

An Account of Global Factor Trade

By DONALD R. DAVIS AND DAVID E. WEINSTEIN*

A half century of empirical work attempting to predict the factor content of trade in goods has failed to bring theory and data into congruence. Our study shows how the Heckscher-Ohlin-Vanek theory, when modified to permit technical differences, a breakdown in factor price equalization, the existence of nontraded goods, and costs of trade, is consistent with data from ten OECD countries and a rest-of-world aggregate. (JEL F1, F11, D5)

A central objective of international economic research has been to account for the factor content of trade. There are two principal reasons. The first is that economists want to trace the effects of international influences on relative and absolute factor prices within a country. The Heckscher-Ohlin model and its variants, with their emphasis on trade arising from differences in the availability of productive factors, provide a natural setting for such investigations.¹

The second reason for the focus on the factor content of trade is that it provides a concrete prediction against which to measure how well our models work. The models are extraordinary in their ambition. They propose 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 have made these theories irresistible to empirical researchers.

In recent years, empirical research has focused on the Heckscher-Ohlin-Vanek (HOV) theorem because of the sharp predictions that this theory has for the links between trade, technology, and endowments. 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 Leontief's (1953) "paradox" is widely regarded as the first blow against the empirical validity 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 et al. [henceforth BLS] (1987) report that country net factor service exports are no better predicted by measured factor abundance than by a coin flip. In a major contribution, Daniel Trefler (1995) documented the "mystery of the missing trade"—that measured factor service trade is an order of magnitude smaller than that predicted based on national endowments. This point is underscored in work by Xavier Gabaix (1997). Davis et al. (1997) [henceforth DWBS] do report positive results for a variant of the HOV model. However they accomplish this by restricting the sample for which factor price equalization (FPE) is assumed to hold to regions of Japan and by

* Department of Economics, Columbia University, New York, NY 10027, and National Bureau of Economic Research. 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. Davis thanks colleagues at the Federal Reserve Bank of New York, where part of the work on this paper was completed. The views expressed in this paper do not necessarily reflect those of the Federal Reserve. We have benefited 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.

¹ One measure of the importance and freshness accorded this issue is that the *Journal of International Economics* devoted much of its February 2000 Jubilee Issue to a series of articles by Edward E. Leamer et al., debating precisely this question of when and how one can convert from measured factor contents to implied effects on national factor prices.

remaining agnostic about the degree to which the FPE framework can be extended across nations.

Our current research follows the recent literature in asking if parsimonious amendments allow the model to match the data. 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.

The results are striking. The step-by-step introduction of our key hypotheses yields corresponding improvement in measures of model fit. Countries export their abundant factors and in approximately the right magnitudes. The results are remarkably consistent across variations in weighting schemes and sample. As in prior studies, the simple HOV model is strongly rejected by the data. However a model that allows for technical differences, a breakdown of factor price equalization, the existence of nontraded goods, and costs of trade, is consistent with data for ten OECD countries and a rest-of-world aggregate.

I. Theory

Various hypotheses have been advanced to account for the divergence of theory and data, such as technical differences and divergences in demand structure. These have so far proved wholly insufficient to bridge the gap 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.

A successful account should provide a parsimonious and plausible set of departures from 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 assumptions one at a time. The novel elements of the theory are two. The first is to show, in an approximate FPE world, how

aggregation may systematically bias measured factor trade downward relative to the theoretically appropriate measure. The second is to develop factor content predictions for the case with no FPE that recognize the crucial role played in this case by the nontraded sector. These theoretical departures are developed in this section and 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 countries have identical, constant returns to scale production functions. Markets for goods and 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. Avinash K. Dixit and Victor Norman, 1980). Then factor prices will be equalized, so all producers will choose the same techniques of production. Let the matrix of total (direct plus indirect) factor inputs for country c be given by \mathbf{B}^c . This has dimension equal to the number of factors times the number of goods. The foregoing implies that for all countries c :

$$\mathbf{B}^c = \mathbf{B}^{c'} = \mathbf{B} \quad \forall c, c'.$$

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

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

where \mathbf{V}^c is the endowment vector for country c and \mathbf{Y}^c is the net output vector for country 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 equaliz-

ing traded goods prices and FPE equalizing nontraded goods prices, the demand in a country will be proportional to world net output:

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

where Y^W is net output for the world. Premultiplying this by the matrix of total factor inputs converts this to the factor contents:

$$B^{c'} D^c = s^c B^{c'} Y^W = s^c V^W.$$

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):

$$B^{c'} Y^c = V^c \text{ for a specified common technology matrix } B^{c'}.$$

Trade Specification (T1):

$$B^{c'} T^c = B^{c'} (Y^c - D^c) = V^c - s^c V^W \forall c.$$

Here and in subsequent specifications of both production and trade it will be convenient to distinguish the left- and right-hand sides of the equation respectively as the *measured* and *predicted* factor contents of trade.

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 that for country c the measured technology matrix is given as:

$$\ln B^c = \ln B^\mu + \varepsilon^c$$

where each element of $\ln B^\mu$ is the natural logarithm of the corresponding element of the true technology matrix and ε^c is a matrix of

normal error terms. This gives rise to our second set of tests:

Production Specification (P2):

$$B^\mu Y^c = V^c.$$

Trade Specification (T2):

$$B^\mu T^c = V^c - s^c V^W \forall c.$$

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., Dale W. Jorgenson and Mashiro 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. BLS (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. This can be characterized via country-specific technology shifts λ^c :

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

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 United States as a base, and U.S. factors are twice as productive as Italian factors, then $\lambda^{Italy} = 2$. In general, we can express a country’s endowments in efficiency terms:

$$V^{cE} = \frac{1}{\lambda^c} V^c \forall c.$$

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):

$$B^\lambda T^c = V^{cE} - s^c V^{WE} \forall c.$$

All succeeding models and the associated empirical specifications will be in efficiency units,

although we will henceforth suppress the superscript E for simplicity.

D. Dornbusch-Fischer-Samuelson Model

So far we have allowed differences in input coefficients across countries only as a Hicks-neutral shift, which we refer to as "adjusted FPE." 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. David Dollar et al. (1988) estimated cross-country differences in capital to labor usage and found this was correlated with country capital abundance.

From a theoretical perspective, the correlation between country factor abundance and industry factor usage could arise for either of two reasons. First, as Dollar et al. argued, it could reflect a breakdown in factor price equalization. We consider this possibility in the next subsection. Before doing so, however, we want to consider a simpler hypothesis. This is the hypothesis that FPE is *approximately* correct and that the differences in input usage by industry that we observe across countries arises almost exclusively due to aggregating goods of heterogeneous factor content within industry categories.

In a two-country Dornbusch-Fischer-Samuelson-type model with approximate FPE and this type of aggregation problem, we would expect to find that industry input usage is correlated with country capital abundance for tradables, but not for nontradables.² The reason is that the necessity for each country to produce its own nontradables, in combination with the standard assumptions on preferences, implies that input ratios will differ at most trivially for the nontraded sectors.

There is one additional assumption that is useful in thinking about why previous studies measured the factor content of trade to be so

low. This is that the empirical industrial aggregates cover a very wide range of goods, some nontraded, some traded, only some of the latter produced in any one country. Since virtually all previous studies have used the capital-abundant country's (in practice the United States) technology coefficients to measure the factor content of trade for all countries, there will arise in this framework a downward bias in the factor content of trade for both countries. The reasoning is as follows. If all goods have the same trade costs, then nontraded goods will tend to be those of intermediate factor intensity. Exports from the capital-abundant country in a given industry are more capital intensive than total production in that industry because of the presence of these nontraded goods. Thus by employing average input coefficients, there is a downward bias in the measured capital content of the capital-abundant country's exports. Measured capital intensity in the importable industry will likewise be a weighted average of the traded and nontraded goods produced in that empirical industry in the capital-abundant country. But this strongly understates the labor intensity of imports, because the goods actually imported are noncompeting, so do not even figure into the calculation of the input coefficients. In this two-country framework, the factor content of trade will be downward biased for both countries.³

These insights motivate a specification which explicitly recognizes that (1) The factor content of production in tradable industries varies systematically with country capital abundance; and (2) The factor content of absorption must be measured bilaterally with the producer country's input coefficients.⁴ This gives rise to the following specifications:

³ The situation is a little more complicated in a many-country world. What we can say is that the strength of the downward bias will be greatest for those countries with extreme factor abundance. As predicted factor contents get smaller, the degree of downward bias would be expected to decline. However, since it is the countries with little measured factor service trade and extreme factor abundances that are most important in generating the mystery of the missing trade, this bias is likely to be important.

⁴ It is worth noting that this ameliorates the aggregation problem in tradable sectors across countries that arises from the fact that they may produce distinct goods; however when we turn to the data exercises, we will not be able to wholly eliminate the aggregation problem caused by the contrast between average and exportable input coefficients.

² These ideas are developed in greater detail in a Dornbusch-Fischer-Samuelson-type setting in Davis and Weinstein (1998). See also Yingfeng Xu (1993).

Production Specification (P4):

$$\mathbf{B}^{cDFS}\mathbf{Y}^c = \mathbf{V}^c.$$

Trade Specification (T4):

$$\begin{aligned} \mathbf{B}^{cDFS}\mathbf{Y}^c - \left[\mathbf{B}^{cDFS}\mathbf{D}^{cc} + \sum_{c' \neq c} \mathbf{B}^{c'DFS}\mathbf{M}^{cc'} \right] \\ = \mathbf{V}^c - s^c\mathbf{V}^W. \end{aligned}$$

where the superscript in \mathbf{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.

E. Case Without Factor Price Equalization

Elhanan Helpman (1999) 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.⁵ One consequence of this is that the common set of nontraded goods will be produced using different techniques. In turn, this will affect our HOV factor content predictions. We now consider the implications.⁶

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 re-

⁵ In the theory, the failure of FPE is taken to be due to endowment differences. Unfortunately the analysis does not carry over precisely for the case in which the failure of FPE is due to nonnegligible trade costs. The key difficulty is that in general one would no longer expect the factor content of consumption of tradables to be $s^c\mathbf{V}^{WT}$. If one were willing to place enough structure on the problem, including assuming a symmetric geography, then one could get a very similar expression with a correction for the overconsumption of home goods and underconsumption of goods from the remaining countries. However we are not happy dealing in this way with the (important) fact that trade costs matter. See the approach of trade test (T7) below.

⁶ Adrian 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.

strict the number of nontraded 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 $\mathbf{B}^{cH} = [\mathbf{B}^{cHN}\mathbf{B}^{cHT}]$, where the partition is between nontradables and tradables. Let output be similarly divided, so $\mathbf{Y}^c = \begin{bmatrix} \mathbf{Y}^{cN} \\ \mathbf{Y}^{cT} \end{bmatrix}$. Then the factor content of production, by factor market clearing, is $\mathbf{B}^{cH}\mathbf{Y}^c = \mathbf{V}^c$. If we separate out nontradables and rearrange, we get $\mathbf{B}^{cHT}\mathbf{Y}^{cT} = \mathbf{V}^c - \mathbf{B}^{cHN}\mathbf{Y}^{cN}$. Let us call the expression on the right $\mathbf{V}^c - \mathbf{B}^{cHN}\mathbf{Y}^{cN} \equiv \mathbf{V}^{cT}$, so that $\mathbf{B}^{cHT}\mathbf{Y}^{cT} = \mathbf{V}^{cT}$. With no FPE, the prices of nontraded goods in terms of tradables will typically differ across countries. Assume that preferences in all countries between tradables and nontradables are similar 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 $\mathbf{D}^{c'cT} = s^c\mathbf{Y}^{c'T}$. The factor content of this absorption, using the factors actually engaged in production of the good, is $\mathbf{B}^{c'T}\mathbf{D}^{c'cT} = s^c\mathbf{B}^{c'HT}\mathbf{Y}^{c'T} = s^c\mathbf{V}^{c'T}$. Define $\mathbf{V}^{WT} \equiv \sum_c \mathbf{V}^{cT}$ and note that for $c \neq c'$, $\mathbf{D}^{c'cT} \equiv \mathbf{M}^{c'c'}$ (imports). Then it follows that:

$$\begin{aligned} \mathbf{B}^{cHT}\mathbf{Y}^{cT} - \left[\mathbf{B}^{cHT}\mathbf{D}^{c'cT} + \sum_{c' \neq c} \mathbf{B}^{c'HT}\mathbf{M}^{c'c'} \right] \\ = \mathbf{V}^{cT} - s^c\mathbf{V}^{WT}. \end{aligned}$$

That is, under the conditions stated above, we get something very much 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 nontradables 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 nontradables because the same goods are produced with different factor proportions. Let us expand the equation above:

$$\begin{aligned} & \mathbf{B}^{cHT}\mathbf{Y}^{cT} - \left[\mathbf{B}^{cHT}\mathbf{D}^{ccT} + \sum_{c' \neq c} \mathbf{B}^{c'HT}\mathbf{M}^{cc'} \right] \\ &= \mathbf{V}^{cT} - s^c \mathbf{V}^{WT} \\ &= [\mathbf{V}^c - \mathbf{B}^{cHN}\mathbf{Y}^{cN}] \\ &\quad - s^c \left[\sum_{c'} \{ \mathbf{V}^{c'} - \mathbf{B}^{c'HN}\mathbf{Y}^{c'N} \} \right]. \end{aligned}$$

The right-hand side of this equation can be rearranged to be:

$$\begin{aligned} &= [\mathbf{V}^c - s^c \mathbf{V}^W] \\ &\quad - \left[\mathbf{B}^{cHN}\mathbf{Y}^{cN} - s^c \sum_{c'} \mathbf{B}^{c'HN}\mathbf{Y}^{c'N} \right]. \end{aligned}$$

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

Production Specification (P5):

$$\mathbf{B}^{cH}\mathbf{Y}^c = \mathbf{V}^c.$$

Trade Specification (T5):

$$\begin{aligned} & \mathbf{B}^{cHT}\mathbf{Y}^{cT} - \left[\mathbf{B}^{cHT}\mathbf{D}^{ccT} + \sum_{c' \neq c} \mathbf{B}^{c'HT}\mathbf{M}^{cc'} \right] \\ &= [\mathbf{V}^c - s^c \mathbf{V}^W] - [\mathbf{V}^{cN} - s^c \mathbf{V}^{WN}]. \end{aligned}$$

where the superscript in \mathbf{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 right-hand side in (T5) is the standard HOV prediction, while the second is an adjustment that accounts for departures in

factor usage in nontradable goods from the world average. Note, for example that a capital-abundant country will have high wages, inducing substitution in nontradables toward capital. The second bracketed term 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 nontradables 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.

In trade specification (T6) we explore how our estimates of ROW ("Rest of the World") technology affect the results. We will discuss this in greater detail later, but in theoretical terms (T6) is the same as (T5). We now turn to results focused on the absorption side of the model.

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 [John McCallum (1995); Charles Engel and John H. Rogers (1996)]. 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 the measured factor trade in previous studies 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, nontariff 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. James Harrigan (1993) shows that the measures we have for nontariff barriers show very little impact on trade volumes in the

OECD relative to tariffs, which are themselves low. Nonetheless, as McCallum (1995) has 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. David L. Hummels (1999a, b) has undertaken the important task of improving our measures of trade costs and thinking hard about the role they play. However, we do not believe that the literature is yet sufficiently developed to pursue this direct approach.

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 James E. Anderson (1979), Jeffrey H. Bergstrand (1990), and Alan V. 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.

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' . Leaving aside the issue of market clearing for a moment, 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:

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

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 normal error term. The parameter estimates can then be used to generate predicted imports, $\hat{M}_{cc'}$. 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 solve the problem with a two-step procedure. Bilateral import demand is generated from the parameters of the gravity equation. We then set demand for domestically produced goods equal to the difference between total country demand in sector i and total imports predicted from the gravity equation. This implementation method enables us to impose the condition that the predicted total demand equals actual total demand. Nonetheless, deviations caused by the failure of the gravity equation to describe real trade flows could still cause problems with our demand approach.

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 $s^c \mathbf{V}^{WT}$. 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):

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

We wish to emphasize that the approach reflected in trade specification (T7) represents a significant departure from that in earlier specifications. It incorporates directly fitted values for import demand and complementary measures of own demand to generate the factor content of consumption. In doing so, it takes a larger step away from conventional HOV predictions, which generate the predicted factor content of consumption based directly on endowment data. The reason for pursuing this direction, this shortcoming notwithstanding, is that it will provide insight into how much further gain can be expected from a more complete model that does a

better job than standard HOV in predicting aggregate trade volumes.⁷

G. Summary of Specifications

The seven different specifications can be grouped in a variety of ways. In order to help keep the assumptions underlying these specifications clear, we summarize the four key groupings below:

	Yes	No
I. All countries share a common technology matrix (Absolute FPE)	(T1)–(T2)	(T3)–(T7)
II. FPE (Absolute, Adjusted, or Approximate)	(T1)–(T4)	(T5)–(T7)
III. Industry capital to labor ratios identical across countries	(T1)–(T3)	(T4)–(T7)
IV. Identical, homothetic preferences with zero trade costs	(T1)–(T6)	(T7).

II. 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 the 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 Organization

⁷ The fact that the gravity model is generally known to perform well in predicting trade volumes may make it seem likely that an improvement will result. Hence it is reasonable to ask what additional insight is gained from this experiment beyond a reaffirmation that the gravity model of demand fits well. We think that we do continue to learn a lot. First, while industry-level gravity regressions do a good job of predicting trade volumes, with typical R^2 's of approximately 0.7 in our runs, a great deal of this comes from size differences that may mask large proportional prediction errors. Second, nothing precludes the possibility that the errors may be systematic in a way that disrupts our HOV prediction. Finally, there is little doubt that there are a large number of considerations that potentially could be introduced to the analysis (e.g., industry-level scale economies). It is revealing to see how much incremental explanatory power is attained by better predictions of trade volumes and also revealing to see how much remains to be explained by the myriad factors excluded from the analysis.

for Economic Cooperation and Development's Input-Output Database [OECD (1995)]. This database provides input-output tables, gross output, net output, intermediate input usage, domestic absorption and trade data for 34 industries in ten OECD countries.⁸ Significantly, all of the data are designed to be compatible across countries. We constructed the country endowment data and the matrices of direct factor input requirements using the OECD's International 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:

$$(1) \quad \mathbf{B}^c \mathbf{Y}^c \equiv \mathbf{V}^c.$$

$$(2) \quad \mathbf{Y}^c - \mathbf{D}^c \equiv \mathbf{T}^c.$$

We also have data for 20 other countries that we aggregate and refer to as the "Rest of the World" or ROW.⁹ Data on capital is derived from the Robert Summers and Alan W. Heston (1997) 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 are taken from the United Nation's *Industrial Statistics Yearbook*, as modified by DWBS (1997). 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 were drawn from Robert C. Feenstra et al. (1997) and scaled so that bilateral industry import totals match country totals from the input-output (IO)

⁸ Australia, Canada, Denmark, France, Germany, Italy, Japan, Netherlands, United Kingdom, and the United States. These are the ten available in the IO database.

⁹ 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.

tables. Country c 's bilateral imports from country c' in nonmanufacturing sector i is set equal to the share of country c 's total manufacturing imports that came from country c' times total nonmanufacturing imports in sector i .¹⁰ ROW absorption was then set to satisfy condition (2) above.¹¹ Bilateral distance data comes from Shang-Jin Wei (1996).

In sum, this data set provides us with ten 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 United States) and no

¹⁰ This is not ideal, but given that the median ratio of imports to gross output in sectors outside of agriculture, mining, and manufacturing for our sample of countries is 1 percent, this is not likely to introduce large errors.

¹¹ 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 [1997] for a discussion of problems with the endowment data). It is quite difficult to match U.N. data with OECD IO data because of aggregation issues, varying country industry definitions, and various necessary imputations (see DWBS [1997] for details on what calculations were necessary). As a result, the output and absorption numbers of any individual country in the ROW are 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. In view of all this, we decided that we did not want to pollute a high quality data set with a large number of poorly measured observations.

observations whatsoever on 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.¹² 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 are 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 $\mathbf{B}^c \mathbf{Y}^c \equiv \mathbf{V}^c$.¹³

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 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

¹² Helpman (1999).

¹³ See the discussion in BLS (1987) of related difficulties. The only exception to this is the ROW where we were forced to use an estimated \mathbf{B} . See Section III for details.

abundant in it. This suggests that one appropriate interpretation is that our variable labor is a very rough proxy for unskilled or semiskilled 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, the United States 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 (1987) and Treffer (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 unavailability 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 ten 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 fall short of our ideal, we do not believe that this introduces any bias toward favorable results.

III. Statistical Tests on Technology and Absorption

Our principal statistical tests will work directly with the data on technology and absorp-

tion. We consider a variety of assumptions regarding technology and absorption suggested by theory and let the data select a preferred set. In Section IV we will 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 $\mathbf{B}^c = \mathbf{B}^{c'}$. 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 U.S. industry. While this may indicate different production techniques, the fact that net output in U.S. 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 \mathbf{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 United States tend to be small abroad, then evaluating the factor content of foreign production using the U.S. 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 United States.

A simple solution to this problem is to postulate that all countries use identical technologies but each measured \mathbf{B}^c is drawn from a random distribution centered on a common \mathbf{B} . If we postulate that errors are distributed normally, this relationship can be estimated by running the following regression:

$$(\hat{P}2) \quad \ln B_{fi}^c = \beta_{fi} + \varepsilon_{fi}^c.$$

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. This arises from the fact that we have better information about what the average unit input requirements are when there is more output. 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

$$\frac{\bar{B}_{fi} \sqrt{\ln(VA_{ic})}}{\bar{B}_f}$$

where \bar{B}_{fi} corresponds to the average across all countries of the unit factor input requirement in sector i , \bar{B}_f is the average of B_{fi} factor intensity across all sectors, and VA_{ic} refers to value added in sector i in country c .¹⁴

We noted earlier that 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) that are common across sectors, as developed in Section I, subsection C. If we denote these differences by λ^c , then we can econometrically identify these technical differences by estimating:

$$(\hat{P}3) \quad \ln B_{fi}^c = \theta^c + \beta_{fi} + \psi_{fi}^c$$

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 the possibility that production might be characterized by a continuum of goods DFS model in which industry input coefficients in tradables depend on country capital

¹⁴ The few sectors with nonpositive value added were given a zero weight in our regressions.

abundance (P4). These models can be easily implemented. We postulate that input coefficients are characterized by the following equation:

$$(\hat{P}4)$$

$$\ln B_{fi}^c = \theta^c + \beta_{fi} + \gamma_f^T \ln\left(\frac{K^c}{L^c}\right) TRAD_i + \phi_{fi}^c$$

where $TRAD_i$ is a dummy variable that takes on a value of one if the sector is tradable and ϕ_{fi}^c is an error term. If we wish to require that capital to labor ratios should not affect the aggregate productivity of the country, we must also require that¹⁵

$$\sum_f \gamma_f^T = 0.$$

If FPE breaks down and countries are in different production cones (P5), this will affect production coefficients in nontraded sectors as well. This hypothesis can easily be nested with (P4) by estimating

$$(\hat{P}5) \quad \ln B_{fi}^c = \theta^c + \beta_{fi} + \gamma_f^T \ln\left(\frac{K^c}{L^c}\right) TRAD_i + \gamma_f^{NT} \ln\left(\frac{K^c}{L^c}\right) NT_i + \phi_{fi}^c$$

where $NT_i = (1 - TRAD_i)$ is a dummy that is one for nontraded goods sectors. Here our assumption that capital to labor ratios do not affect productivity levels requires us to impose the following restriction on the estimates:

$$\sum_f (\gamma_f^T + \gamma_f^{NT}) = 0.$$

¹⁵ A word of caution is in order. For certain specifications of technical differences (i.e., all sectors within a country share a common, country-specific, Hicks-neutral shift term), our measure of productivity is identical to total factor productivity. Within this simple framework, our normalization scheme makes sense. However, our aggregate productivity number is subject to all of the standard critiques of any index number. In particular, the weights chosen for each sector, factor, and country will affect the measured aggregate productivity of the country.

TABLE 1—TESTS OF TECHNOLOGICAL VARIATION

Model	Measurement error	Hicks-neutral technical differences (HNTD)	Continuum model with HNTD and FPE	Helpman no-FPE model with HNTD	Unrestricted Helpman no-FPE model with HNTD	Implied λ^c
	($\hat{P}2$)	($\hat{P}3$)	($\hat{P}4$)	($\hat{P}5$)	($\hat{P}5'$)	
θ^{Aus}	—	0.531 (0.035)	0.531 (0.035)	0.530 (0.035)	0.528 (0.035)	1.7
θ^{Can}	—	0.381 (0.035)	0.381 (0.035)	0.380 (0.035)	0.381 (0.034)	1.5
θ^{Den}	—	0.508 (0.036)	0.504 (0.036)	0.508 (0.036)	0.508 (0.034)	1.7
θ^{Fra}	—	0.494 (0.034)	0.493 (0.034)	0.494 (0.034)	0.492 (0.035)	1.6
θ^{Ger}	—	0.112 (0.034)	0.111 (0.034)	0.112 (0.034)	0.111 (0.035)	1.1
θ^{Italy}	—	0.709 (0.034)	0.707 (0.034)	0.709 (0.034)	0.704 (0.036)	2.0
θ^{Japan}	—	0.431 (0.033)	0.430 (0.033)	0.431 (0.033)	0.430 (0.034)	1.5
θ^{Neth}	—	0.057 (0.035)	0.057 (0.035)	0.056 (0.035)	0.058 (0.035)	1.1
θ^{UK}	—	0.520 (0.034)	0.516 (0.034)	0.520 (0.034)	0.542 (0.040)	1.7
θ^{US}	—	0	0	0	0	1.0
γ^{KT}	—	—	0.408 (0.046)	0.364 (0.061)	—	
γ^{KN}	—	—	—	0.493 (0.071)	—	
γ^{LT}	—	—	-0.408 (0.046)	-0.449 (0.060)	—	
Number of parameters	68	77	78	80	144	
-Log L	-1741.5	-934.4	-855.7	-802.8	-740.7	
Schwarz criterion	-1963.3	-1185.5	-1110.1	-1063.7	-1210.3	

Notes: 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.

Finally, we can consider a more general specification in which we do not force country capital to labor ratios to have the same effects across sectors. This can be estimated by relaxing the assumption that the γ_f 's are the same across sectors, and estimating:

$$(\hat{P}5') \quad \ln B_{fi}^c = \theta^c + \beta_{fi} + \gamma_{fi} \ln \left(\frac{K^c}{L^c} \right) + \phi_{fi}^c$$

where now we simply require that the γ_{fi} 's sum to zero.

In each specification, we have 68 equations: one for each factor-industry pair. In specifications (P3) and higher we have cross-equation restrictions that require us to estimate the equations jointly. Ideally, we would use a seemingly

unrelated regressions procedure and exploit the cross-equation correlations in the error terms, but this would have used up more degrees of freedom than we had available. We therefore estimated these equations as a system of seemingly unrelated regressions with cross-equation restrictions but imposed a diagonal variance-covariance matrix on the residuals.

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 United States 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 Schwarz model selection criterion favors the model with neutral technical differences and no factor price equalization (P5).¹⁶ 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 1-percent increase in a country's capital to labor ratio typically raises each industry capital to labor ratio 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.¹⁷

Furthermore, we can use this approach to test an implication of factor price equalization within our model.¹⁸ 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 nontraded goods sector there should be no systematic variation in factor input ratios. Our estimates indicate that a 1-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 nontradables. 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 struc-

ture 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:

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

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 in specification (T7). One of the problems that we faced in implementation, however, was how to calculate the distance of any individual OECD country to the ROW. In all specifications we calculated this distance as the GDP-weighted average distance from a particular country to the countries in the ROW. In some sectors we found large systematic errors in predicting trade with the ROW. This may be the result of mismeasurement 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 and to save space we do not present them. 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).

We have already noted in the theoretical section that relying on a gravity specification to

¹⁶ 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 tended to work against models (P3)–(P5). Second, when we tried industry-by-industry estimation of $B_{K_i}^c/B_{L_i}^c$ 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.

¹⁷ 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 Schwarz criterion and only marginally better than (P2) and (T2) in the production and trade tests below. In short, in combination with the dependence of industry input ratios on country capital abundance, the efficiency adjustment is playing an important role in improving the performance of the HOV model.

¹⁸ Note that in this paper we do not examine factor prices, so do not test FPE itself.

generate import demands and the complementary own demands represents a substantial step away from the conventional HOV framework. Future work would do well to develop a more fully specified model relating production costs, trade costs, and trade volumes. Nonetheless, we believe that important insights can be gained from the gravity framework even in our more circumscribed setting.¹⁹

IV. Implications for Net Factor Trade

We have shown how a series of both conventional and novel assumptions about how trade links production technology and the choice of techniques can be tested with data on the technology matrices of countries. This allows us to select as the preferred model within this set production model (P5), based on a continuum of goods with no FPE. In turn, this would select trade specifications (T5) or (T6) as the preferred trade specifications. If one is willing to take a larger step away from the conventional HOV assumptions determining the underlying structure of demand, as reflected in our gravity estimates, then the preferred model is (T7).

This identification of preferred models notwithstanding, we now turn to provide various measures of how closely the production and trade specifications conform to the restrictions imposed on them by theory for all of our models. There are two reasons for doing this. First, by its nature, our procedure had to identify a

¹⁹ We outline here a few reasons we chose this path. 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 similarly 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 empirical results on technology strongly endorse a model with specialization in tradables, precisely the setting in which the gravity model is appropriate. Fourth, 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. Finally, we should note that there are, in principle, many things that could still be wrong with our model. Using the gravity demand specification will at least provide some insight about how much will remain to be explained once we have a fundamentals model which gets average trade volumes correct.

preferred specification. The reader, though, will wonder whether the preferred specification does well in an absolute sense. An examination of plots and simple regression statistics provides insight into this question. Second, it is reasonable to ask which of the assumptions are crucial for any improvements that result as we move between models. Because the production models are nested and feature a sensible progression from wholly standard to more novel approaches, it is straightforward to identify the impact of individual assumptions given the base of other cumulated assumptions.

The key specifications were developed in Section I and are summarized in Table 2. We 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.²⁰ 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).²¹ For example, in specification (P1) this involves a regression of $\mathbf{B}^c \mathbf{Y}^c$ on \mathbf{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

²⁰ In each case, efficiency units correspond to the estimated productivity parameters in the specification being considered. As one can see from Table 1, estimated productivity parameters differ only in the third digit as we move across specifications. Hence, the fact that country endowments measured in efficiency units change slightly across specifications has no significant quantitative impact on our results.

²¹ Measured Factor Content refers to factor content calculations obtained by applying a technology matrix to production or trade data. Predicted Factor Content refers to factor content calculations based on endowment data.

TABLE 2—KEY SPECIFICATIONS

	Key assumption	Production specifications		Trade specifications
(P1)	Conventional HOV with U.S. technology	$\mathbf{B}^{US}\mathbf{Y}^c = \mathbf{V}^c$	(T1)	$\mathbf{B}^{US}\mathbf{T}^c = \mathbf{B}^{US}(\mathbf{Y}^c - \mathbf{D}^c) = \mathbf{V}^c - s^c\mathbf{V}^W$
(P2)	Average technology matrix	$\hat{\mathbf{B}}^\mu\mathbf{Y}^c = \mathbf{V}^c$	(T2)	$\hat{\mathbf{B}}^\mu\mathbf{T}^c = \mathbf{V}^c - s^c\mathbf{V}^W$
(P3)	Hicks-neutral efficiency adjustment	$\hat{\mathbf{B}}^\lambda\mathbf{Y}^c = \mathbf{V}^{cE}$	(T3)	$\hat{\mathbf{B}}^\lambda\mathbf{T}^c = \mathbf{V}^{cE} - s^c\mathbf{V}^{WE}$
(P4)	Continuum model: Different input ratios in traded goods and H-N efficiency	$\hat{\mathbf{B}}^{cDFS}\mathbf{Y}^c = \mathbf{V}^c$	(T4)	$\hat{\mathbf{B}}^{cDFS}\mathbf{Y}^c - [\hat{\mathbf{B}}^{cDFS}\mathbf{D}^{cc} + \sum_{c' \neq c} \hat{\mathbf{B}}^{c'DFS}\mathbf{M}^{cc'}] = \mathbf{V}^c - s^c\mathbf{V}^W$
(P5)	Helpman no-FPE model, different input ratios in all, H-NE Forces ROW production model to work Adds gravity-based demand determination	$\hat{\mathbf{B}}^{cH}\mathbf{Y}^c = \mathbf{V}^c$	(T5)	$\hat{\mathbf{B}}^{cH}\mathbf{Y}^{cT} - [\hat{\mathbf{B}}^{cH}\mathbf{D}^{cET} + \sum_{c' \neq c} \hat{\mathbf{B}}^{c'H}\mathbf{M}^{cc'}] = [\mathbf{V}^c - s^c\mathbf{V}^W] - [\mathbf{V}^{cN} - s^c\mathbf{V}^{WN}]$
			(T6)	As above
		(T7)	$\hat{\mathbf{B}}^{cH}\mathbf{Y}^c - [\hat{\mathbf{B}}^{cH}\mathbf{D}^{cc} + \sum_{c' \neq c} \hat{\mathbf{B}}^{c'H}\mathbf{M}^{cc'}] = \mathbf{V}^c - [\hat{\mathbf{B}}^{cH}\hat{\mathbf{D}}^{cc} + \sum_{c' \neq c} \hat{\mathbf{B}}^{c'H}\hat{\mathbf{M}}^{cc'}]$	

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

absolute prediction error as a proportion of the predicted factor content of production. For example, for (P1) this is $|\mathbf{B}_f^{US}\mathbf{Y}^c - \mathbf{V}_f^c|/\mathbf{V}_f^c$.

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 $sign(\mathbf{B}_f^{US}\mathbf{T}^c) = sign(\mathbf{V}_f^c - s^c\mathbf{V}_f^W)$. 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 $\mathbf{B}'\mathbf{T}^c$ on $(\mathbf{V}^{fc} - s^c\mathbf{V}^{fW})$. 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.²²

B. *The Simple HOV Model Employing U.S. 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 3, specification (P1) fails miserably, but in an interesting way. A plot of (P1) for all countries appears as Figure 1. The United States 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 U.S. 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 4. 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 2. 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!²³ The

²² Not all of these tests have been run in prior work. However, the prior work does yield a simple bottom line: All prior studies on international data have fared poorly by at least one of these measures.

²³ At the 7-percent level of significance.

TABLE 3—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		
Standard error	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		
Observations	20	22	22	22	22		
Trade tests: Dependent variable MFCT							
	(T1)	(T2)	(T3)	(T4)	(T5)	(T6)	(T7)
Predicted	-0.002	-0.006	-0.05	0.17	0.43	0.59	0.82
Standard error	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
Variance ratio	0.0005	0.0003	0.008	0.07	0.19	0.38	0.69
Observations	22	22	22	22	22	22	22

Notes: The theoretical coefficient on “predicted” is unity. The theoretical value of the sign test is unity (100-percent correct matches). The variance ratio is $\text{Var}(\text{MFCT})/\text{Var}(\text{PFCT})$ and has a theoretical value of unity.

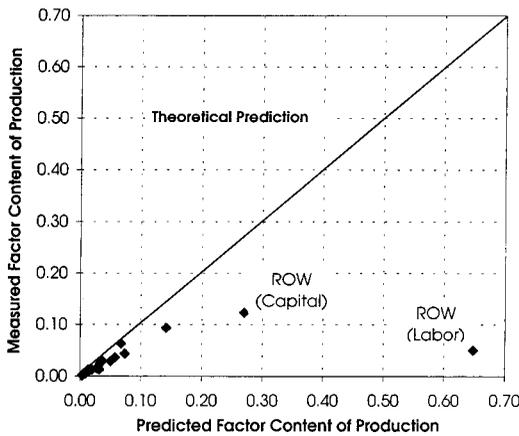


FIGURE 1. PRODUCTION WITH COMMON TECHNOLOGY (US)
(P1)

variance ratio is 0.0005, indicating that the variance of the predicted factor content of trade is about 2,000 times that of measured. This is missing trade big time! And the slope coefficient is zero (actually -0.0022 , $s.e. = 0.0048$). Moreover, under this specification our data reveal a Leontief “paradox” in which the United States is measured to be a net importer of capital services and a net exporter of labor.

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 U.S. 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). A plot of predicted and measured factor content of production appears as Figure 3. When all data points are included, the slope is 0.33. If we focus only on a regression based on our ten OECD countries (so excluding

TABLE 4—PRODUCTION AND TRADE TESTS
CAPITAL

	Production tests: Dependent variable MFCP					Trade tests: Dependent variable MFCT						
	(P1)	(P2)	(P3)	(P4)	(P5)	(T1)	(T2)	(T3)	(T4)	(T5)	(T6)	(T7)
Predicted	0.77	0.99	0.87	0.90	0.99	0.06	0.02	-0.03	0.25	0.57	0.65	0.87
Standard error	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: Dependent variable MFCP					Trade tests: Dependent variable MFCT						
	(P1)	(P2)	(P3)	(P4)	(P5)	(T1)	(T2)	(T3)	(T4)	(T5)	(T6)	(T7)
Predicted	0.07	0.12	0.92	0.87	0.94	-0.008	-0.008	-0.07	0.14	0.42	0.59	0.81
Standard error	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

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

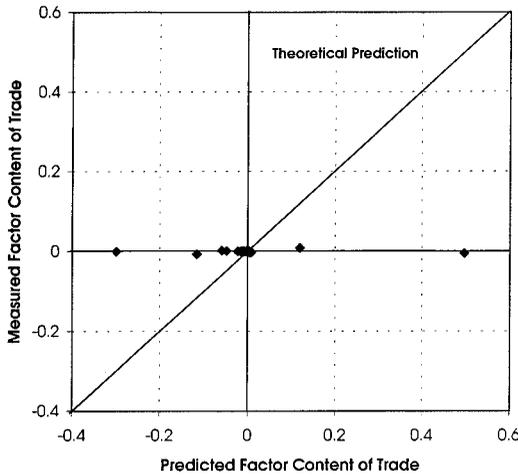


FIGURE 2. TRADE WITH COMMON TECHNOLOGY (US) (T1)

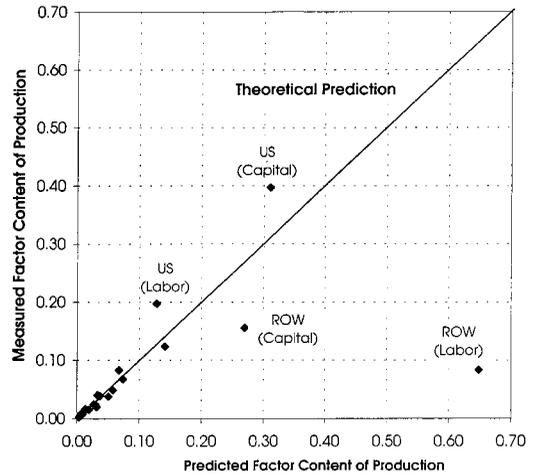


FIGURE 3. PRODUCTION WITH COMMON TECHNOLOGY (AVERAGE) (P2)

ROW), the slope in the production test rises sharply to 1.27, reflecting most strongly the influence of high productivity in the United States. If we exclude the United States as well, the slope falls to about 0.90. The R² 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 United States,

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. The plot of predicted and measured net factor trade looks essentially identical to Figure 2, indicating massive missing trade. 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.²⁴ The estimation of these efficiency differences is discussed above in Section III and here we view the implementation. A plot of (P3) appears as Figure 4. There continue to be substantial prediction errors, the largest by far being for the ROW, but also sizable ones for the United Kingdom 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

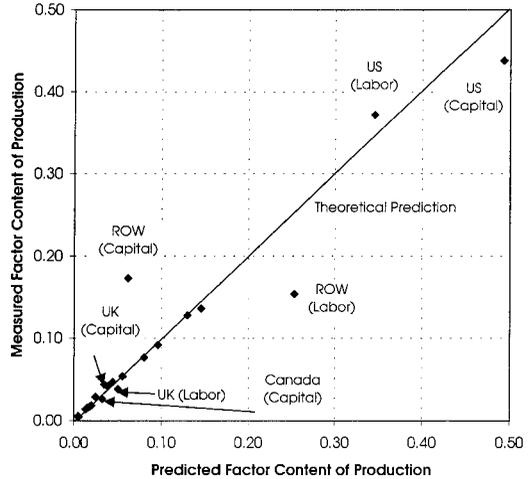


FIGURE 4. PRODUCTION WITH HICKS-NEUTRAL TECHNICAL DIFFERENCES (P3)

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. This high R^2 is a result of the important size effects present when comparing measured and actual factor usage across countries.

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.²⁵ These systematic biases are exactly what one would expect to find when using a common or neutralized technology matrix in the presence of

²⁴ 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 III. λ^{ROW} was set equal to the average productivity of labor and capital or:

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

In specifications (T6) and (T7), when we force the technology to fit exactly for the ROW, we pick two λ^{ROW} 's such that for each factor:

$$\lambda_f^{ROW} = \frac{f^{ROW}}{\hat{\mathbf{B}}_f^{ROW} \mathbf{Y}^{ROW}}.$$

²⁵ If we normalize the U.S. capital to labor ratio to one, then the capital to labor ratios of the remaining countries in descending order are Australia (0.95), Canada (0.92), Netherlands (0.92), France (0.88), Germany (0.84), Japan (0.79), Italy (0.71), Denmark (0.62), and the United Kingdom (0.48). For the four most capital-abundant countries we on average underestimate the capital intensity of their production by 10 percent and overestimate their labor intensity by 8 percent. For Denmark and the United Kingdom we overestimate their capital intensity by 25 percent and underestimate their labor intensity by 16 percent. In addition to the pattern we observe among the six countries discussed in the text, for the ROW (0.17), we also overestimate the capital content of ROW production and underestimate its labor content.

a continuum of goods. Moreover these biases are not small. Quite often these biases in over- or underpredicting 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. The plot of (T3) looks essentially the same as Figure 2, again indicating massive missing trade. 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 DFS Continuum Model with Industry Variation in Factor Employment: (P4) and (T4)

As we discussed in Section III, subsection A, 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 DFS (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 5. 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 6. 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

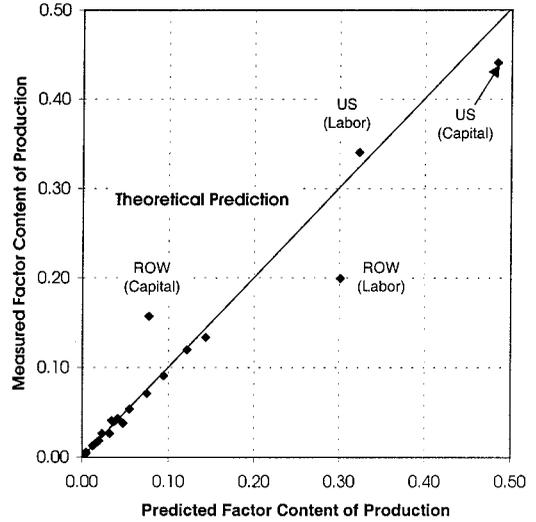


FIGURE 5. PRODUCTION WITH CONTINUUM OF GOODS MODEL AND FPE (P4)

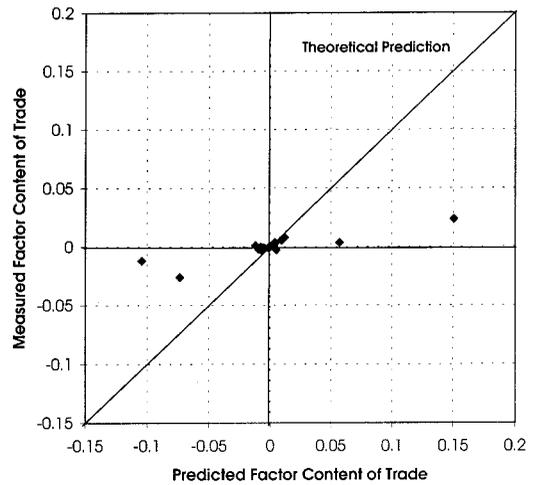


FIGURE 6. TRADE WITH CONTINUUM OF GOODS MODEL AND FPE (T4)

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 U.S. data for both capital and labor.²⁶ The most impressive statistic is the slope coefficient of 0.17,

²⁶ This type of Leontief paradox is absent in all subsequent tests.

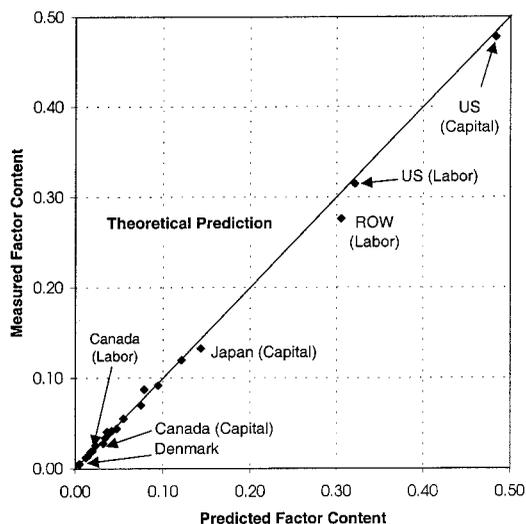


FIGURE 7. PRODUCTION WITHOUT FPE (P5)

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 Nontraded 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 nontradables 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 III. 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 nontradables arising from the failure of FPE.

This is our best model so far. Plots of production and trade specifications (P5) and (T5) appear in Figures 7 and 8. The production slope coefficient rises to 0.97, with an R^2 of 0.997. 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

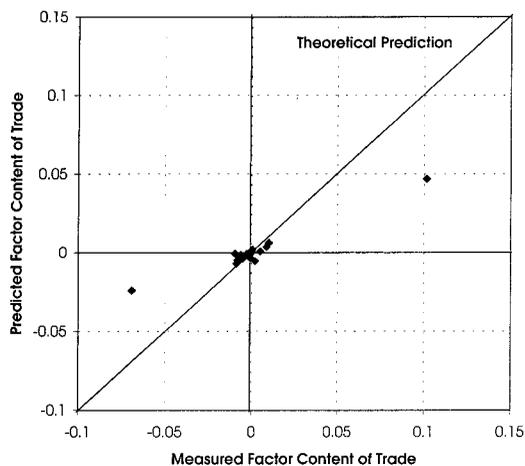


FIGURE 8. TRADE WITH NO-FPE, NONTRADED GOODS (T5)

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.²⁷

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, these errors need not immediately call for a revision of our framework.

There is one case, however, in which a closer

²⁷ Implementing production model (P5') (i.e., not forcing all sectors to have identical γ 's) yields results that are almost identical to model (P5), and so we do not report them.

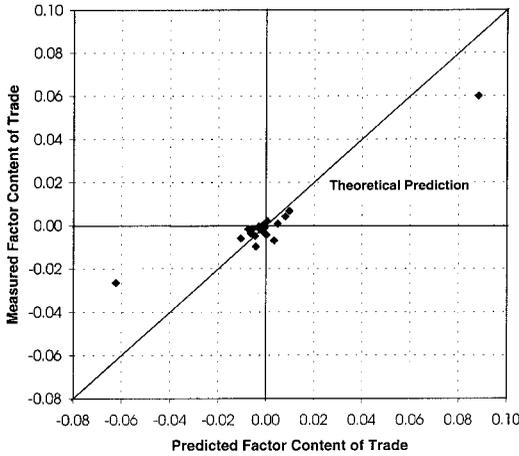


FIGURE 9. TRADE WITH NO-FPE, NONTRADED GOODS, AND ADJUSTED ROW (T6)

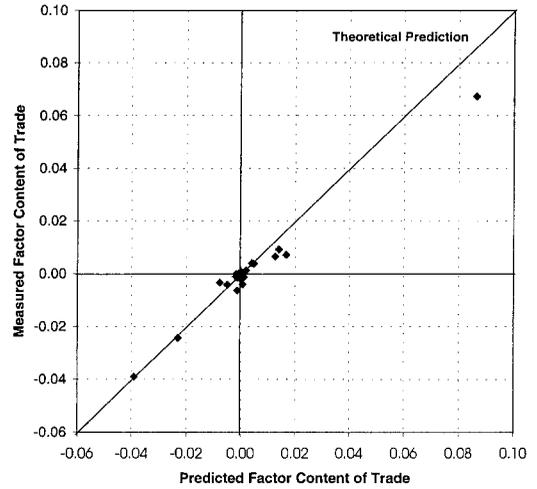


FIGURE 10. TRADE WITH NO-FPE, GRAVITY DEMAND SPECIFICATION, AND ADJUSTED ROW (T7)

review is appropriate. For the ten OECD countries, we have data on technology which enters into our broader estimation exercise. But this 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 countries and the ROW is large, there is a large amount of uncertainty about the adequacy of this projection. As it turns out, the prediction errors for ROW are large: the estimated technology matrix underpredicts labor usage by 9 percent, and overpredicts 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., $\mathbf{B}^{ROW}\mathbf{Y}^{ROW} \equiv \mathbf{V}^{ROW}$. A plot appears as Figure 9.²⁸ 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 countries to be importing well over half of the traded goods they absorb. Simple inspection of the data reveals this to be a wild overestimate of actual import levels.

We now take a larger step away from the standard HOV framework by estimating the log form of the gravity equation introduced earlier. This provides us with estimates of bilateral import flows in a world of perfect specialization with trade costs. We then use these estimates of import and implied own demand in order to generate factor service trade predictions. The results are presented in column (T7) and illustrated in Figure 10. By every measure, this is our best model of net factor trade. In moving from (T6) to (T7), the slope coefficient rises from 0.59 to 0.82. That is, measured factor

²⁸ To maintain consistency with the foregoing, we report the results here and in (T7) with all 11 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.

TABLE 5—TRADE TESTS
EXCLUDING ROW AND THE UNITED STATES
ALL FACTORS

Trade tests: Dependent variable MFCT							
	(T1)	(T2)	(T3)	(T4)	(T5)	(T6)	(T7)
Prediction	-0.05	-0.04	0.05	0.30	0.35	0.42	0.64
Standard error	0.01	0.01	0.07	0.07	0.11	0.14	0.09
Sign test	0.28	0.39	0.56	0.61	0.83	0.78	0.89
Variance ratio	0.00	0.01	0.09	0.09	0.30	0.49	0.54
R^2	0.47	0.33	0.03	0.51	0.41	0.36	0.76

Notes: The theoretical coefficient on “predicted” is unity. The theoretical value of the sign test is unity (100-percent correct matches). The variance ratio is $\text{Var}(\text{MFCT})/\text{Var}(\text{PFCT})$ and has a theoretical value of unity.

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 and trade frictions in limiting trade volumes and net factor contents.

I. Robustness Checks

We have undertaken a variety of additional checks of the robustness of our results. We examined the consequences of dropping influential points such as the ROW and the United States. The results of this exercise appear as Table 5 and indicate the robustness of our results to these exclusions. We have also considered alternative heteroskedasticity corrections similar to those of Trefler (1995) and Gabaix (1997).²⁹ While there are minor variations, the important point is that all of these checks yield results that strongly endorse the conclusions in our central case.

An additional question is why even in (T7) the slope coefficient is only 0.82 while the theoretical coefficient is unity. Several paths

for exploration seem reasonable. We did look at the possibility of attenuation bias due to measurement error. In the specifications that perform best, the high R^2 's leave little room for attenuation bias to be a major factor. A second factor at work is that our estimation can look only at adjustments to the capital to labor ratio in the average production matrix, not the export production matrix. Hence, even our estimates may understate the true factor content of trade. Third, some error surely arises from the fact that we have assumed that intermediates used in the production of exportables are all of national origin. We suspect these errors are not large, but they are worthy of more careful examination. Finally, and most obviously, trade barriers and home biases in demand 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.

V. Conclusion

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 replicate the feature of prior studies that the standard HOV theory performs miserably. However, our results also provide strong support for a model with a small number of amendments. Conditional on these amendments, countries export their abundant factors and they do so in approximately the right magnitudes. 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. Technical differences matter, even among the rich OECD countries.

²⁹ A full description of these results are available in Davis and Weinstein (1998).

Rich countries of the OECD cannot even be considered part of an efficiency-adjusted FPE set. This comes through most clearly in our results about systematic differences in industry factor usage in nontraded sectors. This failure of FPE promotes specialization in tradables and factor substitution in nontradables. The former implies that previous studies have underestimated the actual factor content of trade, while the latter implies that they have overestimated

the relative abundance of factors across countries that are available for production of tradables. Finally, there is significantly less trade than is implied by a frictionless model with perfect specialization in tradables. 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.

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) 1960–1995.

For production, demand, and trade:

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

Countries:

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

Australia, Canada, Denmark, France, Germany, Italy, Japan, Netherlands, United Kingdom, and United States.

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), United Kingdom (1984).

Sectors:

Data for each of the ten 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 Revision 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 [not elsewhere classified]
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 nonmanufacturing sectors, data were taken from ISDB ITD, which contains information on gross fixed capital formation in 1990 PPP prices in U.S. 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 in 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 nonmanufacturing sector j , GFCF was calculated as:

$$GFCF_j = ITD_j^{ISDB} * \frac{P^{US,1985}}{P^{US,1990}} * \frac{VA_j^{IO}}{VA_j^{ISDB}}.$$

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 nonmanufacturing sectors would be consistent if ISDB or STAN consistently under- or overreport 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:

$$GFCF_i = GFCF_i^{STAN} * \frac{ITD_i^{ISDB}}{IT_i^{ISDB}} * \frac{P^{US,1985}}{P^{US,1990}} * \frac{GFCF^{ISDB, Total Mfg.}}{GFCF^{STAN, Total Mfg.}} * \frac{VA_i^{IO}}{VA_i^{STAN}}.$$

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 U.S. PPP dollars and then followed our standard procedure.

Japanese capital formation was therefore calculated as follows:

$$GFCF_{i, Japan} = GFCF_i^{STAN} * \frac{1}{CGPD} * \frac{ITD_i^{ISDB}}{ITV_i^{ISDB}} * \frac{P^{US,1985}}{P^{US,1990}} * \frac{GFCF^{ISDB, Total Mfg.}}{GFCF^{STAN, Total Mfg.}} * \frac{VA_i^{IO}}{VA_i^{STAN}}.$$

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 U.S. 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:

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

For each nonmanufacturing 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 U.S. dollars.

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	(3–15 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)
United Kingdom	(5 ² , 14 ² , 23 ² , 27 ^{3,5} , 28 ^{3,5} , 31 ^{3,5} , 32 ^{3,5} , 35 ^{1,5})
United States	(20 ^{2,4} , 21 ^{2,4} , 27 ^{2,4} , 28 ^{2,4} , 35 ^{1,3,5}).

- 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), United Kingdom (35), United States (35).
- 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), United Kingdom (5, 14, 23), United States (20, 21, 27, 28).
- The following sectors have missing ISDB value added (VA) data:
France (35), Italy (32), United Kingdom (27, 28, 31, 32), United States (35).
- 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), United States (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), United Kingdom (27, 28, 31, 32, 35), United States (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 nonmanufacturing 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:

$$\tilde{X}_i^r = (X_i^r)^{STAN} * \frac{(X^{r,Total Mfg.})^{IO}}{(X^{r,Total Mfg.})^{STAN}}.$$

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), United Kingdom (35).

For all other sectors we set $\tilde{X}_i^c = X_i^c$.

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.

$$\tilde{K}_i^r = \frac{\sum_{c \neq r} K_i^c}{\sum_{c \neq r} \tilde{X}_i^c} * \tilde{X}_i^r.$$

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), United Kingdom (5, 14, 23, 27, 28, 31, 32), United States (20, 21, 27, 28, 35).

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

$$\tilde{L}_i^r = \frac{\sum_{c \neq r} L_i^c}{\sum_{c \neq r} \tilde{X}_i^c} * \tilde{X}_i^r.$$

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), United Kingdom (27, 28, 31, 32), United States (20, 21, 27, 28, 35).

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

This was done in: Denmark (21, 22, 35), France (35), United Kingdom (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 in the ISDB TET. Thus capital and labor for each sector were scaled as follows:

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

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

Production values were not rescaled. These were the final values (of capital, labor, and production) used in this paper. Once we had these variables, we then constructed the matrix of direct factor input requirements by dividing the amount of a given factor employed in a sector by gross output in that sector.

Construction of the Matrix of Intermediate Input Usage, Demand, and Trade Data:

1. The matrix of intermediate input usage, (henceforth the \mathbf{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 \mathbf{A} matrix were replaced by the average value of other countries whose corresponding elements have no

problem. That is,

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

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

$$\text{Let } \mathbf{A}^R \mathbf{X}_{IOPL} = \begin{cases} \mathbf{A}^R \mathbf{X}_{IO} & \text{For values of } \mathbf{X}_{IO} \text{ that were not constructed} \\ \tilde{\mathbf{A}}^R \tilde{\mathbf{X}} & \text{Otherwise.} \end{cases}$$

Let $\mathbf{A} \mathbf{X}_{IOPL}$ be the vector whose rows are composed of $\mathbf{A}^R \mathbf{X}_{IOPL}$
 Find λ such that

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

Then $\tilde{\tilde{\mathbf{A}}} = \lambda \tilde{\mathbf{A}}$ was used as the final \mathbf{A} matrix. We then postmultiplied our matrix of direct factor input requirements by $(\mathbf{I} - \tilde{\tilde{\mathbf{A}}})^{-1}$ to obtain the matrix of direct and indirect factor input requirements.

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:

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

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:

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

where

$$\tilde{E}_i^R = \frac{(E_i^R)^{STAN}}{(E^{M,Total Mfg.})^{STAN}} * (E^{M,Total Mfg.})^{IO}$$

$$\tilde{M}_i^R = \frac{(M_i^R)^{STAN}}{(M^{M,Total Mfg.})^{STAN}} * (M^{M,Total Mfg.})^{IO}.$$

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:

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

REFERENCES

- Anderson, James E.** "A Theoretical Foundation for the Gravity Equation." *American Economic Review*, March 1979, 69(1), pp. 106–16.
- Bergstrand, Jeffrey H.** "The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis and the Determinants of Bilateral Intra-industry Trade." *Economic Journal*, December 1990, 100(403), pp. 1216–29.
- Bowen, Harry P.; Leamer, Edward E. and Sveikauskas, Leo.** "Multicountry, Multifactor Tests of the Factor Abundance Theory." *American Economic Review*, December 1987, 77(5), pp. 791–809.
- Davis, Donald R. and Weinstein, David E.** "An Account of Global Factor Trade." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 6785, November 1998.
- Davis, Donald R.; Weinstein, David E.; Bradford, Scott C. and Shimpo, Kazushige.** "Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works." *American Economic Review*, June 1997, 87(3), pp. 421–46.
- Deardorff, Alan V.** "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" in Jeffrey Frankel, ed., *Regionalization of the world economy*. Chicago: University of Chicago Press and NBER, 1998.
- Dixit, Avinash K. and Norman, Victor.** *Theory of international trade*. Cambridge: Cambridge University Press, 1980.
- Dollar, David; Wolff, Edward N. and Baumol, William J.** "The Factor-Price Equalization Model and Industry Labor Productivity: An Empirical Test across Countries," in Robert C. Feenstra, ed., *Empirical methods for international trade*. Cambridge, MA: MIT Press, 1988, pp. 23–47.
- Dornbusch, Rudiger; Fischer, Stanley and Samuelson, Paul A.** "Heckscher-Ohlin Trade Theory with a Continuum of Goods." *Quarterly Journal of Economics*, September 1980, 95(2), pp. 203–24.
- Engel, Charles and Rogers, John H.** "How Wide Is the Border?" *American Economic Review*, December 1996, 86(5), pp. 1112–25.
- Feenstra, Robert C.; Lipsey, Robert E. and Bowen, Harry P.** "World Trade Flows, 1970–1992, with Production and Tariff Data." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5910, January 1997.
- Gabaix, Xavier.** "The Factor Content of Trade: A Rejection of the Heckscher-Ohlin-Leontief Hypothesis." Mimeo, Harvard University, 1997.
- Harrigan, James.** "OECD Imports and Trade Barriers in 1983." *Journal of International Economics*, August 1993, 35(1–2), pp. 91–111.
- Helpman, Elhanan.** "The Structure of Foreign Trade." *Journal of Economic Perspectives*, Spring 1999, 13(2), pp. 121–44.
- Hummels, David L.** "A Geography of Trade Costs." Mimeo, University of Chicago Graduate School of Business, 1999a.
- _____. "Transportation Costs and International Integration in Recent History." Mimeo, University of Chicago Graduate School of Business, 1999b.
- Jorgenson, Dale W. and Kuroda, Masahiro.** "Productivity and International Competitiveness in Japan and the United States, 1960–1985," in Charles R. Hulten, ed., *Productivity growth in Japan and the United States*. Chicago: University of Chicago Press, 1990, pp. 29–55.
- Leontief, Wassily.** "Domestic Production and Foreign Trade: The American Capital Position Re-Examined." *Proceedings of the American Philosophical Society*, 1953, 97, pp. 332–49.
- Maskus, Keith E.** "A Test of the Heckscher-Ohlin-Vanek Theorem: The Leontief Commonplace." *Journal of International Economics*, November 1985, 19(3–4), pp. 201–12.
- McCallum, John.** "National Borders Matter: Canada-U.S. Regional Trade Patterns." *American Economic Review*, June 1995, 85(3), pp. 615–23.
- Organization of Economic Cooperation and Development.** "The OECD Input-Output Database." Diskette, OECD Economic and Statistics Division, 1995.
- _____. "International Sectoral Database 1960–1995 Version 96." Diskette, OECD Statistics Directorate, 1996.
- _____. "The OECD STAN Database 1970–1995." Diskette, OECD Economic and Statistics Division, 1997.

- Summers, Robert and Heston, Alan W.** "Penn World Table, Version 5.6." <http://pwt.econ.upenn.edu/>, 1997.
- Trefler, Daniel.** "The Case of the Missing Trade and Other Mysteries." *American Economic Review*, December 1995, 85(5), pp. 1029–46.
- _____. "The Structure of Factor Content Predictions." Mimeo, University of Toronto, 1998.
- Wei, Shang-Jin.** "Intra-National versus International Trade: How Stubborn Are Nations in Global Integration?" National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5531, 1996.
- Wood, Adrian.** *North-South trade, employment and inequality: Changing fortunes in a skill-driven world*. Oxford: Oxford University Press, 1994.
- Xu, Yingfeng.** "A General Model of Comparative Advantage with Two Factors and a Continuum of Goods." *International Economic Review*, May 1993, 34(2), pp. 365–80.