NBER WORKING PAPER SERIES

THE HECKSCHER-OHLIN-VANEK MODEL OF TRADE: WHY DOES IT FAIL? WHEN DOES IT WORK?

Donald R. Davis
David E. Weinstein
Scott C. Bradford
Kazushige Shimpo

Working Paper 5625

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1996

We wish to thank Alan V. Deardorff, J. David Richardson, and Daniel Trefler for comments and suggestions. Jie Wang provided tremendous help entering the data for this paper. The authors blame each other for any errors. This paper is part of NBER's research program in International Trade and Investment. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1996 by Donald R. Davis, David E. Weinstein, Scott C. Bradford and Kazushige Shimpo. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

NBER Working Paper 5625
June 1996

THE HECKSCHER-OHLIN-VANEK MODEL OF TRADE: WHY DOES IT FAIL? WHEN DOES IT WORK?

ABSTRACT

The Heckscher-Ohlin-Vanek model of factor service trade is a central construct in international economics. Empirically, though, it is a flop. This warrants a new approach. Using Japanese regional data we are able to test the HOV model by independently examining its component production and consumption elements. The strict HOV model performs poorly because it cannot explain the international location of production. However, relaxing the assumption of universal factor price equalization yields a dramatic improvement. We also solve most of what Trefler (1995) calls the "mystery of the missing trade." In sum, the HOV model performs remarkably well.

Donald R. Davis
Department of Economics
Harvard University
Cambridge, MA 02138
and NBER

David E. Weinstein
Department of Economics
Harvard University
Cambridge, MA 02138

Scott C. Bradford
Department of Economics
Harvard University
Cambridge, MA 02138

Kazushige Shimpo
Department of Business and
and Commerce
Keio University
Tokyo 108
JAPAN

THE HECKSCHER-OHLIN-VANEK MODEL OF TRADE:

WHY DOES IT FAIL? WHEN DOES IT WORK?

1.0 Introduction

Starting with the classic "paradox" of Leontief (1953), and continuing through the influential work of Bowen, Leamer and Sveikauskas (1987), the Heckscher-Ohlin model has consistently performed poorly in empirical tests. This led Trefler (1993) to aver that "its predictions are always rejected empirically." In spite of these empirical failures, the Heckscher-Ohlin model remains ubiquitous in theory, empirics and policy analysis. In part, this reflects an *a priori* belief that the model embodies fundamental general equilibrium links between primary factors and production structure that we believe will be part of any fully articulated and empirically relevant theory.[1] Moreover, we will argue that when applied to regions within a country, the theory may do quite well as a simple description of the data.

An important departure in our analysis is to focus separately on the Heckscher-Ohlin predictions concerning the location of production and the pattern of consumption, rather than directly considering the pattern of trade. We argue that it is the location of production that is the heart of the Heckscher-Ohlin framework, and show that the model performs admirably in describing Japanese regional patterns of production. Moreover, when certain specifications of Heckscher-Ohlin fail, we are able to directly identify the reasons for the failure rather than, as in previous work, rely on indirect inferences.

A second important departure in our study is that we show how to derive *exact* predictions for the net factor content of trade in a world in which only a subset of regions share factor price

---

[1] The commitment of the profession to the Heckscher-Ohlin framework in the face of contrary evidence was highlighted in this paper's original title: "Interregional and International Trade: Woody Allen was Right!" This referred to an anecdote Allen tells at the close of his movie, *Annie Hall*, regarding a man whose brother thinks he's a chicken. Asked by his psychiatrist why he doesn't inform his brother that he is no fowl, Allen replies that he would, but the family needs the eggs. Just so, the profession has needed a general equilibrium framework, as in Heckscher-Ohlin, linking endowments, technology, and trade, contrary evidence notwithstanding.

equalization. This allows us to forego the heroic assumption of universal factor price equalization, continue to embed this in a full world general equilibrium, yet derive exact predictions to compare with the data.[2]

The results of this study should be heartening for international trade economists. The study does not provide evidence that the Heckscher-Ohlin-Vanek framework can be blithely and blindly applied to international data. However, it does validate the use of the underlying general equilibrium structure as an excellent description of national data. Insofar as an important concern of trade economists is to trace through the impact on national economies of international disturbances, this should be helpful. For example, augmentation of national factor supplies via international factor movements can be expected to have conventional impacts on output supplies, as per the Rybczynski theorem. One should even be able to trace the impact of this disturbance to the regional level. While our study does not contemplate any comparative statics involving international prices, the validation of the underlying general equilibrium production structure gives some support to the link between the international terms of trade and national factor returns posited in the Stolper-Samuelson theorem. While we develop these results in a regional setting, our approach can be extended to a cross-country study even if only a subset of countries share the assumptions underlying factor price equalization.

We also solve most of what Trefler (1995) refers to as the "mystery of the missing trade." Using international data, Trefler graphed the residuals from the Heckscher-Ohlin-Vanek predictions against predicted net factor trade. Rather than being centered around zero, the residuals instead closely follow a line reflecting zero net factor trade, hence the mystery. Seeking to account for this anomaly, he then implemented a hypothesis testing approach earlier employed by Bowen, Leamer and Sveikauskas (1987). Using maximum likelihood estimates of key parameters, he selects a model with neutral technological differences and a home bias in consumption as the

---

[2] Bowen, Leamer, and Sveikauskas (1987) effectively impose universal factor price equalization. Horiba (1992) and Grimes and Prime (1993) fail to consider the full world general equilibrium. The work of Brecher and Choudhri (1988) is closest to our approach. They derive exact predictions for the factor content of consumption per unit of expenditure. The full world equilibrium, though, enters their model only indirectly through its impact on relative spending shares for the countries considered. By contrast our approach also incorporates information on the full world production structure.

preferred account from a variety of nested hypotheses. We replicate his "missing trade" results for the Japanese regions under the maintained assumption of universal factor price equalization. However, under our preferred specification, the mystery of the trade that didn't embark largely disappears. Our approach has two advantages. First, the solution that we propose uses data and theory only. Because all theoretical parameters can be calculated directly from the data based on the theory, we are not obliged to *estimate* productivity differences or taste differences that allow the data to fit the model. Second, because we use separate and direct measures of production and absorption, we need not rely on econometric specifications to identify which is responsible for failures of the theory and hence what revisions are required to obtain a good fit with the data. As we will see, this leads to a substantively different result than that suggested by Trefler.

Our results also throw a revealing light on the recent literature on economic geography [Krugman (1991)]. This literature, born of the "new" trade theory of the 1980s, interacts scale economies with transport costs. Externalities, both pecuniary and non-pecuniary, have played a central role in this literature, helping to explain regional concentration. One consequence of these externalities is the appearance of significant differences across regions in average input coefficient matrices [cf. Krugman (1991)]. By contrast, in the conventional Heckscher-Ohlin model, both average and marginal input coefficients will be common across regions. In fact, our work on the Heckscher-Ohlin theory of the location of regional production can be looked at as a rough test of the equality of average input coefficients across the regions. As we will see, the Heckscher-Ohlin model provides an excellent fit for the data. This suggests that, at least for Japan, geography models that emphasize regional differences in average input coefficients will add little to our understanding of regional net factor trade, a traditional concern of Heckscher-Ohlin. By contrast, the inability of the simple model to account for the cross-national pattern of production leaves open the possibility that these concerns may have more salience for international production patterns.

Finally, our study provides the first evidence regarding the Courant-Deardorff (1992) conjecture that lumpiness – highly uneven regional distribution of national endowments – may be an important determinant of international trade patterns. The evidence showing an excellent fit of

the basic Heckscher-Ohlin model for our regional data suggests that, at least for Japan, lumpiness is not an important determinant of the structure of national production, hence trade.

1.1 The Importance of Heckscher-Ohlin to International Economics

It would be difficult to exaggerate the importance of the Heckscher-Ohlin model to international economics.[3] The basic model appears quite restrictive, requiring a long list of special assumptions. Yet the framework has proven remarkably flexible, particularly in the versions due to Vanek (1968), Helpman (1981), and Helpman and Krugman (1985), which emphasize implicit trade in factor services. It is consistent with trade in final goods and intermediates, as well as with the existence of nontraded goods. It is robust to a shift from constant returns to scale to some patterns of increasing returns. It can incorporate technical differences when the number of goods exceeds the number of factors, perhaps as the result of local or international spillovers of learning by doing [see, for example, Davis (1995)]. It is consistent with market structures as diverse as perfect competition, monopolistic competition, oligopoly, and contestable markets. It has also been a fundamental input into much of the new international growth theory, as in Grossman and Helpman (1992). In addition to its importance to theorists, Heckscher-Ohlin has provided the framework for countless empirical studies, as for example in the recent controversy over the influence of trade on US wages [Lawrence and Slaughter (1993)].

In short, the Heckscher-Ohlin framework continues to command interest among those concerned with international exchange from a wide variety of methodological and theoretical viewpoints. Moreover, this continued interest is no atavistic affectation. It reflects the enduring power of the model as an intellectual organizing framework, providing access to such powerful results as Factor Price Equalization, Stolper-Samuelson, Rybczynski, as well as the Heckscher-

---

[3] Strictly, one should distinguish between the Heckscher-Ohlin theory and the specific version known as Heckscher-Ohlin-Vanek (HOV). The former encompasses a broad class of competitive general equilibrium models which view cross-country differences in endowments as central to accounting for trade patterns. It is best known in the two good, two factor version articulated by Samuelson. The latter is a specific version developed in a many good, many factor framework that focuses on implicit trade in factor services. Often the poor empirical results of the latter have been taken as a repudiation of the entire framework. Our approach will bridge the narrow and broad theories by examining cases in which the more specific requirements of HOV are satisfied only by a subset of countries. In the text we will distinguish the versions explicitly only as required for clarity.

Ohlin theorem itself. The continued interest also reflects the absence of a clear and simple alternative that does better. Recent work, notably Trefler (1993, 1995), has sought to remedy this by examining simple alterations that improve the explanatory power of the model, and identifying regular features in the data inconsistent with the model.

A troubling fact concerning much of the evidence against Heckscher-Ohlin is that it does not identify why the model's predictions fail or how deeply the critique cuts. Are deviations from the model's predictions largely the result of a failure of factor price equalization, differing technologies, non-identical preferences, or other elements that may differ across countries? Or does the critique penetrate even more deeply, so that the Rybczynski and Stolper-Samuelson theorems are unreliable even on the national level? Our study starts us on the road to delimiting the nature of the empirical failures and successes of the model.

2.0 Theory and Tests

Here we lay out the theoretical foundation of our work on Heckscher-Ohlin-Vanek. We do so with three aims in mind. The first is to articulate the theoretical case for using the HOV theory to examine the total[4] trade of regions within a country. The second aim is to make clear how the component theories of the structure of production and absorption mesh in the derivation of the standard HOV equation. We do this because an important advance in our work is to disentangle these elements in considering the failure of HOV in some of the trade tests. The third aim is to articulate the implications of the framework as we relax some of the strictures of the standard model. As we will see, the strictest version of our model will perform poorly, so it is natural to consider what relaxations of assumptions matter most in getting the model to work well.

One important feature of our work is that we will examine the HOV framework in terms of its implications for the trade of regions of Japan. Here we will discuss the theoretical justification for such an approach (while the exact equations will be derived below). The standard HOV model

[4] By total trade we mean a region's trade with the rest of the world not just trade with other Japanese regions.

is developed in a setting in which there is factor price equalization (FPE) for the world as a whole. As argued by Dixit and Norman (1980), the very essence of FPE is that under certain conditions trade in goods alone replicates what is known as an "integrated equilibrium." This is a hypothetical world economy in which both goods and factors are perfectly mobile. When the relevant conditions are satisfied, the division into countries has no consequence for the real equilibrium that the world attains. It follows that the fundamental HOV relations hold at *any* level of aggregation or disaggregation of the world endowments that one may want to consider, whether by countries or by regions.

The HOV theory does not predict the pattern of goods trade, but rather the net factor content of trade. As net goods trade is simply the difference between output and absorption, net factor trade will equal the difference in their factor content. While our overall objective will be to examine the net factor content of trade, a major advance in our work is that the data we employ will allow us to examine separately the accuracy of the theory of production structure from that of absorption.

We begin by establishing some notation. Let $c \in W$ be an index of countries in the world, and $r \in J$ be an index of regions in Japan. Then we can begin by describing the technology. The matrix of direct factor inputs is given by $\boldsymbol{B}^k$, for $k = r$ or c. The input-output matrix is given by $\boldsymbol{A}^k$, for $k = r$ or c. In the standard HOV case of identical technologies and factor price equalization for the world, both $\boldsymbol{B}^k$ and $\boldsymbol{A}^k$ are common for the entire world. In fact, previous empirical work on HOV has held as a maintained assumption the equality of $\boldsymbol{B}^k$ and $\boldsymbol{A}^k$ across all countries in the world. Data limitations oblige us to continue to treat $\boldsymbol{A}^k$ as common across all countries in some of our tests (see Section 3.0 for more discussion of this point). However, allowing for the possibility that $\boldsymbol{B}^k$ may vary across countries will be an important element in our story.

Assumptions on Technology:

Maintained:	$\boldsymbol{A}^k = \boldsymbol{A}^J$	for $k = r$ or c	
World FPE:	$\boldsymbol{B}^k = \boldsymbol{B}^J$	for $k = r$ or c	(FPE-W)

FPE for Japan: $\boldsymbol{B}^r = \boldsymbol{B}^J$ for $r \in J$ (FPE-J)

The first condition merely asserts that it is a maintained assumption of this study that the Japanese input-output matrix is common to all countries in the world. The second and third represent alternative assumptions that will be considered in the course of our work.

Production Theory:

Let $\boldsymbol{X}^k$ be the gross output vector, and $\boldsymbol{V}^k$ be the factor endowment vector for $k = r \in J$, $c \in W$. Then the constraint that HOV places on the relation between output, technology, and endowments is very simple. It is summarized by two types of production tests, corresponding to two assumptions on the geographical extent to which FPE holds. When FPE holds throughout the world, all countries throughout the world (including Japan) use the same techniques.

P-I: $\boldsymbol{B}^J\boldsymbol{X}^c = \boldsymbol{V}^c \quad c \in W$ (FPE-W)

Even if factor price equalization fails to hold for the world as a whole, if it holds for the regions of Japan, then we can still write:

P-II: $\boldsymbol{B}^J\boldsymbol{X}^r = \boldsymbol{V}^r \quad r \in J$ (FPE-J)

Consumption Theory:

The HOV theory of absorption holds that preferences are identical and homothetic across the whole world (IHP-W). Assuming that trade is universally free and that transport costs are zero insures that all locales face the same vector of goods prices. Let $\boldsymbol{D}^k$ be absorption, $\boldsymbol{Y}^k$ be net output, and s^k be the share in world spending for $k = r \in J$, $c \in W$. Then for any region $r \in J$, we can state the standard HOV prediction of the goods content of absorption as:

C-I: $\boldsymbol{D}^r = s^r\boldsymbol{Y}^W \quad r \in J$ (IHP-W)

Several authors, such as Hunter and Markusen (1989), Hunter (1991), and Trefler (1995), have argued that it is failures in equation C-I that may account for much of the differences between HOV prodictions and reality. Therefore, an important question that we will address in this paper is the

role of failures of IHP-W in accounting for failures in our trade tests. Thus we will consider a hypothesis which makes much weaker claims. In effect, it abandons trying to explain the pattern of Japanese *aggregate* absorption, and instead examines whether the assumption of identical homothetic preferences is plausible for the regions of Japan alone. This may be expressed as:

C-II: $\boldsymbol{D}^r = (s^r / s^J)\, \boldsymbol{D}^J \qquad r \in J \qquad$ (IHP-J)

Trade Theory:

The theory of absorption directly concerns goods, while the HOV theory concerns net trade in factors. So, in order to incorporate the consumption theory into the the HOV trade theory, the consumption theory must be converted from a theory about consumption of goods to one of factors. The conversion to the direct and indirect factor content is accomplished via premultiplication by the term $\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}$. We can also note that under our maintained assumption that the input-output matrices are common to the world, $(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{Y}^W = \boldsymbol{X}^W$. Then one can always premultiply the C-I equation by $\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}$, to yield:

$$\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{D}^r = s^r\boldsymbol{B}^J\boldsymbol{X}^W \qquad r \in J \qquad \text{(IHP-W)}$$

If, *in addition*, the condition FPE-W is satisfied, it will also be true that $\boldsymbol{B}^J\boldsymbol{X}^W = \boldsymbol{V}^W$. In this case, this gives the implicit factor content of absorption as:

$$\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{D}^r = s^r\,\boldsymbol{V}^W \qquad r \in J \qquad \text{(IHP-W, FPE-W)}$$

If we believe neither in identical homothetic preferences nor FPE for the world, but instead, just identical homothetic preferences within Japan, then the implied factor content of absorption is:

$$\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{D}^r = (s^r / s^J)\,\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{D}^J \qquad r \in J \qquad \text{(IHP-J)}$$

Given these three forms of the factor content of absorption, we can pair these with production test II above to derive three tests of the factor content of trade of the regions of Japan. Recall that gross and net output are related by $\boldsymbol{X}^r = (\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\,\boldsymbol{Y}^r$, and that net trade is given as $\boldsymbol{T}^r = \boldsymbol{Y}^r - \boldsymbol{D}^r$. Taking the relevant differences, we find:

T-I: $\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\ [\boldsymbol{Y}^r - \boldsymbol{D}^r] = \boldsymbol{V}^r - s^r\boldsymbol{V}^W \qquad r \in J \qquad$ (IHP-W, FPE-W)

If FPE fails to hold for the world as a whole, then we cannot work with the above equation. If we believe that FPE still holds for the regions of Japan, the appropriate test is:

T-II: $\boldsymbol{B}^J(\boldsymbol{I}-\boldsymbol{A}^J)^{-1}\ [\boldsymbol{Y}^r-\boldsymbol{D}^r]=\boldsymbol{V}^r-s^r\boldsymbol{B}^J\boldsymbol{X}^W \qquad r\in J \qquad$ (IHP-W, FPE-J)

If, as well, we assume that identical homothetic preferences hold for Japan but not for the world as a whole, then the relevant test is:

T-III: $\boldsymbol{B}^J(\boldsymbol{I}-\boldsymbol{A}^J)^{-1}\ [\boldsymbol{Y}^r-\boldsymbol{D}^r]=\boldsymbol{V}^r-(s^r/s^J)\,\boldsymbol{B}^J(\boldsymbol{I}-\boldsymbol{A}^J)^{-1}\boldsymbol{D}^J \qquad r\in J$ (IHP-J, FPE-J)

The equations above are the basis for all tests in this paper. The theory establishes, for each country or region, a vector equality between what we term the *measured* (left hand side) and *predicted* (right hand side) net factor content of trade.[5] It is convenient to group these equations for the K regions (countries) into matrices for our tests. Thus gather the equations for measured and predicted factor content of production, which will be denoted:

$$\boldsymbol{M}_{\Pi}=\boldsymbol{P}_{\Pi} \quad \text{(Dimension } F\times K\text{)}$$

$$\left[\boldsymbol{B}^{fJ}\boldsymbol{X}^k\right]=\left[V^{fk}\right]$$

where there are F factors and $\boldsymbol{B}^{fJ}$ corresponds to the fth row of the $\boldsymbol{B}^J$ matrix. Similarly, we gather the equations for consumption of the K regions (countries) and N goods:

$$\boldsymbol{M}_D=\boldsymbol{P}_D \quad \text{(Dimension } N\times K\text{)}$$

$$\left[D^{nk}\right]=\left[s^kD^{nW}\right]$$

Finally, we gather the equations for the net factor content of trade of the K regions:

$$\boldsymbol{M}_T=\boldsymbol{P}_T \quad \text{(Dimension } F\times K\text{)}$$

$$\left[\boldsymbol{B}^{fJ}(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{T}^k\right]=\left[V^{fk}s^k\boldsymbol{B}^{fJ}(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{Y}^W\right]$$

$$=\left[V^{fk}-s^kV^{fW}\right] \qquad \text{Under FPE}$$

[5] Caution in interpreting the distinction between measured and predicted is warranted. Both are rooted in data. We think of the distinction largely as the orientation of our interest. For example, we think of endowments as predicting the net factor content of trade. Thus, data on endowments establishes the prediction, and data on trade and technology yields the measured factor content.

Having set out the theory for the variety of hypotheses that we will consider, we turn now to sketch the tests that we employ. In this, we follow the lead of the seminal work of Maskus (1985) and Bowen, Leamer and Sveikauskas (1987). As the nature of the tests for the various cases is very similar, we will outline the tests for the full trade theory, and then note amendments relevant to the cases of production and absorption. The theoretical trade equations (T-I to T-III above) establish an exact link between technology, output, absorption, and endowments. Of course, these exact relations are too much to hope for with real data. So we consider less exacting tests.

First, we will consider rank tests. If corresponding cells of the matrices are supposed to be identical, then one should expect that when comparing corresponding rows or columns, there should be high raw and rank correlations. For example, when examining the full theory of the net factor content of trade, one test compares corresponding rows, i.e. considering a single factor across regions:

$$\boldsymbol{M}_T^f = \boldsymbol{P}_T^f \qquad \text{(Dimension } 1 \times K\text{)}$$

The other compares corresponding columns, i.e. considering a single region across factors:

$$\boldsymbol{M}_T^k = \boldsymbol{P}_T^k \qquad \text{(Dimension } F \times 1\text{)}$$

A complement to the rank test, relevant only for testing the full trade theory, is the sign test. The idea is that if corresponding entries are supposed to be identical, one would hope that they would at least have the same sign. For a typical element:

$$\text{sign}\left\{\boldsymbol{B}^{fJ}(\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{T}^k\right\} = \text{sign}\left\{V^{fk} - s^k V^{fW}\right\}$$

A sign match implies that the country in fact is a net exporter or importer of the factors that theory so predicts. One can calculate the proportion of correct sign matches by factor (across countries), by country (across factors), or for the matrix as a whole. One can then test whether the theoretical model does better than a coin flip.

Since these tests do not specify a clear null hypothesis, they merely give us an indication of how consistent the data is with the theory. Even if the two sets of rank orderings matched perfectly, it still would be possible for the assumptions of the model to be violated in the real

world. In this case, our test would simply tell us that the real world deviations are not sufficient in magnitude to change the basic predictions of the theory regarding which regions are net exporters of which factors. It is in this spirit that we "test" the Heckscher-Ohlin model. When the model fits the data well, we conclude that relaxing the basic assumptions will not greatly enhance our understanding of the factor content of trade; when the model fits poorly, we conclude that there may be substantial gains from considering alternative specifications.

2.1 Previous Work: A String of Empirical Rejections

The seminal empirical critique of Heckscher-Ohlin is due to Leontief (1953).[6] He used data on input requirements and US trade to measure capital to labor ratios in US imports and exports separately. To universal surprise, widespread dismay, and scattered consternation, he showed that US imports were more capital intensive than US exports. This suggested that the US is relatively labor abundant – a result ever after known as the "Leontief Paradox." Leamer (1980) showed, however, that Leontief applied a conceptually inappropriate test of the Heckscher-Ohlin hypothesis. When he re-examined the same data in a conceptually correct way, the paradox vanished. Nonetheless, this paradox refused to perish. Brecher and Choudhri (1982) pointed out that one (counterfactual) implication of Leamer's approach is that US expenditure per worker would have to be lower than for the world as a whole. Stern and Maskus (1981) applied Leamer's (1980) approach to US data for both 1958 and 1972, finding the Leontief paradox held in the former but not in the latter year. Extensive surveys of previous work on Heckscher-Ohlin can be found in Deardorff (1984) and Leamer and Levinsohn (1995).

The theory of trade reflected in equations (1) through (3) above has previously been examined by Maskus (1985) and BLS (1987). Their results have severely undermined confidence in the robustness of the Heckscher-Ohlin framework. Maskus (1985) carried out both sign and rank tests on data for two time periods (1958 and 1972), and for three high quality factors

[6] It is worth noting that Leontief did not view his own work as contradicting the Heckscher-Ohlin model. Rather he saw it as evidence against the notion that the US was capital abundant. Later authors, however, came view his work as providing "the single biggest piece of evidence against the factor proportions theory." [Krugman and Obstfeld (1994), p. 77]

(professional, unskilled labor, and capital). He reports results only for the United States, perhaps because the Leontief paradox had focused on it. The sign test is correct for only one factor in 1958, but for all three in 1972. This might be seen to suggest that the Heckscher-Ohlin-Vanek relations fare well, at least in the latter period. However, the test lacks power. As Maskus notes, if we consider the alternative that the signs were determined randomly, we will have two or fewer sign failures out of six tries 34.4 percent of the time. Moreover, even if we limit ourselves to the 1972 data, under the same alternative, there will be no sign violations (as in his data) one in eight times. The results were, if anything, worse in the rank test. The direct measures of US factor abundance relative to the rest of the world were stable, with physical capital most abundant, professionals second, and unskilled labor least abundant. However, the trade-imputed measures of factor abundance in 1958 suggested the US was most abundant in unskilled labor, and least abundant in physical capital! The 1972 trade-imputed measure of factor abundance showed unskilled labor shifting dramatically to be least abundant, and reverses the relative abundance of physical capital and professionals. A repeat of the tests, restricted to OECD data, yielded no improvement. As Maskus notes, "paradoxical outcomes may be the rule rather than the exception."

BLS likewise report results widely viewed as undercutting Heckscher-Ohlin-Vanek. An important contribution was extending the test to a much broader set of countries (27) and factors (12). Thus, whereas the Maskus test was based on a matrix of only three cells for each time period, the BLS matrix had 324 entries. Because of the greater dimensionality of the matrix, it became possible to conduct sign and rank tests not only for a single country across factors (as in Maskus (1985)), but also for a single factor across countries. The sign test was correct more than half of the time for eleven of the twelve factors, but was correct over 70 percent of the time for only four in twelve. The sign matches were correct more than half the time for 18 of 27 countries, but over 70 percent of the time for only 8 of the 27. Only 61 percent of the total sign matches were correct. They note that independence between the signs of corresponding entries can be rejected at the 95 percent level for only one factor in twelve, and for only four of the 27 countries. In effect,

in determining which factors' services would on net be exported or imported, Heckscher-Ohlin did little better than a coin-flip.

The rank proposition fares no better. BLS report both rank correlations and the proportion of correct rankings when entries are compared two at a time. A zero correlation is rejected for only four of the 12 factors and eight of the 27 countries. Moreover, one factor and five countries have the wrong sign on the correlation. While the pairwise comparisons get over 50 percent correct rankings for 22 of the 27 countries, the same is true for only three of the 12 factors (all land variables). In sum, BLS note that the sign and rank propositions yield the Heckscher-Ohlin-Vanek model "relatively little support."

The pessimism regarding Heckscher-Ohlin was partly relieved by Trefler (1993), only to be revived by Trefler (1995). They represent alternative approaches to resolving the problems identified by BLS. The former follows up on Leontief's suggestion that the failure of Heckscher-Ohlin may be due to factor-based differences in efficiency. Trefler chooses the efficiency factors so that the Heckscher-Ohlin-Vanek equations fit exactly. He then shows that the implied productivity differentials correlate nicely with evidence on cross country differences in wages and rentals, suggesting a version of adjusted factor price equalization.

Trefler (1995) returned to the simple Heckscher-Ohlin-Vanek framework using the US $\boldsymbol{B}$ matrix, $\boldsymbol{B}^{US}$. In one exercise, he graphed the net factor trade residuals, $\boldsymbol{\varepsilon_T} = \boldsymbol{B}^{US}(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{T} - (\boldsymbol{V}^c - s^c\boldsymbol{V}^W)$, against the predicted net factor trade, $\boldsymbol{V}^c - s^c\boldsymbol{V}^W$. Theory would predict that these should be centered around the line $\boldsymbol{\varepsilon_T} = \boldsymbol{0}$. Instead they closely followed the line $\boldsymbol{\varepsilon_T} = -(\boldsymbol{V}^c - s^c\boldsymbol{V}^W)$, or equivalently, $\boldsymbol{B}^{US}(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{T} = \boldsymbol{0}$. This says that measured net factor trade is approximately zero, to which he applied the colorful moniker "the case of the missing trade." This is certainly a serious difficulty for an endowments-based model of trade. After considering a variety of alternatives, he identifies neutral technological differences and a home bias in consumption as the most likely culprits for the missing trade. However, by working with trade data – rather than production and absorption data – he can only test this indirectly. The technical differences are estimated rather than

observed. Thus, while these papers provide some glimmerings of hope for Heckscher-Ohlin, they also serve to underscore the serious empirical shortcomings of the simple model.

Recently, a few authors have studied Heckscher-Ohlin using United States regional data, as in Horiba (1992), and Grimes and Prime (1993).[7] The rationale for such studies is that the conventional assumptions of identical technologies and preferences, factor price equalization, and free trade are more likely to hold for regions of a single country than across different countries. However, an important flaw underlies both studies – the United States is treated as a closed economy, which is inherently anti-Heckscher-Ohlin. They assume that the factor content of regional consumption is proportional to the US – rather than the world – endowment. If this were so, then there would be no reason for the US to trade with the rest of the world. More importantly, if one ignores the rest of the world, the model does not have determinate predictions about the factor content of trade. The general model yields predictions for any region *vis à vis* the rest of the world, but not relative to any subset of regions or countries in the world. In addition, there are problems in interpreting the data. As Grimes and Prime note, their data treats US regional exports to the rest of the world as though they are consumed at the port of exit, wherever they may come from, and wherever they are destined. They also adopt a curious measure of some of their endowments, one closer to a measure of output than input. For example, the measure of "agricultural endowment" is the "value of agricultural products sold by farms in millions of dollars." This comes close to reducing the production side of the Heckscher-Ohlin model to a tautology, and so testing only homothetic preferences. The authors deserve credit for attempting to implement a regional Heckscher-Ohlin model. However, data problems and methodological flaws make interpretation of their results difficult.

2.2 Our Approach

Trade is simply the difference between production and consumption (absorption). Hence a theory of international trade must join two theories: a theory of the location of production and a

[7] An earlier set of papers sought to examine Heckscher-Ohlin with regional data, as in Horiba (1973). However, these primarily relied on the method of Leontief (1953), which is vulnerable to the criticisms of Leamer (1980).

theory of consumption. The theory of consumption in the conventional version is supplied by the assumption that preferences across countries are identical and homothetic. The real intellectual capital of the theory, however, is staked on the theory of the location of production (and its underlying assumptions). Three of the principal theorems – Rybczynski, Factor Price Equalization, and Stolper-Samuelson – make no use of the consumption theory. The fourth – the Heckscher-Ohlin theorem – is essentially a corollary to the Rybczynski theorem once we have added the assumption on preferences.

This suggests, paradoxically, that the most penetrating approach to investigating Heckscher-Ohlin may be to ignore trade data *per se*, and to focus instead on the two components of trade, particularly the location of production. We investigate this using Japanese regional data. This is justified as the underlying assumptions of the model are more likely to hold for regions of a country than for the full set of countries in the world. Moreover, in contrast to previous efforts at examining Heckscher-Ohlin in a regional framework, we embed this in a theoretically appropriate model of the full world general equilibrium.

3.0 Implementation

Leamer has frequently emphasized that the Heckscher-Ohlin-Vanek theory relates three elements: endowments, technology, and trade. We suggest that it is useful to extend this to four, by breaking the last element into its components: production and absorption. Ideally, a test of the theory would incorporate direct, appropriate and independent observations of the four elements. This is not an easy task.

Consider, first, the problem of technology. A maintained assumption of BLS is that all countries have access to the same technology. Thus one could, in principle, use the technical coefficients taken from any country. In practice, their assumption was that US technical coefficients were employed by all countries. We will consider three hypotheses. The first, following the standard model, holds that all countries use Japanese input coefficients. The second requires only that this be true for the regions of Japan that we include in our tests. The final

hypothesis that we consider, in our implementation of the consumption theory, is that even if not all countries share the same matrix of direct factor inputs, $\boldsymbol{B}^J$, they may yet share a common input-output matrix, $\boldsymbol{A}^J$. This is a necessary assumption to implement one of our consumption tests. Moreover, it may have some plausibility. For example, a car may be produced with varying degrees of substitution of capital for labor across countries. Yet the same car may still need a certain amount of steel, plastic, rubber and glass.

Consider, next, the problem of measuring endowments. One option is to choose a coarse partition of endowments into so-called "good" factors – like college-educated labor, non-college-educated labor, and capital – that are good in the sense that the factors have clear definitions and are conceptually close to the notion of endowment. Or one can choose a much finer partition into "bad" factors, for example, distinguishing labor categories such as professional/technical, clerical, sales, and service. The finer partition is appealing in that it provides additional data points against which to test the theory. However, the occupational categories associated with the bad factors may fail to be independent of the output measure. In addition, there is some danger of arbitrary disaggregation causing the Heckscher-Ohlin-Vanek theorem to fail. The former introduces a bias in favor of the theorem, the latter against. *Ex ante* it is difficult to know which will dominate. Maskus worked with the coarse partition into good factors, while BLS used the fine partition into bad factors. We prefer the coarse partition but also consider the alternate case.[8]

There are two other concerns regarding our measure of endowments. The first relates to the work of Trefler (1993). He showed that introducing neutral factor-based productivity differences in the model allowed a dramatic improvement in the model's ability to predict cross-national differences in factor returns. This means that efficiency units differ from measured units of factors, either due to differences in productivity or classification. This suggests, then, that a model using unadjusted measures of world endowments may fare poorly, for the same reasons. The second concerns the predictions of the theory when we discard the assumption that the rest of the world

[8] Lack of data on factors, unlike goods, is not a serious problem for the tests. One can imagine that we have all the data for our calculations, but are simply unable to inspect the rows for missing factors. This does not, though, disturb the relations for the other factors.

has the same technology and factor price equalization with the Japanese regions. In this case, as noted above, one wants to work with direct measures of endowments for the Japanese regions, but imputed factors for the rest of the world. The latter correspond to the factors that would have been necessary to produce the world output vector if production had taken place with the Japanese technical coefficients.

4.0 Results

4.1.1 International Production Evidence: Bad Results from Good Factors

We first test the Heckscher-Ohlin theory by examining its consistency with the location of world production (Equation P-I). We assume that all countries use the Japanese direct input requirements, $\boldsymbol{B}^J$. This implies that we meet the assumptions of identical CRS technology and factor price equalization or that these coefficients are technologically fixed. In effect, the test is to see whether $\boldsymbol{B}^J\boldsymbol{X}^c = \boldsymbol{V}^c$ for the various countries c. Our $\boldsymbol{X}^c$ vector for each country contained thirty sectors comprising total output, and the elements of $\boldsymbol{B}^J$ are the factor usages per unit output. For countries that only reported value added by sector, we used the Japanese ratio of total output to value added for the sector to estimate total output. This enabled us to assemble a sample of 35 countries, but limitations on the availability of education data forced us to reduce our sample to 21.[9]

Figure 1 plots the imputed number of people in each country that have not gone to college in a country ($\boldsymbol{B}^{fJ}\boldsymbol{X}^c$) against the actual number of people (V^{fc}). As the graph reveals, there is not a very close relationship between the levels of non-college graduates that each country should have, given its structure of production and the actual levels. In fact, the correlation coefficient between the two series is only 0.290. Figures 2 and 3 repeat this experiment for college graduates and for capital stocks. Over the full sample, the correlation coefficient for both of these factors is quite high: 0.962 for college graduates and 0.922 for capital. At first glance, it might appear that the theory seems to work quite well for these factors. However, much of the variance for both of these

[9] See the Data Appendix for details. Because $\boldsymbol{B}^J\boldsymbol{X} = \boldsymbol{V}^J$ for Japan by construction, we dropped Japan from our sample for international comparisons.

factors comes from the United States which has by far the largest number of college graduates and the largest capital stock. Dropping the US from the sample (Figure 2b and Figure 3b) reveals that for college educated workers the Heckscher-Ohlin model of production does not predict factor contents very well for most of the world (ρ = 0.271). Interestingly, the model seems to work much better for capital. Even after leaving the US out of the sample, the correlation coefficient between actual and imputed levels of capital only falls to 0.628.

Undoubtedly, one of the major reasons for the high correlations in the data is the fact that as long as the marginal product of each factor is positive there is likely to be a positive correlation between country output and factor abundance. Indeed, considering the enormous size differentials across our sample, one would need to have tremendous production distortions in order to have a negative correlation between factor abundance and production. These size effects can mask fairly large predictive errors. For example, despite the fact that inclusion of the US in the sample greatly increases the fit of the capital relationship, the point estimate for the US capital stock is off by a factor of two. One way to reduce the influence of extreme observations is to calculate rank correlations between the actual and predicted values. This method moves big outliers towards the mean but also causes points that are very close together to be further apart. Recalculating the correlation coefficients using country ranks reveals that, on average, the rank correlation between actual country endowments and imputed levels tends to be around 0.6 [see Table 1]. Whether this number seems high or low depends on one's priors. Setting a null that there should be no correlation between output and factor supply could be rejected in most cases, but little comfort can be obtained by rejecting such an absurd proposition. Indeed, considering that it would require a tremendous error to erroneously rank Singapore or Holland ahead of, say, India, we feel that the existence of modest correlations between the two matrices across countries lends little support for the Heckscher-Ohlin theory of production.

A better way to examine how well the model fits international data without having to worry about country size is to compare actual and imputed levels of factors within countries. Since it would be meaningless to compare the amount of capital in one country with the number of non-

college educated workers, we normalized each country's factor endowment by dividing by the world endowment of that factor (i.e. compare whether $(M^{1c}/P^{1W},\ldots, M^{Fc}/P^{FW}) = (P^{1c}/P^{1W},\ldots, P^{Fc}/P^{FW})$). Hence we compared the imputed share of each factor with the actual physical share. Table 2 presents the results from this exercise. Overall, in almost one half of the sample, the correlation between predicted and actual factor shares had the wrong sign. The average correlation was only 0.268, indicating that the Heckscher-Ohlin model, on average, only explained a small part of the variance. The rank correlations were slightly better but still hovered around 0.4 on average. It is interesting to note, however, that although the Japanese $\boldsymbol{B}^J$ matrix did not work well for every wealthy country, the set of countries with correlations above 0.7 (Canada, Finland, Germany, Ireland, Netherlands, New Zealand, Norway, and the US) are all countries with per capita incomes that are close to Japan. This suggests that these countries may produce within the same cone or that the quality of endowments does not vary significantly among these countries.

These results are entirely consistent with those presented in BLS, but provide us with different information. Since BLS were using trade data, they could not directly separate failures in the model due to implausible assumptions about production or consumption. Regression analysis let them infer that differences in technology likely played an important part in the explanation of the poor performance of the Heckscher-Ohlin model, but reliance on trade data made it difficult to test this directly. The fact that there are very low correlations between actual and predicted factor contents of production suggests that predictions regarding trade patterns based on a strict interpretation of the Heckscher-Ohlin model are likely to be at odds with the data for a broad sample of countries.

4.1.2 Regional Production Evidence: Good Results from Good and Bad Factors

The simple Heckscher-Ohlin theory, we have seen, does a poor job of describing the international location of production. However, as we saw in Section 2.0, even if FPE fails for the world, it is possible that it works within countries. In other words, equation P-II may hold even if P-I fails. In this section, we use regional Japanese data to demonstrate that in a case in which there

are minimal barriers to trade and factor mobility, the Heckscher-Ohlin model of production performs exceedingly well. [10]

It is important to realize that our tests which failed on international data could likewise fail on regional data. Such a failure could arise from a variety of influences. For example, it could arise from regional differences in employment or utilization rates. As well, anything that gave rise to different input coefficient matrices would lead to failure. This could include a failure of factor price equalization, technological differences, imperfect mobility of goods and factors, distortions in goods or factor markets, economies of scale, local external effects, etc. So the success of the regional test is far from a foregone conclusion.

Japanese regional data is well suited for this type of analysis because it is collected by the central government and therefore all categories are consistent across regions. Theoretically, we could have used data from all 48 Japanese prefectures, but in practice we were forced to aggregate many of these prefectures into 10 larger regions. This aggregation was necessary because some of the labor data was only reported for aggregated regions and also because it is very common for people who work in, say, Tokyo to live in an adjacent prefecture. These regions differ quite substantially in size. For example, Okinawa, the smallest region, has only one seventieth the number of college graduates and one fortieth the amount of capital as the largest region, Kanto.

We now re-examine the Heckscher-Ohlin theory of the location of production by restricting our sample to ten Japanese regions. That is, we check whether $\boldsymbol{B^J X^r = V^r}$. Even if the pattern of production in the rest of the world fails to satisfy these relations, they should continue to hold for the various Japanese regions. Figure's 4 through 6 relate actual and imputed levels of our three factors. In every case, imputed ranks match the actual ranks and the correlation coefficients exceed 0.99. In Table 3, we examine the relationship across prefectures and find that the theoretical

[10] There is one notable difference between the conventional international Heckscher-Ohlin model and the regional version. The conventional model assumes no cross-national factor mobility, while obviously there is at least some mobility across the regions of Japan. We believe this poses no fundamental problem for the regional model. It will remain true that the Heckscher-Ohlin relations must hold *ex post*. Alternatively, given that the Heckscher-Ohlin relations do hold, the actual regional distribution of factors belong to a factor price equalization set associated with an integrated equilibrium. Thus if immobility were now imposed, it would have no impact on the resulting equilibrium [cf. Helpman and Krugman (1985)].

predictions are close to the data. Furthermore, as Table 4 shows, the average correlation within regions and across factors between theoretical factor contents of production and actual endowments (as a percentage of national endowment levels) was 0.8890. The results are strikingly better with regional than international data.

Previous tests of the Heckscher-Ohlin theory have often used more than three factors of production despite the problems that arise from having factors that are closely related to industry categories. Despite the risk of tautologically true relationships between, for example, the number of agricultural workers and agricultural output, using more factors reduces the probability that our results are due to chance. While we would have liked to have broken capital into various components and used various types of land as factors, disaggregated capital stocks are not available at the regional level and land usage rates are not available for the service sectors. Detailed labor data were employed that broke up employment into nine categories – professional and technical workers, managers, office workers, sales workers, service workers, miners, transport and communication workers, production workers, security employees, and agriculture, forest and fishery workers. In practice the results for mining and miners were unreliable due to the treatment of oil inventories so this category was dropped from our sample.[11]

We repeat our test at the regional level, again checking that $\boldsymbol{B}^J\boldsymbol{X}^r = \boldsymbol{V}^r$, but now using the expanded set of ten factors. Tables 5 and 6 present the results from repeating our experiment using ten factors. Once again we found that the both the raw correlations and the rank correlations across regions always exceeded 0.95. The correlation across factors, however, was somewhat lower: the raw correlation averaged 0.77 for the whole sample. Interestingly, the average rank correlation was only 0.48. The difference between these two numbers is especially striking for Tohoku which had a raw correlation of 0.9 and a rank correlation of only 0.03. The reason why the rank correlation tended to be lower than the raw correlation is that the rankings tended to pull apart points that were

[11] The problem is that while oil imports are not counted as regional production, changes in inventories are counted as production. Since Japanese mining output is extremely small, changes in crude oil inventories tended to dominate regional mining output resulting in implausibly large or small output levels for certain regions. While this should technically affect all of our results, the fact that mining is such a small sector in all Japanese regions meant that errors in mining inventories had little effect on our overall results.

quite close together. For example, in Tohoku all of the actual factors except agriculture, forest, and fishery workers were within 20% of each other. While the raw correlation was high because the model correctly predicted an abundance of these workers, the rank correlation was low because relatively minor differences in factor supplies led to quite different rank orders. In general, the model seemed to be able to rank factors correctly when their relative abundance differed by more than 15%, but was not accurate for smaller differences.

These results compare quite favorably to those of BLS. In contrast to our rank correlation of 0.97, BLS found that the rank correlations across countries averaged only 0.21 and only exceeded 0.5 for one factor: arable land. Even our lower rank correlation of 0.49 across factors looks relatively good in comparison to BLS's result of 0.27. Indeed, in the BLS study only one quarter of the countries in their sample had a rank correlation across factors that exceeded 0.49 and five out of their 27 countries had negative rank correlations while none of ours did. To some degree, our results are not directly comparable to BLS's because those authors were examining trade rather than production data. Furthermore, our regional data are richer than BLS's in labor categories but do not include the various land categories available for countries as a whole.[12] The omission of land categories, however, probably understates the strength of our results. In fact, if BLS had not used land in their sample, their average rank correlation would have *fallen* from 0.21 to 0.12. The magnitudes of their correlations seem very much in line with our cross-country results and imply that the underlying assumption of identical techniques across a broad sample of countries is likely to be a major problem in applying the Heckscher-Ohlin model across countries.[13]

On the whole, we take our results to be an important validation of the Heckscher-Ohlin theory of the location of production across regions.

---

[12] We were unable to use land categories because we did not have information on land usage in the services sector for Japan.

[13] See also Trefler (1993).

4.2 Consumption

The assumption of homothetic preferences also plays a crucial role in the Heckscher-Ohlin-Vanek theory of trade, and it seems reasonable to ask how accurate this is. Hunter and Markusen (1989) and Hunter (1990) have shown, on international data, that there is good reason to believe that the consumption of goods is not homothetic. It stands to reason that poor countries are likely to spend greater shares of their income on food products than rich countries. In fact, Hunter (1990) has suggested that these non-homotheticities in consumption may explain as much as 25% of world trade. While the homotheticity assumption is likely to cause problems in international data because of tremendous income disparities across countries, in regional data one should expect these disparities to be less severe.

To test the assumption of homotheticity in Japanese regions, we used the *Family Income and Expenditure Survey* to collect data on regional consumption expenditures for 42 different product categories. Our first test of homotheticity was whether Japanese regional household consumption was proportional to national household consumption or whether equation C-II held (i.e. $\boldsymbol{D}^r = (s^r/s^J)\boldsymbol{D}^J$). In structuring our test this way, we implicitly allow two possible reasons why the theory might differ from the data. First, it may be the case that consumers in different regions have different preferences for different commodities, and second, there may be interregional differences in the prices of goods. In this sense, we are not testing homotheticity *per se*, but rather we are examining the joint assumption that prices do not significantly differ across regions and preferences are identical and homothetic.[14]

Tables 7 and 8 present the results from comparing predicted and actual regional consumption. Looking at the cross regional and cross factor results, the data seem to be consistent with the assumption of homotheticity. Only consumption of liquefied fuels seems to be an outlier. This is because Hokkaido, the northernmost region, consumes proportionally more fuel than the rest of the country. All of the remaining rank and raw correlation coefficients exceed 0.9,

[14] Alternatively, preferences may be non-homothetic if there is little to no variation in the distribution of income within regions.

indicating that one may reasonably assume that Japanese regions consume as if their constituent households had identical and homothetic preferences.[15]

In order for the international version of the Heckscher-Ohlin model to hold, it must be the case that Japanese final goods absorption should be proportional to world final goods absorption (i.e. equation C-I must hold).[16] Hunter and Markusen's (1989) work on international data suggests important deviations from this theory. While we do not pretend to replicate their work, it is interesting to examine how closely Japanese consumption follows world production. It is inappropriate to compare Japanese consumption with the sum of world production because a large share of output is used up as intermediate products in the production of other goods. Unfortunately, we do not have good input-output data for most countries in our sample and so it is impossible to know how much of each country's gross output is used as intermediates in other industries. In order to adjust world consumption by intermediate input demand, we multiplied world production by $(\boldsymbol{I}-\boldsymbol{A}^{J})$ (where $\boldsymbol{A}^{J}$ is the Japanese input-output matrix) to generate world net production. We then compared this number to Japanese household, investment, and government consumption. This is therefore a joint test of two assumptions: the assumption of homotheticity and the assumption that intermediate input usage (not including primary factors) is identical across the world.

Figure 7 plots Japanese final goods absorption against the imputed level derived from world production and the Japanese input-output matrix. The raw correlation between these two vectors is 0.80 and the rank correlation is 0.82. Looking at the graph reveals that one of the largest outliers is "other services" which is heavily consumed in Japan but not very heavily consumed in the world. Most probably, this represents a classification error resulting from a discrepancy between Japanese and international categories. In fact, if "other services" is dropped, the raw correlation jumps to 0.89. Interestingly, the homotheticity assumption tends to work even better

---

[15] Most likely these results arise from the closeness of per capita income within Japan. Indeed, Japan has one of the lowest Gini coefficients in the world.

[16] In all subsequent tests and tables we used absorption data rather than household consumption data. The absorption data also includes government consumption, investment, and business consumption. See the data appendix for details. Final goods absorption is total absorption less intermediate input absorption.

for tradables than for non-tradables. Figure 8 presents the results for just the tradable goods sectors and here the correlation between the two series is 0.90.

These results suggest that the assumption that consumption of tradables is homothetic is not likely to cause big problems reconciling Japanese consumption and world production. Furthermore, it suggests that while there are very large problems associated with assuming that primary factors enter into international production functions identically across countries, these problems are not particularly severe for intermediate products, possibly because the latter are often tradable, relatively homogeneous in quality, and less substitutable.

4.3 Putting Production and Consumption Together: Trade

4.3.1 Regional Trade Under Strict HOV Assumptions

We now turn to examine the full Heckscher-Ohlin-Vanek theory of the net factor content of trade. We do not apply this test to the full complement of countries. Given our negative results on the theory's ability to account for the location of production, any positive results would have to be spurious. We do, though, begin with a test predicated on the assumption that all countries use the Japanese input coefficient matrices, $\boldsymbol{A^J}$ and $\boldsymbol{B^J}$. This corresponds to testing the strict form of HOV as expressed by equation T-I. The reason for implementing this is to demonstrate that gains in explanatory power are not tied solely to a shift from international to regional comparisons, but rely also on our assumptions concerning cross-country comparability of technology.

Our test, then, is strict in assuming that all countries use the Japanese input-output matrices, $\boldsymbol{A^J}$ and $\boldsymbol{B^J}$. Our tests focus on the ten regions of Japan, examining the relation $\boldsymbol{B^J\,(I - A^J)^{-1}\,[Y^r - D^r] = V^r - s^rV^W}$. Ultimately, we want to see how well the theories of production and consumption can be integrated into a coherent theory of trade. Our first test was to examine how the strictest version of the Heckscher-Ohlin model fit the data using actual world endowment data. The top panel of Table 9 and the first column of Table 10 present the results of this comparison. As one might have guessed from our previous results trying to map production into factor endowments, the overall fit of the regional data into the strict version of the Heckscher-

Ohlin model is poor. Overall, out of 30 possible sign matches, we were only able to correctly predict factor flows for 19. This cannot be statistically distinguished even at the 10% level from the expected number of sign matches that one could have gotten by simply flipping a coin to determine relative factor abundances.[17] In terms of rank correlations, while the strict version of the model did work well for college educated workers, it performed considerably worse for non-college graduates and capital. Looking at the results across factors yielded similar results. The rank correlations averaged only 0.26 across all prefectures and only 0.35 across factors. These results are comparable to those of BLS, who also found very low rank correlations in their trade tests. While BLS found that cross-national factor flows could not be predicted by a strict version of the Heckscher-Ohlin model, we demonstrate that this problem extends to regional total net trade patterns as well.

We can also investigate whether the problem that Trefler (1995) identified on international data as the "case of the missing trade" also appears in the Japanese data. We begin with summary measures for Japan as a whole. If we add across regions, $\boldsymbol{V}^J - s^J\boldsymbol{V}^W$ is the predicted net factor content of trade for Japan as a whole. Scaling this by national endowments of the respective factors, we predict that Japan's net imports of non-college factor services will equal 31% of the national endowment. Predicted net exports of the services of Japan's college and capital endowments equal 37% and 40% of the national endowment respectively. In short, the model predicts that a huge proportion of national endowments will on net be imported or exported. However, when we turn to the measured net factor content of trade, $\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{T}^J$, the picture looks very different. Again scaling by the national endowment, Japan is a net exporter of 3.6% of the non-college endowment, 7.4% of the college endowment, and 2.3% of the non-college endowment. That is, the measured factor content of trade is an order of magnitude smaller than that predicted. At the aggregate level, the problem of missing net factor trade is very much in evidence.

[17] The significance levels were based on a binomial distribution that was conditioned on the assumption that the expected number of negative signs equaled the proportion of negatives in each vector. This conditioning was necessary because the chances of getting matches rises if the underlying statistical process tends to generate more of one sign than the other.

We can also look at this in the regional context. Following Trefler, we graph the net factor trade residuals, $\varepsilon_T = \boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{T}^r - (\boldsymbol{V}^r - s^r\boldsymbol{V}^W)$, against the predicted net factor trade, $\boldsymbol{V}^r - s^r\boldsymbol{V}^W$. Theory tells us that this should be a horizontal line at zero. Instead, as we see in Figure 9, this is very close to having a slope of minus one, indicating that measured net factor trade, $\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{T}^r$, is small relative to predicted net factor trade, $\boldsymbol{V}^r - s^r\boldsymbol{V}^W$. Thus, we also encounter the "missing trade" in the regional data. However, the fact that we are working separately with production and absorption allows us to probe more deeply into the causes of the missing trade. It is convenient to define separate residuals for production, $\varepsilon_\pi = \boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{Y}^r - \boldsymbol{V}^r$, and for absorption, $\varepsilon_D = -[\boldsymbol{B}^J(\boldsymbol{I} - \boldsymbol{A}^J)^{-1}\boldsymbol{D}^r - s^r\boldsymbol{V}^W]$. Since $\varepsilon_T = \varepsilon_\pi + \varepsilon_D$, a graph of the three sets of residuals against $\boldsymbol{V}^r - s^r\boldsymbol{V}^W$ vertically decomposes the net factor trade errors into the component parts. This is depicted in Figure 10. As is evident from the plot, the production residuals ε_π very closely follow the theoretical prediction of a horizontal line at zero. We take this as an important confirmation of the Heckscher-Ohlin model of the location of production.

The plot in Figure 10 further reveals that essentially the entire problem of missing trade arises from the absorption errors, ε_D. Does this evidence some failure in the consumption theory? Although the failure turns up in the factor content of absorption, the answer is no. Recall that the Heckscher-Ohlin-Vanek theory of consumption requires two stages. The first assumes identical homothetic preferences, and equal goods prices everywhere. This yields an equation in terms of goods absorption that we saw in Figures 7 and 8 works quite nicely. The second step requires that the goods demands be translated to the implicit demand for factors. This requires universal factor price equalization to insure $\boldsymbol{V}^W$ can be substituted for $\boldsymbol{B}^J\boldsymbol{X}^W$, so that the implicit absorption of factors will be $s\boldsymbol{V}^W$. Yet we have seen in Tables 1 and 2 that this simple model of international production does not work at all well.

Significantly, the errors in the factor content of absorption, ε_D, are systematic. All ten errors for non-college endowment are negative; nine of ten errors on absorption of college endowment are positive; and all ten errors on absorption of capital endowment are positive. One might expect this pattern to arise if $\boldsymbol{V}^W$ systematically overestimates effective world supplies of

non-college endowment, and underestimates the world's college and capital endowment (all measured in terms of Japanese productivity). In fact, this is what a direct comparison of V^W and BX^W reveals. The former is 40% larger for non-college endowment, and 37% and 41% smaller for college and capital endowment respectively. This suggests the value of pursuing a version of our model which does not require universal factor price equalization, a task to which we now turn.

4.3.2 Regional Trade Without Universal Factor Price Equalization

The failure of the strict version of the Heckscher-Ohlin model leads us to consider what types of relaxations would make the model fit better. Clearly, simply using regional data is not enough of a departure from earlier studies to generate an improvement in the fit of the HOV equations. Fortunately, the previous analysis pinpointed the locus of a major failure of the model: $B^J X^W \neq V^W$. Theory tells us that if the condition of world factor price equalization has been violated, then the appropriate trade test is T-II, not T-I. In other words, we will continue to require that the B^J and A^J matrices are identical for the various regions of Japan, but no longer require that the B^J matrix is identical in other countries of the world. In constructing the theoretical Japanese factor content of consumption, this requires that we drop information on the actual world endowments, V^W, and instead use the imputed endowments, $B^J X^W$. We continue to assume that Japanese absorption is proportional to world production. Thus for each of the Japanese regions, we ask whether $B^J(I - A^J)^{-1}[Y^r - D^r] = V^r - s^r B^J X^W$. The second panel of Table 9 and the second column of Table 10 display the results from this exercise. Once the assumption of identical techniques is relaxed, our results improve dramatically, with the average rank correlation across factors and regions jumping to 0.75 or higher. Here, the sign test reveals that the Heckscher-Ohlin model is correct in 23 out of 30 cases, which is significant at the 1% level. Relaxing the assumption about identical techniques makes the Heckscher-Ohlin model a very good predictor of net factor trade flows.

We return now to Trefler's (1995) problem of "missing trade." We continue to assume that factor price equalization holds for the Japanese regions, but not necessarily for other countries.

Again we begin with summary measures for Japan as a whole. Recall that the measured factor content of trade as a share of Japanese national endowment is 3.6%, 7.4% and 2.3% for non-college, college, and capital factors respectively. This was an order of magnitude smaller than our earlier predictions. Now we repeat the experiment, using $\boldsymbol{V^r - s^r B^J X^W}$ as the Japanese predicted factor content of trade, and again scaling each factor by the national endowment. The new predictions for Japan as a whole are 5.4%, -0.2%, and -1.5% for non-college, college, and capital respectively. That is, the predictions are of the right order of magnitude to fit measured net factor trade. Moreover, the change in the model has shifted us from a prediction of huge net factor trade flows (an absolute average of 36% of the national endowment of the three factors) to much more modest flows (an absolute average under 2.4% of national endowments). In such a case, a finding that measured net factor trade is very small is no longer a puzzle.

The regional data confirms this basic result. Recall that the only alteration that we have made is in the factor content of absorption, which shifts from $s^r V^W$ to $s^r B^J X^W$. As a result, the production residuals discussed above, ε_π, have not changed. The errors in net factor trade, ε_T, change only due to changes in the absorption residual, which is now $\varepsilon_D = -[B^J(I - A^J)^{-1}D^r - s^r B^J X^W]$. The magnitude of the errors is sharply reduced. The median error declines by more than two-thirds. Fully two-thirds of the errors are cut by more than half. An example is the case of Kanto (which includes Tokyo). The error on non-college endowment is only 15% as large; on college endowment, only 12% as large; and on capital endowment, only 3% as large. In sum, the fit is greatly improved.

A similar story emerges when we look at the corresponding errors in the factor content of trade, ε_T. As Figure 11 reveals, there continues to be a negative relation between the errors in net factor trade and predicted levels. However, this relation is greatly diminished under the new hypothesis. However, the big story is the dramatic decline in the magnitude of the errors. Of thirty errors, twenty are cut by half or more, while only three rise by a factor of two or more. Again, we can consider the case of Kanto. The trade error for non-college labor is only 15% as large; for

college endowment, only 26% as large; and for capital, only 10% as large. Again, the fit for the net factor content of trade is greatly improved.

One question that naturally arises at this point is whether our positive results are just the product of using regional rather than national data. Recall from section 4.3.1 that the HOV hypothesis failed just as miserably with regional data as it had in earlier studies with national data. In fact, the pattern of failure was very similar, likewise exhibiting Trefler's "missing trade." This shows that a shift to regional data does not insure success. A related question is whether the results on the regions should be looked on as distinct observations, rather than really as one national observation iterated ten times. This concern would be particularly apt if the regions were in effect scaled down versions of the national economy. However, these concerns are misplaced. First, the regional results are based on a rich data base (see Appendix) that reflects genuine regional variation in production patterns, endowments, and absorption. Second, the regions do not just look like scaled down versions of the national economy. One window on this problem is to ask whether there is typically a strong correlation between the factors which are on net exported by the regions and by the nation. Since the nation is the aggregate of the regions, one may expect some positive correlation. But if the regional observations are to count as distinct observations, one would not want this correlation to be too strong. How strong is it? Recall that we can make these comparisons for either measured or predicted factor contents of trade, and that we can do this for ten regions and three factors. For measured net factor content of trade, the sign at the regional level was the same as the national level in only 15 out of 30 cells. For predicted net factor content, the sign at the regional level matched that at the national level only 20 of 30 times. Given that the nation is the aggregate of the regions, and thus that we might expect a positive correlation, the surprising fact is how independent the signs at the regional level are from the national aggregate. That these regional variations reflect more than just noise is confirmed in our sign tests on regional net factor trade, which reject at the 99% level the independence of the signs of measured and predicted net factor trade. In sum, we believe that these observations validate the use of the regional approach to testing the HOV theory.

4.3.3 Regional Trade With Japan-Specific Demand

In the final test, we continue to focus on Japanese regions, but drop the assumption that Japanese regional absorption is proportional to world production, adopting instead the stronger assumption that it is proportional to Japanese national absorption. This corresponds to testing equation T-III. In the last panel of Table 9 and the last column of Table 10, we present results from estimating the model for the case in which $\boldsymbol{B^J(I - A^J)^{-1}\,[Y^r - C^r]} = \boldsymbol{V^r - (s^r/s^J)B^J(I - A^J)^{-1}}$ $\boldsymbol{D^r}$. What is striking about this adjustment is that there is little or no improvement in the overall fit of the model. Just as in the case where the identical technology assumption had been relaxed, we get 23 sign matches and the rank correlations are, if anything, a bit lower. In fact every element of the net factor content matrix calculated with world production data has the same sign as the corresponding element in the matrix calculated with consumption data. This suggests that most of the failure of the Heckscher-Ohlin model stems from its assumption about identical technologies while its assumption about homothetic consumption, at least as far as Japanese data are concerned, seems to be reasonably valid.

The failure to find an improvement in prediction with the Japan-specific demand contains important information regarding a result of Trefler (1995). His preferred specification for a revised HOV model includes a home bias in demand. This can be written as $\boldsymbol{D^r = (s^r/s^J)[(1 - \lambda)s^J Y^W + \lambda Y^J\,]}$. The standard HOV model has $\lambda = 0$, while a pure home bias model sets $\lambda = 1$. As the degree of home bias is unobservable, it must be estimated. Suppose, then, that we have selected λ by whatever criterion we find appropriate. Within the context of this model, the most favorable result possible is that the term in brackets, $[(1 - \lambda)s^J Y^W + \lambda Y^J\,]$ will exactly equal $\boldsymbol{D^J}$, so that our predicted consumption becomes $\boldsymbol{D^r = (s^r/s^J)D^J}$. However, this is the model that we have just seen yields no improvement over the standard HOV consumption model. If under the best possible circumstances we find no improvement, then we can conclude that allowing for home bias in demand does not contribute to an explanation of net factor trade patterns, at least for the regions of Japan.

We believe this is an important advance. The mystery of "missing trade" is resolved here by use of data and theory alone, without a need to estimate technological or demand diffences that allow the theory to fit the data. We are able to directly examine the source of the problem, as we use separate data for the theories of production and absorption. And we arrive at a substantively distinct result from that of Trefler. In addition to international technology differences, he suggested that a home bias in consumption may be required to eliminate the missing trade. We have shown that it is not necessary to make any changes to the demand model to greatly improve the fit.

5.0 Conclusion

The Heckscher-Ohlin-Vanek model has been central to the analysis of international trade. It provides the intellectual framework for a remarkable array of literature – theoretical, applied and policy. As a result, the empirical failure of the HOV model in prominent studies has been profoundly disappointing. Despair, though, is unwarranted. Very likely the empirical failure owes importantly to examining the theory in its most general – and least realistic – form.

Recent work, notably by Trefler (1993), has sought to identify simple modifications of this general model that fit the data well. Our work is in the same spirit. We make two important modifications. The first is to abandon the notion that the technology implemented is identical everywhere. This leads us to focus on the implications of the model for regions of Japan for which the assumption of identical technologies is more plausible. In contrast to other regional models of Heckscher-Ohlin trade, we do this while continuing to embed our model in a theoretically appropriate way within the full world equilibrium. The second modification is to ignore trade data as such, and instead focus on the components – production and absorption. This allows us to test separately the Heckscher-Ohlin theories of the location of production and pattern of consumption. By examining them separately, we know the source of any deviations that arise when we assemble the pair to a full test of the trade theory.

We first consider the Heckscher-Ohlin theory of the pattern of production under the assumption that all countries in the world utilize the same input coefficients. Our results find little

support for this version of Heckscher-Ohlin, confirming earlier studies. The results improve dramatically, though, under the more modest assumption that all Japanese regions share a common set of input coefficients. In this formulation, Heckscher-Ohlin is quite successful in accounting for the location of production according to the standard tests. We also show that this result is robust to altering the degree of aggregation in the definition of factors.

We then turn to the Heckscher-Ohlin theory of the pattern of consumption. We examine this first by considering Japanese regional absorption, which the theory suggests should be proportional to world net output. To implement this test requires an ability to move from gross to net output figures, hence an assumption that the Japanese input-output matrix is common to the world. The Heckscher-Ohlin model of proportional absorption does surprisingly well under this assumption. We next consider the milder hypothesis that Japanese regional consumption is proportional to national consumption, a result strongly supported in our model. In all, the Heckscher-Ohlin theory of consumption stands up remarkably well as a simple description of the data, at least for the regions of Japan.

We then assemble this information for a full test of the Heckscher-Ohlin theory of the net factor content of trade. Our earlier results showed that the theory could not account for the international pattern of production. Hence no point is served by looking at the implied net factor content of trade of the various countries, as positive results would have to be spurious. Instead, we focus on accounting for the net factor trade of the Japanese regions. Three approaches are developed. The first establishes a benchmark. It uses data on actual world factor endowments, implicitly assuming again that all countries use the same input coefficients. We show that the model performs poorly and replicate Trefler's (1995) "mystery of the the missing trade." In the next two cases, we examine this using the endowments imputed to the world, given their measured output, as if they had used the Japanese input coefficients. In the first of these, we assume that Japanese regional absorption is proportional to world net output. This model is a marked improvement over that based on measured world endowments, and largely eliminates the missing trade. In the last case, we consider this under the assumption that Japanese regional consumption is proportional to

national consumption. This yields no further improvement, demonstrating that changing the consumption model is not necessary to better the model's fit.

An important question is whether the results we have obtained using regional data have important implications for the relevance of Heckscher-Ohlin-Vanek to the international economy. We believe that they do. First, note that, although we have principally been concerned with the predictions for regional data, we have consistently applied models that encompass the full world general equilibrium. Second, we have replicated in the regional framework the most serious problems that have afflicted the Heckscher-Ohlin-Vanek model when applied in its simple form to international data. Third, once we have made alterations in our specification that are both sensible and consistent with the spirit of the model, we have had excellent results. We believe that the approach that we have developed in the regional framework holds great promise when appropriately applied to cross-country data.

In sum, we find that the Heckscher-Ohlin model under the conventional restrictive assumptions is a poor predictor of the international pattern of production, hence of net factor trade. However, this changes markedly when applied to predictions for regions of Japan. Given the long string of empirical failures of Heckscher-Ohlin, it is surprisingly successful as a theory of the location of production and the pattern of consumption – hence the net factor content of trade – of these regions.

6.0 Bibliography

Bowen, Harry P.; Leamer, Edward E.; and Sveikauskas, Leo; "Multifactor, Multicountry Tests of the Factor Abundance Theory," *American Economic Review*, December 1987, 77, 791-809.

Brecher, Richard A., and Choudhri, Ehsan U., "The Leontief Paradox: Continued," *Journal of Political Economy*, August 1982, 90, 820-823.

Brecher, Richard A. and Choudhri, Ehsan U. (1988), "The Factor Content of Consumption in Canada and the United States: A Two-Country Test of the Heckscher-Ohlin-Vanek Model," in Feenstra, Robert C., *Empirical Methods for International Trade*, Cambridge: MIT Press.

Courant, Paul N. and Deardorff, Alan V., "International Trade with Lumpy Countries," *Journal of Political Economy*, February 1992, 100, 198-210.

Davis, Donald (1995), "Intraindustry Trade: A Heckscher-Ohlin-Ricardo Approach," *Journal of International Economics,* 39: 201-26.

Deardorff, Alan V., "Testing Trade Theories and Predicting Trade Flows," in Ronald W. Jones and Peter B. Kenen eds. *Handbook of International Economics*, Vol. I, New York, North Holland, 1984.

Dixit, Avinash and Norman, Victor, *Theory of International Trade: A Dual, General Equilibrium Approach*. Cambridge: Cambridge University Press, 1980.

Grimes, Donald R. and Prime, Penelope B., "A Regional Multifactor Test of the Heckscher-Ohlin-Vanek Theorem," mimeo, University of Michigan and Kenesaw State College, 1993.

Grossman, Gene and Helpman, Elhanan., *Innovation and Growth in the Global Economy*. Cambridge: MIT Press, 1992.

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

Helpman, Elhanan, and Krugman, Paul R., *Market Structure and Foreign Trade*, Cambridge: MIT PR., 1985.

Horiba, Yutaka, "On the Empirical Relevance of the Heckscher-Ohlin Trade Model," mimeo, Tulane University, 1992.

Horiba, Yutaka, "Factor Proportions and the Structure of Inter-Regional Trade: The Case of Japan," *Southern Economic Journal*, 1973.

Hunter, L. and J. Markusen, (1989) "Per-Capita Income as a Determinant of Trade," in Feenstra, *Empirical Methods for International Trade*, 89-109.

Hunter, L. (1991) "The Contribution of Nonhomothetic Preferences to Trade," *Journal of International Economics*, May , 345-358.

Krugman, Paul, "Increasing Returns and Economic Geography," *Journal of Political Economy*, (1991).

Krugman, Paul and Maurice Obstfeld (1994), *International Economics: Theory and Policy*, Third Edition, New York: Harper-Collins.

Lawrence, Robert Z. and Slaughter, Matthew J., "Trade and US Wages: Great Sucking Sound or Small Hiccup?" *Brookings Papers on Economic Activity: Microeconomics*, 1994.

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

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

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

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

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

Trefler, Daniel, "International Factor Price Differences: Leontief Was Right!" *Journal of Political Economy*, December 1993, 101, 961-987.

Trefler, Daniel, (1995) "The Case of the Missing Trade and Other Mysteries," *The American Economic Review*, 85: 1029-46.

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

Table 1

INTERNATIONAL PRODUCTION TESTS

Correlations between imputed and actual endowments across countries

	Raw	Rank
Non-College		
Full Sample	0.2897	0.6338
without U.S.	**0.2579**	**0.5805**
College		
Full Sample	0.9622	0.6688
without U.S.	**0.2709**	**0.6165**
Capital		
Full Sample	0.9219	0.6282
without U.S.	**0.6161**	**0.5456**

Table 2

INTERNATIONAL PRODUCTION TESTS

Correlations Across Factors

	Raw	Rank
Argentina	-0.4673	-0.5
Austria	-0.0119	0.5
Canada	0.9931	1
Finland	0.994	1
Germany	0.63385	1
India	-0.3254	0.5
Indonesia	-0.1148	0.5
Ireland	0.5763	0.5
Israel	0.31895	0.5
Italy	-0.139	0.5
Mexico	0.17019	0.5
Netherlands	0.79682	0.5
New Zealand	0.73001	0.5
Norway	0.9402	1
Philippines	0.39929	0.5
Portugal	-0.2647	-0.5
Singapore	-0.3198	-0.5
Spain	-0.7568	-0.5
Thailand	-0.4546	-0.5
Turkey	0.39704	0.5
U.S.A.	0.79585	0.5
Average	0.2678	0.3864

Table 3

REGIONAL PRODUCTION TESTS

Correlations between imputed and actual endowments across prefectures ("Good" Factors)

	Raw	Rank
Non-College	0.9967	1
College	0.9951	1
Capital	0.9989	1
Average	0.9969	1

Table 4

REGIONAL PRODUCTION TESTS

Correlations between imputed and actual endowments across factors ("Good" Factors)

	Raw	Rank
Hokkaido	0.9900	1.0
Tohoku	0.8612	0.5
Kanto	0.9186	0.5
Hokuriku	0.8992	1.0
Tokai	0.9745	1.0
Kinki	0.9588	1.0
Chugoku	0.7297	0.5
Shikoku	0.8816	0.5
Kyushu	0.9032	0.5
Okinawa	0.7734	1.0
Average	0.8890	0.7500

REGIONAL PRODUCTION TESTS

Table 5

Correlations between imputed and actual endowments across prefectures ("Bad" Factors)

	Raw	Rank
Professional/Technical	0.9978	0.9879
Managerial	0.9932	0.9879
Office	0.9992	0.9879
Sales	0.9971	0.9879
Ag/For/Fish	0.9749	0.9515
Transport/Communication	0.9894	0.9758
Production/Blue Collar	0.9964	0.9879
Security	0.9971	1
Service	0.9942	1
Capital	0.9989	1
Average	0.9938	0.9867

Table 6

Correlations between imputed and actual endowments across factors

	Raw	Rank
Hokkaido	0.6466	0.6606
Tohoku	0.9088	0.0303
Kanto	0.7929	0.4182
Hokuriku	0.6102	0.2606
Tokai	0.8430	0.5636
Kinki	0.9141	0.4424
Chugoku	0.4207	0.4667
Shikoku	0.8500	0.3333
Kyushu	0.9551	0.7333
Okinawa	0.7855	0.8667
Average	0.7727	0.4776

Table 7

REGIONAL CONSUMPTION TESTS

Correlations between actual and predicted consumption

	Raw	Rank
Rice, bread, and other cereals	0.9974	1.0000
Meat	0.9811	0.9394
Fish	0.9922	0.9879
Milk, cheese, and eggs	0.9976	1.0000
Oil and fat	0.9952	0.9879
Fruits, vegetables, potatoes, other tubers	0.9992	1.0000
Sugar, other foods	0.9989	0.9879
Coffee, tea, cocoa	0.9984	1.0000
Soft drinks	0.9970	1.0000
Alcoholic drinks	0.9908	0.9758
Cigarettes, tobacco	0.9949	0.9879
Clothing	0.9997	1.0000
Repairs to clothing, footwear; household services	0.9989	1.0000
Footwear	0.9997	1.0000
Rents	0.9877	0.9273
Water charges	0.9993	1.0000
Electricity	0.9950	0.9515
Gas	0.9972	0.9879
Liquefied fuels, other fuels	0.5993	0.5394
Furniture, fixtures, carpets, household appliances	0.9955	0.9879
Repair of furniture, fixtures, carpets, etc.	0.9903	0.9879
Household textiles and textile furnishings	0.9963	0.9879
Glassware, tableware, utensils	0.9983	0.9758
Nondurable household items	0.9992	1.0000
Domestic services	0.9728	0.9515
Pharmaceutical products, medical supplies, etc.	0.9977	0.9879
Medical and health care services	0.9984	1.0000
Transportation equipment	0.9839	0.9879
Tires, tubes , parts, motor fuels, oils	0.9858	0.9879
Purchases transport services	0.9958	0.9758
Postage, telephone, telegraph	0.9981	0.9879
Radios, TVs, cameras; parts and repairs.	0.9973	0.9879
Other semi-durable recreational goods	0.9982	0.9758
Services of recreation, culture;	0.9987	0.9879
Books, magazines, newspapers	0.9993	1.0000
Educational fees	0.9976	0.9879
Services of hairdressers, etc.	0.9994	1.0000
Toilet articles and cosmetics	0.9938	0.9879
Jewelry, watches; writing equipment and supplies	0.9916	0.9758
Restaurants	0.9980	1.0000
Hotel and lodging services	0.9852	0.9636
Travel tours	0.9910	0.9515
Average	0.9852	0.9740

Table 8

REGIONAL CONSUMPTION TESTS

Correlation between actual and predicted consumption across commodities

	Raw	Rank
Hokkaido	0.9809	0.9300
Tohoku	0.9710	0.9562
Kanto	0.9940	0.9929
Hokuriku	0.9727	0.9548
Tokai	0.9929	0.9937
Kinki	0.9959	0.9885
Chugoku	0.9901	0.9831
Shikoku	0.9844	0.9846
Kyushu	0.9892	0.9921
Okinawa	0.9474	0.8953
Average	0.9819	0.9671

Table 9

TRADE TESTS

Strict HO Production Model

	Raw Correlations Across Regions	Rank Correlation Across Regions	Percent of Correct Pairwise Comparisons
Non-College Graduates	-0.4048	0.1394	0.4889
College Grads	0.9840	0.8788	0.8667
Capital	0.2019	0.3818	0.6667
Average	0.2604	0.4667	0.6740

Number of correct Sign Matches: 19/30

Consumption is proportional to World Net Production

	Raw Correlations Across Regions	Rank Correlation Across Regions	Percent of Correct Pairwise Comparisons
Non-College Graduates	0.6040	0.7697	0.8444
College Grads	0.9810	0.9030	0.8889
Capital	0.7994	0.5758	0.7333
Average	0.7948	0.7495	0.8222

Number of correct Sign Matches: 23/30

Consumption is proportional to Japanese Consumption

	Raw Correlations Across Regions	Rank Correlation Across Regions	Percent of Correct Pairwise Comparisons
Non-College Graduates	0.4598	0.5030	0.7111
College Grads	0.9873	0.9273	0.9111
Capital	0.8185	0.6121	0.7778
Average	0.7552	0.6808	0.8000

Number of correct Sign Matches: 23/30

Table 10

TRADE TESTS

Raw Correlations Across Factors*

	Assumption		
	Strict HO Production Model	Consumption is proportional to World Net Production	Consumption is proportional to Japanese Consumption
Hokkaido	0.1443	0.9995	0.9938
Tohoku	0.5993	0.8692	0.8927
Kanto	0.6634	0.9992	0.9899
Hokuriku	0.7094	0.9373	0.9707
Tokai	0.9996	0.2909	0.6624
Kinki	0.4966	1.0000	0.9790
Chugoku	0.8064	0.2797	0.4200
Shikoku	-0.0084	0.9201	0.9407
Kyushu	-0.0365	0.9815	0.9946
Okinawa	-0.8628	0.8978	0.9060
Average	0.3511	0.8175	0.8750

Rank Correlations Across Factors

	Assumption		
	Strict HO Production Model	Consumption is proportional to World Net Production	Consumption is proportional to Japanese Consumption
Hokkaido	-0.5	1.0	1.0
Tohoku	0.5	0.5	0.5
Kanto	0.5	1.0	0.5
Hokuriku	0.5	1.0	1.0
Tokai	1.0	0.5	0.5
Kinki	0.5	1.0	0.5
Chugoku	0.5	0.5	0.5
Shikoku	0.5	0.5	0.5
Kyushu	-0.5	1.0	1.0
Okinawa	-1.0	0.5	0.5
Average	0.2	0.8	0.7

*We do not report pairwise comparisons because with three numbers, the information is the same as in the rank correlation. A rank correlation of 1 implies 3 out of 3 comparisons were correct; A rank correlation of 0.5 means 2 out of 3 were correct; a rank correlation of -0.5 means only 1 was correct; and a rank correlation of -1 means none were correct.

Figure 1

INTERNATIONAL PRODUCTION TESTS:

ACTUAL VS. IMPUTED NON-COLLEGE ENDOWMENT

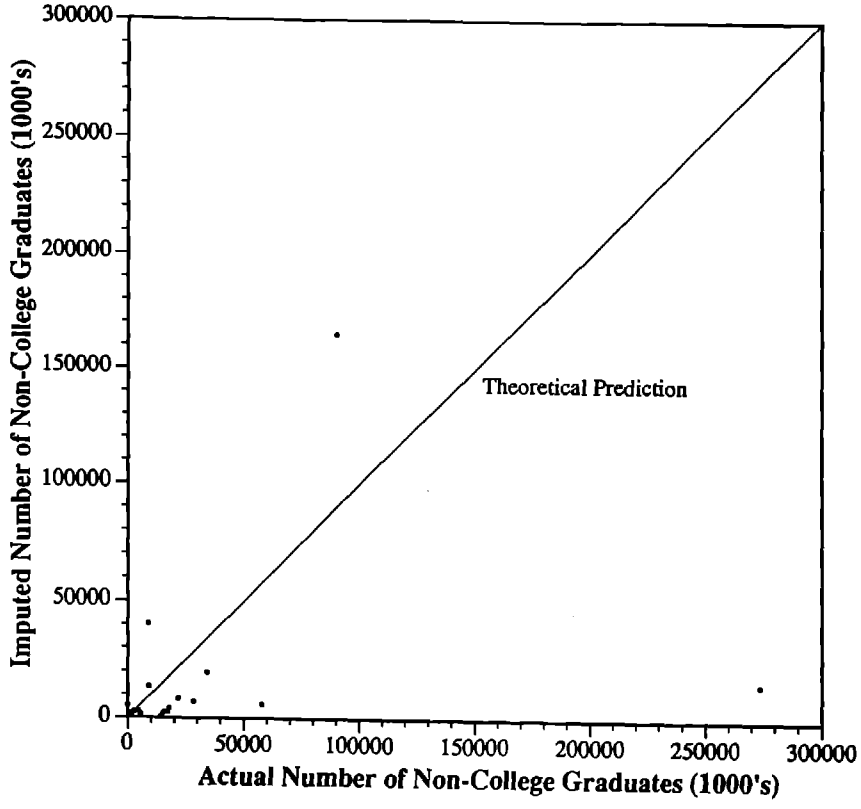


Figure 2a

INTERNATIONAL PRODUCTION TESTS:

ACTUAL VS. IMPUTED COLLEGE ENDOWMENT

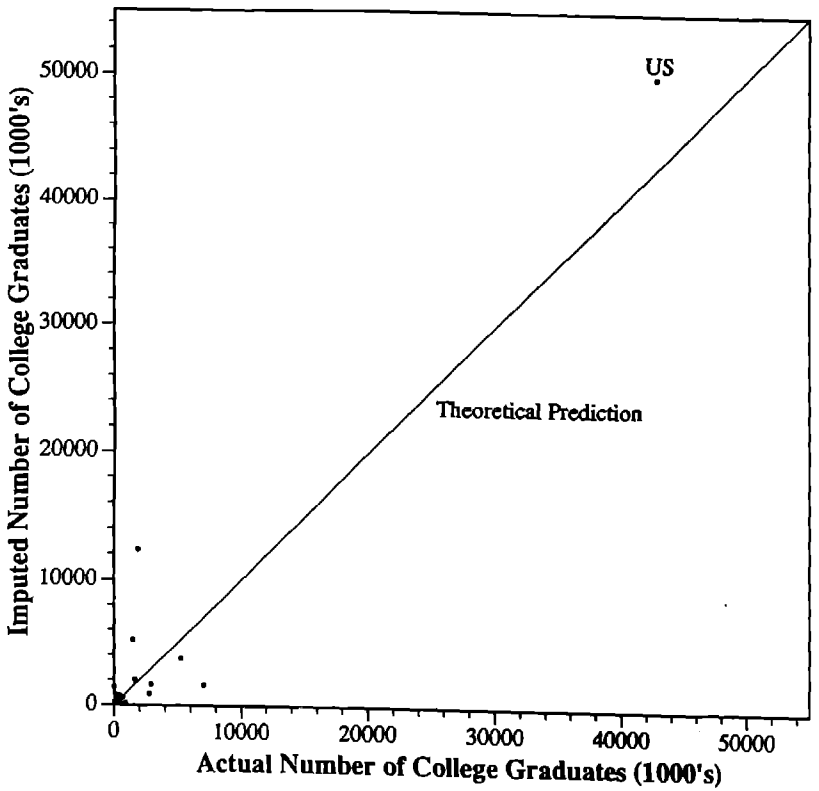


Figure 2b

INTERNATIONAL PRODUCTION TESTS:

ACTUAL VS. IMPUTED COLLEGE ENDOWMENT (DROPPING US)

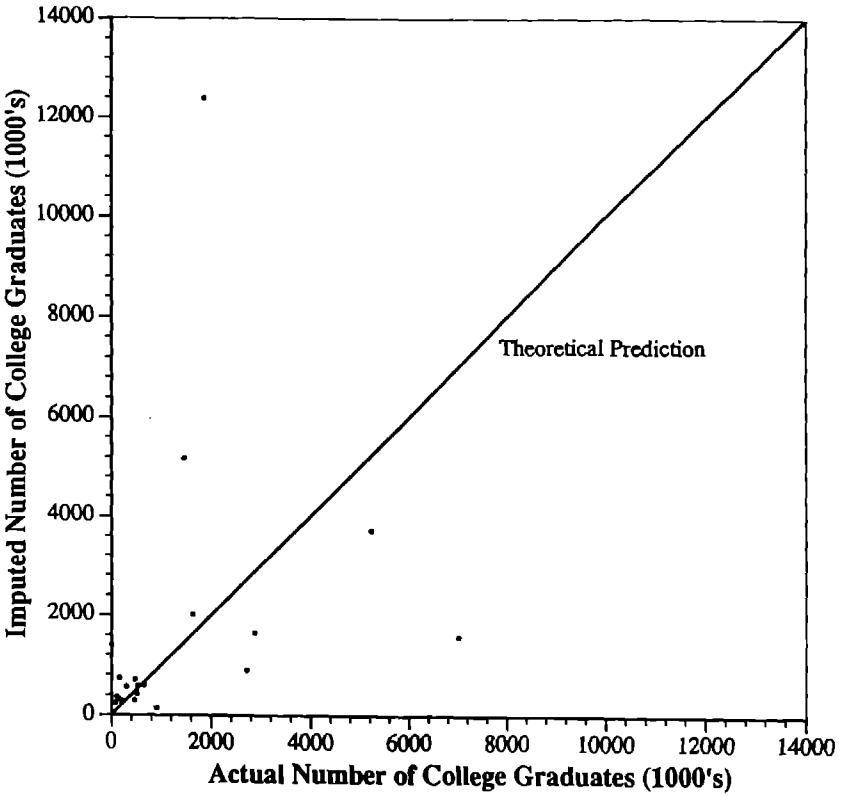


Figure 3a

INTERNATIONAL PRODUCTION TESTS:

ACTUAL VS. IMPUTED CAPITAL STOCK

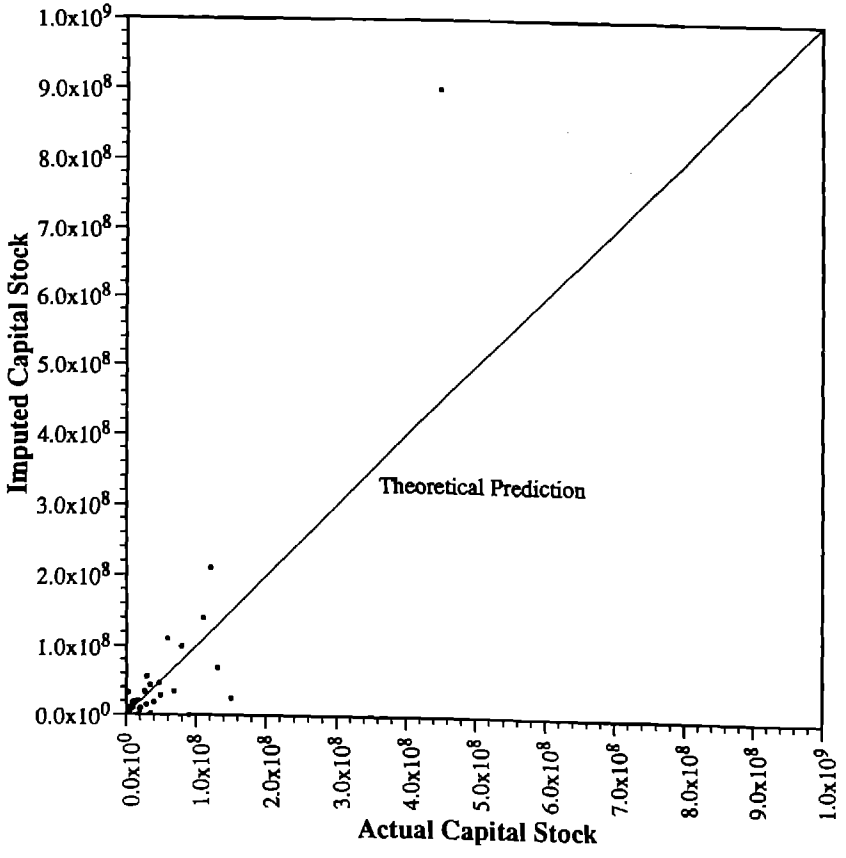


Figure 3b

INTERNATIONAL PRODUCTION TESTS:

ACTUAL VS. IMPUTED CAPITAL STOCK (DROPPING US)

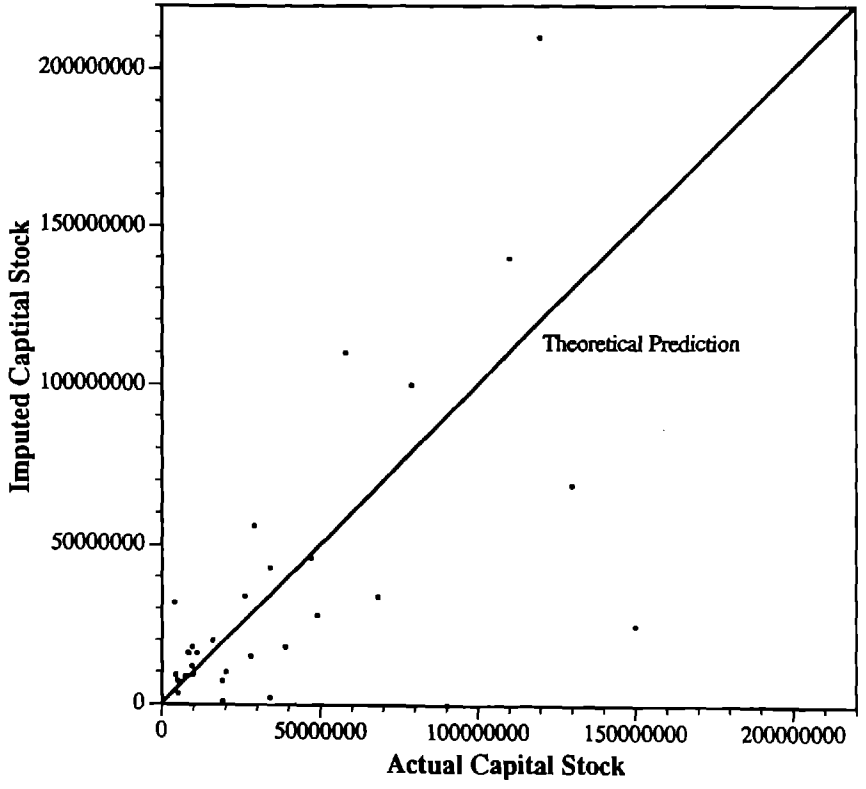


Figure 4

REGIONAL PRODUCTION TESTS

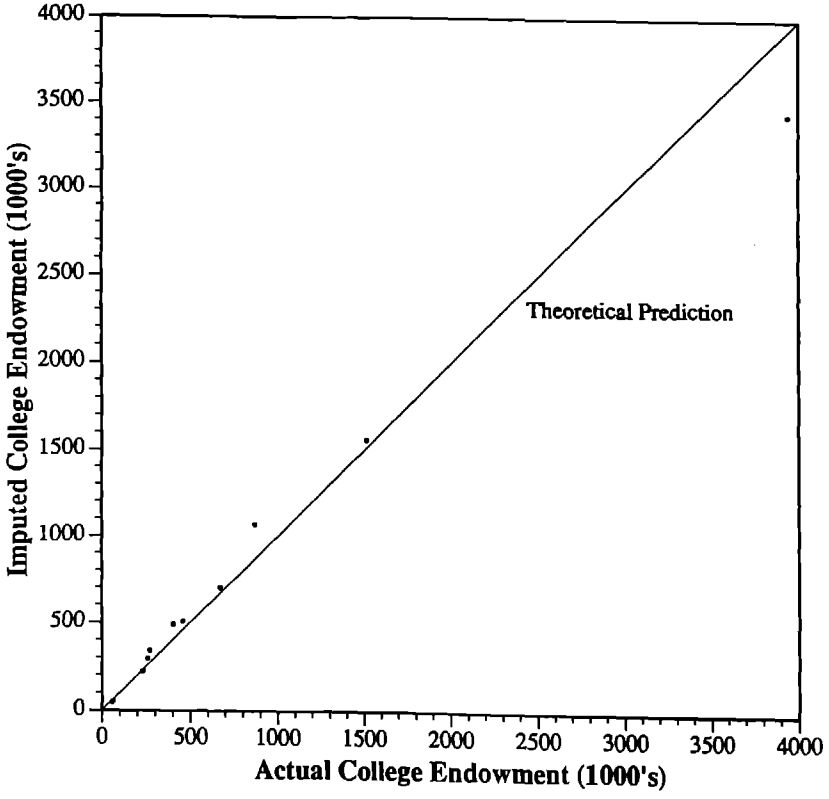


Figure 5

REGIONAL PRODUCTION TESTS

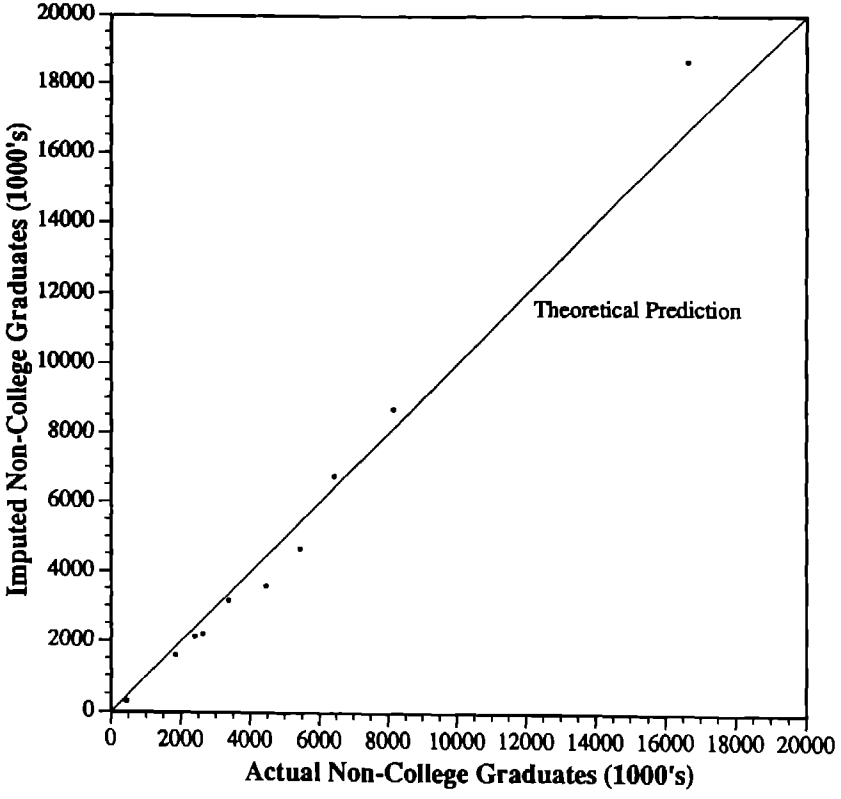


Figure 6

REGIONAL PRODUCTION TESTS

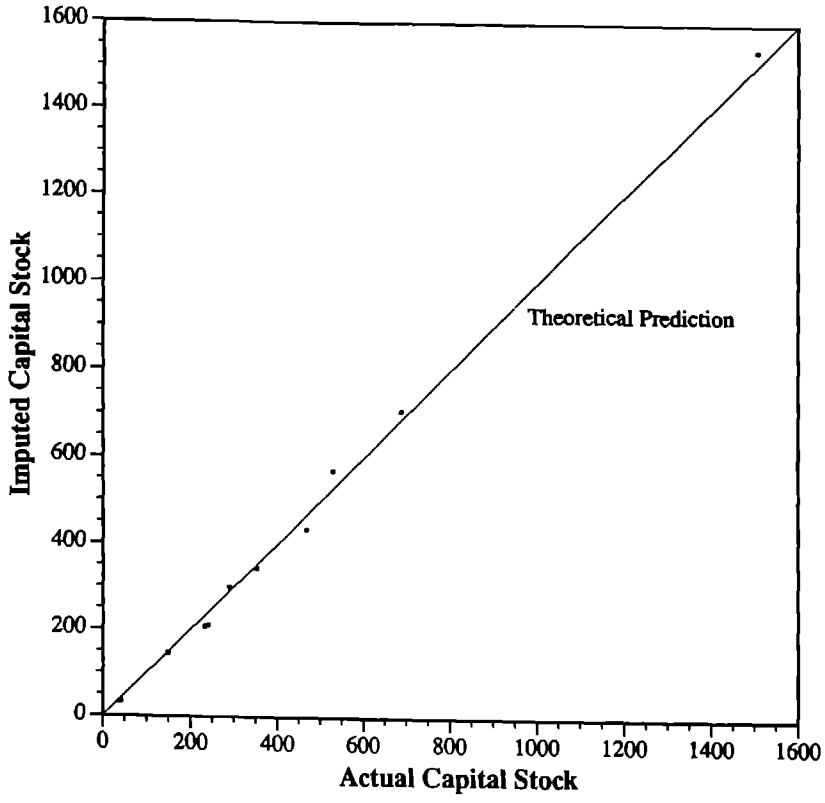


Figure 7

JAPANESE CONSUMPTION TEST
(¥ BILLIONS)

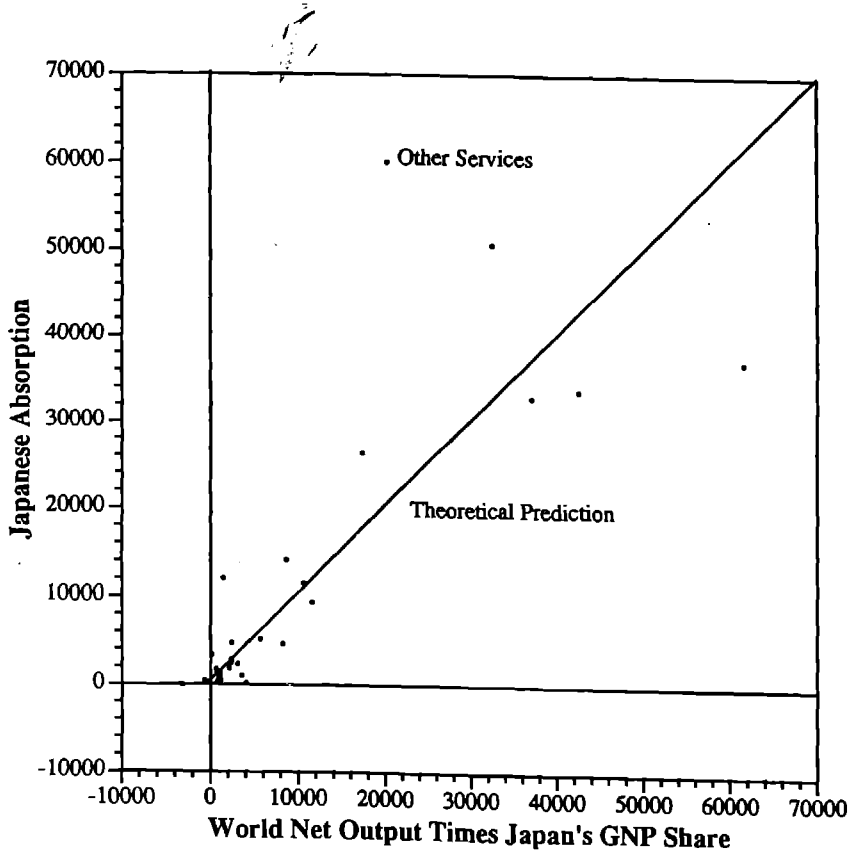


Figure 8

JAPANESE CONSUMPTION TEST: ONLY TRADABLES
(¥ BILLIONS)

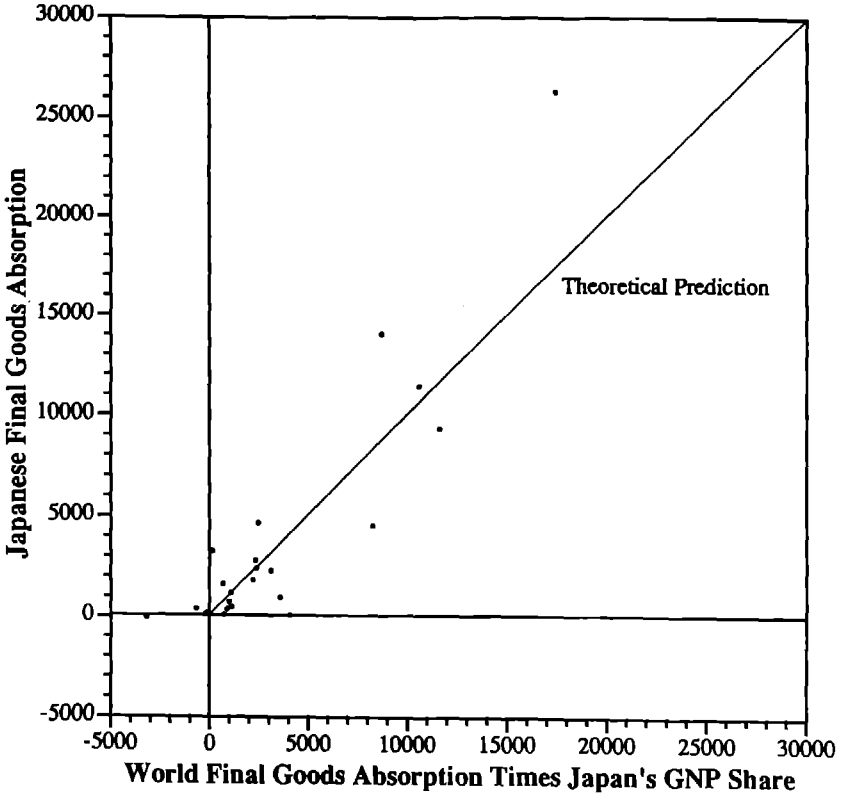


Figure 9

Trade Errors Versus Factor Abundance

(All Factors are Expressed Relative to World Endowment, V^W)

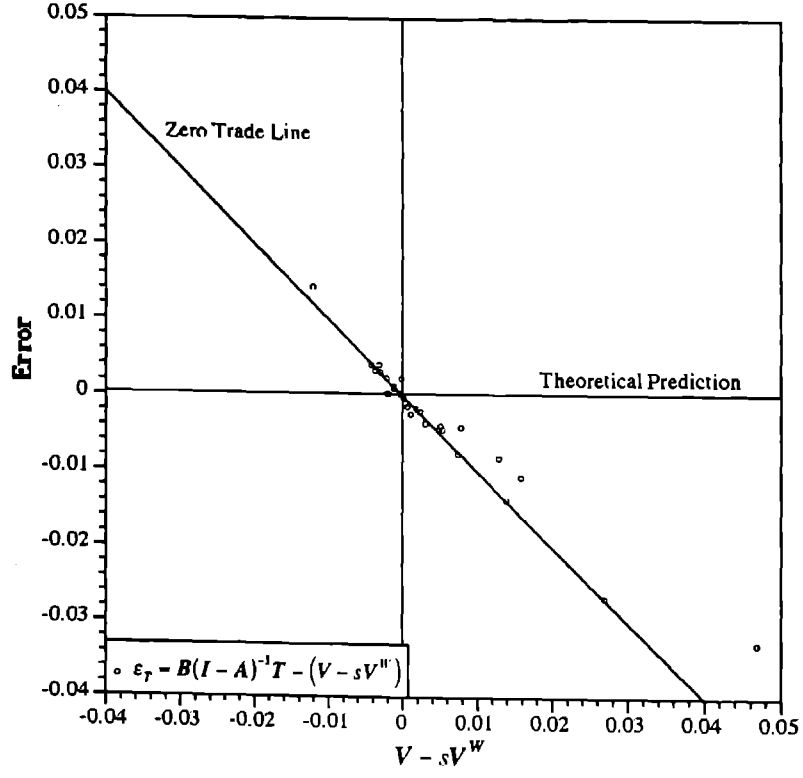


Figure 10

Errors Versus Factor Abundance

(All Factors are Expressed Relative to World Endowment, V^W)

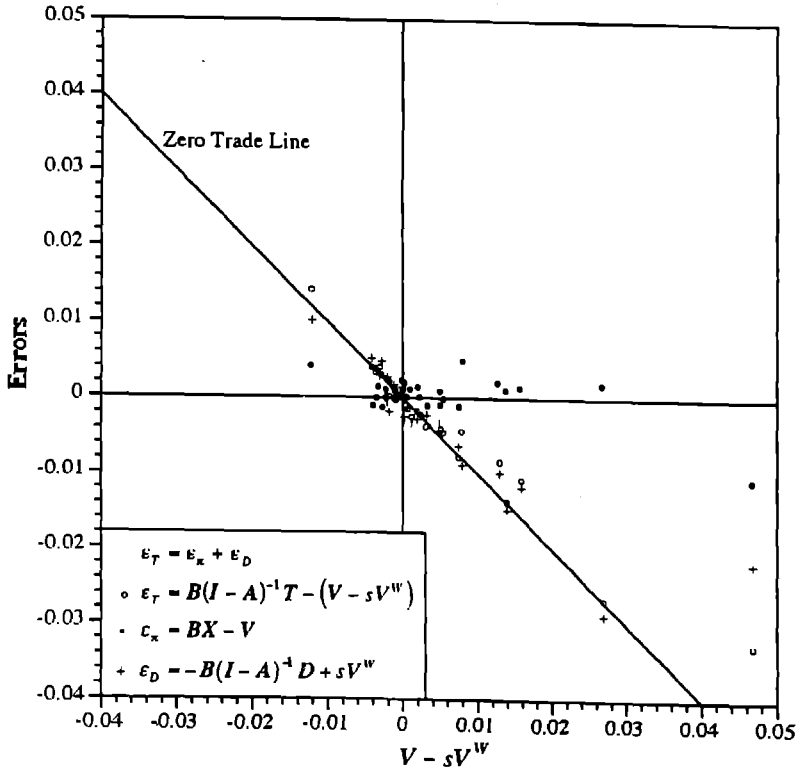
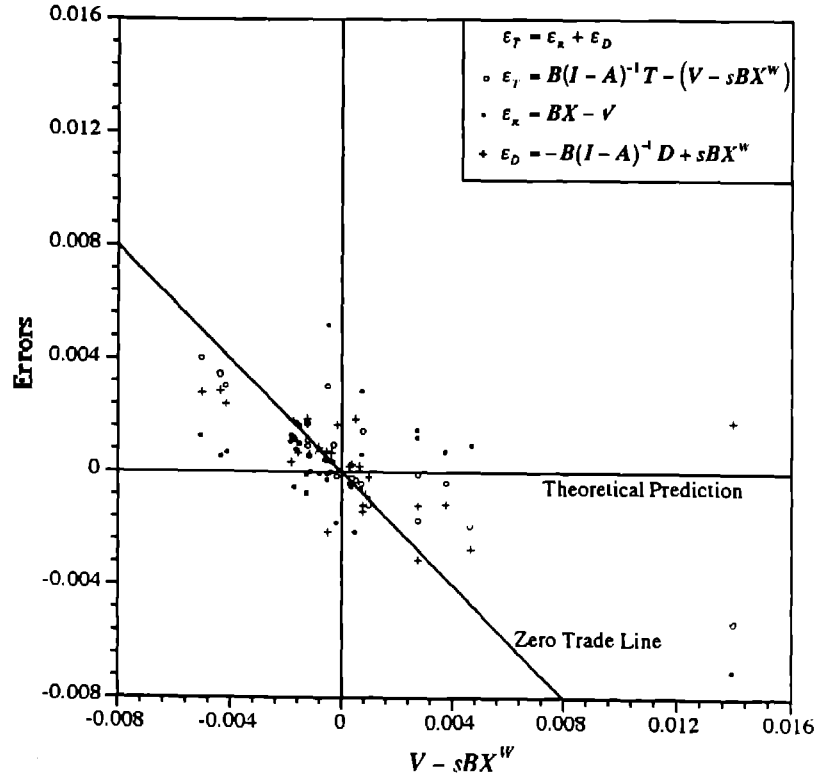


Figure 11

Errors Versus Factor Abundance

(All Factors are Expressed Relative to Imputed World Endowment, BX^W)



Data Appendix

REGIONAL ENDOWMENTS

The numbers of college-educated workers and non-college-educated workers ("good factors") were entered by prefecture directly from the *Employment Status Survey* of 1987 (*Shugyo Kozo Kihon Chosa Hokoku*) and then summed to get regional totals. The numbers of 10 kinds of workers ("bad factors") were also entered by prefecture directly from the *Employment Status Survey* of 1987 and then summed to get regional totals. The capital stocks by region were imputed from prefectural investment data. Japan's yearly *Prefectural Accounts* (*Kenmin Keizai Keisan Nempo*) give investment flows for each prefecture from 1975 to 1985. These flows were used to impute capital stock levels for each prefecture in 1985, using capital goods price deflators from the *National Accounts* and a rate of depreciation of 0.133 (This was the same rate of depreciation used by BLS). Each year's flow was deflated using a capital deflator from the *National Accounts*. The 1985 imputed levels were then aggregated to get regional capital stocks for 1985.

WORLD ENDOWMENTS

World endowments of capital stocks were calculated using ten years of investment data from the Summers and Heston (1988) data set. World endowments of labor force by educational level were taken from the *UNESCO Statistical Yearbook*. Once again we had a scaling problem arising from the fact that the Summers and Heston numbers and the *UNESCO* numbers did not match the Japanese numbers exactly. We therefore scaled each country's capital stock by the ratio of our calculated Japanese capital stock to the Summers and Heston value for the Japanese capital stock. The imputed international labor endowments were similarly scaled by the proportional difference between the *UNESCO* numbers and the actual Japanese endowments.

REGIONAL PRODUCTION

The gross output of 20 manufacturing sectors in each prefecture was taken from the Japanese *Census of Manufactures* for 1985. The gross output of 10 non-manufacturing sectors in each prefecture was taken from the *Prefectural Accounts* for 1985. These numbers for all 30 sectors were then aggregated to get regional production totals. Finally, these totals were scaled so that the 10-region total for each sector exactly matched the total Japanese output as reported in the *1985 Input-Output Table of Japan*. Thus, in effect, the data from the *Census of Manufactures* and from the *Prefectural Accounts* was used in order to distribute total Japanese output for each sector across the 10 regions as accurately as possible.

WORLD PRODUCTION

Data on international levels of production came from the United Nation's *National Accounts Statistics: Main Aggregates and Detailed Tables, 1985*. These numbers differed slightly from the numbers reported in the Japanese IO table, so the output of each sector in every country's output was scaled by the factor necessary to make the international data on Japan match the IO data. The 35 countries used to calculate total world production were Argentina, Australia, Austria, Belgium, Brazil, Canada, China, Denmark, Finland, France, India, Indonesia, Iran, Italy, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Philippines, Poland, Portugal, Saudi Arabia, Singapore, South Africa, Spain, Sweden, Switzerland, Thailand, Turkey, United Kingdom, USA, and West Germany. Output was converted into Yen using exchange rate data from the *IMF*'s *International Financial Statistics*.

ABSORPTION AND NET TRADE

Since the production data is gross output rather than value added, we had to subtract off five categories of absorption in order to calculate the net trade matrix: intermediate use by producers, business consumption (in Japanese data this category is largely entertainment expenses), household (final) consumption, investment, and government purchases.

Intermediate input use in each region was calculated using the Japan 30x30 IO matrix for 1985. Thus, $\boldsymbol{INPUT^c} = \boldsymbol{AX^c}$, where $\boldsymbol{INPUT^c}$ is intermediate consumption in region c, $\boldsymbol{A^J}$ is the IO matrix, and $\boldsymbol{X^c}$ is gross output in region c. Both $\boldsymbol{INPUT^c}$ and $\boldsymbol{X^c}$, therefore, are 30x1 vectors. The 10 $\boldsymbol{INPUT^c}$ vectors together form a 30x10 intermediate consumption matrix. The basic source for the final consumption data was the *Household Expenditure Survey* for 1987. This survey lists household expenditures by region for 56 spending categories. This data was then aggregated up to 42 categories, producing a 42x10 matrix of final consumption by region. This matrix was then premultiplied by a 30x42 bridging matrix from the Economic Planning Agency (there is no 30x56 bridging matrix) which mapped the 42 commodities into the 30 core sectors which we used in our analysis. The survey data was based, of course, on consumer prices, so the bridge matrix was specially constructed to translate the consumption expenditure into producer prices. Most of the difference between consumer and producer prices results from wholesale and retail markups and from transportation costs, so the mapping shifted portions of spending on each final good into the wholesale/retail trade and transportation sectors, reflecting the fact that to consume anything bought retail is to consume the wholesaling, retailing, and transportation services which brought the product to the store. Without this adjustment, the data would have greatly underestimated final consumption of wholesale/retail and transportation services and would have shown each region exporting far more of these services than is plausible.

There are no investment figures broken down for 30 sectors and 10 regions, so these numbers were imputed using IO Table investment data. The IO Table breaks down investment into the 30 sectors for Japan as a whole. This vector was then distributed across regions, using as weights each region's share in total investment for Japan as a whole. Thus, $\boldsymbol{INV^c} = (TI^c/TI^{Japan})\boldsymbol{INV^{Japan}}$, where $\boldsymbol{INV^c}$ is a 30x1 investment vector for region c, TI^c is the total investment for that region in 1985 (taken from the prefectural accounts), TI^{Japan} is Japan's total investment for 1985, and $\boldsymbol{INV^{Japan}}$ is the 30x1 investment vector taken from the IO Table. These 10 $\boldsymbol{INV}$'s therefore formed a 30x10 investment matrix. The government consumption matrix was constructed in a similar fashion, using the 30x1 IO government vector, except that the weights used were not each region's share in Japan's total investment but each region's share in Japan's total GDP. Business consumption was added to the data in a similar fashion.

Net trade for each region was then calculated by subtracting the following from gross output: intermediate consumption, household consumption, investment, and government consumption. Thus, $\boldsymbol{T} = \boldsymbol{X} - \boldsymbol{AX} - \boldsymbol{BC} - \boldsymbol{INV} - \boldsymbol{G}$, where $\boldsymbol{T}$ is net exports, $\boldsymbol{X}$ is gross output, $\boldsymbol{BC}$ is business and household consumption, and each of these symbols represents a 30x10 matrix. It is worth noting that by construction, the sum of net trade across regions will equal Japan's net trade vector.

TECHNOLOGY

Each element of the 3x30 "good factor" technology matrix $\boldsymbol{B^J}$ was calculated by dividing Japanese total output for the 30 sectors into the number of each factor present in each sector. Most of the data on college and non-college workers in each sector came from the 1988 *Wage Census*. There were some gaps in this data as follows: 1) There was no data for college and non-college workers for agriculture, forestry, and fisheries or for government. These numbers were taken from the 1987 *Employment Status Survey*. 2) There was also no data for the petroleum/coal and leather industries. Total employment for each of these sectors was taken from the 1985 *Census of Manufactures*. The number of college workers per unit output for each was then imputed by assuming that petroleum/coal has the same fraction of college workers as the chemicals sector and that leather has the same fraction as manufacturing overall. The data on the "bad" 9 categories of labor came from the IO Table, which reports the numbers of each of 9 kinds of workers in the 30 sectors. The capital stocks in each of the 30 sectors were imputed from investment numbers, using the *Annual Report of the Corporation Survey* for non-manufacturing and the *Census of Manufactures* for manufacturing.