RESEARCH SEMINAR IN INTERNATIONAL ECONOMICS

School of Public Policy
The University of Michigan
Ann Arbor, Michigan 48109-1220

Discussion Paper No. 435

An Account of Global Factor Trade

Donald R. Davis
Harvard University, Federal Reserve Bank of New York, and NBER

and

David E. Weinstein
University of Michigan and NBER

October 16, 1998

An Account of Global Factor Trade

by

Donald R. Davis
Harvard University
Federal Reserve Bank of New York
and NBER

and

David E. Weinstein
University of Michigan
and NBER

Draft 1.3
October 16, 1998

ABSTRACT: A half-century of empirical work on the factor proportions theory has identified “paradoxes” and “mysteries,” but has failed to devise simple amendments that bring theory and data into reasonable congruence. Our study considers standard and novel hypotheses regarding the failures of the Heckscher-Ohlin-Vanek formulation and is the first to examine these directly on the technology and absorption data of interest. We show how a few simple and plausible amendments, verified directly by this data, suffice for a striking confirmation of the HOV theory. Countries export the services of abundant factors and in approximately the right magnitude. HOV works.

We are grateful to Alan Deardorff, Gene Grossman, and Elhanan Helpman for very helpful comments. Much-appreciated support for this project has come from the National Science Foundation, the Harvard Institute for International Development, and the Clark Fund. We have benefitted from excellent research assistance by Paris Cleanthous, Joshua Greenfield, William Powers, and especially Pao-Li Chang. Shang-Jin Wei graciously provided us with distance data. The views expressed in this paper are those of the authors and do not express any official view of the Federal Reserve system.

Correspondence to David E. Weinstein, University of Michigan Business School, Ann Arbor, MI 48109
Phone: Davis (212) 720-8401, Weinstein (734) 936-2866, e-mail: ddavis@fas.harvard.edu, weinstei@umich.edu

I. Global Factor Service Trade

Theory strives to be simple, rich, and robust. When it succeeds, it attains extraordinary influence. No doubt this explains the ubiquity of Heckscher-Ohlin theory in the field of international trade, particularly in the Factor Price Equalization (FPE) version of Samuelson (1947).[1] Countless theoretical and empirical studies have built on this foundation.[2]

The FPE theory's most impressive feature is its extraordinary ambition. It proposes to describe, with but a few parameters, and in a unified constellation, the endowments, technologies, production, absorption, and trade of all countries in the world. This juxtaposition of extraordinary ambition and parsimonious specification has made the theory irresistible to empirical researchers.[3]

---

[1] Since Samuelson's work, theorists in the FPE tradition have continued to make important advances. They have explored the robustness of the original insight and re-interpreted the model in a manner suitable for empirical implementation. Samuelson (1953) extends the theory to multiple factors and goods. Jaroslav Vanek (1968) re-interprets the model as one of trade in factor services. Avinash K. Dixit and Victor A. Norman (1980) formalize an allegory from Samuelson (1949) to provide the deep economic intuition of the FPE model. Elhanan Helpman (1981) and Helpman and Paul R. Krugman (1985) demonstrate that the essential prediction for trade in factor services is robust to a variety of specifications with scale economies and imperfect competition.

[2] Just in the last few years it has been the preferred framework for a vigorous discussion of the impact of international trade on wages and employment [Krugman (1996), Leamer (1995), Lawrence and Slaughter (1993), Davis (1998)]. Similarly, O'Rourke and Williamson (1994) have used it to study the consequences of Anglo-American commodity price convergence for factor price convergence.

[3] This ambition stands in stark relief when compared to that of the two other approaches that have dominated research in empirical trade. Leamer (1984) explored the empirical validity of the H-O model with FPE in the so-called "square" case. This has contributed importantly to our understanding of trade. However Leamer cautions that, in contrast to HOV empirics, these do not provide a "complete" test of the model, since it does not employ any data on technology. The importance of this caution is underscored by the recent work of Bernstein and Weinstein (1997), which confirms that the estimated parameters of the square model fail to have the structural interpretation theory imposes. The other principal approach to empirical trade is the gravity equation. While there are multiple general equilibrium theories that yield a gravity equation,

In recent years, empirical research has focused on a relatively robust version of the theory, embodied in the Heckscher-Ohlin-Vanek (HOV) theorem. The HOV theorem yields a simple prediction: The net export of factor services will be the difference between a country's endowment and the endowment typical in the world for a country of that size. The prediction is elegant, intuitive, and spectacularly at odds with the data.

Wassily A. Leontief's (1953) "paradox" is widely regarded as the first blow against the empirical veracity of the factor proportions theory. Confirmation of the paradox in later work led Keith E. Maskus (1985) to dub it the "Leontief commonplace." In one of the most widely-cited and seemingly-damning studies, Harry P. Bowen, Edward E. Leamer, and Leo A. Sveikauskas [henceforth BLS] (1987) report that the factor services a country will on net export are no better predicted by measured factor abundance than by a coin flip.

The work of Daniel Trefler (1993,1995) played an extremely important role in reviving interest in tests of HOV and exploring the dimensions in which the model was failing. Of particular importance is Trefler's identification of several HOV "mysteries." The latter paper follows up on BLS to consider a variety of departures from the simple FPE version of HOV. While some progress is made, Trefler's results fail to bring the theory and data into reasonable congruence. This point is underscored in work by Gabaix (1997).

Donald R. Davis, David E. Weinstein, et al. (1997) do report positive results for the HOV model. However they accomplish this by restricting the sample for which FPE is assumed to hold

---

empirical implementation typically makes no use whatsoever of the underlying production theory. Instead it predicts the bilateral trade levels (aggregate or by industry) for *given* levels of output across countries.

to regions of Japan and by remaining agnostic about the degree to which the FPE framework can be extended across nations.

In sum, a half-century of empirical work has failed to find simple amendments that allow the theory to provide a unified description of the international data. Nevertheless, the effort has been instructive. A key contribution of Trefler (1995) is his identification of systematic discrepancies between the theory and the international data. Chief among these is the so-called "mystery of the missing trade." In simple terms, the mystery is that measured factor service trade is an order of magnitude smaller than predicted factor service trade based on national endowments. To date, the mystery remains one of the great challenges in understanding the international data (cf. Gabaix 1997).

The salient feature of the recent research is to ask if parsimonious amendments allow the model to match the data. The research focuses on two classes of amendments: technology and absorption. The technological assumptions considered include cross-country differences, either Hicks-neutral or factor-augmenting, and industry-level economies of scale [Trefler 1993, 1995; Gabaix 1997; Antweiler and Trefler 1997]. The assumptions about absorption introduce non-homotheticities, most prominently a home-bias in demand [BLS, Trefler 1995].

The search for parsimonious amendments that allow the model to work is precisely the right research strategy. However, the existing literature has one major drawback. The hypothesized amendments concern technology and absorption. Yet the empirical tests draw on only a single direct observation on technology (typically that of the United States) and no observations whatsoever concerning absorption.[4] Hence even if these hypotheses improve the model's performance by selected statistical measures, it remains uncertain if the estimated parameters have a structural interpretation in terms of the economic fundamentals.[5]

In the present study, we likewise search for parsimonious amendments that allow the HOV model to work. However, in contrast to all prior work, we have sufficient data on technology and absorption to estimate the structural parameters directly. Having estimated these directly from the data of interest, we then impose the resulting restrictions in our tests of the HOV model. By starting with the basic model and relaxing one assumption at a time, we see precisely how improvements in our structural model translate into improvements in the fit of the HOV predictions.

---

[4] An exception is Hakura (1997). She has technology matrices for five EC countries and reports positive results for HOV on tests that may appear similar to ours, but which are in fact quite distinct. Hakura works with specifications that draw on and extend Staiger, Deardorff and Stern (1987). She runs three key tests. The first employs the standard HOV assumption of identical technologies (here using the German coefficients). The second test uses each country's *true* technology matrix. The third test purges the data of the influence of foreign intermediates. Hakura's tests are valuable in providing some suggestion that getting a better handle on the nature of technological differences will be an important part of getting HOV to work. Unfortunately, her approach also has some major drawbacks. First, the Staiger, et al. approach differences the factor content predictions. This means that departures from the HOV predictions which are systematic and proportional to country size will be masked. While this has its value, the fact that it is done for five relatively similar countries means that even as a test of the identical technologies version of HOV, it is very weak. Second, what fits identically as a matter of data construction cannot meaningfully be tested. Hence the production side of the HOV model — what we think of as the core intellectual content — fits by construction and cannot be tested in her framework. Finally, her most positive result is that as much as 94 percent of the sign tests are correct if you exclude pairs with Italy. However, with only 5 countries, there are only 10 bilateral relations. Italy will figure in 4 of these 10. Hence excluding Italy, for which no real reason is provided, amounts to omitting 40 percent of the observations.

[5] Cf. Helpman (1998).

corresponding improvement in measures of model fit. Countries export their abundant factors and in approximately the right magnitude. The results are remarkably consistent across variations in

receives powerful support in our study.

II. Theory

standard form, it does not describe the world that actually exists. Various hypotheses have been advanced to account for the divergence of theory and data, such as technical differences and

between theory and data. Nonetheless, they are likely to be part of a complete account. We advance several new hypotheses with the hope of providing a first successful match.

the standard model. In order to understand the role played by each of the assumptions, it is important, both in the theory and empirics, to begin with the standard model, relaxing the

implemented empirically in the following section.

A. *The Standard HOV Model*

We begin by developing the standard HOV model from first principles. Assume that all factors are perfectly competitive. There are no barriers to trade and transport costs are zero. The number of tradable goods is at least as large as the number of primary factors. We assume that the distribution of these factors across countries is consistent with the world replicating the integrated equilibrium (cf. Helpman and Krugman 1985). Then factor prices will be equalized, so all producers will choose the same techniques of production. Let the matrix of total factor inputs for country c be given by B^{c}. The foregoing implies that for all countries c:

$$B^{c} = B^{\quad\prime} \qquad \forall\, c, c^{\prime}$$

These assumptions enable us to use a single country's technology matrix (in prior studies, typically that of the US) in order to carry out all factor content calculations. We now can relate endowments and production:

$$B^{c} Y^{c} = V^{c} = B^{c^{\prime}} Y^{c}$$

The first equality is effectively a factor market clearing condition, while the second embodies the assumption of FPE.

The standard demand assumption is based on identical and homothetic preferences across countries. With free and costless trade equalizing goods prices and FPE equalizing non-traded goods prices, the demand in a country will be proportional to world net output:

$$D^{c} = s^{c} Y^{W}$$

Pre-multiplying this by the matrix of total factor inputs converts this to the factor contents:

$$B^{c^{\prime}} D^{c} = s^{c} B^{c^{\prime}} Y^{W} = s^{c} V^{W}$$

The first equality follows simply from the assumption of identical homothetic preferences and common goods prices. The second relies on the fact that FPE insures that all countries use the common technology matrix $B^{c'}$.

Collecting terms, we can state the two key tests of the standard HOV model:

Production Specification (P1) $B^{c'} Y^c = V^c$ for a specified common technology matrix $B^{c'}$.

Trade Specification (T1) $B^{c'} T^c = B^{c'}(Y^c - D^c) = V^c - s^c V^W \quad \forall c$

B. *A Common Technology Matrix Measured With Error*

The foregoing assumes that both the true and measured technology matrices are identical across countries. A glance at the measured technology matrices reveals this is not the case. Before we pursue more elaborate hypotheses on the nature of actual technological differences, it is worth investigating the case in which the technology matrices are measured with error. Assume, then, that the measured technology matrix for country c is given as:

$$B^c = B^\mu \epsilon^c$$

where ϵ^c is a matrix of errors. This gives rise to our second set of tests:

Production Specification (P2) $B^\mu Y^c = V^c$

Trade Specification (T2) $B^\mu T^c = V^c - s^c V^W \quad \forall c$

C. *Hicks-Neutral Technical Differences*

A wide body of literature, both in productivity and in trade, suggests that there are systematic cross-country differences in productivity, even among the richest countries [e.g.

Jorgenson and Kuroda (1990)]. This is very likely an important reason why Trefler (1995) found that the data suggests poor countries are "abundant" in all factors and vice versa for the rich countries. Bowen, et al. (1987) and Trefler (1995) have focused attention on Hicks-Neutral technical differences as a parsimonious way to capture these effects. Under this hypothesis, the technologies of countries differ only by a Hicks-neutral shift term.[6] This can be characterized via country-specific technology shifts λ^c:

$$B^c = \lambda^c B^\lambda \quad \forall c$$

In order to implement an amended HOV equation, it is convenient to think of the productivity differences as reflecting efficiency differences of the factors themselves (rather than technology per se). For example, if we take the US as a base and US factors are twice as productive as Italian factors, then $\lambda^{\text{Italy}} = 2$. In general, we can express a country's endowments in efficiency terms:

$$V^{cE} = \frac{V^c}{\lambda^c} \quad \forall c$$

The standard HOV equation then holds when the endowments of each country are expressed in efficiency units:

Production Specification (P3) $B^\lambda Y^c = V^{cE}$

Trade Specification (T3) $B^\lambda T^c = V^{cE} - s^c V^{WE} \quad \forall c$

All succeeding models and the associated empirical specifications will be in efficiency units, although we will henceforth suppress the superscript E for simplicity.

---

[6] Bhagwati (1964) first considered a combined Heckscher-Ohlin and Ricardian model.

D. *Dornbusch-Fischer-Samuelson Model*

So far we have allowed differences in input coefficients across countries only as a Hicks-neutral shift. For cases of adjusted FPE, this implies that capital to labor ratios are fixed by industry across countries. However, there is good reason to believe this is not the case. The simple Rybczynski relation suggests that countries with a relatively large stock of capital should have an output mix shifted toward relatively capital intensive goods, but with FPE they should not use different input coefficients within any individual sector. Dollar, Wolff, and Baumol (1988) estimated cross-country differences in capital to labor usage and found this was correlated with country capital abundance. They interpreted this as evidence against the FPE model, although they recognized that aggregation might be a problem. We develop this insight in a model that accounts for the positive relation between country and industry capital to labor usage, yet preserves (approximate) FPE and the simple HOV prediction. This is valuable in that it will provide a first set of theoretical insights that help us to understand why the mystery of the missing trade might arise in previous data exercises, even if the HOV prediction is being met. In the following section we will go on to consider the question of how to pursue the problem if indeed FPE has broken down.

In order to make our discussion compact, we will provide only a sketch of the model that provides the essential insights. Consider as a starting point the Dornbusch-Fischer-Samuelson (1980) continuum Heckscher-Ohlin model. Goods are arrayed on the unit interval with continuous and strictly increasing capital to labor ratios by sector. We consider first the integrated equilibrium (cf. Helpman and Krugman 1985). The FPE set is depicted in Figure 1 as a “Deardorff lens,” reflecting the factor intensities and usages for the corresponding sectors.[7] Assume that the point dividing the world endowments between the two countries lies within the FPE set. The factor content of production for each country is its endowment V^c. With identical homothetic preferences, common goods prices, and production with the integrated equilibrium techniques, the factor content of absorption is $s^c V^W$. Together these yield the standard HOV prediction for the net factor content of trade: $V^c - s^c V^W$. However we also know that with more goods than factors, the pattern of goods production, so also the pattern of goods trade, is not determinate.[8]

In order to make the trade and production patterns determinate, we resort to an artifice originally introduced by Samuelson (1954) and considered within the continuum framework by Xu (1993). Imagine that all goods have iceberg transport costs, so that if $\tau > 1$ units are shipped, only one unit arrives. We will think of these trade costs as being strictly positive but arbitrarily small. Goods prices will be arbitrarily close to those of the integrated equilibrium, as will absorption and so the net factor content of trade. However, as Samuelson suggested, trade will be arranged so as to minimize trade costs. Since all goods are assumed to have the same proportional costs, this is equivalent to minimizing the volume of trade subject to achieving (approximately) the HOV-required net factor content. This problem has a very simple solution: insofar as possible, the capital abundant country will concentrate its exports among the very most capital intensive goods (call them the *X*-goods), the labor abundant country will concentrate its exports among the

---

[7] See Deardorff (1994).

[8] The relative number of goods versus factors may appear to be an esoteric, nearly imponderable concept. Not so. Bernstein and Weinstein (1997) show that a framework in which the number of goods exceeds the number of factors is a very useful way to think about the determinacy of production patterns in regional versus international data.

most labor-intensive goods (call them the *Y*-goods). Goods of intermediate factor intensity (*N*-goods) will not be traded. Thus the real equilibrium features a pattern of perfect specialization in the goods that are (in equilibrium) traded. The capital abundant home country produces only *X* (its export) and *N*, while the labor abundant country produces only *Y* (its export) and *N*.

Of course, one will find such extreme production specialization nowhere in real data. So we must discuss how the empirical industries in our data sets match up with the real equilibrium described above. It is well-known that the industrial classification system was not designed with the concerns of Heckscher-Ohlin researchers in mind. Hence actual industrial classification, in contrast to our theoretical industries, includes goods of very heterogeneous capital to labor ratios. Consider two industries, 1 and 2. Assume that on average industry 1 has a tendency to include the more capital intensive goods, but that it actually includes goods from *Y*, *N*, and *X*. Similarly, assume industry 2 tends to include more goods in the labor intensive sectors, but also includes goods from *Y*, *N*, and *X*. For simplicity, assume the densities for sectors 1 and 2 are uniform over each of the intervals *Y*, *N*, and *X* (taken separately). A schematic representation appears in Figure 2.

Think now about how previous tests have been implemented. Call the capital abundant country the US. Prior tests have used the US technology matrix to measure the factor content of trade. Consider how the input coefficients are constructed for the empirical US industry 1. Let B_X be the column of average input coefficients for goods in the *X* sector and B_N be the column of average input coefficients for goods in the the *N* sector. Then the measured input coefficients for sector 1 will be:

$$B_1 = \pi_1 B_X + (1 - \pi_1)B_N + 0\ B_Y$$

The weight π_1 is determined by the *X*-sector's weight in US output in sector 1 and we include the zero-weighted term B_Y to emphasize that it does not figure at all into calculation of the US technical coefficients. Note that the coefficients so estimated are a weighted average of the goods that the US actually exports (X) and goods with much more labor-intensive coefficients (N). That is, the estimated technology matrix will tend to understate the capital content and overstate the labor content of US exports. The consequence is to bias our measures of net factor trade toward zero. A parallel calculation for industry 2 would reveal the same downward bias in the US net factor content.

Now consider what happens if we apply the coefficients B_1 taken from the US to exports by the labor abundant country. Again, B_1 is a weighted average of US input coefficients in N and X. But the labor abundant country exports only Y goods — which are more labor intensive than *either X* or N. That is, use of the measured US technology matrix will strongly overstate the capital content of the labor abundant country's exports, while underestimating the labor content. Use of the US technology matrix biases measures of the factor content of trade in *both* countries toward zero.

While the theoretical model is special in some respects, it does highlight two insights that we believe are more general than this example. The first is a pointed reminder that goods produced in different countries that are classified in the same industrial categories need not be the same goods at all. When we ignore this fact, we may well miss an important component of net factor trade. Second, insofar as trade in factor services is one motive for trade, when there are many goods that could embody this factor service trade, there will be an incentive to focus exports among those goods most intensively using the abundant factors. Hence average input coefficients for any country are likely to understate the true factor content of trade.[9]

How would one know whether these theoretical problems are a real feature of the data? One consequence would be that industry factor usage will vary systematically with country capital abundance. We will explore this more fully below when we estimate the extent to which this affects factor ratios by industry across countries. The consequence here is twofold. First, we have to recognize that the technology matrices will differ systematically by country capital abundance, and so construct technology matrices that reflect this. Second, we will likewise need to recognize that the factor content of absorption must be measured bilaterally with the producing country's technology matrix. With these two points in mind, it is relatively simple to derive the key expressions:

Production Specification (P4) $$B^{cDFS} Y^{c} = V^{c}$$

Trade Specification (T4) $$B^{cDFS} Y^{c} - \left[B^{cDFS} D^{cc} + \sum_{c \neq c'} B^{c'DFS} M^{cc'} \right] = V^{c} - s^{c} V^{W}$$

where the superscript in B^{cDFS} reflects the fact that in the continuum of goods, Dornbusch-Fischer-Samuelson model, the unit input requirements *in the tradable goods sectors* will vary in accordance with the country's capital to labor ratio.

---

[9] In the test that follows, we will focus on the resulting specialization. With the present data we are not able to examine directly the difference between average and marginal capital intensity of the relevant sectors. We do hope to look at this further based on microeconomic data on inputs and trade behavior at the firm level.

E. *Case Without Factor Price Equalization*

Helpman (1998) proposes an account of the missing trade in the same spirit as the continuum model but which focuses on more substantial departures from FPE and the existence of specialization "cones" of production in tradables. One consequence of this is that the common set of non-traded goods will be produced using different techniques. In turn, this will affect our HOV factor content predictions. We now consider the implications.[10]

To arrive at a definite result, we need to apply a little more structure on demand than is standard. Consider a world with any number of countries, two factors (capital and labor) and in which the extent of differences in endowments is sufficient that at least some countries do not share factor price equalization. We do not restrict the number of non-traded goods, although we assume that the number of traded goods is sufficiently large that we can safely ignore boundary goods produced by countries in adjoining production cones with different production techniques. Define a country c's technology matrix at equilibrium factor prices in the Helpman no-FPE model as $B^{cH} = [B^{cHN}\ B^{cHT}]$, where the division is between non-tradables and tradables. Let output be similarly divided, so $Y^c = \begin{bmatrix} Y^{cN} \\ Y^{cT} \end{bmatrix}$. Then the factor content of production, by factor market clearing, is $B^{cH} Y^c = V^c$. If we separate out non-tradables and rearrange, we get $B^{cHT} Y^{cT} = V^c - B^{cHN} Y^{cN}$. Let us term the expression on the right $V^c - B^{cHN} Y^{cN} \equiv V^{cT}$, so that $B^{cHT} Y^{cT} = V^{cT}$. With no FPE, the price of non-traded goods in terms of tradables will typically differ across countries. Assume that preferences in all countries between tradables and non-tradables are similar

---

[10] Wood (1994) likewise emphasizes that input coefficients differ within the same industry for traded goods produced in a developing country as opposed to a developed country. His work addresses the consequence of this for studies of wage changes linked to the factor content of imports rather than using it to think about tests of HOV.

and Cobb-Douglas, so feature fixed expenditure shares. Let s^c be country c's share of world income (in units of tradables). Then it follows that s^c is also c's share of world spending on tradables.

Assume that preferences across countries for tradables are identical and homothetic. The absorption by country c of tradable goods produced in c' is then $D^{cc'T} = s^c Y^{c'T}$. The factor content of this absorption, using the factors actually engaged in production of the good, is $B^{c'T} D^{cc'} = s^c B^{c'TH} Y^{c'T} = s^c V^{c'T}$. Define $V^{WT} \equiv \sum_c V^{cT}$ and note that for $c \neq c'$, $D^{cc'T} \equiv M^{cc'}$ (imports). Then it follows that:

$$B^{cHT} Y^{cT} - [B^{cHT} D^{ccT} + \sum_{c' \neq c} B^{c'TH} M^{cc'}] = V^{cT} - s^c V^{WT}.$$

That is, under the conditions stated above, we get something very like the simple HOV equation so long as we restrict ourselves to world endowments devoted to tradable production and weight absorption according to the actual coefficients employed in production.

We now need to contemplate the implications of this model for what we will observe in the data. We know that input coefficients both in tradables and in non-tradables will differ across countries. The failure of FPE plays a role in both cases, but they have important differences. The input coefficients differ in tradables because the failure of FPE has led the countries to specialize in different goods. They differ in non-tradables because the same goods are produced with different factor proportions. Let us expand the equation above:

$$B^{cHT} Y^{cT} - [B^{cHT} D^{ccT} + \sum_{c' \neq c} B^{c'HT} M^{cc'}] = V^{cT} - s^c V^{WT} = [V^c - B^{cHN} Y^{cN}] - s^c [\sum_{c'} \{V^{c'} - B^{c'N} Y^{c'N}\}]$$

The RHS of this equation can be re-arranged to be:

$$= [V^c - s^c V^W] - [B^{cHN} Y^{cN} - s^c \sum_{c'} B^{c'HN} Y^{c'N}]$$

If we denote by V^{cN} the resources devoted in country c to production of non-tradable goods (and correspondingly for the world), then our production and trade tests can be written as:

Production Specification (P5) $B^{cH} Y^{c} = V^{c}$

Trade Specification (T5):

$$B^{cHT} Y^{cT} - [B^{cHT} D^{ccT} + \sum_{c' \neq c} B^{c'HT} M^{cc'}] = [V^{c} - s^{c} V^{W}] - [V^{cN} - s^{c} V^{WN}]$$

where the superscript in B^{cH} reflects the fact that in the no-FPE model, all input coefficients in a country's technology matrix will vary according to the country's capital to labor ratio.

The first term on the RHS in T5 is the standard HOV prediction, while the second is an adjustment that accounts for departures in factor usage in non-tradable goods from the world average. Note, for example that a capital abundant country will have high wages, inducing substitution in non-tradables toward capital. The second term in brackets will typically be positive then for the case of capital, meaning that the simple HOV prediction overstates how much trade there really ought to be in capital services. In the same case, the actual labor usage in non-tradables is less than the world average, and so the simple HOV equation will tend to overstate the expected level of labor service imports. In both cases, the new prediction for factor service trade will be less than that of the simple HOV model.

Trade specification (T6) will in theoretical terms be the same as (T5); as we will see, it will differ only in how technology for the ROW is implemented.

F. *Demand, HOV, and Gravity*

Among the more outlandish simplifications in the HOV model is the assumption that international trade is wholly costless. This is false on its face and overwhelmingly refuted by the data [McCallum (1995), Engel and Rogers (1995)]. The consequence for understanding net factor trade is that the frictionless model wildly overstates the expected volume of trade and so also overstates the opportunities for arbitrage of factor price differences. This may well be important in understanding why in previous studies measured factor trade is below that predicted by the frictionless model.

A key question, then, is how to incorporate trade frictions into our framework. One approach would be to take direct measures of costs of trade such as tariffs, non-tariff barriers, transport costs, etc. and to combine this with information about import demand elasticities to calculate departures from the predictions of the frictionless model. While this would likely push in the right direction, it has severe shortcomings. There are good reasons to believe that these overt measures of costs of trade are inadequate to the task. Harrigan (1993) shows that the measures we have for NTBs show very little impact on trade volumes in the OECD relative to tariffs, which are themselves low. Nonetheless, as McCallum (1995) and Helliwell (1998) have shown, the international trade volumes are much too low to be accounted for by the level of tariff barriers and direct trade costs we observe.

Since these direct measures of trade costs are likely to prove problematic, we turn to the model of trade volumes known as the gravity equation, which uses distance as a proxy for costs of trade. This model of trade volumes is known to be very successful. As appears in Anderson (1979), Bergstrand (1985), and Deardorff (1998), a gravity equation arises very simply when there is a high degree of production specialization, similarity of preferences, and costs of trade. The gravity model has not appeared previously in empirical tests of the HOV factor content predictions. The reason is that the bilateral trade relations posited in the gravity model are not typically well defined in a many-country HOV model [see Deardorff (1998) and Trefler (1998)]. However, they are well defined in the production model we have developed precisely because all countries feature perfect specialization in tradables.[11]

In this case, the demand for imports bilaterally has to be amended to account for bilateral distance. Let $d_{cc'}$ be the distance between countries c and c′. Then a simple way to introduce trade costs is to posit that import demand in country c for products from c′ takes the form of a standard gravity equation:

$$\ln\left(M_i^{c\,c'}\right) = \alpha_{0i} + \alpha_{1i}\ln\left(s_i^{Tc}\, X_i^{c'}\right) + \delta_i \ln\left(d_{c\,c'}\right) + \ln\left(\zeta_i^{c\,c'}\right)$$

where s_i^{Tc} is total domestic absorption (of final and intermediate goods) as a share of world gross output, $X_i^{c'}$ is gross output in sector i in country c', the α's and the δ are parameters to be estimated, and ζ is a log normal error term. We can then estimate the log form of this equation to obtain parameter estimates. These can then be used to generate predicted imports, $\hat{M}_{cc}$. Reversing the sign of this matrix then gives us predicted bilateral exports. While the gravity model is quite successful at predicting bilateral import flows, own demand seems to be determined by a quite different process (see McCallum (1995)). Rather than trying to model this process directly, we decided to solve the problem with a two-step procedure. First we assumed that total demand in each sector is equal to a country's share of world demand for final goods times world demand in that sector. We then set demand for domestically produced goods as equal to the difference between its total demand and predicted imports from the gravity equation.

---

[11] Rather than provide the full derivation here, readers interested in learning more about how the gravity can be integrated into our framework should see Anderson (1979). Anderson provides a cogent and clear analysis of how a gravity equation with a negative coefficient on distance can be derived in a world with perfect specialization and Cobb-Douglas utility.

Measured net factor trade will be exactly the same as in T5 above. However, in this case the predicted factor content of absorption of tradables is no longer $s^c\, V^{WT}$. Instead the predictions for bilateral absorption must be those generated by the gravity specification, weighted by the factor usage matrices appropriate to each partner. Let carets indicate fitted values. This gives rise to:

Trade Specification (T7)

$$B^{cH} Y^c - \left[B^{cH} D^{cc} + \sum_{c' \neq c} B^{c'H} M^{cc'} \right] = V^c - \left[B^{cH} \hat{D}^{cc} + \sum_{c' \neq c} B^{c'H} \hat{M}^{cc'} \right]$$

III. Data Sources and Issues

A. *Data Sources*

An important contribution of our study is the development of a rich new data set for testing trade theories. This has been a major project on its own. We believe that the data set we develop is superior to that available in prior studies in numerous dimensions. The mechanics of construction of the data set are detailed in a Data Appendix. Here we provide a brief description of the data and a discussion of the practical and conceptual advances.

The basis for our data set is the OECD's Input-Output Datase [OECD (1995)]. This database provides input-output tables, gross output, net output, intermediate input usage, domestic absorption and trade data for ten OECD countries.[12] Significantly, all of this data is designed to be compatible across countries. We constructed the country endowment data and the

---

[12] Australia, Canada, Denmark, France, Germany, Italy, Japan, Netherlands, United Kingdom, and the United States. These are the ten available in the IO database.

matrices of direct factor input requirements using the OECD's Inter-Sectoral Database and the OECD's STAN Database. Hence for all countries, we have data on technology, net output, endowments, absorption, and trade. By construction, these satisfy:

1) $B^c Y^c \equiv V^c$.

2) $Y^c - D^c \equiv T^c$

We also have data for 20 other countries that we refer to as the "Rest of the World" or ROW.[13] Data on capital is derived from the Summers and Heston Database while that for labor is from the International Labor Organization. For countries that do not report labor force data for 1985 we took a labor force number corresponding to the closest year and assumed that the labor force grew at the same rate as the population. Gross output data is taken from the UN's Industrial Statistics Yearbook, as modified by DWBS. Net output is calculated by multiplying gross output by the GDP weighted average input-output matrix for the OECD and subtracting this from the gross output vector.

Bilateral trade flows for manufacturing between each of our ten OECD countries as well as between each country and the ROW was drawn from Feenstra, Lipsey, and Bowen (1997) and scaled so that bilateral industry import totals match country totals from the IO tables. Bilateral imports for non-manufacturing sectors are set equal to the share of manufacturing imports from

---

[13] Argentina, Austria, Belgium, Finland, India, Indonesia, Ireland, Israel, Korea, Mexico, New Zealand, Norway, Philippines, Portugal, Singapore, South Africa, Spain, Sweden, Thailand, and Turkey. These are the countries for which either gross output or value added is available for all sectors.

that country times total non-manufacturing imports in that sector.[14] ROW absorption was then set to satisfy condition 2.[15]

In sum, this data set provides us with 10 sets of compatible technology matrices, output vectors, trade vectors, absorption vectors, and endowment vectors. In addition we have a data set for the ROW that is comparable in quality to that used in earlier studies.

B. *Data Issues*

We would emphasize several characteristics of the data to underscore its advantage over prior data sets. The first draws on the nature of the tests considered. The prior work is uniform in rejecting the simplest HOV model. Hence the most interesting work has gone on to consider alternative hypotheses. Importantly, the most prominent of these theories concern alterations in assumptions about technological similarity across countries (e.g. Hicks-neutral technical differences) and the structure of absorption (e.g. a home bias in demand). Yet typically these studies have only a single observation on technology (that of the US) and no observations

---

[14] This is not ideal, but given that the median ratio of imports to gross output in non-manufacturing for our sample of countries is 1 percent, this is not likely to introduce large errors.

[15] It is reasonable to ask why we aggregated the ROW into one entity rather than working with each country separately. A major strength of this paper is that our data are compatible and of extremely high quality. Unfortunately, the output and endowment data for the ROW countries are extremely noisy (See Summers and Heston for a discussion of problems with the endowment data). It is quite difficult to match UN data with OECD IO data because of aggregation issues, varying country industry definitions, and various necessary imputations (see DWBS for details on what calculations were necessary). As a result, the output and absorption numbers of any individual country in the ROW is measured with far more error than OECD data. To the extent that these errors are unbiased, we mitigate these measurement errors when we aggregate the ROW. Ultimately we decided that we did not want to pollute a high quality data set with a large number of poorly measured observations.

whatsoever about the structure of absorption. The technological and absorption parameters are chosen to best fit the statistical model, but these yield little confidence that they truly do reflect the economic parameters of interest.[16] Our construction of the technology matrices allows us to test the theories of technological difference directly on the relevant data and similarly for our hypotheses about absorption. This ability to directly test the cross-country theories of interest greatly enhances our confidence that the estimated technology and absorption parameters indeed do correspond to the economic variables of interest.

A second issue is the consistency with which the data is handled. In part this corresponds to the fact that we are able to rely to a great extent on data sources constructed by the OECD with the explicit aim to be as consistent as practicable across sources. In addition, the OECD has made great efforts to insure that the mapping between output data and trade categories is sound. Finally, the consistency extends also to conditions we impose on the data which should hold as simple identities, but which have failed to hold in previous studies because of the inconsistencies in disparate data sets. These restrictions include that each country actually uses its own raw technology matrix, reflected in $B^c Y^c \equiv V^c$.[17]

We would also like to note, though, that the desire to bring new data sources to bear on the problem has carried a cost. Specifically, the factors available to us for this study are limited to capital and aggregate labor. We would very much have liked to be able to distinguish skilled and

---

[16] Helpman (1998).

[17] See the discussion in BLS of related difficulties. The only exception to this is the ROW where we were forced to use an estimated *B*. See section IV for details.

unskilled workers, but unfortunately the number of skilled and unskilled workers *by industry* is not available for most countries.

We would like to note how the reader should think about this factor “aggregate labor,” and why we do not believe this presents too great a problem for our study. There are at least a couple of interpretations that can be given. A first fact about our labor variable is that under most specifications the OECD countries are judged scarce in labor while the ROW is abundant in it. This suggests that one appropriate interpretation is that our variable labor is a very rough proxy for unskilled or semi-skilled labor. Note, though, that in most of our later implementations, labor is converted to efficiency units. If this is an appropriate way to merge skilled and unskilled, then the fact that these OECD countries are scarce in it suggests that this is true, even when we convert all labor to common efficiency units. We have little doubt that if it were possible to distinguish highly-skilled labor separately for our study that the US and some of the other OECD countries would be judged abundant in that factor.

These reservations notwithstanding, we believe that there are good reasons to believe that choice of factors does not confer an advantage to us over prior studies. Many of the factors we omit are land or mineral factors, which were the best performers for BLS and Trefler (1995). Hence their omission should only work against us. As we will see below, the factors that we do include exhibit precisely the pathologies (e.g. “mystery of the missing trade” and the Leontief paradox) that have characterized the data in prior studies. Finally, in our study of the net factor trade of Japanese regions, DWBS (1997), we were able both to include more factors and to distinguish between skilled and unskilled workers. The HOV theory performed admirably in these circumstances. These points suggest that, if anything, the availability of factor data for our study may make it more difficult to find positive results for HOV, not easier.

In sum, we have constructed a rich new data set with compatible data for 10 OECD countries across a wide range of relevant variables. Importantly, we introduce to this literature direct testing on technology and absorption data of the central economic hypotheses in contest. Finally, although in some respects the available data fell short of our ideal, we do not believe that this introduces any bias toward favorable results.

IV. Statistical Tests on Technology and Absorption

The principal hypotheses that distinguish alternative implementations of HOV concern technology and absorption. In prior work, researchers have selected technological and absorption parameters designed to allow the model of net factor trade to work as well as possible. Little data on technology (only that of the US) and no data on absorption were employed.

By contrast, our principal statistical tests will work directly with the data on technology and absorption. We consider a variety of models of technology and absorption suggested by theory and select a preferred model for each. In this respect, the formal statistical tests in this paper will be complete once we have selected the preferred models. Nonetheless, both because the principal concern of trade economists here is in measures of net factor trade, and also for direct comparability with prior studies, in Section V we will go on to implement each of the models of technology and absorption. In doing so, we will gain a rich view of the role played by each change in improving the working of the HOV model. For reference, we will indicate the production specification associated with the distinct models of technology.

A. *Estimating Technology*

Our first model of technology (P1) is the standard starting point in all investigations of HOV: it postulates that all countries use identical production techniques in all sectors. This can be tested directly using our data. For any countries c and c′, it should be the case that $B^c = B^{c'}$. We reject this restriction by inspection.

One possible reason for cross-country differences in measured production techniques is simple measurement error (P2). The Italian aircraft industry is four times as capital intensive as the US industry. While this may indicate different production techniques, the fact that net output in US aircraft is approximately 200 times larger than in Italy raises the question of whether the same set of activities are being captured in the Italian data. This raises a more general point that is readily visible in the data. Namely extreme outliers in measured B^c tend to be inversely related to sector size. In tests of trade and production theory this is likely to produce problems when applying one country's technology matrix to another country. If sectors that are large in the US tend to be small abroad, then evaluating the factor content of foreign production using the US matrix is likely to magnify measurement error. Large foreign sectors are going to be precisely the ones that are measured with greatest error in the US.

A simple solution to this problem is to postulate that all countries use identical technologies but each measured B^c is drawn from a random distribution centered on a common B. If we postulate that

$$B^c = B\epsilon^c$$

where we assume that ϵ^c is distributed log normally. This can relationship can be estimated by running the following regression:

$$\ln B_{fi}^{c} = \beta_{fi} + \epsilon_{fi}^{c}$$

here β_{fi} are parameters to be estimated corresponding to the log of common factor input requirement for factor *f* in sector *i*. We can contemplate two sources of heteroskedasticity. The first arises because larger sectors tend to be measured more accurately than smaller sectors. The second arises because percentage errors are likely to be larger in sectors that use less of a factor than sectors that use more of a factor. In order to correct for this heteroskedasticity, in all regressions we weighted all observations by the square root of the log of value added multiplied by $\overline{B}_{fi} / \overline{B}_{f}$ where $\overline{B}_{fi}$ corresponds to the average factor intensity in sector *i* and $\overline{B}_{f}$ corresponds to the average factor intensity across all sectors.[18]

As we noted earlier, there is good reason to believe that there are efficiency differences, even among the rich countries. A convenient specification is to allow for Hicks-Neutral technical differences (P3). If we denote these differences by λ^{c}, then we can econometrically identify these technical differences by estimating:

$$\ln B_{fi}^{c} = \theta^{c} + \beta_{fi} + \psi_{fi}^{c}$$

where $\exp(\theta^{c}) = \lambda^{c}$. Estimation of this specification requires us to choose a normalization for the θ^{c}. A convenient one is to set θ^{US} equal to zero (or equivalently $\lambda^{US} = 1$).

We have also suggested that it might be possible that production might be characterized by a continuum of goods DFS model in which industry input coefficients in tradables depend on

---

[18] Sectors with no value added were given a zero weight in our regressions.

country capital abundance (P4). The latter feature may arise also if FPE breaks down and countries are in different production cones (P4), in which case this will affect production coefficients in non-traded sectors as well. These models can be easily implemented. We postulate that input coefficients are characterized by the following equation:

$$\ln B_{fi}^{c} = \theta^{c} + \beta_{fi} + \gamma_{fi} \ln\left(\frac{K_c}{L_c}\right) + \phi_{fi}^{c}$$

Once again we need to choose a normalization. A convenient choice is that country capital-to-labor ratios should not affect country productivity levels. This is tantamount to requiring that

$$\sum_{fi} g_{fi} = 0$$

.

This last specification can be estimated either in an unpooled specification in which we allow each sector to have a different γ_{fi} or in a pooled specification in which we only allow the γ_{fi} to vary by factor and whether the good is traded or non-traded.

Table 1 presents the results of estimating these equations. As one can see in all specifications the θ^{c}'s (λ^{c}'s) are estimated very precisely and seem to have plausible values. The US is the most productive country with a λ^{c} of unity and Italy is the least productive with a λ^{c} of about two. Interestingly, one also sees that industry capital-to-labor ratios seem to move in concert with country capital-to-labor ratios. Our Schwartz model selection criterion clearly favors the pooled model with neutral technical differences and no factor price equalization (P5).[19]

---

[19] Implicitly, we are assuming no problems arising from the fact that we had to impute certain elements of our technology matrix. There are two points to bear in mind on this point. First, since our imputation method replaced missing values with average international values, this

The continuum model seems to be not only the statistical model of choice, but very important economically as well. The estimated coefficients indicate that a one percent increase in the country's capital-to-labor ratio typically raises each industry ratios by about 0.85 percent. Given that capital-to-labor ratios move by a factor of two across the ten countries for which we have IO data, this translates into large systematic movements in unit input coefficients across countries and within industries.[20]

Furthermore, we can use this approach to test for factor price equalization within our model. If (approximate) FPE holds, then specialization in traded goods may give rise to the observed differences in input coefficients within industries across countries. However, in the non-traded goods sector there should be no systematic variation in factor input ratios. Our estimates indicate that a one percent increase in a country's capital-to-labor ratio corresponds to a 0.8 percent increase in capital intensity in tradables and a 0.9 percent increase in non-tradables.[21]

---

tended to work against models P3-P5. Second, when we tried industry by industry estimation of B^c_{Ki}/B^c_{Li} on a constant and K^c/L^c we found a very strong positive relationship between industry and country capital intensity in almost every sector even when we dropped all constructed data.

[20] Although we do not report it in our tables, we also examined versions of P4 and P5 that do not allow for neutral technical shifts. The idea was to identify the specific role played by the efficiency adjustments relative to that of adjusting industry input ratios. These specifications perform only slightly better than P1 by the Schwartz criterion and only marginally better than P2 and T2 in the production and trade tests. In short, the efficiency adjustment is playing an important role, in combination with the dependence of industry input ratios on country capital abundance, in improving the performance of the HOV model.

[21] The evidence here that capital to labor ratios employed in non-traded production rise systematically with country capital abundance strongly suggests the existence of underlying differences in wage to rental ratios. This provides an interesting counterpoint to the results of Repetto and Ventura (1997).

Furthermore in both sectors we can reject the hypothesis that input coefficients are independent of country capital-to-labor ratios.

Hence, the technology data strongly support the hypothesis that the OECD production structure can be best explained by a model of specialization in tradables with Hicks-neutral technical differences and no factor price equalization (P5).

B. *Estimating Demand*

In the theoretical section we introduced our gravity model:

$$\ln\left(M_i^{c\,c'}\right) = \alpha_{0i} + \alpha_{1i}\ln\left(s_i^{\,Tc}\, X_i^{c'}\right) + \delta_i \ln\left(d_{c\,c'}\right) + \ln\left(\zeta_i^{c\,c'}\right)$$

In a zero trade cost world with perfect specialization, we have the following parameter restrictions, $\alpha_{0i} = \delta_i = 0$. If there are trade costs that increase with distance, these parameter restrictions cease to hold. We can statistically test for the existence of trade costs simply by estimating this equation and testing whether $\alpha_{0i} = \delta_i = 0$ for all *i*. Not surprisingly, the data resoundingly reject this hypothesis.

We therefore decided to use a gravity model as the basis for our demand predictions. One of the problems that we faced in implementation, however, was how to calculate the distance of any individual country to the ROW. In all specifications we calculated this distance as the GDP-weighted average distance from a particular country to all the other countries in the ROW. In some sectors we found large systematic errors in predicting trade with the ROW. This may be the result of mis-measurement of distance or the fact that the true ROW is some multiple of our sample of countries. We therefore added a dummy variable corresponding to the exporting country being the ROW and a dummy corresponding to the importing country being the ROW.

Other than this, the results of our estimation of the gravity model are entirely conventional. Typically α_{1i} is close to one in most specifications and δ_i is significant and negative in all sectors. We statistically reject the hypothesis of costless trade. We will incorporate the new gravity-based absorption model into trade specification T7.

One concern is that by employing a gravity specification, we are allowing the data to generate the "prediction." We share this concern but believe that on balance our approach is sensible for several reasons. First, costs of trade are an obviously important feature of the world which cannot be ignored if there is to be any hope of matching theory and data. Second, it is inevitable that any manner of considering the consequences of trade costs will have to use the data if only to calculate import demand elasticities and relate these to primitive measures of trade costs — approaches which have serious drawbacks of their own. Third, both the theory and our results on technology strongly endorse a gravity specification as the appropriate way to introduce trade costs precisely because they have led us to a model with specialization in tradables. Finally, each industry gravity regression has 110 observations of bilateral imports, which are used to estimate just five parameters. In short, we have deliberately treated the data with a light hand in order to avoid unduly prejudicing the results.

V. Implications for Net Factor Trade

If one takes a narrow statistical approach to the data, our work is done. Trade is a model of production and absorption. The production model has been tested using the technology matrices and the evidence clearly favors production hypothesis P5, which in turn underlies trade models T5 through T7. Furthermore, anyone familiar with the log form of the a gravity model knows that the constant term is not zero and that distance enters negatively. Hence, from a statistical point of view, model T7, which postulates a continuum model, no-FPE, and trade costs, is preferred. Why bother reading on?

The reason is that the trade literature is replete with proposed amendments to the HOV model that in the end do not help us to understand actual factor service flows. We have already evaluated the hypotheses statistically, so in this section we examine the extent to which these hypotheses help us to understand real world factor trade flows. In order to understand the economic significance of our models, we conduct tests of the HOV model of production and trade under a variety of specifications, as developed in Section II and summarized in Table 2. Here our tests are designed not for model selection, but rather to help us see the economic implications for the HOV model of each of the hypotheses that we have considered. We will begin by working primarily on the production side. Once we have made the major improvements we anticipate in that area, we move on to consider an amendment to the absorption model.

A. *Production and Trade Tests*

For each specification, we provide two tests of the *production* model. In all cases the technology matrices that we use are based on the fitted values obtained in the previous section. Furthermore, we express both measured and predicted factor content numbers as a share of world endowments in efficiency units. This adjustment eliminates the units problem and enables us to plot both factors in the same graph. The production **Slope Test** examines specifications P1 to P5 by regressing the measured factor content of production (MFCP) on the predicted factor content of production (PFCP). For example, in specification P1 this involves a regression of $B^f Y^c$ on V^{fc}. The hypothesized slope is unity, which we would like to see measured precisely and with good fit. The **Median Error Test** examines the absolute prediction error as a proportion of the predicted factor content of production. For example, for P1 this is $|B^{US} Y^c - V^c| / V^c$.

We provide three tests of the *trade* model. The first is the **Sign Test**. It asks simply if countries are measured to be exporting services of the factors that we predict they are exporting, i.e. is $sign(MFCT) = sign(PFCT)$? For example, in trade specification T1, it asks if $sign(B^f T^c) = sign(V^{fc} - s^c V^{fW})$. The statistic reported is the proportion of sign matches. The trade **Slope Test** examines specifications T1 to T5 by regressing the MFCT on the PFCT. For example, in specification T1 this involves a regression of $B^f T^c$ on $(V^{fc} - s^c V^{fW})$. The hypothesized slope is again unity, which we would like to see measured precisely and with good fit. The **Variance Ratio Test** examines the ratio Var(MFCT)/Var(PFCT). One indicator of “missing trade” is when this ratio is close to zero, whereas if the model fit perfectly the variance ratio would be unity. We also consider several robustness checks.

Before turning to our own results, it is well to have in mind how the HOV model has fared under these tests in prior work. Results from the most relevant studies are summarized in Table 3. The results lend themselves to a simple bottom line: *All prior studies on international data have fared very poorly by at least one of these measures.*

We now turn to tests of our various production and trade specifications.

B. *The Simple HOV Model Employing US Technology: P1 and T1*

We have the same point of departure as prior studies: an assumption that all countries share a common technology matrix and an implementation that uses that of the United States. However, our study is the first to examine directly the production component of this model. As one can see in Table 4, specification P1 fails miserably, but in an interesting way. A plot of P1 for all countries appears as Figure 3. The US is excluded, since it fits perfectly by construction. A glance at the plot reveals two key facts. First, for all countries and factors, measured factor content of production is always less than predicted. Second, this gap is most severe for ROW. This carries a simple message: if these countries used the US technology matrix to produce their actual output, they would need much less of each factor than they actually employ. The slope coefficient of measured on predicted factor trade is only 0.24. Excluding the ROW raises the slope coefficient to 0.67, still well short of the theoretical prediction of unity. The results by factor are presented in Table 5. The median prediction error is 34 percent for capital and 42 percent for labor. Thus our direct data on production suggest strongly that adjusting for productivity differences will be an important component in getting HOV to work.

Now consider trade specification T1. A plot appears as Figure 4. Factor abundance correctly predicts the sign of measured net factor trade only 32 percent of the time. This is significantly worse than relying on a coin flip! [22] The variance ratio is 0.0005, indicating that the variance of the predicted factor content of trade is about two-thousand times that of measured. This is missing trade big-time! And the slope coefficient is zero (actually −0.0022, s.e. = 0.0048).

---

[22] This does worse than a coin flip at the 7 percent level of significance.

Moreover, under this specification our data reveal a Leontief "paradox" in which the US is measured to be a net importer of capital services and exporter of labor services.

Since the production specification P1 performs so poorly, it is perhaps no surprise that the trade specification T1 is likewise a debacle. Nonetheless, this provides an extremely important baseline for our study precisely because it reveals that our data exhibit all of the pathologies that plague prior studies. Hence we can rule out that changes in the country sample, aggregation of many countries into a composite ROW, or the selection of productive factors suffice to account for positive results that may follow.

C. *An Average Technology Matrix: P2 and T2*

Examination of specification P1 strongly suggested that the US technology matrix is an outlier. Is it useful to think of there being an average technology matrix B^{μ} that is a good approximation to a common technology? That is the question explored in specifications P2 and T2. If we focus first on regressions based on our ten OECD countries, the slope rises sharply to 1.27, reflecting most strongly the influence of high productivity in the US. If we exclude the US as well, the slope falls to about 0.90. The R^2 in each case is respectably above 0.9. Also, in both cases, the median production errors are approximately 20 percent. The ROW continues to be a huge outlier, given its significantly lower productivity. These results suggest that use of an average technology matrix is a substantial improvement over using that of the US, since median production errors fall by one-third to one-half. Nonetheless, the fact that prediction errors are still on the order of 20 percent for the OECD group, and much larger for the ROW, suggests that there remains a lot of room for improvement.

Examination of T2 can be brief. The sign test correctly predicts the direction of net factor trade only 45 percent of the time. The variance ratio continues to be essentially zero, again indicating strong missing trade. The Slope Test coefficient is −0.006. In short, factor abundance continues to provide essentially no information about which factors a country will be measured to export. These statistics are reinforced by the pictures in Figure 5 and Figure 6. Overall, this model is a complete empirical failure.

D. *Hicks-Neutral Technical Differences: P3 and T3*

Specifications P3 and T3 are predicated on the existence of Hicks-neutral differences in efficiency across countries.[23] The estimation of these efficiency differences is discussed above in Section IV and here we view the implementation. A plot of P3 appears as Figure 7. There continue to be substantial prediction errors, the largest by far being for the ROW, but also sizable ones for the UK and Canada. Nonetheless, the median prediction error falls to about one-third of its previous level, now around 7 percent. The slope coefficient varies somewhat according to the inclusion or exclusion of the ROW, although typically it is around 0.9. When all data points are included, the R^2 is about 0.9. When we exclude ROW, the R^2 rises to 0.999.

---

[23] In this and all subsequent specifications we were forced to calculate ROW endowments in efficiency units. Since we did not have a technology matrix for the ROW we were forced to estimate this matrix based on our parameter estimates generated in section IV. We then set:

$$\lambda^c = \frac{1}{2}\,\frac{L}{(\hat{B}_L^{ROW} Y^{ROW})} + \frac{1}{2}\,\frac{K}{(\hat{B}_K^{ROW} Y^{ROW})}\,.$$

Later, when we force the technology to fit exactly for the ROW, we pick two λ's such that:

$$\lambda_f^c = f/(\hat{B}_f^{ROW} Y^{ROW})\,.$$

There is an additional pattern in the production errors. If we define capital abundance as capital per worker, then for the four most capital abundant countries, we underestimate the capital content of production and overestimate the labor content. The reverse is true for the two most labor abundant countries. These systematic biases are exactly what one would expect to find when using a common or neutrally-adjusted technology matrix in the presence of a continuum of goods. Moreover these biases are not small. Quite often these biases in over- or under-predicting the factor content of production were equal to 20 percent of a country's endowment. Thus, while allowance for Hicks-neutral efficiency differences substantially improves the working of the production model, prediction errors remain both sizable and systematic.

We have seen that the Hicks-neutral efficiency shift did give rise to substantial improvements for the production model. Will it substantially affect our trade results? The answer is definitely not. A plot of T3 appears as Figure 8. The sign test shows that factor abundance correctly predicts measured net factor trade exactly 50 percent of the time. The trade variance ratio is 0.008, indicating that the variance of predicted factor trade still exceeds that of measured factor trade by a factor of over 100. The slope coefficient is essentially zero. In sum, while the adjustment for efficiency differences is useful in improving the fit of the production model, it has done next to nothing to resolve the failures in the trade model.

E. ***The D-F-S Continuum Model with Industry Variation in Factor Employment: P4 and T4***

As we discussed in the section on estimating the technologies, there is a robust feature of the data that has been completely ignored in formal tests of the HOV model: capital to labor input ratios by industry vary positively with country factor abundance. We consider this first within the framework of the Dornbusch-Fischer-Samuelson (1980) continuum model, as this allows us to conserve yet a while longer the assumption of (approximate) factor price equalization.

Consider production specification P4, as in Figure 9. The production slope coefficient remains at 0.89, but the median production error falls slightly to 5 percent. What is most surprising is how the continuum model affects the trade specification T4. A plot appears as Figure 10. The proportion of correct sign tests rises sharply to 86 percent (19 of 22) — significantly better than a coin flip at the 1 percent level. The variance ratio remains relatively low, although at 7 percent it is much higher than in any of the previous tests. T4 is the first specification that eliminates the Leontief paradox in the US data for both capital and labor.[24] The most impressive statistic is the slope coefficient of 0.17, where all of the previous trade slopes were zero. Clearly, allowing country capital to labor ratios to affect industry coefficients is moving us dramatically in the right direction.

F. *A Failure of FPE and Factor Usage in Non-Traded Production: P5 and T5*

Our next specification considers what happens if the endowment differences are sufficiently large to leave the countries in different cones of production. In such a case, FPE will break down and non-tradables will no longer be produced with common input coefficients across countries. This specification of the production model was preferred in our statistical analysis of technology in Section IV. Our trade tests now require us to focus on the factor content of tradables after we have adjusted the HOV predictions to reflect the differences in factor usage in non-tradables arising from the failure of FPE.

---

[24] The Leontief paradox is absent in all subsequent tests.

This is our best model so far. Plots of production and trade specifications P5 and T5 appear in Figures 11 and 12. The production slope coefficient rises to 0.97, with an R^2 of essentially unity. The median production error falls to just 3 percent. We again achieve 86 percent correct matches in the sign test. The variance ratio rises to 19 percent. The slope coefficient is 0.43 for all factors, and 0.57 and 0.42 for capital and labor respectively. Again, the slopes still fall well short of unity. But this must be compared to prior work and specifications T1 to T3, all of which had a zero slope, and T4, which had a slope that is less than half as large. Under specification T5, for example, a rise of one unit in Canadian "excess" capital would lead to the export of nearly 0.6 units of capital services. The amended HOV model is not working perfectly, but given the prior results, the surprise is how well it does.[25]

G. *Corrections on ROW Technology: T6*

We have seen that production model P5 works quite well for most countries. There are a few countries for which the fit of the production model is less satisfying. There are relatively large prediction errors (ca. 10 percent) for both factors in Canada, for capital in Denmark, and for labor in Italy. Given the simplicity of the framework, the magnitude of these errors is not surprising. Since we would like to preserve this simplicity, neither do these errors immediately call for a revision of our framework.

There is one case, however, in which a closer review is appropriate. For the ten OECD countries, we have data on technology which enters into our broader estimation exercise. But this

---

[25] Implementing production model P5' (i.e. not pooling across sectors) yields results that are almost identical to model P5, and so we do not report them.

is not the case for ROW. The technology for ROW is projected from the OECD data based on the aggregate ROW endowments and the capital to labor ratio. Because the gap in capital to labor ratios between the ten and the ROW is large, there is a good measure of uncertainty about the adequacy of this projection. As it turns out, the prediction errors for ROW are large: the estimated technology matrix under-predicts labor usage by 9 percent, and over-predicts capital usage by 12 percent. Moreover, these errors may well matter because ROW is predicted to be the largest net trader in both factors and because its technology will matter for the implied factor content of absorption of all other countries.

Hence we will consider specification T6, which is the same as T5 except that we force the technology for ROW to match actual ROW aggregate endowments, i.e. $B^{ROW} Y^{ROW} \equiv V^{ROW}$.[26] A plot appears as Figure 13. This yields two improvements over specification T5. The slope coefficient rises by over one-third to 0.59 and the trade variance ratio doubles to 0.38. This suggests that a more realistic assessment of the labor intensity of ROW production materially improves the results.

H. *Adding Gravity to the HOV Demand Model: T7*

As we note in the theory section, one of the more incredible assumptions of the HOV model is costless trade. With perfect specialization and zero trade costs, one would expect most

---

[26] To maintain consistency with the foregoing, we report the results here and in T7 with all eleven countries. Because the move to T6 forces the production model of ROW to fit perfectly, we will want to consider below whether excluding the ROW points affects the main thrust of these results. We will see that it does not.

countries to be importing well over half of all goods they absorb. Simple inspection of the data reveals this to be a wild overestimate of actual import levels.

By estimating the log form of the gravity equation introduced earlier, we can obtain estimates of bilateral import flows in a world of perfect specialization with trade costs. We then use these estimates of import and own demand in order to generate the HOV factor service predictions. The results are presented in column T7 and illustrated in Figure 14. By almost every measure, this is our best model of net factor trade. The slope coefficient rises from 0.59 under T6 to 0.82 under T7. That is, measured factor trade is over 80 percent of that predicted. The standard errors are small and the R^2 is 0.98. Signs are correctly predicted over 90 percent of the time. The variance ratio rises to nearly 0.7. The results look excellent for each factor considered separately, and especially for capital, which has a slope coefficient of 0.87 and correctly predicts the direction of net factor trade in all cases. These results strongly endorse our use of the gravity equation to account for the role of distance or trade frictions in limiting trade volumes and net factor contents.

I. *Robustness Checks*

There are a variety of robustness checks that we would like to make. The first notes that specifications T6 and T7 have included the ROW point even though both force the ROW production model to fit perfectly. We have already provided reasons for believing that adjustment of the ROW technology is appropriate. Nonetheless, it would be troubling if the steady improvement in the model owed solely to inclusion of the ROW points once this adjustment is made. Our check on this is to return to models T4 through T7, excluding ROW in each case. The results are presented in Table 6. Exclusion of ROW does tend to reduce the slope coefficients in each case. And the improvement of T6′ over T5′ seems somewhat less substantial than that of T6 over T5. Nonetheless, the key observation is that the results are broadly consistent across the two sets of tests. Most importantly, the slope coefficient and the trade variance ratio rise consistently across both sets of tests, beginning and culminating at very similar levels. Even if we exclude ROW, the model correctly predicts the direction of net factor trade 90 percent of the time and the measured factor trade is over three-fourths the level predicted. Thus the results are highly robust to exclusion of ROW.

A second robustness check is to note that by sheer size, not only the ROW, but also the US, frequently provides influential data points. The US is a major exporter of capital and importer of labor while the reverse is true for the ROW. While these countries are extremely important to include in the analysis because they contribute so much variance, it would be troubling if our results were only a result of their inclusion. In Table 7 we drop the US and ROW and repeat our experiments. The slope coefficient in T7 rises to 0.64 and is precisely measured with an R^2 of 0.76. The overall pattern is very similar to the tests including the US and ROW. Specifications T1 through T3 show little improvement in the HOV predictions. The movement to T4 provides a very substantial improvement, those to T5 and T6 somewhat smaller improvements, and finally a substantial improvement in the move to T7. Hence the amended HOV model works quite well even when we drop two points that contribute a great deal to the variance.

A third robustness check is to consider the various ways that previous papers on factor service trade have weighted the data in order to account for heteroskedasticity. Up to now we have been focusing on untransformed data because all graphs and regressions have a clear interpretation in terms of actual factor service flows when these units are used. However, it is reasonable to ask whether our results are fragile when we shift weighting schemes.

The first weighting scheme that we try is one suggested by Trefler (1995). In that paper, Trefler deflates the data by the square root of a country's absorption share multiplied by the standard deviation of the predicted factor service flows (expressed in natural units).[27] This weighting scheme reduces the importance of large countries and factors with substantial variation in country abundance. In Table 8 we repeat our trade results obtained above and also present our results when recast in Trefler units. The switch to Trefler units matters little. Now the coefficient on predicted factor trade actually rises from 0.82 to 0.88. Our variance ratio test statistic falls a little but overall the same basic picture emerges. Clearly our results are robust to this specification.

Xavier Gabaix (1997) has suggested a second weighting scheme for evaluating factor content studies. If one deflates both sides of the HOV trade equation by the country's share of absorption, one eliminates all size-based variation from the data. This adjustment is tantamount to projecting each country's endowment point on to the same iso-income line. The results also appear in Table 8. Once again we see a steady rise in the slope coefficient as we move from T3 to T7. The final specification has a slope coefficient of 0.83, almost identical to our primary specification.

We conclude that our results are robust to a wide variety of weighting schemes. It appears that relaxing neutral technical differences, FPE, and allowing for non-traded goods results

---

[27] Instead of the standard deviation of predicted factor service flows in natural units, Trefler actually uses F_{fc} - (V_{fc} - $s_c V_{fW}$), where F_{fc} is his measured factor trade. However, since F_{fc} in his data is essentially zero, this weighting scheme is essentially the same as the one we implement.

in dramatic improvements in the HOV model regardless of the units chosen. Furthermore, accounting for the influence of trade costs on bilateral trade volumes results in further strong improvements.

An additional remaining question regarding our results is why in T7 we obtain a coefficient of only 0.82 when theory says it should be unity. There are three basic reasons. The first is attenuation bias due to measurement error. By conducting the reverse regression of predicted factor trade on measured, we can obtain maximum likelihood bounds for the effects of measurement error. Under specification T7, the high R^2 leaves little room for measurement error to matter, with an ML upper bound for the coefficient of 0.84. Under specification T7′, measurement error places the ML upper bound on the coefficient at 0.89.

The second reason is that our adjustments apply only to the impact that country capital to labor ratios have on average technology matrices, not export technology matrices. Theory suggests that this will still under-measure factor service trade because exported goods within an industry use more extreme factor proportions than goods which in equilibrium are non-traded in the same industry. While our analysis has adjusted for the fact that average input coefficients shift with country factor proportions, we have not adjusted for differences in factor intensity between export and average sectors. This will tend to result in apparent missing trade.

Finally, and most obviously, trade barriers, demand irregularities, and non-neutral technological differences really do exist. Hence, it would be astonishing if we could ignore all of these and describe global factor trade flows perfectly. The real surprise is just how well we do.

VI. Conclusion

The empirical validity of the factor proportions theory has been a focus of research for nearly one-half century. In the process, researchers have accumulated a great deal of experience that has informed our work. Leontief's (1953) seminal work provided the first true factor content study. The work of Maskus (1985) and Bowen, Leamer and Sveikauskas (1987) is extremely important not only for the methodological contributions, but also for the extraordinary energy they brought to their studies. The same could be said of the work of Trefler (1993, 1995, 1997), which (among other contributions) provides extremely lucid characterizations of anomalies in the data. These important contributions notwithstanding, this half-century of empirical research failed to produce a set of simple departures that allow the theory to match the salient features of the international data.

Our study starts from a simple premise. Since the principal hypotheses of the nature of HOV's failures in prior work concern technological differences and absorption patterns, it is crucial to address these directly on the relevant technological and absorption data. We develop a small set of hypotheses, some traditional, some novel, of why prior tests of HOV fail. We then estimate the crucial parameters directly from the relevant data and impose these restrictions on our empirical implementation of the HOV theory.

Our results provide striking support for the HOV theory, suitably amended. Countries export their abundant factors and they do so in approximately the right magnitude. The results are extraordinarily consistent across specifications and are robust to changes in the sample.

Perhaps the most exciting feature of our results is the simple and unified picture they draw of the global economy. No doubt much is left out of our account. Yet it is startling that such a plausible and simple set of departures from the conventional model allows us to so accurately match the international data.

References

Anderson, James, "A Theoretical Foundation for the Gravity Equation," *American Economic Review*; 69(1), March 1979, pages 106-16.

Antweiler, Werner; and Trefler, Daniel (1998) "Increasing Returns and All That: A View From Trade," mimeo, U. of Toronto.

Bergstrand, Jeffrey H. (1990) "The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis and the Determinants of Bilateral Intra-industry Trade," *Economic Journal*, 100(403), pp. 1216-29.

Bernstein, Jeffrey and David E. Weinstein (1998) "Do Endowments Predict the Location of Production? Evidence from National and International Data," University of Michigan, mimeo.

Bhagwati, J.N. "The Pure Theory of International Trade: A Survey," *Economic Journal*, 74 (1964) 1-84.

Bowen, Harry P., Edward Leamer, and Leo Sveikauskas (1987) "Multicountry, Multifactor Tests of the Factor Abundance Theory," *American Economic Review*, v. 77.

Davis, Donald R. (1997) "Critical Evidence on Comparative Advantage? North-North Trade in a Multilateral World." *Journal of Political Economy*, October.

Davis, Donald R. (1998) "Does European Unemployment Prop up American Wages? national Labor Markets and Global Trade," American Economic Review, June.

Davis, Donald R., David E. Weinstein, Scott C. Bradford and Kazushige Shimpo (1997) "Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works," *American Economic Review*, v. 87.

Deardorff, Alan V. (1994) "The Possibility of Factor Price Equalization, Revisited," Journal of International Economics; 36(1-2), February, pages 167-75.

Deardorff, Alan V. (1998) "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" in Jeffrey Frankel, ed., *Regionalization of the World Economy*, Chicago: U. of Chicago and NBER.

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

Dixit, Avinash K., and Norman, Victor (1980) *Theory of International Trade*, Cambridge: Cambridge University Press.

Dollar, David, Edward Wolff, and Baumol, William "The Factor-Price Equalization Model and Industry Labor Productivity: An Empirical Test across Countries,"(1988) in Feenstra, Robert ed., "Empirical methods for international trade," Cambridge, Mass. and London: MIT Press, 1988.

Dollar, David and Edward N. Wolff (1993) *Competitiveness, Convergence, and International Specialization*, Cambridge: The MIT Press.

Dornbusch, Rudiger, Stanley Fischer and Paul A. Samuelson (1980) "Heckscher-Ohlin Trade Theory with a Continuum of Goods," *Quarterly Journal of Economics*, 95, 203-224.

Frankel, Jeffrey A.; Stein, Ernesto; Wei, Shang-Jin, "Regional Trading Arrangements: Natural or Super-Natural?" National Bureau of Economic Research Working Paper: 5431, January 1996.

Gabaix, Xavier (1997) "The Factor Content of Trade: A Rejection of the Heckscher-Ohlin-Leontief Hypothesis," Harvard University, mimeo.

Hakura, Dalia, (1997) "A Test of the General Validity of the Heckscher-Ohlin Theorem for Trade in the European Community," mimeo IMF June.

Harrigan, James (1993), "OECD Imports and Trade Barriers in 1983," *Journal of International Economics*; 35(1-2), August.

Harrigan, James (1995) "Factor Endowments and the International Location of Production: Econometric Evidence from the OECD, 1970-1985," *Journal of International Economics*.

Harrigan, James (1997) "Technology, Factor Supplies and International Specialization: Testing the Neoclassical Model," forthcoming *American Economic Review*.

Helpman, Elhanan (1981) "International Trade in the Presence of Product Differentiation, Economies of Scale and Monopolistic Competition: A Chamberlin-Heckscher-Ohlin Approach," *Journal of International Economics*, 11:3.

Helpman, Elhanan (1984) "The Factor Content of Foreign Trade," *The Economic Journal*, March 84-94.

Helpman, Elhanan (1998) "The Structure of Foreign Trade," mimeo, Harvard University.

Helpman, Elhanan and Krugman, Paul (1985) *Market Structure and Foreign Trade*, Cambridge: MIT Press.

Jorgenson, Dale and Masahiro Kuroda (1990) "Productivity and International Competitiveness in Japan and the United States, 1960-1985," in Hulten, Charles R. ed. *Productivity Growth in Japan and the United States*, Chicago: University of Chicago Press.

Krugman, Paul R., "But For, As If, and So What? Thought Experiments on Trade and Factor Prices," mimeo, MIT, November 1996.

Lawrence, Robert Z., and Slaughter, Matthew J. "International Trade and American Wages in the 1980s: Giant Sucking Sound or Small Hiccup?" *Brookings Papers on Economic Activity, Microeconomics*, 1993, pp. 161-210.

Leamer, Edward (1980) "The Leontief Paradox Reconsidered," *Journal of Political Economy*, 88, 495-503.

Leamer, Edward E., "In Search of Stolper-Samuelson Effects Between Trade and US Wages" mimeo 1995, Yale and UCLA.

Leamer, Edward and James Levinsohn (1995) "International Trade Theory: The Evidence" in Gene Grossman and Kenneth Rogoff, eds., *Handbook of International Economics*, vol. 3, New York: Elsevier.

Leontief, Wassily (1953) "Domestic Production and Foreign Trade: The American Capital Position Re-Examined," *Proceedings of the American Philosophical Society*, 97, 332-349.

Linder, Staffan Burenstam (1961) *An Essay on Trade and Transformation*, NY: John Wiley and Sons.

McCallum, John (1995) "National Borders Matter: Canada-US Regional Trade Patterns," *American Economic Review*, 85:3.

Markusen, James R. (1986) "Explaining the Volume of Trade: An Eclectic Approach," *American Economic Review*, 76, 1002-1011.

Markusen, James R. and Wigle, Randall M. (1990) "Explaining the Volume of North-South Trade," *Economic Journal*, 100, 1206-1215.

Maskus, Keith E. (1985) "A Test of the Heckscher-Ohlin-Vanek Theorem: The Leontief Commonplace," *Journal of International Economics*, November, 9, pp. 201-212.

O'Rourke, Kevin; Williamson, Jeffrey G. (1994) “Late Nineteenth-Century Anglo-American Factor-Price Convergence: Were Heckscher and Ohlin Right?” *Journal of Economic History*; 54(4), December, pages 892-916.

Samuelson, Paul A. (1949) “International Factor Price Equalization Once Again,” *Economic Journal*, Vol.LIX, No. 234, pp. 181-197.

Samuelson, Paul A. (1953) “Prices of Goods and Factors in General Equilibrium,” *Review of Economic Studies*, 21, No. 54.

Samuelson, Paul A. (1954) “The Transfer Problem and Transport Costs: Analysis of Effects of Trade Impediments,” *Review of Economics and Statistics*, 46, No. 2.

Staiger, Robert; Alan V. Deardorff; and Robert Stern (1987) “An Evaluation of Factor Endowments and Protection as Determinants of Japanese and American Foreign Trade,” *Canadian Journal of Economics*, No. 3, August.

Stern, Robert M. and Maskus, Keith E. (1981) “Determinants of the Commodity Structure of US Foreign Trade, 1958-1976,” *Journal of International Economics*, 11, pp. 207-224.

Trefler, Daniel (1993) “International Factor Price Differences: Leontief was Right!”, *Journal of Political Economy*, 101.

Trefler, Daniel (1995) “The Case of the Missing Trade and other Mysteries,” *American Economic Review.*

Trefler, Daniel (1998) “The Structure of Factor Content Predictions,” mimeo, U. of Toronto.

Vanek, Jaroslav, "The Factor Proportions Theory: The N-Factor Case," Kyklos, October 1968, 21, 749-755.

Wei, Shang-Jin (1996) “Intra-National versus International Trade: How Stubborn are Nations in Global Integration?” NBER Working Paper #5531.

Wood, Adrian (1994) *North-South Trade, Employment and Inequality: Changing Fortunes in a Skill-Driven World,* Oxford: Oxford University Press.

Xu, Yingfeng (1993) “A General Model of Comparative Advantage with Two Factors and a Continuum of Goods,” International Economic Review, 34, May, pp. 365-380.

Table 1

Tests of Technological Variation

Model	Measurement Error P2	Hicks Neutral Technical Differences (HNTD) P3	Continuum Model with HNTD and FPE P4	Pooled Helpman No-FPE Model with HNTD P5	Unpooled Helpman No-FPE Model with HNTD P5'	Implied λ^c
θ^{Aus}	–	0.531	0.531	0.530	0.528	1.7
		(0.035)	(0.035)	(0.035)	(0.035)	
θ^{Can}	–	0.381	0.381	0.380	0.381	1.5
		(0.035)	(0.035)	(0.035)	(0.034)	
θ^{Den}	–	0.508	0.504	0.508	0.508	1.7
		(0.036)	(0.036)	(0.036)	(0.034)	
θ^{Fra}	–	0.494	0.493	0.494	0.492	1.6
		(0.034)	(0.034)	(0.034)	(0.035)	
θ^{Ger}	–	0.112	0.111	0.112	0.111	1.1
		(0.034)	(0.034)	(0.034)	(0.035)	
θ^{Italy}	–	0.709	0.707	0.709	0.704	2.0
		(0.034)	(0.034)	(0.034)	(0.036)	
θ^{Japan}	–	0.431	0.430	0.431	0.430	1.5
		(0.033)	(0.033)	(0.033)	(0.034)	
θ^{Neth}	–	.057	.057	.056	.058	1.1
		(0.035)	(0.035)	(0.035)	(0.035)	
θ^{UK}	–	0.520	0.516	0.520	0.542	1.7
		(0.034)	(0.034)	(0.034)	(0.040)	
θ^{US}	–	0	0	0	0	1.0
γ^{KT}	–	–	0.408	0.364	–	
			(0.046)	(0.061)		
γ^{KN}	–	–	–	0.493	–	
				(0.071)		
γ^{LT}	–	–	-0.408	-0.449	–	
			(0.046)	(0.060)		
No. of Param.	68	77	78	80	144	
- Log L	-1741.5	-934.4	-855.7	-802.8	-740.7	
Schwartz Crit.	-1963.3	-1185.5	-1110.1	-1063.7	-1210.3	

Standard errors are reported in parentheses. $\gamma^{LN} = - \gamma^{LT} - \gamma^{KT} - \gamma^{KN}$. There is very little variation in the θ's as we move across specifications because of the constraint that capital to labor ratios cannot affect productivity.

Table 2
Key Specifications

	KEY ASSUMPTION	**PRODUCTION SPECIFICATIONS**		**TRADE SPECIFICATIONS**
P1	Conventional HOV with US Technology	$B^{US} Y^c = V^c$	**T1**	$B^{US} T^c = B^{US} (Y^c - D^c) = V^c - s^c V^W$
P2	Average Technology Matrix	$\hat{B}^{\mu} Y^c = V^c$	**T2**	$\hat{B}^{\mu} T^c = V^c - s^c V^W$
P3	Hicks-Neutral Efficiency Adjustment	$\hat{B}^{\lambda} Y^c = V^{cE}$	**T3**	$\hat{B}^{\lambda} T^c = V^{cE} - s^c V^{WE}$
P4	Continuum Model: Different Input Ratios in Traded Goods and H-N Efficiency	$\hat{B}^{cDFS} Y^c = V^c$	**T4**	$\hat{B}^{cDFS} Y^c - \left[\hat{B}^{cDFS} D^{cc} + \sum_{c' \neq c} \hat{B}^{c'DFS} M^{cc'} \right] = V^c - s^c V^W$
P5	Helpman No-FPE Model, Diff Input Ratios in All, H-NE	$\hat{B}^{cH} Y^c = V^c$	**T5**	$\hat{B}^{cH} Y^{cT} - \left[\hat{B}^{cH} D^{ccT} + \sum_{c' \neq c} \hat{B}^{c'H} M^{cc'} \right] = \left[V^c - s^c V^W \right] - \left[V^{cN} - s^c V^{WN} \right]$
	Forces ROW Prod. Model to Work		**T6**	As above
	Adds Gravity-Based Demand Determ.		**T7**	$\hat{B}^{cH} Y^c - \left[\hat{B}^{cH} D^{cc} + \sum_{c' \neq c} \hat{B}^{c'H} M^{cc'} \right] = V^c - \left[\hat{B}^{cH} \hat{D}^{cc} + \sum_{c' \neq c} \hat{B}^{c'H} \hat{M}^{cc'} \right]$

Hats (^) indicate fitted values from estimation of technology and absorption.

Table 3

Prior Production and Trade Tests

Production Slope and Median Error Tests

Neither Bowen, et al. (1987) nor Trefler (1995) conduct production tests. Davis, Weinstein, et al. (1997) do report international production tests based on three factors and the Japanese technology matrix. The results are disappointing, featuring very large prediction errors (frequently over 100 percent).

Trade Sign Tests

BLS report trade sign tests for 12 factors and 27 countries. They conclude that measured factor abundance provides no more insight than a coin flip in identifying which factor services a country will export. Trefler (1995) reports two types of sign tests for 9 factors and 33 countries. The first is, as above, a simple proportion of sign matches. He also reports a weighted-sign test, where the weights are given by an observation's absolute share in the measured factor content of trade. The simple HOV model is correct 50 percent of the time under the simple sign test or 71 percent under the weighted-sign test, which he terms "uncomfortably close" to a coin flip. Trefler's preferred specification, with a home bias in demand and Hicks-neutral technical differences, correctly predicts the sign 93 percent of the time in the weighted-sign test (unweighted not reported).

Trade Variance Ratio Tests

Trefler (1995) is the only prior paper that reports the variance ratio test (theoretical prediction is unity). Under the simple HOV model, the ratio is 0.032. Under his preferred specification, the ratio is 2.226. While this remains far from unity, Trefler concludes that this is an improvement over the simple HOV model.

Trade Slope Test

Neither BLS nor Trefler report a slope test. Gabaix (1997), using the same data as Trefler (1993, 1995), reports several variants of a slope test scaled by country size (theoretical prediction is unity). The model performs reasonably well for the resource-based factors cropland and pasture. However, for the factors that dominate incomes — production labor, aggregate labor, and capital — the results are poor. In almost all specifications and restrictions of the sample, the estimated coefficients are insignificant, or where significant, *negative*.

Table 4

Production and Trade Tests
All Factors

Production Tests: Dependent Variable MFCP					
	P1	P2	P3	P4	P5
Predicted	0.24	0.33	0.89	0.89	0.97
se	0.09	0.11	0.06	0.05	0.01
R^2	0.27	0.29	0.92	0.94	1.00
Median Error	0.34	0.21	0.07	0.05	0.03
obs.	20	22	22	22	22

Trade Tests: Dependent Variable MFCP							
	T1	T2	T3	T4	T5	T6	T7
Predicted	-0.002	-0.006	-0.05	0.17	0.43	0.59	0.82
se	0.005	0.003	0.02	0.02	0.02	0.04	0.03
R^2	0.01	0.14	0.31	0.77	0.96	0.92	0.98
Sign Test	0.32	0.45	0.50	0.86	0.86	0.82	0.91
Var. Ratio	0.0005	0.0003	0.008	0.07	0.19	0.38	0.69
obs.	22	22	22	22	22	22	22

The theoretical coefficient on “predicted” is unity. The theoretical value of the sign test is unity (100 percent correct matches). The Variance Ratio is var(MFCT)/var(PFCT) and has a theoretical value of unity.

Table 5
Production and Trade Tests
Capital

Production Tests: Dep. Var. MFCP						Trade Tests: Dep. Var. MFCT						
	P1	**P2**	**P3**	**P4**	**P5**	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**
Pred.	0.77	0.99	0.87	0.90	0.99	0.06	0.02	-0.03	0.25	0.57	0.65	0.87
se	0.11	0.15	0.08	0.05	0.01	0.01	0.01	0.02	0.05	0.07	0.15	0.07
R^2	0.84	0.84	0.92	0.94	1.00	0.80	0.28	0.14	0.77	0.88	0.68	0.95
Median Error	0.34	0.20	0.07	0.06	0.02							
Sign Test						0.45	0.64	0.73	0.91	0.77	0.80	1.00

Production and Trade Tests
Labor

Production Tests: Dep. Var. MFCP						Trade Tests: Dep. Var. MFCT						
	P1	**P2**	**P3**	**P4**	**P5**	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**
Pred.	0.07	0.12	0.92	0.87	0.94	-0.008	-0.008	-0.07	0.14	0.42	0.59	0.81
se	0.06	0.09	0.09	0.08	0.02	0.002	0.003	0.03	0.01	0.03	0.05	0.03
R^2	0.15	0.16	0.92	0.93	0.997	0.627	0.529	0.43	0.94	0.96	0.94	0.99
Median Error	0.42	0.22	0.08	0.05	0.05							
Sign Test						0.18	0.27	0.27	0.82	1.00	0.80	0.81

The theoretical coefficient on “predicted” is unity. The theoretical value of the sign test is unity (100 percent correct matches).

Table 6

Trade Tests
All Factors, Excluding ROW

Trade Tests: Dependent Variable MFCP (excluding ROW)				
	T4′	T5′	T6′	T7′
Pred.	0.19	0.35	0.42	0.77
se	0.02	0.03	0.04	0.07
R^2	0.80	0.87	0.86	0.86
Sign Test	0.90	0.90	0.80	0.90
Var. Ratio	0.04	0.14	0.20	0.68
obs.	20	20	20	20

Table 7

Trade Tests
Excluding ROW and the US
All Factors

Trade Tests: Dep. Var. MFCT							
	T1	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**
Pred.	-0.05	-0.04	0.05	0.30	0.35	0.42	0.64
se	0.01	0.01	0.07	0.07	0.11	0.14	0.09
R^2	0.47	0.33	0.03	0.51	0.41	0.36	0.76

The theoretical coefficient on “predicted” is unity. The theoretical value of the sign test is unity (100 percent correct matches). The Variance Ratio is var(MFCT)/var(PFCT) and has a theoretical value of unity.

Table 8

Trade Tests With Various Heteroskedasticity Corrections
All Factors

	Trade Tests: Dependent Variable MFCT						
	T1	**T2**	**T3**	**T4**	**T5**	**T6**	**T7**
Standard Units							
Predicted	-0.002	-0.006	-0.049	0.17	0.43	0.59	0.82
se	0.005	0.003	0.016	0.02	0.02	0.04	0.03
R^2	0.01	0.14	0.31	0.77	0.96	0.92	0.98
Trefler Units							
Predicted	0.02	-0.01	0.005	0.18	0.44	0.55	0.88
se	0.01	0.002	0.022	0.03	0.06	0.10	0.04
R^2	0.12	0.60	0.003	0.72	0.74	0.63	0.95
Gabaix Units							
Predicted	-0.01	-0.013	-0.006	0.21	0.44	0.59	0.83
se	0.01	0.009	0.041	0.05	0.06	0.09	0.06
R^2	0.04	0.110	0.001	0.51	0.71	0.66	0.90

The theoretical coefficient on “predicted” is unity.

Figure 1

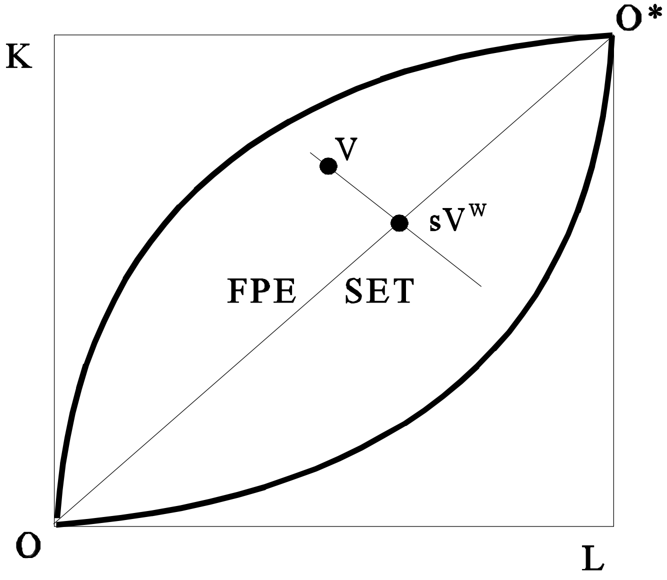


Figure 2

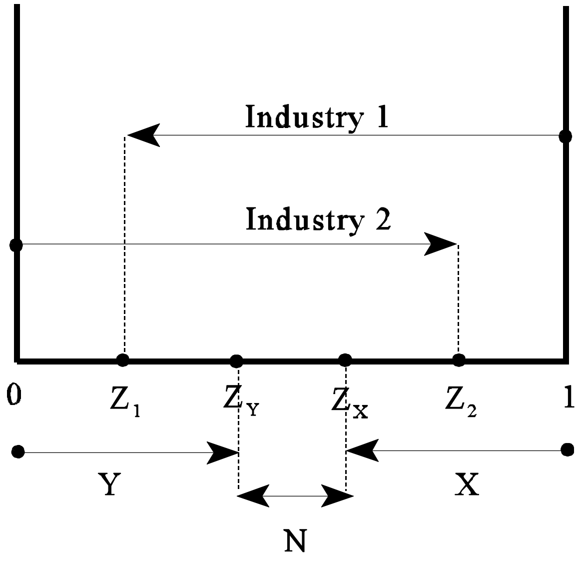


Figure 3
P1
Production with Common Technology (US)

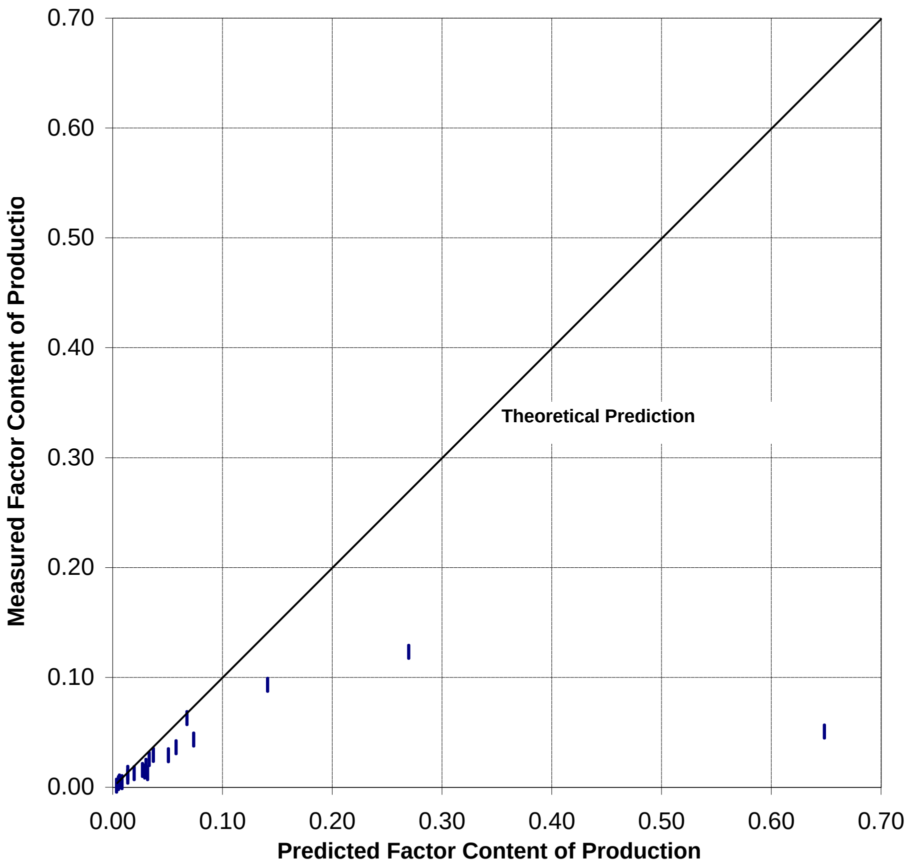


Figure 4
T1
Trade with Common Technology (US)

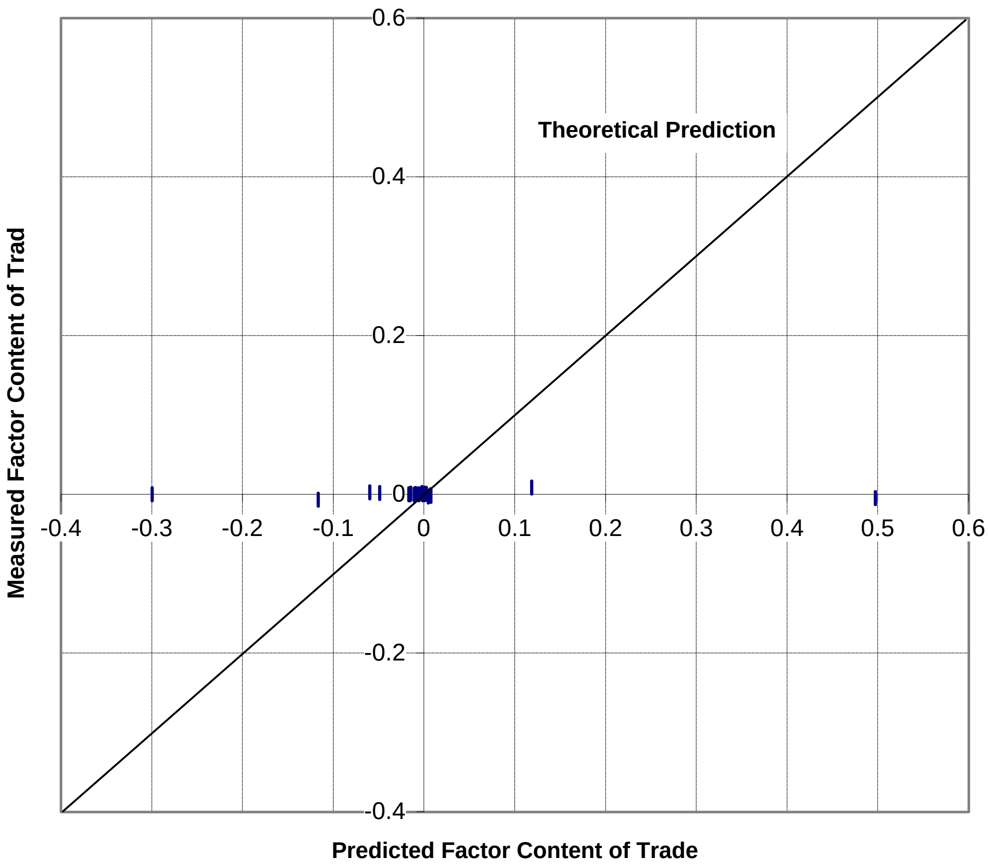


Figure 5
P2
Production with Common Technology (Average)

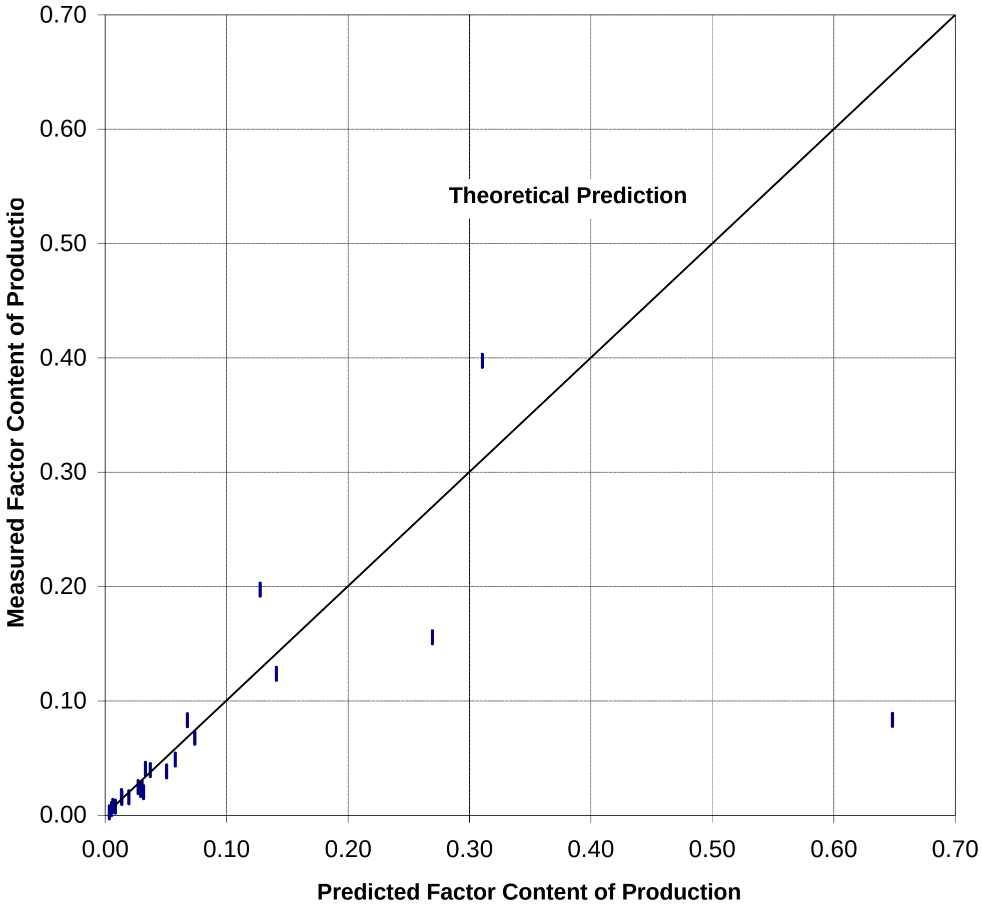


Figure 6
T2
Trade with Common Technology (Average)

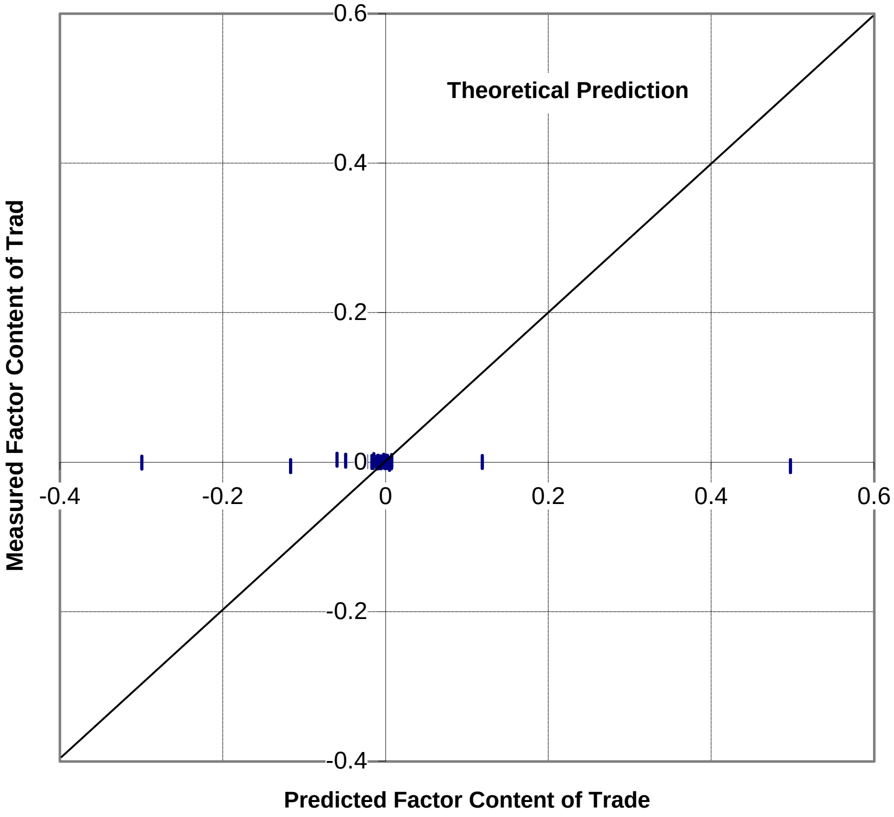


Figure 7
P3
Production with Hicks-Neutral Technical Differences

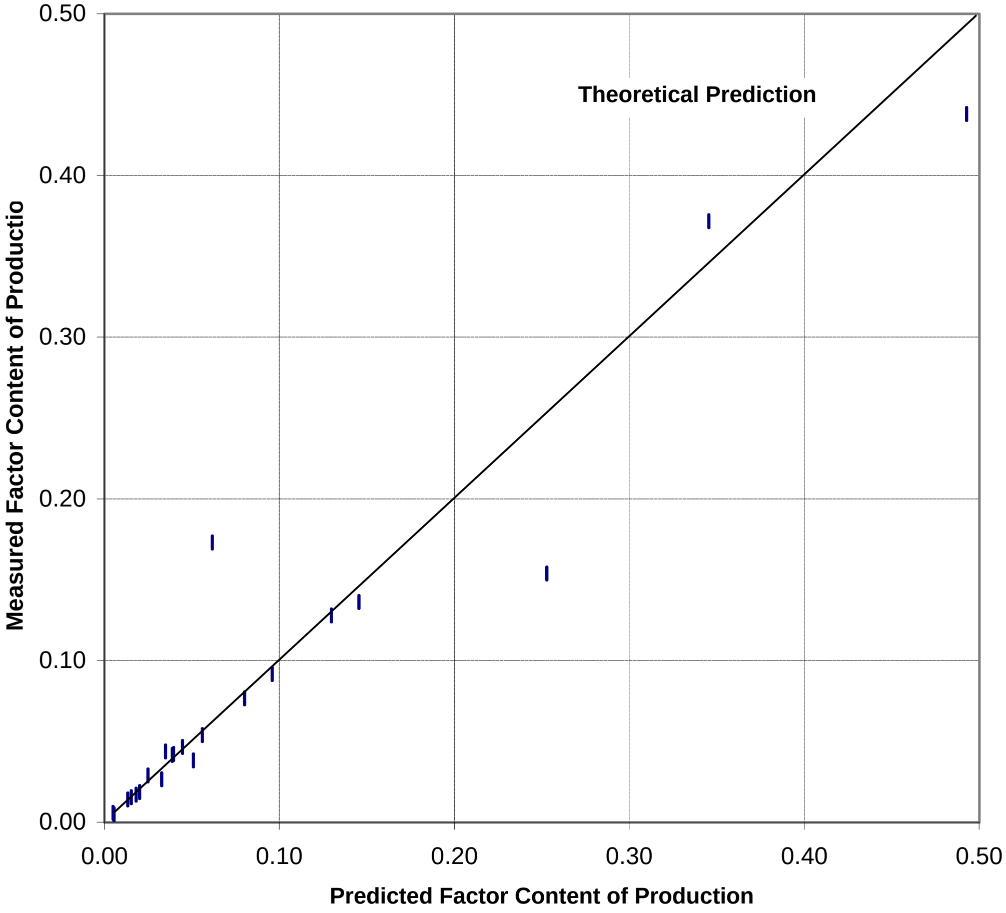


Figure 8
T3
Trade with Hicks-Neutral Technical Differences

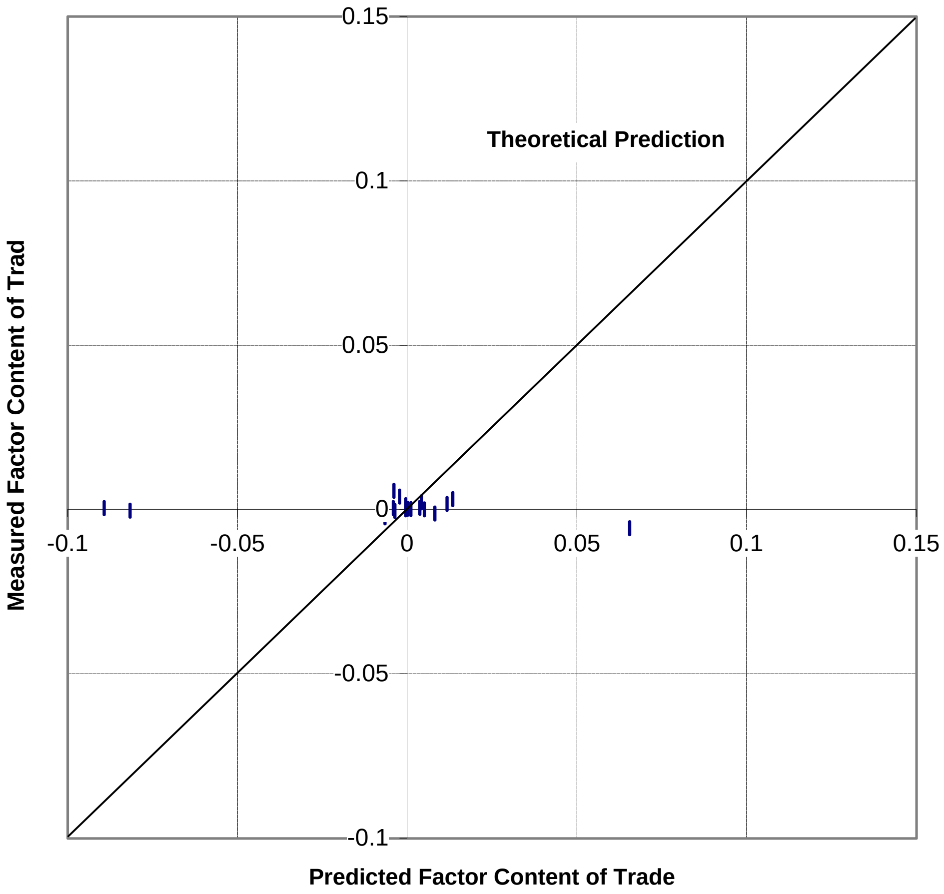


Figure 9
P4
Production with Continuum of Goods Model and FPE

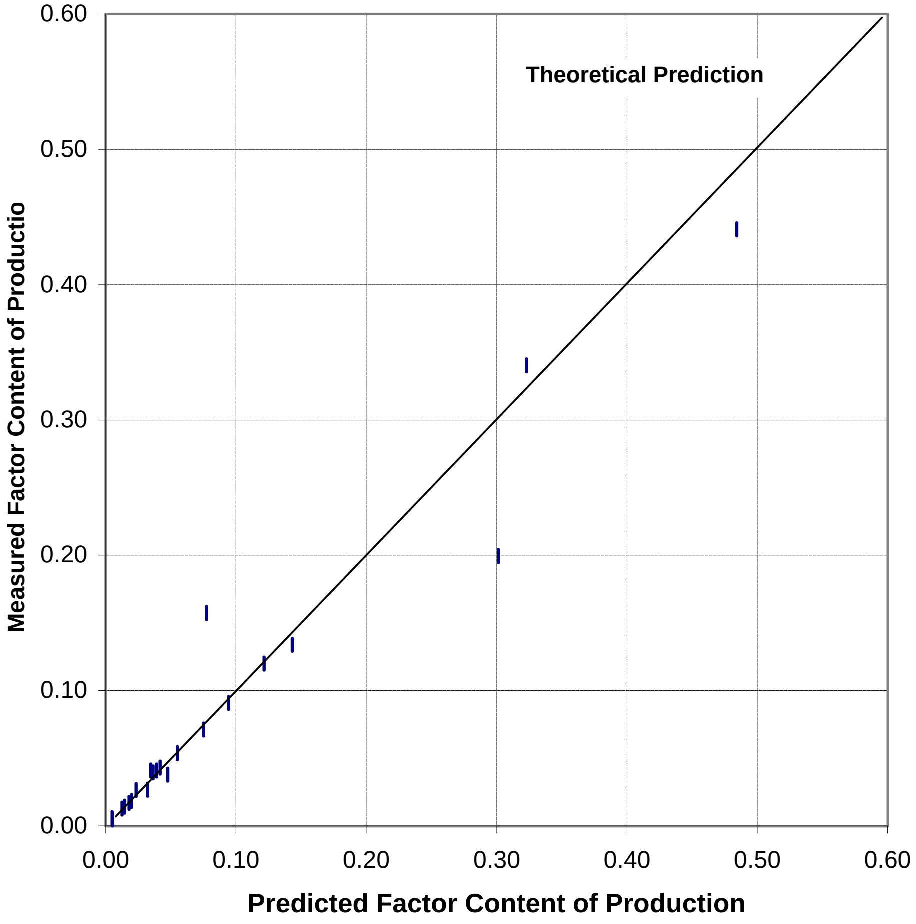


Figure 10
T4
Trade with Continuum of Goods Model and FPE

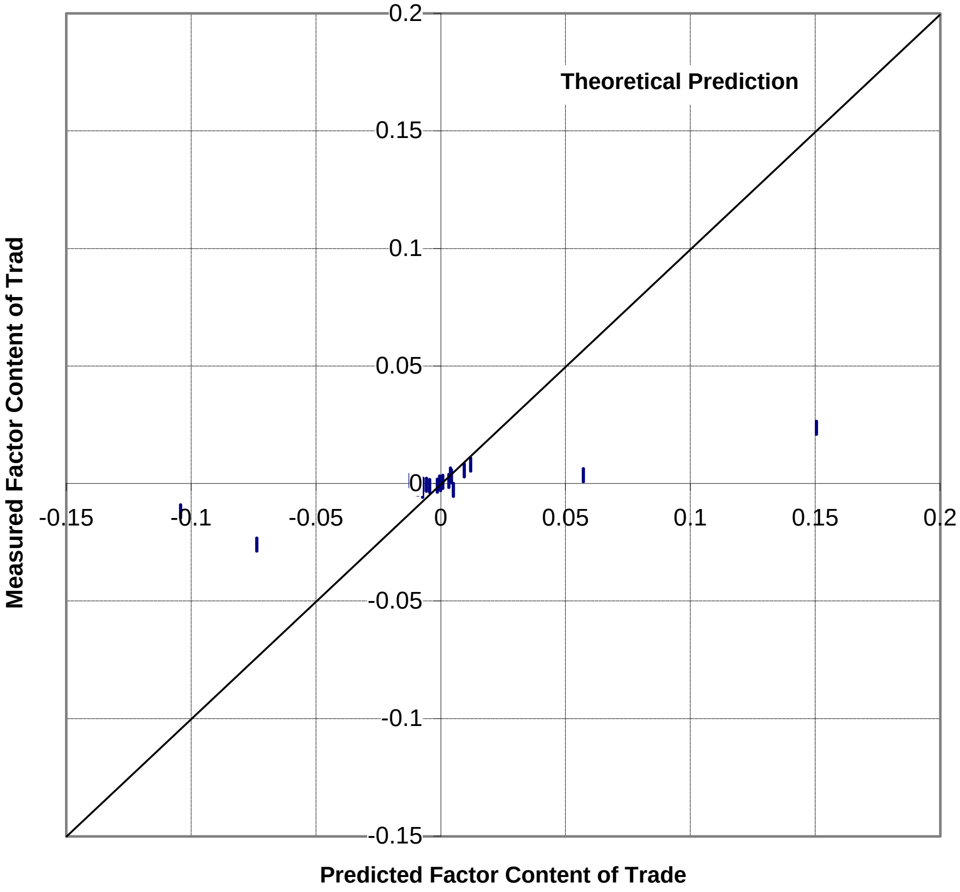


Figure 11
P5
Production without FPE

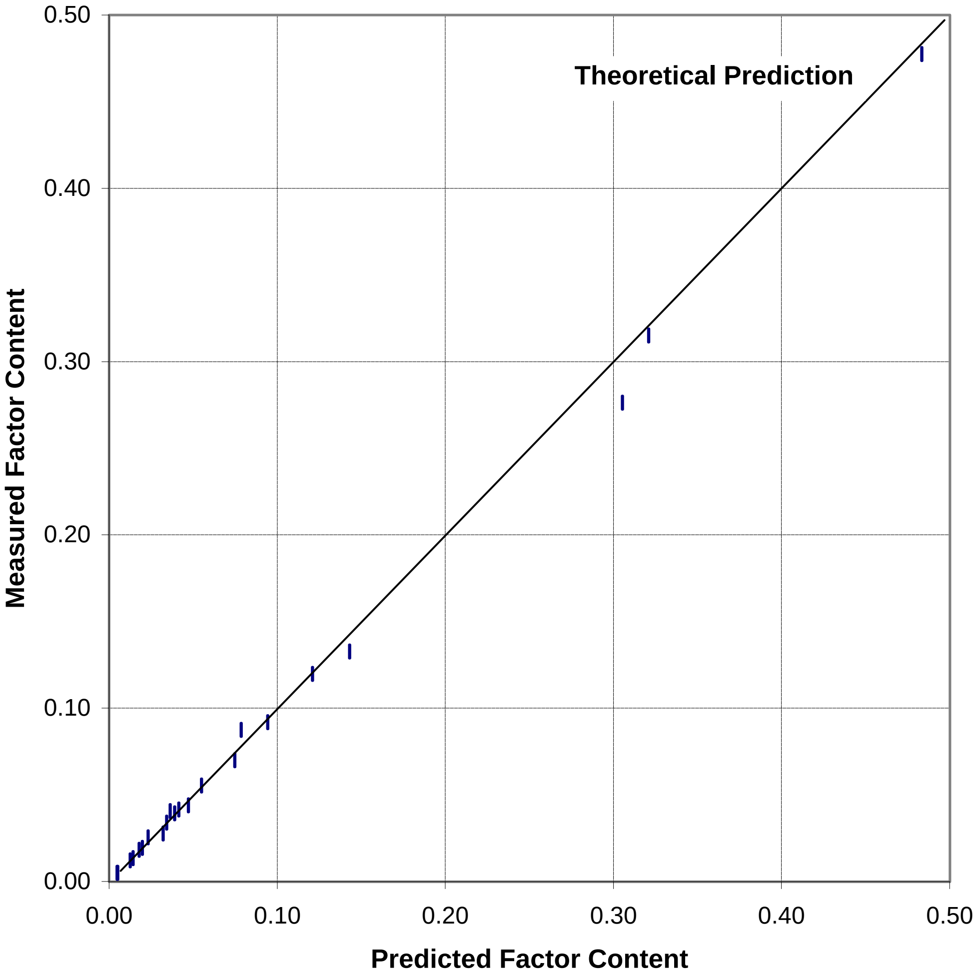


Figure 12
T5
Trade with No-FPE, Non-Traded Goods

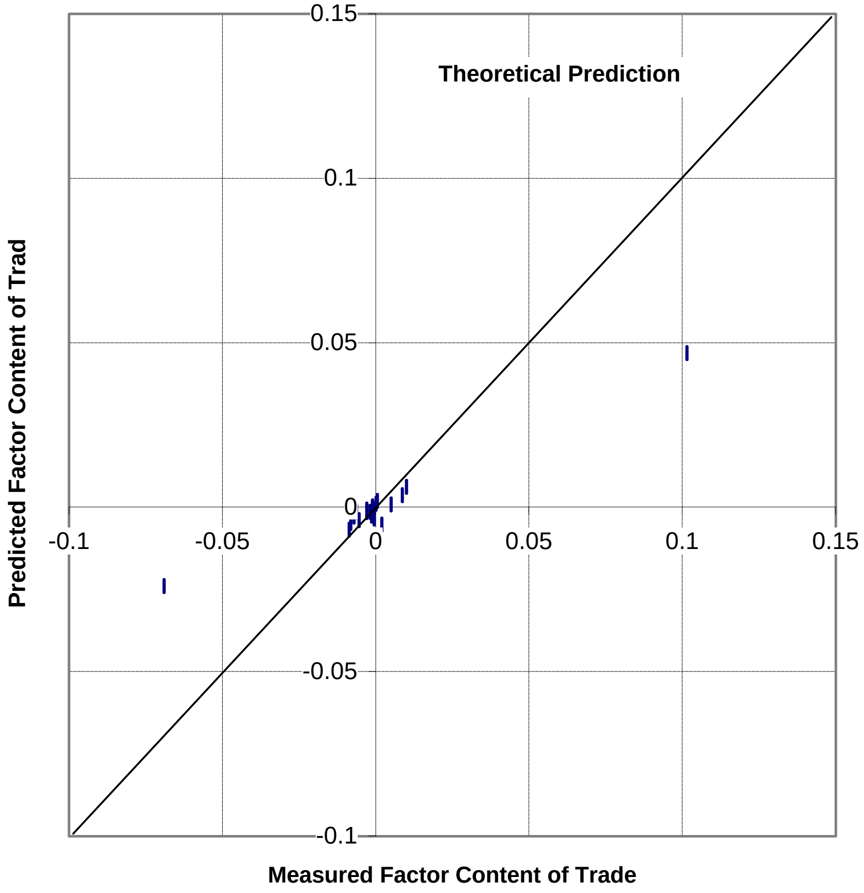


Figure 13

T6

Trade with No-FPE, Non-Traded Goods, and Adjusted ROW

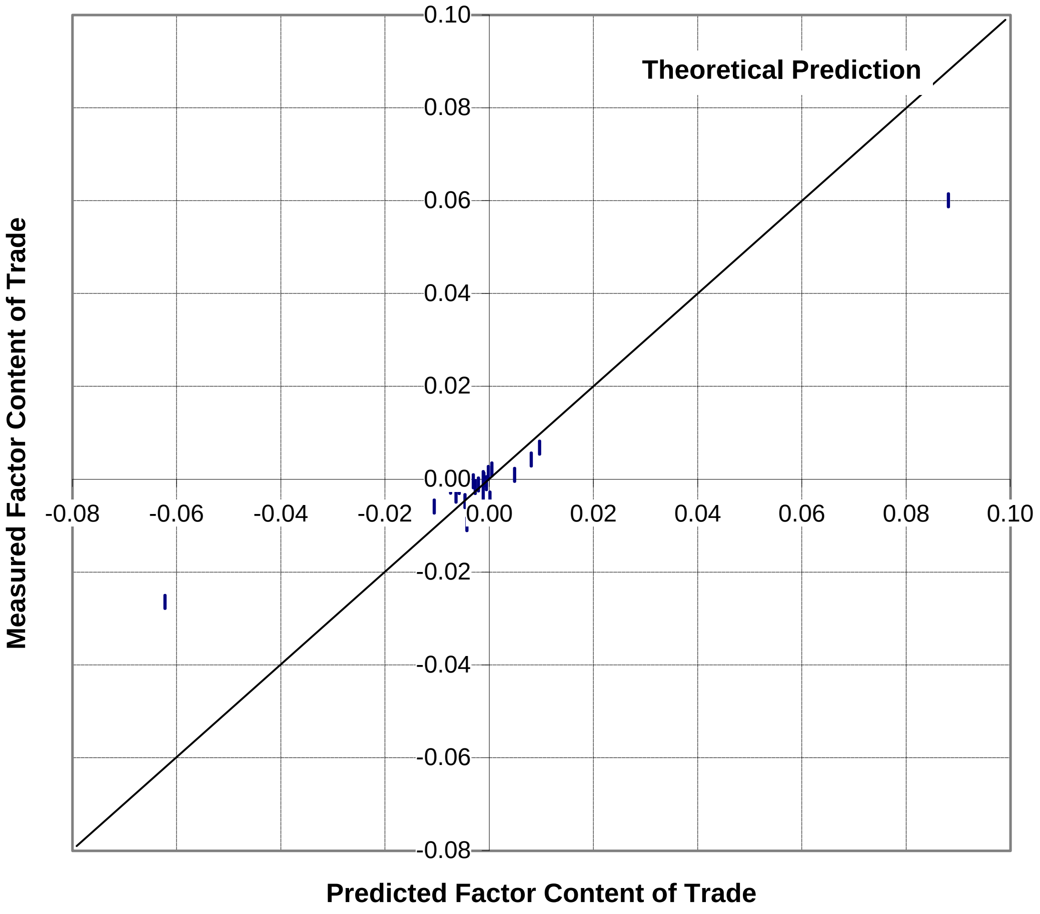
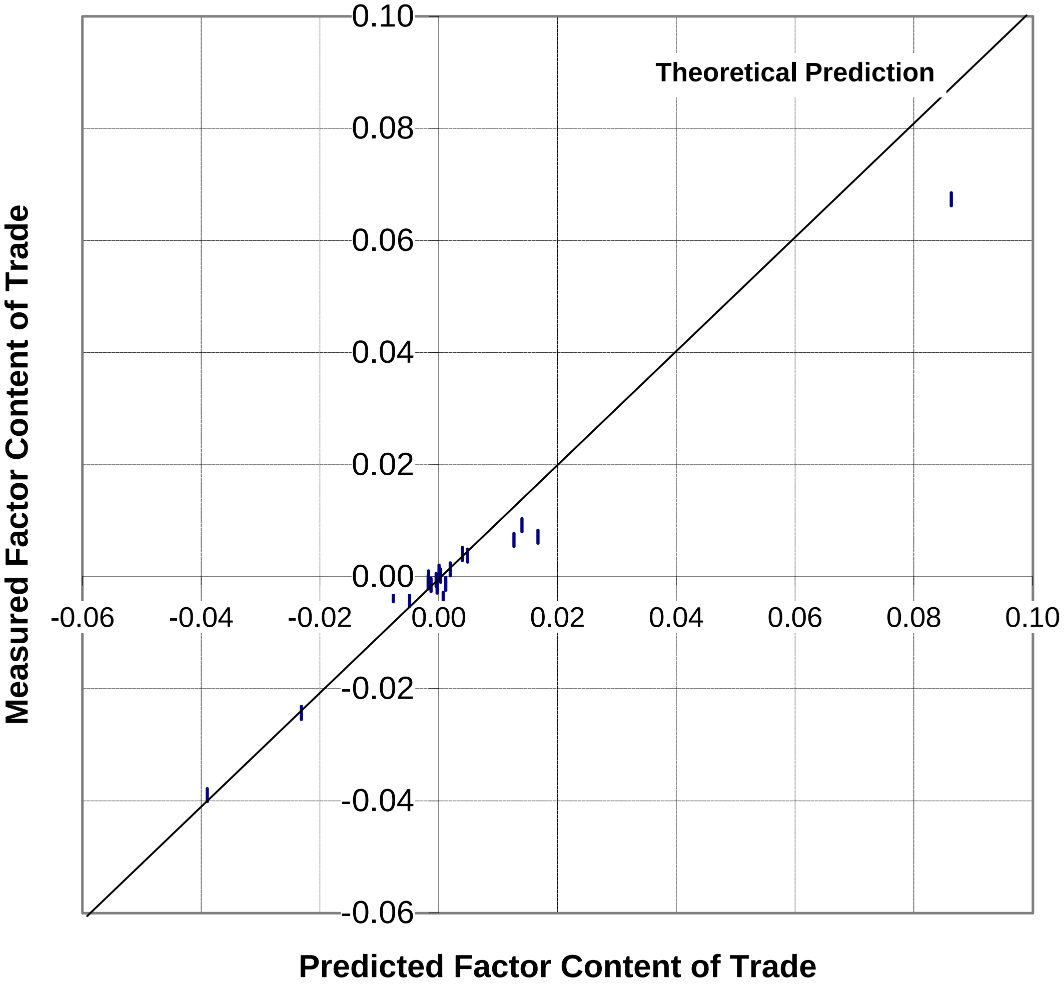


Figure 14

T7

Trade with No-FPE, Gravity Demand Specification, and Adjusted ROW



DATA APPENDIX

Data Sources:

For capital and labor:

Data for manufacturing sectors were taken from the 1997 OECD Structural Analysis (STAN) Industrial Database for years 1970-1995.

Data for other sectors were taken from the 1996 International Sectoral Database (ISDB) for years 1960-1995.

For production, demand and trade:

Data were taken from the 1995 OECD Input-Output Database.

Countries:

We used the 10 countries included in the OECD IO Database:

Australia, Canada, Denmark, France, Germany, Italy, Japan, Netherlands, UK, and US

Some countries did not have an IO table for 1985. We chose the closest year to 1985 for which an IO table existed. These countries and their related years are:

Australia (1986), Canada (1986), Germany (1986), Netherlands (1986), UK (1984)

Sectors:

Data for each of the 10 countries is organized into 34 sectors. All sectors were defined as in the IO tables except for sectors 29 and 30 (ISIC 7100 and 7200) which were aggregated due to the inability to disaggregate ISDB data for these two sectors. The individual sectors and their ISIC Revision 2 codes are given below:

IO Sector	ISIC Rev. 2 codes	Description
1	1	Agriculture, forestry, and fishery
2	2	Mining and quarrying
3	31	Food, beverages, and tobacco
4	32	Textiles, apparel, and leather
5	33	Wood products and furniture
6	34	Paper, paper products, and printing
7	351+352-3522	Industrial chemicals
8	3522	Drugs and medicines
9	353+354	Petroleum and coal products
10	355+356	Rubber and plastic products
11	36	Non-metallic mineral products
12	371	Iron and steel
13	372	Non-ferrous metals
14	381	Metal products
15	382-3825	Non-electrical machinery
16	3825	Office and computing machinery
17	383-3832	Electric apparatus, nec
18	3832	Radio, TV, and communication equipment
19	3841	Shipbuilding and repairing
20	3842+3844+3849	Other transport
21	3843	Motor vehicles
22	3845	Aircraft
23	385	Professional goods
24	39	Other manufacturing
25	4	Electricity, gas, and water
26	5	Construction
27	61+62	Wholesale and retail trade
28	63	Restaurants and hotels
29/30	71+72	Transport and storage, and Communication
31	81+82	Finance and insurance
32	83	Real estate and business services
33	9	Community, social, and personal services
34		Producers of government services
35		Other producers

Capital Stock Data:

Capital stock was calculated using the perpetual inventory method. For non-manufacturing sectors, data were taken from ISDB ITD, which contains information on gross fixed capital formation in 1990 PPP prices in US dollars. All values were then converted to 1985 prices. One compatibility problem that arises in these data is that sometimes the value added in a sector in ISDB is different from that the IO tables. To prevent variation in classification to produce variations in factor intensities we scaled up all investment series by the ratio of value added in the IO tables relative to value added in the same sector as reported in the ISDB.

Formally, for each non-manufacturing sector (j), GFCF was calculated as:

$$\mathrm{GFCF}_j = \mathrm{ITD}_j^{\mathrm{ISDB}} * \frac{P^{\mathrm{US,1985}}}{P^{\mathrm{US,1990}}} * \frac{\mathrm{VA}_j^{\mathrm{IO}}}{\mathrm{VA}_j^{\mathrm{ISDB}}}$$

For manufacturing sectors, the ISDB data was at a higher level of aggregation than we liked. Therefore, data were taken from the STAN database. The investment series we used was Gross Fixed Capital Formation (GFCF), in current prices and national currencies. To convert all data to 1985 PPP prices, the STAN data were multiplied by a capital stock price deflator, derived from the ISDB. Where ISDB sectors contained several STAN sectors, we used the same capital stock price information for each sector. Our price deflator consisted of ISDB ITD/IT, where ISDB IT is investment in current prices and national currencies. We then converted these numbers into 1985 dollars. Manufacturing data were also scaled by the ratio of ISDB to STAN GFCF in total manufacturing. This was done so that the size of manufacturing sectors relative to non-manufacturing sectors would be consistent if ISDB or STAN consistently under- or over-report the size of manufacturing sectors. Finally, all sectors were scaled by the sector ratio of IO to STAN or ISDB Value Added (VA). This was done so that sectors would be weighted more heavily if the sector was larger in the IO table than in STAN or ISDB. Ideally, we would have used ISDB data instead of STAN data for this last adjustment but we could not because the matching between the IO tables and the STAN data was much better for manufacturing.

Hence, for each manufacturing sector (I), for each country, and for each year, GFCF was calculated as:

$$\mathrm{GFCF}_i = \mathrm{GFCF}_i^{\mathrm{STAN}} * \frac{\mathrm{ITD}_i^{\mathrm{ISDB}}}{\mathrm{IT}_i^{\mathrm{ISDB}}} * \frac{P^{\mathrm{US,1985}}}{P^{\mathrm{US,1990}}} * \frac{\mathrm{GFCF}^{\mathrm{ISDB,TotalManuf.}}}{\mathrm{GFCF}^{\mathrm{STAN,TotalManuf.}}} * \frac{\mathrm{VA}_i^{\mathrm{IO}}}{\mathrm{VA}_i^{\mathrm{STAN}}}$$

Note: Japanese ISDB IT data were missing in manufacturing, so a slightly different method was used for each Japanese manufacturing sector (I). An overall capital goods price deflator (CGPD) for each year (from Economic Statistics Annual, Bank of Japan, 1994) was used to first convert all investment levels into 1990 yen prices. We then used the overall capital price deflator from ISDB (ITV/ITD) to convert these prices into 1990 US PPP dollars and then followed our standard procedure.

Japanese capital formation was therefore calculated as follows

$$\mathrm{GFCF}_{i,\mathrm{Japan}} = \mathrm{GFCF}_i^{\mathrm{STAN}} * \frac{1}{\mathrm{CGPD}} * \frac{\mathrm{ITD}^{\mathrm{ISDB}}}{\mathrm{ITV}^{\mathrm{ISDB}}} * \frac{P^{\mathrm{US,1985}}}{P^{\mathrm{US,1990}}} * \frac{\mathrm{GFCF}^{\mathrm{ISDB,TotalManuf.}}}{\mathrm{GFCF}^{\mathrm{STAN,TotalManuf.}}} * \frac{\mathrm{VA}_i^{\mathrm{IO}}}{\mathrm{VA}_i^{\mathrm{STAN}}}$$

After the gross fixed capital formation was calculated for each year and each sector, a permanent inventory method was used to determine capital stocks. Capital formation for each year from 1975 to 1985, inclusive, was used with a depreciation rate of 0.133. Capital formation from 1976-1986 (1974-1984), was used for those countries which have IO tables for 1986 (1984). These capital totals were also converted to 1985 US dollars.

Labor Data:

For manufacturing sectors, data were taken from STAN Number Engaged (NE). For non-manufacturing sectors, the ISDB Total Employment (ET) was used. These labor data include self-employed, owner proprietors, and unpaid family workers. Labor data were taken from the same year as the IO table (1984, 1985, or 1986). Some scaling was also performed on the labor data. All sectors were scaled by the ratio of IO to STAN value added. In addition, manufacturing sectors were scaled by the ratio of ISDB to STAN total manufacturing employment.

For each manufacturing sector (I), in each country, for the year 1984, 1985, or 1986, labor was calculated as:

$$L_i = NE_i^{STAN} * \frac{ET^{ISDB,TotalManuf}}{NE^{STAN,TotalManuf}} * \frac{VA_i^{IO}}{VA_i^{STAN}}$$

For each non-manufacturing sector (j), labor was calculated as:

$$L_j = ET_j^{ISDB} * \frac{VA_j^{IO}}{VA_j^{ISDB}}$$

Production Data:

Data were taken from Gross Output column of the OECD Input-Output table and converted to 1985 US$.

Data Problems

Not all data were available in each database or consistent between databases. The following sectors have data problems of one sort or another. (Superscripts refer to the type of problem, discussed below.)

Australia	$(3\text{-}15^{1}, 16^{7,8}, 17^{1}, 18^{8}, 19^{1}, 20^{5,6,8}, 21^{8}, 22^{6}, 23^{1}, 25^{4}, 28^{1}, 33^{8}, 35^{1,5,8})$
Canada	$(20^{5,6}, 23^{1}, 24^{2,4}, 35^{1,5})$
Denmark	$(14^{1}, 15^{8}, 16^{8}, 17^{1}, 18^{1}, 19^{7}, 20^{2,5,8}, 21^{2,4,8}, 22^{5,6,8}, 23^{1}, 28^{1}, 32^{2}, 33^{2}, 35^{1,8})$
France	$(5^{1,2}, 10^{2}, 20^{2,4}, 23^{2}, 24^{2}, 32^{2}, 33^{2}, 34^{2,4}, 35^{3,4})$
Germany	$(7^{8}, 8^{8}, 17^{8}, 18^{8}, 20^{8}, 32^{2,4}, 33^{2,4})$
Italy	$(2^{2,4}, 5^{1,6}, 7^{2}, 8^{2}, 32^{5}, 33^{2,3,4,8}, 34^{8}, 35^{8})$
Japan	$(5^{1}, 9^{8}, 20^{7,8}, 22^{7}, 24^{8}, 28^{2,4}, 29/30^{2,4}, 31^{2,4}, 32^{2,5}, 33^{2,4}, 35^{1,8})$
Netherlands	$(12^{8}, 13^{8}, 14^{1}, 15^{1}, 16^{1}, 17^{8}, 18^{8}, 19^{2,4}, 20^{2,4}, 21^{1}, 22^{1}, 23^{1}, 31^{2}, 33^{2}, 35^{2})$
UK	$(5^{2}, 14^{2}, 23^{2}, 27^{3,5}, 28^{3,5}, 31^{3,5}, 32^{3,5}, 35^{1,5})$
US	$(20^{2,4}, 21^{2,4}, 27^{2,4}, 28^{2,4}, 35^{1,3,5})$

1. The following sectors have missing ISDB GFCF data or GFCF price data (IT and/or ITD files):

   Australia (3-15, 17, 19, 23, 28, 35), Canada (23, 35), Denmark (14, 17, 18, 23, 28, 35), Italy (5), Japan (5, 35), Netherlands (14-16, 21-23), UK (35), US (35).

2. The following sectors have ISDB or STAN capital data which include or exclude sectors that differ from the IO tables:

   Canada (24), Denmark (20, 21, 32, 33), France (5, 10, 20, 23, 24, 32, 33, 34), Germany (32, 33), Italy (2, 7, 8, 33), Japan (28-33), Netherlands (19, 20, 31, 33, 35), UK (5, 14, 23), US (20, 21, 27, 28).

3. The following sectors have missing ISDB Value Added (VA) data:

   France (35), Italy (32), UK (27, 28, 31, 32), US (35).

4. The following sectors have ISDB or STAN employment data which include different sectors than the IO tables:

   Australia (25), Canada (24), Denmark (21), France (20, 34, 35), Germany (32, 33), Italy (2, 33), Japan (28-31,33), Netherlands (19, 20), US (20, 21, 27, 28).

5. The following sectors have missing ISDB or STAN employment values:

   Australia (20, 35), Canada (20, 35), Denmark (20, 22), Italy (32), Japan (32), UK (27, 28, 31, 32, 35), US (35).

6. The following sectors have completely missing ISDB or STAN GFCF values:

   Australia (20, 22), Canada (20), Denmark (22), Italy (5)

7. The following sectors have ISDB or STAN GFCF values that are missing for some years:

Australia (16), Denmark (19), Japan (20, 22).

8. The following have unusual sectors included or excluded from the IO VA values:

   Australia (16, 18, 20, 21, 33, 35), Denmark (15, 16, 20, 21, 22, 35), Germany (7, 8, 17, 18, 20), Italy (33, 34, 35), Japan (9, 20, 24, 35). Netherlands (12, 13, 17, 18).

These omissions and inconsistencies were dealt with in the following ways:

1. The following sectors had missing GFCF price deflators (ITD/IT), for which the average manufacturing price deflator for the particular country was used.
   Netherlands (14-16, 21-23), Australia (3-15, 17, 19, 23), Canada (23), Denmark (14, 17, 18, 23)

2. Otherwise, see the description below for corrections of other missing data.

Construction of missing data for production, capital, and labor:

In all but a few cases, missing data were replaced by a two-step process. First, we calculated average input coefficients for countries which had output data for the sector. Second, this average was weighted by the size of gross output in the country with the missing sector.

1. For a country (r) with a missing sector (i) in the three non-manufacturing sectors 33-35 (SOC, PGS, and OPR), the production data was calculated as follows:

$$\tilde{X}_i^r = \frac{\sum_c X_i^c}{\sum_c X^{c,\, total}} * X^{r,\, total}$$

   This was done in: Australia (33, 35), Italy (33, 34)

2. To calculate missing or aggregated production data for manufacturing sectors, where STAN data were available, the following formula was used:

$$\tilde{X}_i^r = \left(X_i^r\right)^{Stan} * \frac{\left(X^{r,\, total\ \ manuf}\right)^{IO}}{\left(X^{r,\, total\ \ manuf}\right)^{Stan}}$$

   This was done in: Australia (16, 18, 21), Denmark (15, 16), Germany (7, 8, 17, 18, 20), Netherlands (12, 13, 17, 18)

3. Occasionally, IO, STAN, and ISDB production data were all problematic. In this case, the value of production in these sectors was taken directly from the IO tables without correction. Denmark's sectors 21 and 22 were included in sector 20, and the IO values of zero were used for 21 and 22.

This was done in: Australia (20), Denmark (20, 21, 22)

4. Some countries had data for OPR recorded as zeros, but this was believed to be the correct value.

   These sectors were: Denmark (35), France (35), UK (35)

   For all other sectors we set $\tilde{\mathrm{X}}_{\mathrm{i}}^{\mathrm{c}} = \mathrm{X}_{\mathrm{i}}^{\mathrm{c}}$.

5. For a country (r) with a missing sector (i), the capital stock in sector i was calculated by first finding the average input coefficient in other countries. This average was then multiplied by the total output of the country in the missing sector.

$$\tilde{\mathrm{K}}_{\mathrm{i}}^{\mathrm{r}} = \frac{\sum_{\mathrm{c} \neq \mathrm{r}} \mathrm{K}_{\mathrm{i}}^{\mathrm{c}}}{\sum_{\mathrm{c} \neq \mathrm{r}} \tilde{\mathrm{X}}_{i}^{\mathrm{c}}} * \tilde{\mathrm{X}}_{i}^{\mathrm{r}}$$

   This was done in: Australia (22, 28, 33, 35), Canada (20, 24, 35), Denmark (19, 28, 32, 33), France (5, 10, 20, 23, 24, 32, 33, 34), Germany (32, 33), Italy (2, 5, 7, 8, 32, 33, 34, 35) , Japan (5, 9, 20, 22, 24, 28-33, 35), Netherlands (19, 20, 31, 33, 35), UK (5,14,23, 27, 28, 31, 32), US (20, 21, 27, 28, 35)

6. Sectors with missing labor data were calculated in an identical way.

$$\tilde{\mathrm{L}}_{\mathrm{i}}^{\mathrm{r}} = \frac{\sum_{\mathrm{c} \neq \mathrm{r}} \mathrm{L}_{\mathrm{i}}^{\mathrm{c}}}{\sum_{\mathrm{c} \neq \mathrm{r}} \tilde{\mathrm{X}}_{i}^{\mathrm{c}}} * \tilde{\mathrm{X}}_{i}^{\mathrm{r}}$$

   This was done in: Australia (25, 33, 35), Canada (20, 24, 35), France (20, 34), Germany (32, 33), Italy (2, 32, 33, 34) , Japan (9, 20, 24, 28-33, 35), Netherlands (19, 20), UK (27, 28, 31, 32), US (20, 21, 27, 28, 35)

7. For sectors where production is zero, capital and labor are set equal to zero; K/X & L/X were set to average of other countries' values.

   This was done in: Denmark (21, 22, 35), France(35), UK(35)

8. After recalculating the data by the steps above, the total capital and labor for each country no longer summed to the total value given ISDB TET. Thus capital and labor for each sector were scaled as follows:

$$\tilde{\tilde{K}}_{i}^{r} = \frac{\tilde{K}_{i}^{r}}{\sum_{i} \tilde{K}_{i}^{r}} K^{r,\,\mathrm{ISDB\ TET}}$$

$$\tilde{\tilde{L}}_i^r = \frac{\tilde{L}_i^r}{\sum_i \tilde{L}_i^r} L^{r,\,\mathrm{ISDB\ TET}}$$

Production values were not rescaled.
These were the final values (of capital, labor and production) used in this paper.

Construction of the A matrix, demand, and trade data:

1. The A matrix was constructed by first taking input-output data from the IO tables and then dividing the input used in each sector by the corresponding sector's gross output. Any problematic elements of the A matrix were replaced by the average value of other countries whose corresponding elements have no problem. That is,

   $$\tilde{\mathrm{a}}_{\mathrm{ij}}^{\mathrm{R}} = (\mathrm{a}_{\mathrm{ij}}^{\mathrm{R}})^{\mathrm{avg}}$$

2. Since both the A matrix and production were constructed independently for problematic sectors, $\tilde{\mathrm{A}}^{\mathrm{R}}\tilde{\mathrm{X}}$ did not correspond to the values for total use $\mathrm{A}^{\mathrm{R}}\mathrm{X}_{\mathrm{IO}}$ given in the IO table, where A^{R} is the Rth row of the A matrix. Therefore, the A matrix was further scaled by the following method:

   $$\text{Let } \mathrm{A}^{\mathrm{R}}\mathrm{X}_{\mathrm{IOPL}} = \begin{cases} \mathrm{A}^{\mathrm{R}}\mathrm{X}_{\mathrm{IO}} & \text{if X in this sector was not constructed.} \\ \tilde{\mathrm{A}}^{\mathrm{R}}\tilde{\mathrm{X}} & \text{otherwise.} \end{cases}$$

   Let $\mathrm{AX}_{\mathrm{IOPL}}$ be the matrix whose rows are composed of $\mathrm{A}^{\mathrm{R}}\mathrm{X}_{\mathrm{IOPL}}$

   Find λ such that

   $$\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \tilde{\mathrm{A}}\tilde{\mathrm{X}} = \mathrm{AX}_{\mathrm{IOPL}}$$

   Then $\tilde{\tilde{\mathrm{A}}} = l\tilde{\mathrm{A}}$ was used as the final A matrix.

3. Demand data were taken from the IO table as the sum of Private Domestic Consumption, Government Consumption, GFCF and Changes in Stocks.

   For problematic sectors of SOC, PGS or OPR, the demand data were constructed as:

   $$\tilde{D} = (I - \tilde{\tilde{A}})\tilde{X}$$

   This was done in Australia (33,35), Italy (33,34) because we believed there to be very little trade in these sectors.

   For sectors where export data were missing from the IO table due to aggregation problems but present in STAN, the demand data were constructed as:

   $$\tilde{D} = (I - \tilde{\tilde{A}})\tilde{X} - (\tilde{E} - \tilde{M}),$$

   where

   $$\tilde{E}_i^R = \frac{\left(E_i^R\right)^{\mathrm{Stan}}}{\left(E^{\mathrm{M.\ total\ \ manuf}}\right)^{\mathrm{Stan}}} * \left(E^{\mathrm{M,\ total\ \ manuf}}\right)^{\mathrm{IO}}$$

   $$\tilde{M}_i^R = \frac{(M_i^R)^{\mathrm{Stan}}}{(M^{\mathrm{M,\ total\ manuf}})^{\mathrm{Stan}}} * (M^{\mathrm{M,\ total\ manuf}})^{\mathrm{IO}}$$

   This was done in Australia (16, 18, 21), Denmark (16, 17), Germany (7, 8, 17, 18, 20), Netherlands (13, 14, 17, 18).

4. Trade data were then constructed in the following way:

   $$\tilde{T} = (I - \tilde{\tilde{A}})\tilde{X} - \tilde{D}$$