leading to relatively high import parameter estimates. Therefore, this finding need not contradict the earlier result in the text as much as it underlines the importance of the model identification issue. This point cannot explain the pattern of $\emptyset(M4)$ in case (b), though.

Second, comparing the pure IRS (M2) with the generalized H-O model (M3) for the same cases (a) and (b), we find the following: [29]

		GL=1	GL=2	GL=3	GL=4	GL=5
	$\emptyset(M2)$	0.077	0.081	0.087	0.109	0.131
Case (a)	$\emptyset(M3)$	0.287	0.298	0.304	0.323	0.358
	$\frac{APC(M2)}{APC(M3)}$	1.41	1.28	1.28	1.22	1.24
	$\emptyset(M2)$	0.026	0.031	0.042	0.075	0.145
Case (b)	$\emptyset(M3)$	0.323	0.33	0.332	0.351	0.377
	$\frac{APC(M2)}{APC(M3)}$	2.43	2.0	1.88	1.48	1.24

The table shows that for comparing the pure IRS with the generalized H-O model, we obtain the same results with a cut-off values of $GL = 0.075$ (case (a)) and $GL = 0.033$ (case (b)) as with the value of $GL = 0.05$ as employed in the text: First, the import parameter estimate

---

[29]The following results are based on $R = 10$ and $\sigma_\chi^2 = 0.1$.

of the pure IRS model is rising as GL is rising. Secondly, also the average import parameter of the generalized H-O model is rising as GL is rising. Thirdly, the generalized H-O model is less preferred for higher values of GL.

Summarizing, five out of six inferences made in the text continue to hold even if different reasonable prior beliefs on the presence of IRS-based trade in relation to the *GL* index are adopted. With different priors, the importance of factor proportions differences in determining trade volumes is evidenced primarily by the our findings in support of the generalized H-O model M3.