The Case of the Missing Trade and Other Mysteries

By DANIEL TREFLER *

The Heckscher-Ohlin-Vanek (HOV) theorem, which predicts that countries will export products that are made from factors in great supply, performs poorly. However, deviations from HOV follow pronounced patterns. Trade is missing relative to its HOV prediction. Also, rich countries appear scarce in most factors and poor countries appear abundant in all factors, a fact that squares poorly with the HOV prediction that abundant factors are exported. As suggested by the patterns, HOV is rejected empirically in favor of a modification that allows for home bias in consumption and international technology differences. (*JEL* F11, F14)

What is known about international trade in factor services? Theoretically, the Heckscher-Ohlin theorem states that a capital-abundant country exports the capital-intensive good. Its generalization, the Heckscher-Ohlin-Vanek (HOV) theorem, states that a capital-abundant country exports capital services (see Eli F. Heckscher, 1919; Bertil G. Ohlin, 1933; Paul A. Samuelson, 1948; James R. Melvin, 1968; Jaroslav Vanek, 1968; Edward E. Leamer, 1980). Empirically, the HOV theorem has been repeatedly rejected over the years and rightfully so: it performs horribly. Factor endowments correctly predict the direction of factor service trade about 50 percent of the time, a success rate that is matched by a coin toss. Since the HOV theorem extends to a variety of models displaying increasing returns to scale and imperfect competition (Elhanan Helpman and Paul R. Krugman, 1985), this poor performance has distressing implications for these trade theories as well. In other fields of economics, the poor performance of a major theory leads to more careful consideration of the data and to new theories that can accommodate the anomalies. Yet years of research into why the HOV theorem performs poorly has only produced conjectures. It has not provided a deeper understanding of factor service trade, nor has it identified an alternative hypothesis that performs better. These two failings are the subject of this paper.

First, almost nothing is known about the features of factor service trade that are inconsistent with the HOV theorem. An exception is the Leontief paradox. However, it deals with only two of many factors in only one of many countries; that is, the United States exports too much labor and too little capital. Also, it is not a paradox (Leamer, 1980), and it disappeared from the data at least 20 years ago (Robert M. Stern and Keith E. Maskus, 1981). Thus, with the exception of a few laconic and outdated references, nothing is known. A goal of this paper is to demonstrate that the HOV theorem is rejected because factor service trade departs from its endowments-based prediction in *systematic* and informative ways.

Understanding trade in factor services rather than trade in goods is not simply an academic exercise; it is central to the conduct of trade policy. For example, the number of cars the

* Harris Graduate School of Public Policy Studies, University of Chicago, 1155 East 60th St., Chicago, IL 60637, and Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON Canada, M5S 1A1. To obtain an electronic copy of the data, e-mail the author at tref@cicero.spc.uchicago.edu or trefler@epas.utoronto.ca. I am indebted to Werner Antweiler, Jr. for jointly pursuing the empirical work in an earlier version of this paper. Harry Bowen, Edward Leamer, and Leo Sveikauskas encouraged and helped in my critique of their work. I am grateful to the anonymous referees for suggesting substantive improvements. I benefited from seminar presentations at the NBER, the University of British Columbia, the University of Chicago, the University of Minnesota, MIT, the University of Pennsylvania, and the University of Wisconsin, especially from the comments of Robert Staiger. Research was supported by the Social Sciences and Humanities Research Council (SSHRC) of Canada.

United States imports from Japan is uninteresting in and of itself. It takes on importance because of its factor-market consequences: the U.S. jobs displaced and the effect on wages, the supplanted investment, and the effect on rates of return to capital. These are concerns about factor service trade.

Second, is there a general equilibrium model of factor service trade that is known to perform better than the HOV theorem? The answer is no. The HOV theorem has frequently been rejected in favor of statistical hypotheses such as a zero correlation (e.g., Maskus, 1985; Richard A. Brecher and Ehsan U. Choudhri, 1988; Robert W. Staiger, 1988). This is valuable for showing how poorly the HOV theorem performs but cannot be used to identify economically meaningful models that perform better than the theorem. Harry P. Bowen et al. (1987) pioneered a method for testing the HOV hypothesis against economic alternatives, only to arrive at a negative conclusion: "The Heckscher-Ohlin model does poorly, but we do not have anything that does better. It is easy to find hypotheses that do as well or better in a statistical sense, but these alternatives yield economically unsatisfying parameter estimates" (p. 805). A second goal of this paper is to identify economic hypotheses that perform better than the HOV theorem. Adopting the Bowen et al. method, I consider a large number of hypotheses, including ones with capital accumulation, nontradables, trade in services, and linear expenditure demand. The model that clearly dominates the HOV theorem allows for home bias in consumption (Paul S. Armington, 1969) and international differences in technology. The systematic departures from the HOV theorem noted above are used to explain why some hypotheses perform well and others do not. Along the way I explain Staiger's (1988) result that the HOV model is misspecified and overturn the Bowen et al. (1987) result that no simple modification of the HOV theorem performs well.

A number of citations serve to demarcate the area of study. First, this paper provides a logically complete test of the HOV theorem in the Leamer and Bowen (1981) sense that it uses data on technology, trade, and endowments. Much of the literature relating to the theorem only used two of these three (e.g., Wassily W. Leontief, 1953; Leamer, 1984). Second, this paper is related to a previous work of mine (Trefler, 1993) dealing with international factor-price differences. In that work I considered a variant of the HOV model that allows for international productivity differences. The variant necessarily fits the trade and endowments data perfectly, thus ruling out hypothesis-testing. In contrast, hypothesis-testing is central to what follows. Also, in what follows I cover a wide range of alternative hypotheses and use the systematic patterns in the deviations from the HOV theorem to identify many models that perform poorly and two models that perform well. In contrast, my previous work only examined one model.

I. Testing the HOV Theorem Against Statistical Alternatives

Let $c = 1, \ldots, C$ index countries and $f = 1, \ldots, F$ index factors. Let V_{fc} be the endowment of factor f in country c and let $V_{fw} = \Sigma_c V_{fc}$ be world factor endowments. Let F_{fc} be the factor content of net exports, that is, the amount of factor f needed to produce the net exports of country c. Let $s_c = (Y_c - B_c)/Y_w$ be the consumption share of country c where B_c is the trade balance, Y_c is gross national product (GNP), and $Y_w = \Sigma_c Y_c$. The following "HOV equation" is implied by the usual HOV assumptions (see, for example, Leamer [1980] or the proof below of a more general result):

$$(1) \quad F_{fc} = V_{fc} - s_c V_{fw} \qquad f = 1, \ldots, F \qquad c = 1, \ldots, C.$$

It states that if country c is abundant in factor f ($V_{fc}/V_{fw} > s_c$), then it exports the services of factor f ($F_{fc} > 0$).

The following data will be used to investigate this HOV equation. All data are from 1983 unless indicated otherwise. There are 33 countries in the sample which together account for 76 percent of world exports and 79 percent of world GNP. The choice of countries was largely dictated by the availability of trade data at a detailed industry level. There are nine factors: capital, cropland, pasture, and six categories of labor. The labor categories are professional and technical workers, clerical workers, sales workers, service workers, agriculture workers, and production, transport, and unskilled workers.[1] Under the usual HOV assumptions the factor content of trade is $(F_{1c}, \ldots, F_{Fc})' = \mathbf{A}\mathbf{T}_c$ where $\mathbf{T}_c$ is the vector of net commodity exports and $\mathbf{A}$ is the "technology matrix" giving the amount of each factor needed to produce one unit of each commodity. $\mathbf{A}$ was built using the 1983 U.S. input–output total-requirements table and data on factor usage by industry from various 1982 U.S. industry censuses and the 1983 Annual Survey of Manufactures. The usual caveat about using U.S. technology to evaluate the factor content of non-U.S. trade applies here, albeit with less force since below the technology matrix will be modified in a country-specific fashion. The relevant data are detailed in Trefler (1993).

Factors must be expressed in comparable units in order to satisfy the statistical hypothesis of homoscedasticity. To this end let ε_{fc} be the deviations from the HOV theorem:

$$(2) \qquad \varepsilon_{fc} = F_{fc} - (V_{fc} - s_c V_{fw}).$$

Fix f and let σ_f be the standard error of the ε_{fc}: $\sigma_f^2 = \Sigma_c\, (\varepsilon_{fc} - \bar{\varepsilon}_f)^2/(C-1)$ where $\bar{\varepsilon}_f = \Sigma_c\, \varepsilon_{fc}/C$. I scale all data relating to factor f by σ_f so that factors are expressed in statistically comparable units. In addition, to control for country size I scale by $s_c^{1/2}$. Hence, throughout this paper observation (f, c) is scaled by $\sigma_f s_c^{1/2}$. (See part 8 of the Appendix for additional discussion of scaling.)

From equation (1), the simple correlation between F_{fc} and $V_{fc} - s_c V_{fw}$ provides a test of the HOV equation against a statistical alternative. With nine factors and 33 countries there are 297 observations. The resulting correlation is 0.28, which is statistically significant but hardly impressive. An alternative statistic follows from a weaker statement of the HOV theorem: country c exports the services of its abundant factors and imports the services of its scarce factors. That is, $F_{fc} > 0$ if and only if $V_{fc} - s_c V_{fw} > 0$. I call this "sign HOV." Bowen et al. (1987) reported the percentage of observations for which F_{fc} and $V_{fc} - s_c V_{fw}$ have the same sign. Under the sign-HOV hypothesis, the statistic equals 100 percent. In fact, it equals 49.8 percent (148/297), which means that the HOV prediction is about as good as a coin toss. The sign statistic treats all observations equally. An alternative is to attach more weight to observations with large net factor contents of trade, that is, to weight the sign statistic by

$$|F_{fc}| \Big/ \sum_{f,c} |F_{fc}|.$$

The weighted statistic equals 71 percent. That is, the sign-HOV hypothesis is more accurate when net factor service trade flows are large. Nevertheless, the statistic of 71 percent is far from the HOV null of 100 percent and uncomfortably close to the coin-toss alternative of 50 percent. In short, the HOV theorem performs poorly.

II. A View Through the HOV Window

In order to investigate the failure of the HOV theorem, consider its deviations, $\varepsilon_{fc} = F_{fc} - (V_{fc} - s_c V_{fw})$. Surprisingly, plots of factor service trade against endowments have never been reported. Figure 1 plots ε_{fc} against $V_{fc} - s_c V_{fw}$. Points to the right of the vertical line $V_{fc} - s_c V_{fw} = 0$ correspond to abundant factors. The diagonal line is $\varepsilon_{fc} = -(V_{fc} - s_c V_{fw})$ or $F_{fc} = 0$ so that points above it correspond to $F_{fc} > 0$. The sign-HOV theorem predicts that all observations will lie in two of the four demarcated areas, either where $F_{fc} > 0$ and $V_{fc} - s_c V_{fw} > 0$ or where $F_{fc} < 0$ and $V_{fc} - s_c V_{fw} < 0$. Only half of the observations lie in these areas. Under the HOV equation (1), $F_{fc} = V_{fc} - s_c V_{fw}$ or $\varepsilon_{fc} = 0$; that is, all the observations lie on a horizontal line at zero. Nothing like this pattern emerges.

The main feature of Figure 1 is that all the observations lie close to the $F_{fc} = 0$ line (12

[1] Administrative and managerial workers, forests, oil, coal, and minerals have been used in other studies but were not included in this study. The results are similar when all these factors are included in the analysis, except as noted in part 7 of the Appendix. Nevertheless, they were omitted either on theoretical grounds or because of concerns about data quality.

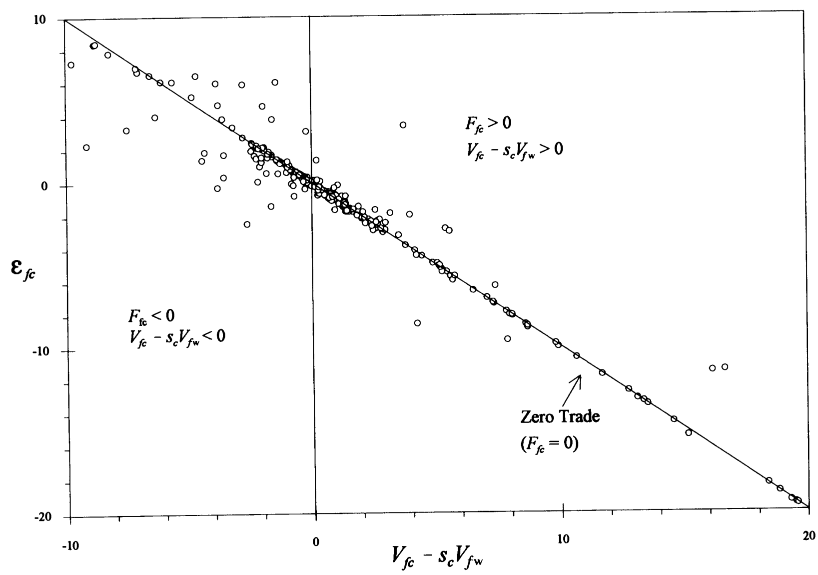

FIGURE 1. PLOT OF $\varepsilon_{fc} = F_{fc} - (V_{fc} - s_c V_{fw})$ AGAINST $V_{fc} - s_c V_{fw}$

observations far from the origin were truncated, but all of these lie close to the $F_{fc} = 0$ line). In absolute values, factor service trade is much smaller than its factor-endowments prediction. I call this phenomenon "the case of the missing trade." A similar phenomenon appears in the Bowen et al. (1987) data for 1966–1967.

Further patterns in the deviations from HOV appear when the data are examined by country. The left panel of Figure 2 displays the number of negative deviations ($\varepsilon_{fc} < 0$) per country. With nine factors this number lies between 0 and 9. Countries are sorted by purchasing-power-parity (PPP) adjusted per capita GDP from the Penn World Tables. Poor countries tend to have negative deviations, and rich countries tend to have positive deviations. The correlation of the number of negative deviations per country with per capita GDP is 0.87. Since F_{fc} is typically "small," results about $\varepsilon_{fc} = F_{fc} - (V_{fc} - s_c V_{fw})$ are likely to be reflected in $V_{fc} - s_c V_{fw}$. The right panel of Figure 2 displays the number of abundant factors per country. Rich countries tend to be scarce in most factors, and poor countries tend to be abundant in all factors. The correlation with per capita GDP is −0.89. I call this phenomenon "the endowments paradox." It appears in the Leontief (1953) data for 1947 (recall that Leamer [1980] showed the United States to be scarce in both labor and capital), in the Bowen et al. (1987) data for 1966–1967, in the Leamer (1984) data for 1958 and 1975, and in the Maskus (1991) data for 1984. It may also underpin Staiger's (1988) observation that country-specific deviations from the HOV theorem are correlated with country-specific data on endowments and size.

Ranking factors in order of abundance for country c, the HOV theorem may be illustrated as in Figure 3. Note that s_c puts a break in the Vanek chain that distinguishes scarce imported factors from abundant exported factors. Figure 2 shows that s_c often lies either to the extreme right or extreme left of this ranking, thus undermining the HOV theorem. This is illustrated in Figure 3 where LDC and DC denote poor and rich countries, respectively. There are two explanations for this: trade imbalances and omitted factors. This follows from the relationship $\Sigma_f\, w_{fc}(V_{fc} - s_c V_{fw}) = B_c$ where w_{fc} is the price of factor f in country c. Hence, if there were no omitted factors and $B_c = 0$, then country c could not

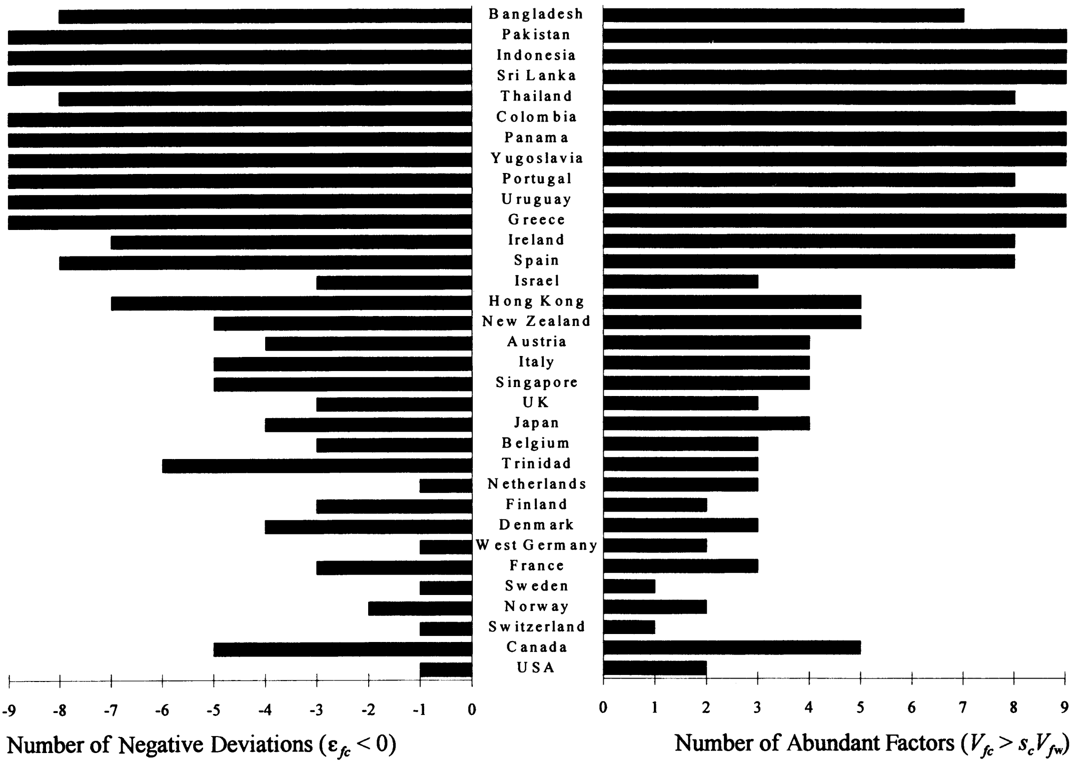

FIGURE 2. DEVIATIONS FROM HOV AND FACTOR ABUNDANCE

be abundant or scarce in all factors. The pattern of Figure 2 would occur if rich countries ran trade deficits ($B_c < 0$) and poor countries ran trade surpluses ($B_c > 0$). If anything, the opposite is true: the rank correlation of B_c with per capita GDP is 0.14. Thus, unless there are omitted factors that are scarce in poor countries, the pattern in the right panel of Figure 2 is inconsistent with the spirit and the letter of a theory whose cornerstone is factor abundance.

III. Economically Meaningful Alternative Hypotheses: Technology

It remains to search for economically meaningful alternatives to the HOV theorem that can account for the case of the missing trade and the endowments paradox. In this section I consider alternatives that modify the technology assumptions of the HOV theorem; in the next I consider alternatives that modify the consumption assumptions.

A. *Theory*

Let f_{ic} be the production function for good i in country c. Let $\mathbf{a}_{ic}$ be a typical $F \times 1$ column of the technology matrix $\mathbf{A}_c$ giving the amounts of each factor needed to produce one unit of good i. By definition, $f_{ic}(\mathbf{a}_{ic}) = 1$. I assume that international technology differences (the c subscript in f_{ic}) are factor-augmenting. That is, $f_{ic}(\mathbf{a}_{ic}) = f_i(\mathbf{\Pi}_c \mathbf{a}_{ic})$ for some internationally common production functions f_i and diagonal matrices $\mathbf{\Pi}_c = \text{diag}(\pi_{1c}, \ldots, \pi_{Fc})$. Since $f_i(\mathbf{\Pi}_c \mathbf{a}_{ic}) = 1$, the larger an element of $\mathbf{\Pi}_c$, the smaller is the corresponding element of $\mathbf{a}_{ic}$. In other words, larger $\mathbf{\Pi}_c$'s correspond to fewer inputs per unit of output or greater productivity. Without loss of generality $\mathbf{\Pi}_{US}$ is taken

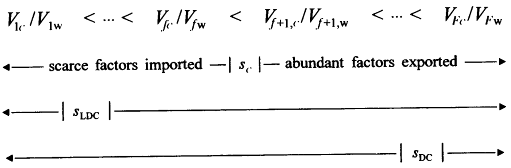

FIGURE 3. FACTOR ABUNDANCE AND SCARCITY

to be the identity matrix so that π_{fc} is the productivity of factor f in country c relative to U.S. productivity.

Let w_{fc} be the price of factor f in country c and $\mathbf{w}_c = (w_{1c}, \ldots, w_{Fc})'$, where a prime denotes matrix transposition. If U.K. labor were half as productive as U.S. labor ($\pi_{L,UK} = \frac{1}{2}$), then one would expect U.K. wages to be half of U.S. wages: $w_{L,UK} = \pi_{L,UK} w_{L,US}$. More generally, assume $w_{fc} = \pi_{fc} w_{fUS}$ or, in matrix notation, $\mathbf{w}_c = \mathbf{\Pi}_c \mathbf{w}_{US}$. Support for this variant of factor price equalization is the main result in Trefler (1993).

Assuming that all goods are produced in all countries, that product prices are the same internationally, and that profits are zero, it follows that unit costs are the same internationally: $c_{iUS}(\mathbf{w}_{US}) = c_{ic}(\mathbf{w}_c)$. Differentiating with respect to $\mathbf{w}_{US}$ yields $\mathbf{A}_{US}(\mathbf{w}_{US}) = \mathbf{\Pi}_c \mathbf{A}_c(\mathbf{w}_c)$.[2] Thus, given knowledge of $\mathbf{\Pi}_c$ and U.S. data, one can infer $\mathbf{A}_c(\mathbf{w}_c) = \mathbf{\Pi}_c^{-1} \mathbf{A}_{US}(\mathbf{w}_{US})$ even if neither $\mathbf{A}_c$ nor $\mathbf{w}_c$ is observed. This fact allows me to place the discussion of factor prices in the background and to simply assume $\mathbf{A}_{US} = \mathbf{\Pi}_c \mathbf{A}_c$.

Under this assumption the HOV equation (1) is replaced by

$$(3) \qquad F_{fc}^{US} = \pi_{fc} V_{fc} - s_c \sum_j \pi_{fj} V_{fj}$$

where $\mathbf{F}_c^{US} = (F_{1c}^{US}, \ldots, F_{Fc}^{US})' = \mathbf{A}_{US}\mathbf{T}_c$ is the factor content of country c's net exports when calculated using U.S. technology.[3] F_{fc}^{US}/π_{fc} answers the important policy question about the quantity of domestic factors embodied in trade had imports been produced domestically.

A problem with equation (3) is that it has as many parameters as observations (FC) and so necessarily fits the data perfectly (i.e., it cannot be tested against the HOV equation). One approach to overfitting restricts the way the π_{fc} vary across factors: $\pi_{fc} = \delta_c$ for all f and c. The δ_c are Hicks-neutral factor-augmenting productivity measures. This yields

$$(4) \qquad F_{fc}^{US} = \delta_c V_{fc} - s_c \sum_j \delta_j V_{fj}$$

where $\delta_{US} = 1$.

A second approach restricts the way the π_{fc} vary across countries. The extreme restriction that $\pi_{fc} = \phi_f$ for all f and c (except $\pi_{fUS} = 1$) performs poorly empirically. An alternative is to divide the countries in the sample into two groups: poor countries that share one set of nonneutrality parameters and rich countries that share a different set of nonneutrality parameters. For example, this allows the French capital–labor ratio in agriculture to be similar to the corresponding German ratio and different from the Bangladeshi ratio. Let C_{DC} be the set of rich countries. This includes the United States with its $\pi_{fUS} = 1$ so that $\pi_{fc} = 1$ for $c \in C_{DC}$. Let C_{LDC} be the set of poor countries: they share a common π_{fc} so that $\pi_{fc} = \phi_f$ for $c \in C_{LDC}$.

A third approach combines the neutral (δ_c) and nonneutral (ϕ_f) technology differences: $\pi_{fc} = \delta_c \phi_f$ for $c \in C_{LDC}$, $\pi_{fc} = \delta_c$ for

[2] To see this, note that $\partial c_{ic}(\mathbf{w}_c)/\partial \mathbf{w}_c = \mathbf{a}_{ic}'(\mathbf{w}_c)$ where a prime denotes matrix transposition. Differentiating $c_{iUS}(\mathbf{w}_{US}) = c_{ic}(\mathbf{w}_c)$ with respect to the vector $\mathbf{w}_{US}$ yields $\mathbf{a}_{iUS}'(\mathbf{w}_{US}) = \mathbf{a}_{ic}'(\mathbf{w}_c)\partial \mathbf{w}_c/\partial \mathbf{w}_{US} = \mathbf{a}_{ic}'(\mathbf{w}_c)\mathbf{\Pi}_c$. Transposing, $\mathbf{a}_{iUS}(\mathbf{w}_{US}) = \mathbf{\Pi}_c \mathbf{a}_{ic}(\mathbf{w}_c)$. Since $\mathbf{a}_{ic}$ is a column of $\mathbf{A}_c$, $\mathbf{A}_{US}(\mathbf{w}_{US}) = \mathbf{\Pi}_c \mathbf{A}_c(\mathbf{w}_c)$, as required.

[3] The proof of equation (3) is as follows. Let $\mathbf{Q}_c$ and $\mathbf{C}_c$ be vectors of production and consumption, respectively; let $\mathbf{V}_c = (V_{1c}, \ldots, V_{Fc})'$, and define $\mathbf{Q}_w = \Sigma_c \mathbf{Q}_c$, $\mathbf{C}_w = \Sigma_c \mathbf{C}_c$, and $\mathbf{V}_w = \Sigma_c \mathbf{V}_c$. By definition, $\mathbf{T}_c = \mathbf{Q}_c - \mathbf{C}_c$ so that $\mathbf{F}_c^{US} = \mathbf{A}_{US}(\mathbf{Q}_c - \mathbf{C}_c)$. From preference homotheticity and world goods-market equilibrium $\mathbf{C}_c = s_c \mathbf{C}_w = s_c \mathbf{Q}_w$ so that $\mathbf{F}_c^{US} = \mathbf{A}_{US}(\mathbf{Q}_c - s_c \mathbf{Q}_w)$. $\mathbf{A}_c = \mathbf{\Pi}^{-1}\mathbf{A}_{US}$ and factor-market clearing, $\mathbf{A}_c \mathbf{Q}_c = \mathbf{V}_c$, imply $\mathbf{A}_{US}\mathbf{Q}_c = \mathbf{\Pi}_c \mathbf{V}_c$ and $\mathbf{A}_{US}\mathbf{Q}_w = \Sigma_j \mathbf{\Pi}_j \mathbf{V}_j$ so that $\mathbf{F}_c^{US} = \mathbf{\Pi}_c \mathbf{V}_c - s_c \Sigma_j \mathbf{\Pi}_j \mathbf{V}_j$. This last is the matrix counterpart of equation (3).

$c \in C_{\mathrm{DC}}$, and $\delta_{\mathrm{US}} = 1$. To distinguish between average or neutral effects (δ_c) and nonneutral effects (ϕ_f) impose the identifying restriction $\Sigma_f \phi_f / F = 1$.[4] Then equation (3) becomes

$$(5)\quad F_{fc}^{\mathrm{US}} = \begin{cases} \delta_c \phi_f V_{fc} - s_c \sum_{j \in C_{\mathrm{LDC}}} \delta_j \phi_f V_{fj} - s_c \sum_{j \in C_{\mathrm{DC}}} \delta_j V_{fj} & c \in C_{\mathrm{LDC}} \\ \delta_c V_{fc} - s_c \sum_{j \in C_{\mathrm{LDC}}} \delta_j \phi_f V_{fj} - s_c \sum_{j \in C_{\mathrm{DC}}} \delta_j V_{fj} & c \in C_{\mathrm{DC}}. \end{cases}$$

To operationalize C_{LDC} fix a constant κ, let y_c be per capita GDP, and let country c be a member of $C_{\mathrm{LDC}}(\kappa)$ if $y_c < \kappa$. Then κ will be estimated along with the δ_c and ϕ_f parameters. In principle there can be as many groupings of countries as is desired, and the groupings need not be along development lines. For example, it is a priori plausible to have three groups: the United States, other rich countries, and poor countries. However, distinguishing the United States from the other rich countries is statistically rejected by the data. Thus, equation (5) is sensible as judged by parsimony, the importance of the technology gap between developed and developing countries, and the correlation of deviations from HOV with per capita GDP.

B. *Results*

Even if equation (5) held exactly, errors in measuring observables (F_{fc}^{US}, V_{fc}, s_c) would obscure this fact and lead one to estimate "best" values for unobservables (δ_c, ϕ_f). There are five potential sources of error: each of F_{fc}^{US}, V_{fc}, and s_c mismeasured, omitted factors, and omitted countries. In parts 1–4 of the Appendix, I show that results are remarkably similar across the five specifications of error. Thus, in this section I simply assume that only F_{fc}^{US} is mismeasured: $F_{fc}^{m} = F_{fc}^{\mathrm{US}} + \mu_{fc}$ where F_{fc}^{m} is the measure and μ_{fc} is measurement error. Then equation (5) becomes

$$(6)\quad F_{fc}^{m} = \begin{cases} \delta_c \phi_f V_{fc} - s_c \sum_{j \in C_{\mathrm{LDC}}} \delta_j \phi_f V_{fj} - s_c \sum_{j \in C_{\mathrm{DC}}} \delta_j V_{fj} + \mu_{fc} & c \in C_{\mathrm{LDC}} \\ \delta_c V_{fc} - s_c \sum_{j \in C_{\mathrm{LDC}}} \delta_j \phi_f V_{fj} - s_c \sum_{j \in C_{\mathrm{DC}}} \delta_j V_{fj} + \mu_{fc} & c \in C_{\mathrm{DC}} \end{cases}$$

where $\delta_{\mathrm{US}} = 1$ and $\Sigma_f \phi_f / F = 1$. I assume that the μ_{fc} are independently and identically distributed normal with mean zero. Adding an intercept ($E\mu_{fc} \neq 0$) makes little difference to the results. Common variance follows from scaling observation (f, c) by $\sigma_f s_c^{1/2}$ (see above).

Table 1 reports statistics for three hypotheses nested within equation (6). The null hypothesis (H_0) is the HOV equation (1). The number of parameters (k_i) is zero. T_1 is the neutral technology-differences model. With 33 countries and $\delta_{\mathrm{US}} = 1$, there are 32 parameters. T_2 is the neutral and nonneutral technology-differences model of equation (6). With nine factors and the restriction $\Sigma_f \phi_f / F = 1$, there are eight ϕ_f's as well as κ for a total of 41 parameters. Appropriately parameterized, T_1 is a linear model and I use ordinary least squares (OLS) to estimate it subject to the linear restriction that $\delta_{\mathrm{US}} = 1$. T_2 is nonlinear, and I use maximum-likelihood estimation.[5] From Table 1, the likelihood-ratio test

[4] If one prefers a normalization other than $\Sigma\, \phi_f / F = 1$, say, $\Sigma\, \omega_f \phi_f = 1$ for some reader-chosen weights ω_f that measure the relative importance of factors, then divide each of the estimated ϕ_f in Table 3 by $\Sigma\, \omega_f \phi_f$ and multiply each of the estimated δ_c ($c \in C_{\mathrm{LDC}}$) in Table 2 by $\Sigma\, \omega_f \phi_f$.

[5] Note that κ is equivalent to an integer indicating how many countries will be in C_{LDC}. This integer nature of κ and the implied nondifferentiability of the likelihood function have implications for the asymptotic standard errors of δ_c and ϕ_f, but I have not explored these. Having found the maximum-likelihood estimate of (κ, δ_c, ϕ_f), I treated the optimal κ as a fixed constant and calculated the standard errors of (δ_c, ϕ_f) from the second derivative of the log-likelihood function with respect to (δ_c, ϕ_f).

TABLE 1—HYPOTHESIS TESTING AND MODEL SELECTION

Hypothesis	Description: Parameters (k_i)	Description: Equation	Likelihood: $\ln(L_i)$	Likelihood: Schwarz criterion	Mysteries: Endowment paradox	Mysteries: Missing trade	Goodness-of-fit: Weighted sign	Goodness-of-fit: $\rho(F, \hat{F})$
Endowment differences								
H_0: unmodified HOV theorem	(0)	(1)	−1,007	−1,007	−0.89	0.032	0.71	0.28
Technology differences								
T_1: neutral	δ_c (32)	(4)	−540	−632	−0.17	0.486	0.78	0.59
T_2: neutral and nonneutral	ϕ_f, δ_c, κ (41)	(6)	−520	−637	−0.22	0.506	0.76	0.63
Consumption differences								
C_1: investment/services/ nontrade.	β_c (32)	(7)	−915	−1,006	−0.63	0.052	0.73	0.35
C_2: Armington	α_c^* (24)	(11)	−439	−507	−0.42	3.057	0.87	0.55
Technology and consumption								
TC_1: $\delta_c = y_c/y_{US}$	(0)	(4)	−593	−593	−0.10	0.330	0.83	0.59
TC_2: $\delta_c = y_c/y_{US}$ and Armington	α_c^* (24)	(12)	−404	−473	0.18	2.226	0.93	0.67

Notes: Here k_i is the number of estimated parameters under hypothesis i. For "likelihood," $\ln(L_i)$ is the maximized value of the log-likelihood function, and the Schwarz-model selection criterion is $\ln(L_i) - k_i \ln(297)/2$. Let $\hat{F}_{fc}$ be the predicted value of F_{fc}. The "endowment paradox" is the correlation between per capita GDP, y_c, and the number of times $\hat{F}_{fc}$ is positive for country c (see Fig. 2). "Missing trade" is the variance of F_{fc} divided by the variance of $\hat{F}_{fc}$ (see Fig. 1). "Weighted sign" is the weighted proportion of observations for which F_{fc} and $\hat{F}_{fc}$ have the same sign. Finally, $\rho(F, \hat{F})$ is the correlation between F_{fc} and $\hat{F}_{fc}$. See Section V for further discussion.

rejects H_0 and T_1 in favor of the unrestricted model T_2. The test statistics are $X^2_{[41]} = 974$ and $X^2_{[9]} = 40$, respectively.[6]

The favored model, T_2, is heavily parameterized. A model-selection criterion that penalizes models with many parameters is the Schwarz criterion, here stated as $\ln(L_i) - k_i \ln(FC)/2$ where L_i is the maximized value of the likelihood function under hypothesis i and $FC = 9 \times 33 = 297$ is the number of observations. The Schwarz criterion favors hypotheses with large (close to zero) values of the criterion, namely, T_1.

Table 2 presents the estimates of δ_c from equation (6) for the unrestricted model T_2 and the neutral technology-differences model T_1. Countries are ordered by per capita GDP, y_c; y_c/y_{US} appears in column (i). The estimate of κ (relevant only for hypothesis T_2) implies that the top 10 countries in the table are included in C_{LDC} and that the remaining countries are included in C_{DC}. The δ_c under the two hypotheses are similar, so that attention is restricted to T_1. I offer three criteria for evaluating the reasonableness of the δ_c's.

First, they must be nonnegative; otherwise, factor inputs yield negative outputs. All of the estimated δ_c's are positive.

Second, the United States is among the most productive countries in the world so that δ_c should be less than unity for most countries, as is the case. Arguably, the δ_c's for countries such as Japan and West Germany are too low. Note though that for the countries with the largest δ_c (Switzerland, 0.79; West Germany, 0.78; France, 0.74; Denmark, 0.73; Netherlands, 0.72; Japan, 0.70), the δ_c's are similar, and the hypothesis that they are the same cannot be rejected. This suggests that it is the United States that has the unusual value of δ_c.

Third, international productivity differences should be reflected in international per capita income differences. For example, if West Germany is only 78 percent as productive as the United States ($\delta_{GER} = 0.78$) then one expects West German per capita

[6] The model with only nonneutral effects (ϕ_f unrestricted and $\delta_c = 1$) is rejected by the data: the likelihood value is −870.

TABLE 2—ESTIMATES OF δ_c FOR 1983

		T_1		T_2	
Country	y_c/y_{US} (i)	δ_c (ii)	t (iii)	δ_c (iv)	t (v)
Bangladesh	0.04	0.03	47.71	0.04	42.28
Pakistan	0.08	0.09	32.10	0.09	34.93
Indonesia	0.11	0.10	39.51	0.13	38.21
Sri Lanka	0.12	0.09	14.85	0.07	20.13
Thailand	0.16	0.17	23.80	0.21	20.17
Colombia	0.21	0.16	18.41	0.29	11.14
Panama	0.23	0.28	3.24	0.24	4.44
Yugoslavia	0.30	0.29	11.35	0.19	18.83
Portugal	0.30	0.14	9.63	0.10	14.78
Uruguay	0.31	0.11	19.46	0.22	9.40
Greece	0.35	0.45	4.63	0.46	4.88
Ireland	0.39	0.55	2.91	0.56	3.08
Spain	0.41	0.42	9.40	0.43	9.88
Israel	0.60	0.49	2.91	0.50	3.03
Hong Kong	0.61	0.40	4.12	0.41	4.35
New Zealand	0.62	0.38	7.89	0.38	8.42
Austria	0.65	0.60	3.03	0.62	3.13
Singapore	0.66	0.48	2.11	0.49	2.20
Italy	0.66	0.60	7.16	0.62	7.38
United Kingdom	0.66	0.58	8.04	0.60	8.30
Japan	0.66	0.70	7.15	0.71	7.25
Belgium	0.67	0.65	2.73	0.66	2.79
Trinidad	0.69	0.47	1.25	0.48	1.30
Netherlands	0.69	0.72	2.66	0.73	2.69
Finland	0.70	0.65	2.17	0.67	2.22
Denmark	0.72	0.73	1.92	0.74	1.94
West Germany	0.73	0.78	3.80	0.80	3.74
France	0.73	0.74	4.84	0.75	4.85
Sweden	0.75	0.57	4.09	0.58	4.25
Norway	0.82	0.69	1.80	0.70	1.83
Switzerland	0.91	0.79	1.41	0.81	1.37
Canada	0.95	0.55	9.82	0.56	10.29
United States	1.00	1.00	—	1.00	—
$\rho(\delta_c, y_c)$:		0.91		0.90	

Notes: Here, t is the asymptotic t statistic for the null hypothesis that $\delta_c = 1$. T_1 and T_2 indicate the restricted ($\phi_f = 1$) and unrestricted estimates, respectively, of equation (6): y_c is per capita GDP; $\rho(\delta_c, y_c)$ is the correlation of δ_c with y_c. The line in column (iv) separates $c \in C_{LDC}$ from $c \in C_{DC}$.

GDP to be 78 percent of U.S. per capita GDP ($y_{GER}/y_{US} = 0.78$). As expected, the δ_c's are highly correlated with y_c, a remarkable 0.91. A related observation is used by Trefler (1993) in discussing international wage differences.

Table 3 reports the estimated ϕ_f's under hypothesis T_2. They are reasonable in three senses. First, all are positive as required by the theory. Second, let a_{fic} be the amount of factor f needed to produce one unit of good i in country c. Then, for example, a_{Kic}/a_{Lic} is the capital–labor ratio in industry i. One expects rich countries to use relatively capital-intensive techniques, that is,

$$a_{Ki,DC} / a_{fi,DC} > a_{Ki,LDC} / a_{fi,LDC}$$

for all i and f ($f \neq K$). Under hypothesis T_2, this reduces to $\phi_K > \phi_f$ for all f ($f \neq$

TABLE 3—NONNEUTRALITY PARAMETERS ϕ_f UNDER T_2

	T_2	
Factor	ϕ_f	t
Capital	1.89	−2.29
Labor		
Professional and technical	1.47	−1.71
Clerical	1.82	−2.45
Sales	0.37	2.41
Service	0.91	0.36
Agriculture	0.05	4.63
Production and transport	0.74	0.88
Land		
Cropland	1.36	−1.48
Pasture	0.38	4.01

Notes: Here, t is the asymptotic t statistic for the hypothesis $\phi_f = 1$. T_2 indicates the unrestricted estimates of equation (6). It allows for neutral (δ_c) and nonneutral (ϕ_f) technology differences.

K).[7] As indicated in Table 3, the estimated ϕ_f's satisfy this condition. Third, one expects rich countries to be more productive (use less factors per unit of output) than poor countries: $a_{fi,\mathrm{DC}} < a_{fi,\mathrm{LDC}}$ for all i and f. Under hypothesis T_2 this is equivalent to $\delta_{\mathrm{DC}} > \delta_{\mathrm{LDC}}\phi_f$ for all f. From Table 2 the largest value of δ_{LDC} is 0.29, and from Table 3 the largest value of ϕ_f is 1.89, so that the condition is $\delta_{\mathrm{DC}} > 0.55$. The ϕ_f's are reasonable in that this condition is satisfied for the richest countries in the sample. Not all the ϕ_f are sensible. In particular, it is not clear why clerical labor and cropland have such large ϕ_f's.

The international-technology-differences model not only does well fitting the data for 1983, it also does well fitting the Bowen et al. (1987) data for 1966–1967. I estimated T_1 using exactly the same methodology as for 1983, except that the choice of countries was dictated by the Bowen et al. data. There are $9 \times 27 = 243$ observations. The results appear in column (ii) of Table 4. As in 1983, the δ_c all lie between zero and unity, and the correlation of δ_c with per capita GDP is a high 0.76.

Bowen et al. (1987) also investigated T_1 but found that it performed poorly. For a related hypothesis, their estimated values of δ_c were poor: only three were between 0 and 1, and the range was (−174, 19). This led Bowen et al. to conclude that "The Heckscher-Ohlin model does poorly, but we do not have anything that does better" (p. 805). This contrasts sharply with my estimates and conclusions. The difference is primarily related to an implementation problem in Bowen et al.'s paper. They estimated variants of

$$F^m_{fc} = \delta_c V_{fc} - s_c \Sigma_j V_{fj} + \mu_{fc}$$

when they should have estimated variants of

$$F^m_{fc} = \delta_c V_{fc} - s_c \Sigma_j \delta_j V_{fj} + \mu_{fc}$$

[i.e., my equation (4)].[8]

IV. Economically Meaningful Alternative Hypotheses: Consumption

In this section, I consider three alternatives to the HOV consumption assumption, two of which are suggested by the missing trade and endowments paradox. Throughout, let $\mathbf{C}_c$ and $\mathbf{Q}_c$ be the vectors of consumption and production in country c, respectively, and define world values $\mathbf{C}_w = \Sigma_c \mathbf{C}_c$ and $\mathbf{Q}_w = \Sigma_c \mathbf{Q}_c$. The HOV homothetic-consumption assumption implies that $\mathbf{C}_c = s_c \mathbf{C}_w$.

A. *Investment, Services, and Nontradables*

The endowments paradox states that poor countries are abundant in most factors and rich countries are scarce in most factors. Let s_{LDC} and s_{DC} be representative consumption shares for poor and rich countries, respectively. As illustrated in Figure 3, the endowments paradox states that s_{LDC} is far left on the Vanek chain and s_{DC} is far right. Let β_c denote the "true" consumption share and suppose

[7] The assumption that $\mathbf{A}_{\mathrm{US}} = \mathbf{\Pi}_c \mathbf{A}_c$ may be rewritten as $\mathbf{\Pi}_{\mathrm{DC}}\mathbf{A}_{\mathrm{DC}} = \mathbf{\Pi}_{\mathrm{LDC}}\mathbf{A}_{\mathrm{LDC}}$ or $\pi_{f,\mathrm{DC}} a_{fi,\mathrm{DC}} = \pi_{f,\mathrm{LDC}} a_{fi,\mathrm{LDC}}$. Under hypothesis T_2, this reduces to $\delta_{\mathrm{DC}} a_{fi,\mathrm{DC}} = \delta_{\mathrm{LDC}} \phi_f a_{fi,\mathrm{LDC}}$ so that $a_{\mathrm{K}i,\mathrm{DC}}/a_{fi,\mathrm{DC}} = (\phi_{\mathrm{K}}/\phi_f) a_{\mathrm{K}i,\mathrm{LDC}}/a_{fi,\mathrm{LDC}}$.

[8] More precisely, Bowen et al.'s (1987) H2 corresponds to my T_1, and the problem can be seen by plugging in their parameter restrictions for H2 (from their table 4) into their equation (14). The problem spills over only to H3, their preferred hypothesis, which underlies the above quotation and estimates of δ_c.

TABLE 4—ESTIMATES OF δ_c AND α_c^* FOR 1966–1967

		T_1		TC_2	
Country	y_c/y_{US} (i)	δ_c (ii)	t (iii)	α_c^* (iv)	t (v)
Korea	0.10	0.07	15.11	0.42	2.73
Philippines	0.11	0.11	17.70	0.20	4.10
Brazil	0.17	0.22	19.68	0.12	9.64
Portugal	0.19	0.22	5.75	0.11	3.33
Yugoslavia	0.20	0.21	12.04	0.18	4.60
Greece	0.24	0.41	4.58	0.15	3.64
Hong Kong	0.28	0.17	5.02	1.09	−0.27
Mexico	0.30	0.33	11.95	0.23	6.38
Argentina	0.32	0.34	13.01	0.37	5.18
Ireland	0.35	0.54	1.92	0.19	1.67
Spain	0.40	0.43	8.14	0.05	6.47
Japan	0.41	0.41	17.95	0.17	16.28
Italy	0.48	0.67	5.56	0.06	14.00
Austria	0.49	0.48	4.49	−0.05	2.10
Finland	0.52	0.55	3.51	0.60	0.53
Belgium	0.58	0.65	3.07	0.26	2.60
Norway	0.59	0.48	3.48	0.50	1.08
United Kingdom	0.59	0.57	8.67	0.18	15.95
France	0.61	0.61	8.33	0.21	5.13
Netherlands	0.61	0.54	4.93	−0.43	4.64
West Germany	0.63	0.57	10.95	0.31	7.31
Denmark	0.70	0.68	2.16	−1.10	3.37
Sweden	0.72	0.65	3.65	0.56	1.54
Australia	0.73	0.13	29.72	0.13	38.68
Canada	0.77	0.59	7.44	0.24	12.89
Switzerland	0.93	0.52	4.41	0.08	6.38
United States	1.00	1.00	—	0.13	33.41
$\rho(\delta_c, y_c)$:		0.76			

Notes: Here y_c is per capita GDP, and t is the t statistic for the null hypothesis that $\delta_c = 1$ or $\alpha_c^* = 1$. T_1 is the restricted ($\phi_f = 1$) version of equation (6). TC_2 refers to equations (12)–(13) with $\delta_c = y_c/y_{US}$. $\rho(\delta_c, y_c)$ is the correlation of δ_c with y_c.

it were different from the measured consumption share s_c. If $s_{LDC} < \beta_{LDC}$ and $\beta_{DC} < s_{DC}$ then the endowments paradox would disappear.

The most obvious source of miscalculated consumption shares is investment: rich countries consume less than is indicated by the income-based measure $s_c = (Y_c - B_c)/Y_w$ because they devote part of their income to investment. Instead of starting the derivation of the HOV equation from $\mathbf{T}_c = \mathbf{Q}_c - \mathbf{C}_c$ and s_c defined by $\mathbf{C}_c = s_c\mathbf{C}_w$, begin with $\mathbf{T}_c = \mathbf{Q}_c - \mathbf{C}_c - \mathbf{Z}_c$, where $\mathbf{Z}_c$ is a vector of investment goods, and define β_c by $\mathbf{C}_c = \beta_c\mathbf{C}_w$. Let $\mathbf{p}$ be the output price vector. Then it is straightforward to show the following. First, $\beta_c = (Y_c - B_c - \mathbf{p}'\mathbf{Z}_c)/(Y_w - \Sigma_j\, \mathbf{p}'\mathbf{Z}_j)$ so that heavily investing rich countries have $\beta_{DC} < s_{DC}$ as required. Second,

$$(7) \qquad F_{fc} = V_{fc} - \beta_c V_{fw} + \mu_{fc} \qquad \text{subject to } \sum_c \beta_c = 1$$

where $(\mu_{1c}, \ldots, \mu_{Fc})' = \mathbf{A}(\mathbf{Z}_c - \beta_c\, \Sigma_j\, \mathbf{Z}_j)$. Since data on $\mathbf{Z}_c$ (investment levels by industry for country c) are not available for poor countries, equation (7) may be treated as a linear regression model with slopes β_c and errors μ_{fc}. It turns out that *exactly* the same analysis goes through with $\mathbf{Z}_c$ interpreted as services

TABLE 5—ESTIMATES OF CONSUMPTION PARAMETERS β_c AND α_c^*

	C_1		C_2		TC_2	
Country	β_c/s_c (i)	t (ii)	α_c^* (iii)	t (iv)	α_c^* (v)	t (vi)
Bangladesh	10.33	−9.07	0.10	66.35	0.42	1.94
Pakistan	5.62	−6.58	0.08	45.15	0.35	2.48
Indonesia	4.42	−7.94	0.00	54.19	0.30	5.32
Sri Lanka	7.24	−3.68	0.14	20.04	0.27	1.27
Thailand	3.25	−3.66	0.21	34.35	0.82	1.35
Colombia	2.57	−2.46	0.11	27.19	0.07	5.86
Panama	2.17	−0.65	0.38	3.42	0.51	0.39
Yugoslavia	2.31	−2.55	−0.01	15.24	−0.52	3.86
Portugal	2.74	−2.18	0.02	12.38	−0.53	3.62
Uruguay	4.62	−2.34	0.03	27.53	0.14	7.93
Greece	1.46	−0.75	0.24	4.54	0.20	1.55
Ireland	1.29	−0.30	0.37	3.79	0.40	1.38
Spain	1.40	−1.36	−0.20**	14.78	0.07	6.52
Israel	1.15	−0.20	0.27	3.11	0.29	1.78
Hong Kong	1.49	−0.70	0.44	3.73	0.56	2.11
New Zealand	1.17	−0.21	0.22	12.75	0.28	7.76
Austria	1.06	−0.12	0.23	2.71	−0.07	2.66
Singapore	1.13	−0.15	0.48	1.97	0.33	1.89
Italy	1.08	−0.40	0.06	10.12	0.50	3.92
United Kingdom	1.04	−0.24	0.16	20.20	0.48	8.39
Japan	0.86	1.30	0.05	47.27	0.30	17.28
Belgium	0.97	0.08	0.23	3.87	0.83	0.64
Trinidad	1.23	−0.15	0.70	0.36	1.07	−0.07
Netherlands	0.74	0.76	−0.41**	8.96	−0.43	7.22
Finland	0.95	0.09	0.29	2.24	0.67	0.79
Denmark	0.84	0.29	0.01	3.41	−0.20	3.11
West Germany	0.82	1.19	0.14	16.12	0.53	6.21
France	0.88	0.73	−0.31**	13.89	−0.02	7.05
Sweden	1.00	0.01	0.24	5.38	0.48	3.04
Norway	0.84	0.29	0.38	2.90	0.39	2.30
Switzerland	0.78	0.56	0.28	4.10	0.44	2.51
Canada	1.18	−0.81	0.34	15.24	0.34	15.95
United States	0.78	4.13	0.02	74.23	0.37	16.87

Notes: In this table, t is the t statistic for the null hypothesis that $\beta_c/s_c = 1$ or the null hypothesis that $\alpha_c^* = 1$. C_1 is the investment equation (7), C_2 indicates the Armington equations (10)–(11), below, and TC_2 indicates the Armington plus technology differences equations (12)–(13), below.

** Coefficient is statistically negative at the 1-percent significance level.

and $\mathbf{C}_c$ as merchandise goods. Then $\beta_{DC} < s_{DC}$ because, empirically, rich countries devote proportionately more of their income to services and less of their income to merchandise goods (i.e., income [s_{DC}] overstates merchandise goods consumption [β_{DC}]). A similar argument can be made for $\mathbf{Z}_c$ interpreted as nontradables.

Column (i) of Table 5 reports the OLS estimates of equation (7). For ease of interpretation, I report β_c/s_c and the t statistics for the hypothesis $\beta_c/s_c = 1$. I also sort countries by per capita GDP to highlight the negative correlation between β_c/s_c and per capita GDP (−0.77). As is apparent from the table, $s_{LDC} < \beta_{LDC}$ and $\beta_{DC} < s_{DC}$, exactly as required to eliminate the endowments paradox. Such good parameter estimates are satisfying, but surprisingly the model barely outperforms the HOV model. (The Schwarz criterion is −1,006, compared to −1,007 for the HOV theorem. See the row for C_1 in Table 1.) It is thus clear that a large class of models—which include investment, services, and nontradables—offer only limited insights into the poor performance of the HOV theorem. This is surprising because the hypotheses were custom-tailored to explain the endowments paradox and because the parameter estimates were as expected.

B. *Armington Home Bias*

It has frequently been observed that consumers display a bias toward domestically produced goods. Whatever the sources of this Armington (1969) behavior, the implication is missing trade relative to the HOV prediction. Consider generalizing the HOV demand assumption $\mathbf{C}_c = s_c\mathbf{C}_w = s_c\mathbf{Q}_w$ by retaining linearity but distinguishing between domestic goods $\mathbf{Q}_c$ and foreign goods $\mathbf{Q}_w - \mathbf{Q}_c$:

$$(8) \quad \mathbf{C}_c = s_c[\alpha_c\mathbf{Q}_c + \alpha_c^*(\mathbf{Q}_w - \mathbf{Q}_c)]$$

where $\alpha_c > 1$ and $\alpha_c^* < 1$ capture home bias.

Premultiplying equation (8) by $\mathbf{p}'$ to impose budget balance yields

$$(9) \quad \alpha_c(Y_c/Y_w) + \alpha_c^*(1 - Y_c/Y_w) = 1.$$

World market-clearing, $\Sigma_c \mathbf{C}_c = \Sigma_c \mathbf{Q}_c$, implies the following:[9]

$$(10) \quad \sum_c s_c\left[(1 - \alpha_c^*)\frac{Y_w}{Y_c} V_{fc} + \alpha_c^* V_{fw}\right] = V_{fw}.$$

Equation (10) imposes F restrictions on the α_c^*'s. Assuming measurement error $F_{fc}^m = F_{fc} + \mu_{fc}$, the variant of the HOV equation implied by the Armington assumption is

$$(11) \quad F_{fc} = V_{fc} - s_c\left[(1 - \alpha_c^*)\frac{Y_w}{Y_c} V_{fc} + \alpha_c^* V_{fw}\right] + \mu_{fc}$$

where the α_c^* satisfy equation (10). Equations (10)–(11) nest the HOV equation (1) for $\alpha_c^* = 1$.

I estimate the α_c^* in equation (11) using OLS subject to the linear restrictions on α_c^* given by equation (10). Column (iii) of Table 5 reports the estimates. Many of them are unexpectedly small. For example, the United States gives only the small weight of $\alpha_c^* = 0.02$ to foreign goods. Also, it is difficult to interpret the negative α_c^*. Nevertheless, all of the α_c^* are less than unity as suggested by Armington home bias. Further, from the row for C_2 in Table 1, the model outperforms the HOV theorem (H_0) when judged by likelihood and Schwarz values.[10]

The strength of the results argues for combining the Armington and technology-difference models. For tractability, consider the neutral technology-difference model T_1. The resulting variant of the HOV equation is

$$(12) \quad F_{fc}^m = \delta_c V_{fc} - s_c\left[(1 - \alpha_c^*)\frac{Y_w}{Y_c}\delta_c V_{fc} + \alpha_c^* \sum_j \delta_j V_{fj}\right] + \mu_{fc}$$

where the α_c^* satisfy

$$(13) \quad \sum_c s_c\left[(1 - \alpha_c^*)\frac{Y_w}{Y_c}\delta_c V_{fc} + \alpha_c^* \sum_j \delta_j V_{fj}\right] = \sum_c \delta_c V_{fc}.$$

To reduce the large number of parameters (δ_c, α_c^*) from 56 to 24, I set δ_c equal to per capita GDP relative to U.S. per capita GDP, $\delta_c = y_c/y_{US}$. Statistics for the Armington model combined with $\delta_c = y_c/y_{US}$ appear in row TC_2 of Table 1. The restriction $\alpha_c^* = 1$ is rejected (from row TC_1 of Table 1, $X_{[24]}^2 = 378$), indicating a complementarity between

[9] To derive equation (10), solve for α_c using equation (9), plug the result into equation (8), sum across countries, premultiply by $\mathbf{A}$, and use $\mathbf{AQ}_c = \mathbf{V}_c$ and $\mathbf{A}\,\Sigma_c\,\mathbf{C}_c = \mathbf{AC}_w = \mathbf{AQ}_w = \mathbf{V}_w$.

[10] A very different consumption model is the linear expenditure system here stated as $\mathbf{C}_c = \boldsymbol{\theta}^0 L_c + \boldsymbol{\theta}^1 s_c$ where L_c is population. This assumption leads to the estimating equation $F_{fc}^m = V_{fc} - s_c V_{fw} - \theta_f(L_c - s_c L_w) + \mu_{fc}$. Estimating the nine θ_f parameters using OLS yields a likelihood value of −806 which, from Table 1, is statistically better than the HOV and investment hypotheses (H_0 and C_1), but worse than the Armington hypothesis (C_2).

home bias and technology differences.[11] The restricted OLS estimates of the α_c^* are reported in column (v) of Table 5. There is an improvement in the estimates as judged by the disappearance of statistically significant negative α_c^*'s and by the many parameters that have increased in size, particularly α_{US}^*.

The exact interpretation of the α_c^*'s is not clear in that the sources of Armington home bias have not been identified. To investigate, consider a simple model implying consumption behavior as in equation (8). There is one good, and there are two countries. Each country produces one variety of the good. Letting an asterisk denote the foreign variety, D_c and D_c^* are the country-c demands for the two varieties, and p_c and p_c^* are the respective prices prevailing in country c. The representative consumer maximizes $\rho_c \ln D_c + (1-\rho_c)\ln D_c^*$, subject to the budget constraint $p_c D_c + p_c^* D_c^* = Y_c - B_c$. Prices are the same internationally except for trade restrictions and transport-cost markups: $p_c^* = \tau_c p_b$ for some constant $\tau_c \geq 1$ and $p_b^* = \tau_b p_c$ for some constant $\tau_b \geq 1$. Since trade is observed for the good, but not for the variety, interest centers on $C_c = D_c + D_c^*$. Solving for the general equilibrium of this model yields equation (8) with the following condition:[12]

$$(14) \quad \alpha_c^* = \frac{1-\rho_c}{(1-\rho_c)s_c + \rho_b(1-s_c)\tau_c}.$$

Since $\partial\alpha_c^*/\partial\rho_c < 0$ and $\partial\alpha_c^*/\partial\tau_c < 0$, the home bias implied by the estimated α_c^* in Table 5 may reflect either primitive preference bias toward the home good or high tariffs and transport costs. Somewhat surprisingly, $\partial\alpha_c^*/\partial s_c$ cannot be signed even under strong assumptions about trade impediments and preference differences. This may explain the weak correlations in Table 5 between the estimated α_c^* and per capita GDP.

Equation (14) suggests a multivariate regression of α_c^* on ρ_c, s_c, and τ_c. While preferences ρ_c have no observable counterpart, my 1983 data together with the Bowen et al. 1966–1967 data form a panel that identifies the ρ_c under the assumption that preferences are stable over time. Letting t index time, s_{ct} is observed, τ_{ct} is measured by c.i.f./f.o.b. factors and average tariffs, and α_{ct}^* is measured by its estimate $\hat{\alpha}_{ct}^*$ under hypothesis TC_2. Since α_{ct}^* is measured with error, write $\hat{\alpha}_{ct}^* = \psi_t + \alpha_{ct}^* + \mu_{ct}$, where μ_{ct} is unsystematic measurement error and ψ_t captures systematic differences in the construction of the 1966–1967 and 1983 data sets. Linearizing equation (14) and substituting in $\hat{\alpha}_{ct}^*$ yields

$$(15) \quad \hat{\alpha}_{ct}^* = \psi_t + \psi_s s_{ct} + \psi_{\text{cif}}(\text{c.i.f./f.o.b.})_{ct} + \psi_{\text{tar}}(\text{tariff})_{ct} + (\rho_c + \mu_{ct})$$

where, because of the short two-year panel, ρ_c is treated as a random effect.

There are 18 countries for which all the necessary data are available.[13] The estimated α_{ct}^*'s for 1966–1967 under hypothesis TC_2 appear in column (iv) of Table 4 and contain no surprises. The first line of Table 6 reports the generalized least-squares (GLS) estimates of

[11] The mix of $\delta_c = y_c/y_{US}$ with the linear expenditure system or C_1 leads to worse likelihood values of -584 and -547, respectively.

[12] The details are as follows. From utility maximization $D_c = (Y_c - B_c)\rho_c/p_c$ and $D_c^* = (Y_c - B_c)(1-\rho_c)/p_c^*$. There are four prices (p_b, p_b^*, p_c, and p_c^*) in the two countries (b and c). The four equilibrium conditions are as follows:

(i) Supply of the country-c variety equals world demand: $Q_c = D_c + D_b^*$.
(ii) Likewise, $Q_b = D_b + D_c^*$.
(iii) $p_c^* = \tau_c p_b$.
(iv) $p_b^* = \tau_b p_c$.

To obtain equation (14), solve for prices, plug prices into D_c and D_c^*, and plug (D_c, D_c^*) into $C_c = D_c + D_c^*$. The implications of $\tau_b \neq 1$ and $\tau_c \neq 1$ for factor price equalization and s_c are discussed in Leamer (1988).

[13] These are the 21 countries that appear in both the 1966–1967 and 1983 data sets less Japan, Hong Kong, and Portugal, for which tariff data are unavailable. The tariff data are from two sources: (i) *IFS Supplement on Government Finance,* 1986 series, "Taxes on International Trade and Transactions (A.6) as Percent of Total Revenue (S.2)" and "Total Revenue (S.2) as Percent of GDP" and (ii) IFS series "GDP (99b)" and "Imports (98c)." (For Bangladesh and Singapore, 71v replaces 98c.) The data do not always match the 1966–1967 and 1983 dates, most often because the tariff data start in the early 1970's. The c.i.f./f.o.b. factors are from the *IFS Supplement on Balance of Payments, 1984.*

TABLE 6—DETERMINANTS OF ARMINGTON HOME BIAS IN CONSUMPTION α^*_{ct}

Dependent variable: $\hat{\alpha}^*_{ct}$	Income s_{ct}	c.i.f./f.o.b. τ_{ct}	Tariff τ_{ct}	Years	n	R^2
Random effects, GLS	0.29	−2.38	−0.24	1966, 1983	36	0.06
	(0.49)	(−1.33)	(−0.16)			
Simple correlation	0.12	−0.39	−0.29	1983	18	
	(0.37)	(0.89)	(0.76)			
OLS	0.34	−4.52	−2.14	1983	18	0.19
	(0.36)	(−1.32)	(−0.75)			

Notes: The table reports estimates of equation (15) using $\hat{\alpha}^*_{ct}$ from TC_2. For the GLS and OLS rows, t statistics are in parentheses, and intercepts are not reported. For the GLS row, the year dummy coefficient is not reported. For the correlation row, p values are in parentheses.

the random-effects model.[14] The model correctly predicts the signs on c.i.f./f.o.b. factors and average tariffs: higher trade barriers lead to "missing trade." The variance of the random effects ρ_c is twice the variance of μ_{ct}, indicating that international preference differences are important. The model makes no prediction about the coefficient sign on s_c; however, the estimated coefficient is positive, indicating that larger economies tend to be inward-oriented. The size and noisiness of the sample (note the statistical insignificance and low goodness-of-fit) argue for a simpler model. Table 6 also reports simple correlations and OLS estimates for the 1983 sample alone. The same sign pattern repeats itself, a fact that lends some confidence to the estimates.

These results point to the benefits of more research into the Armington sources of the case of the missing trade. For present purposes, I am not bothered by this: my main point is that the bias is important and must be confronted theoretically and empirically.

V. Model Selection: Economic Criteria

This section evaluates the performance of each model using the same criteria applied to reject the HOV theorem. The criteria are based on $\hat{F}_{fc}$, the prediction of F_{fc}. For example, under the HOV theorem, $\hat{F}_{fc} \equiv V_{fc} - s_c V_{fw}$; and under TC_2, $\hat{F}_{fc}$ is defined as the right-hand side of equation (12) with $\mu_{fc} = 0$.

1. *Correlation of F_{fc} with $\hat{F}_{fc}$.*—Under the HOV theorem the correlation between F_{fc} and $\hat{F}_{fc}$ is a weak 0.28 and is not visually apparent in Figure 1. The last column of Table 1 reports the correlation for each model. The correlation rises to 0.67 when the neutral technology and Armington models are combined.

2. *Sign-HOV.*—Under the HOV theorem, the sign of $\hat{F}_{fc}$ correctly predicts the sign of F_{fc} for only 148 of 297 observations or slightly worse than a coin-toss prediction. Weighting by the size of the factor content of trade, the statistic rises to 0.71. This appears in the "weighted sign" column of Table 1. The statistic rises to 0.93 when the neutral-technology and Armington models are combined. (The unweighted sign statistics for H_0, T_1, C_2, and TC_2 are 0.50, 0.62, 0.64, and 0.72, respectively, showing comparable improvement.)

3. *Endowments Paradox.*—In Figure 2, the number of times $\hat{F}_{fc}$ is positive for country c was seen to be negatively correlated with per capita GDP. The correlation of −0.89 is reported in the "endowment paradox" column of Table 1. When the neutral-technology and Armington models are combined, a sensible result obtains: there is a weak positive correlation of 0.18 between per capita GDP and abundance.

4. *Case of the Missing Trade.*—Let σ^2 be the variance of F_{fc} and let $\hat{\sigma}^2$ be the variance of $\hat{F}_{fc}$. Under the HOV hypothesis, $F_{fc} = \hat{F}_{fc}$

[14] Weighting by the covariance matrix of the $\hat{\alpha}^*_{ct}$ or using per capita GDP in place of s_c does not alter the conclusions.

so that $\sigma^2 = \hat{\sigma}^2$. From Figure 1, the case of the missing trade can be summarized by noting that $\sigma^2/\hat{\sigma}^2 = 0.032$ (i.e., the variance is off by a factor of 32). This is reported in Table 1 in the "missing trade" column. Under TC_2 the variance ratio is 2.2, so that the second moment is only off by a factor of 2.2. Thus, TC_2 represents a significant improvement over the HOV theorem in dealing with the case of the missing trade.

By all these criteria the HOV theorem is dominated by a model allowing for Armington home bias and neutral technology differences. However, as indicated by the correlations between F_{fc} and $\hat{F}_{fc}$, not all the sample variation is explained. In particular, none of the models does well predicting observations for which $\hat{F}_{fc}$ is small. This suggests a tension between the theory and evidence. $\hat{F}_{fc}$ small means that the endowment of country c is "similar" to the world endowment in a cone-of-diversification sense. Thus, while the theory predicts that the model will do best where endowments are similar ($\hat{F}_{fc}$ small), the model actually does best where endowments are dissimilar ($\hat{F}_{fc}$ large).

VI. Caveats and Conclusions

The HOV theorem performs poorly and, by implication, so do increasing returns to scale and imperfect-competition models that yield the HOV theorem. Yet little is known about the features of national endowments and net factor service trade that lead to this negative result. The only known anomaly, namely, the Leontief "paradox," was reversed in the data at least 20 years ago (Stern and Maskus, 1981).

In other fields of economics, the poor performance of a major theory leads to more careful consideration of the data and to new theories that can accommodate the anomalies. Yet in international economics, such important facts as "the case of the missing trade" and "the endowments paradox" have gone unnoticed.

The first contribution of the present paper was an investigation into the features of the data that lead to the poor performance of the HOV theorem. I identified pronounced patterns in the deviations from the HOV theorem. In presenting them I offered an informative graphical display of the HOV theorem. I view this as the most important contribution of the paper: in place of the countless theoretical conjectures about why the HOV theorem performs poorly, there are now several data patterns around which theoretical analysis can coalesce.

The second contribution of the paper was an examination of alternative hypotheses that could potentially explain the patterns. Models that contributed little to the explanation included those with linear expenditure demand, capital accumulation, nontradables, and trade in services. The importance of the specific alternative models I have advanced is in indicating which models are consistent with the data patterns and which models are not. For example, both the investment model and the Armington model were custom-tailored to fit the data, yet the former was a failure while the latter was a success. The model that performed best combined Armington home bias with neutral international technology differences. The results contrast sharply with those of Bowen et al. (1987) and provide the first rejection of the HOV hypothesis in favor of a satisfying, economically meaningful alternative.

More remains to be done. I am dissatisfied that some alternative hypotheses were data-instigated, because it implies that the reported test statistics overstate the rejection of the HOV theorem. Also, work is needed to explain why endowment similarity is associated with poor predictions. Finally, more detailed data are needed to investigate further the sources of international productivity differences and Armington home bias.

Appendix: Measurement Error and Other Sensitivity Analysis

This appendix summarizes the impact of alternate specifications on estimates of the δ_c parameters for the neutral-technology-differences model T_1. This is a linear model, so that familiar techniques can be brought to bear.

1. *V_{fc} is mismeasured.*—Suppose measured endowments are related to actual endowments by $V^m_{fc} = V_{fc} + \mu^1_{fc}$, where the μ^1_{fc}'s are independently and identically distributed. Then from equation (4)

$$(A1) \quad F^{US}_{fc} = \delta_c V^m_{fc} - s_c \sum_j \delta_j V^m_{fj} + \mu^2_{fc}$$

where $\mu_{fc}^2 = -(\delta_c\mu_{fc}^1 - s_c \Sigma_j \delta_j\mu_{fj}^1)$. Since $E(\mu_{fc}^2 \mid V_{fc}^m) \neq 0$, instrumental-variables (IV) estimation is indicated; however, it seems impossible to imagine a valid set of instruments for a primitive such as endowments.

Alternatively, the V_{fc}^m can be brought to the left-hand side. The next two facts can be proved along the lines of Trefler (1993 [proof of proposition 1]). First, equation (A1) may be written as

$$(\text{A2}) \qquad Y_{fc} = \delta_c + \mu_{fc} \qquad c \neq \text{US}$$

where Y_{fc} is a function of the observables $(F_{fc}^{\text{US}}, V_{fc}^m, s_c)$ and μ_{fc} is a function of the observables and errors μ_{fc}^2. Hence, equation (A2) is a fixed-effects model with fixed effects δ_c. Second, the covariance of μ_{fc} is heteroscedastic block diagonal.

Under the null hypothesis of endowment mismeasurement the OLS estimator of the δ_c in equation (A2) is inefficient, but *not* biased. Although the estimated δ_c for the tiny land-scarce countries of Hong Kong and Singapore are peculiar, when land is omitted all the estimated δ_c's are as good as or better than those reported in the main text. For example, their correlation with per capita GDP rises to 0.95. Further, under the null hypothesis of no mismeasurement, the White test should reject heteroscedasticity (though I caution that the error structure is also block diagonal). In fact, the White test resoundingly rejects heteroscedasticity and hence mismeasurement ($X^2 = 3.2$, which is less than the 5-percent critical value of $\chi^2_{[31]} = 45.0$).

2. *Mismeasurement of* s_c.—There are two measures of s_c: the one from the World Bank used in this study and the one from the Penn World Tables. The estimates of the δ_c in equation (4) are not sensitive to the choice of s_c, indicating that measurement error is not a problem.

3. *Omitted Factors.*—Suppose data on factor g are not available. Inspection of equation (4) reveals that data for factor g do not enter into the equation for factor f ($f \neq g$). Thus, omitting a factor reduces sample size, but does not lead to omitted-variable bias.

4. *Omitted Countries.*—The World Bank data document 137 countries. I have complete data for countries $c = 1, \ldots, 33$, but only incomplete data for countries $c = 34, \ldots, 137$. Define $s_c' = (Y_c - B_c)/\Sigma_{j \leq 137} Y_j$ and $s_c = (Y_c - B_c)/\Sigma_{j \leq 33} Y_j$ and modify equation (4) to read $F_{fc}^{\text{US}} = \delta_c V_{fc} - s_c' \Sigma_{j \leq 137} \delta_j V_{fj}$. The World Bank GNP data show $s_c' = 0.79 s_c$ so that $s_c' \Sigma_{j \leq 137} \delta_j V_{fj} = s_c \Sigma_{j \leq 33} \delta_j V_{fj} + s_c \gamma_f$, where $\gamma_f = 0.79 \Sigma_{33 < j \leq 137} \delta_j V_{fj} - 0.21 \Sigma_{j \leq 33} \delta_j V_{fj}$. Hence, interest focuses on

$$(\text{A3}) \qquad F_{fc}^{\text{US}} = \delta_c V_{fc} - s_c \sum_{j \leq 33} \delta_j V_{fj} - s_c \gamma_f.$$

This is not stochastic, so assume that F_{fc}^{US} is mismeasured and estimate (δ_c, γ_f) in equation (A3) using OLS. The estimates of the γ_f are jointly significant, indicating that omitted countries may be an issue. However, the estimated δ_c from equation (18) are *identical* to those presented in the main text because, given the choice of scaling by $\sigma_f s_c^{1/2}$, each γ_f regressor is orthogonal to each δ_c regressor.

5. *Influential Observations.*—The omission of any single observation does not lead to large changes or sign reversals in the estimated δ_c, nor does omission of any group of observations associated with a single factor. However, there are influential groups of observations associated with some of the poor countries. An interesting pattern presents itself. Order country indexes by per capita GDP so that $y_c < y_{c+1}$. When country c is omitted and it is poor, then it induces a large change in δ_{c+1}, but not in any of the other coefficients.

6. *Fixed Effects.*—Including country or factor fixed effects makes little difference to the performance of the model.

7. *Additional Factors.*—Working with the 14 factors listed in footnote 1 rather than the nine factors that I use does not change the estimates of the δ_c. The case of the missing trade is not present for the oil, coal, and mineral endowments, but this may be because I have mismeasured them as flows when they should be stocks.

8. *Scaling.*—I have scaled each observation (f, c) by $\sigma_f s_c^{1/2}$. Similar results obtain when scaling by V_{fw} and, in a regression setting, when using a GLS correction involving an ancillary log regression of the first-stage squared residuals on (i) V_{fw}, σ_f, or factor-specific components and (ii) s_c or country-specific components. Only the $\log(s_c)$ scaling leads to substantially different results, and this only when its coefficient is large (i.e., when the United States and Japan receive almost no weight).

REFERENCES

Armington, Paul S. "A Theory of Demand for Products Distinguished by Place of Production." *International Monetary Fund Staff Papers,* March 1969, *16*(1), pp. 159–78.

Bowen, Harry P.; Leamer, Edward E. and Sveikauskas, Leo. "Multicountry, Multifactor Tests of the Factor Abundance Theory." *American Economic Review,* December 1987, *77*(5), pp. 791–809.

Brecher, Richard A. and Choudhri, Ehsan U. "The Factor Content of Consumption in Canada and the United States: A Two-Country Test of the Heckscher-Ohlin-Vanek Model," in Robert C. Feenstra, ed., *Empirical methods for international trade.* Cambridge, MA: MIT Press, 1988, pp. 5–17.

Heckscher, Eli F. "The Effect of Foreign Trade on the Distribution of Income" [in Swedish]. *Ekonomisk Tidskrift,* 1919, *21*(2), pp. 1–32; reprinted in *Readings in the theory of international trade.* Homewood, IL: Irwin, 1950, pp. 272–300.

Helpman, Elhanan and Krugman, Paul R. *Market structure and foreign trade: Increasing returns, imperfect competition, and the international economy.* Cambridge, MA: MIT Press, 1985.

Leamer, Edward E. "The Leontief Paradox, Reconsidered." *Journal of Political Economy,* June 1980, *88*(3), pp. 495–503.

______. *Sources of international comparative advantage: Theory and evidence.* Cambridge, MA: MIT Press, 1984.

______. "Cross-Section Estimation of the Effects of Trade Barriers," in Robert C. Feenstra, ed., *Empirical methods for international trade.* Cambridge, MA: MIT Press, 1988, pp. 51–82.

Leamer, Edward E. and Bowen, Harry P. "Cross-Section Tests of the Heckscher-Ohlin Theorem: Comment." *American Economic Review,* December 1981, *71*(5), pp. 1040–43.

Leontief, Wassily W. "Domestic Production and Foreign Trade: The American Capital Position Re-Examined." *Proceedings of the American Philosophical Society,* September 1953, *97*(4), pp. 332–49.

Maskus, Keith E. "A Test of the Heckscher-Ohlin-Vanek Theorem: The Leontief Commonplace." *Journal of International Economics,* November 1985, *19*(3/4), pp. 201–12.

______. "Comparing International Trade Data and Product and National Characteristics Data for the Analysis of Trade Models," in Peter Hooper and J. David Richardson, eds., *International economic transactions: Issues in measurement and empirical research.* Chicago: University of Chicago Press, 1991.

Melvin, James R. "Production and Trade with Two Factors and Three Goods." *American Economic Review,* December 1968, *58*(5), pp. 1249–68.

Ohlin, Bertil G. *Interregional and international trade.* Cambridge, MA: Harvard University Press, 1933.

Samuelson, Paul A. "International Trade and the Equalisation of Factor Prices." *Economic Journal,* June 1948, *58*(230), pp. 163–84.

Staiger, Robert W. "A Specification Test of the Heckscher-Ohlin Theory." *Journal of International Economics,* August 1988, *25*(1/2), pp. 129–41.

Stern, Robert M. and Maskus, Keith E. "Determinants of the Structure of U.S. Foreign Trade, 1958–76." *Journal of International Economics,* May 1981, *11*(2), pp. 207–24.

Trefler, Daniel. "International Factor Price Differences: Leontief was Right!" *Journal of Political Economy,* December 1993, *101*(6), pp. 961–87.

Vanek, Jaroslav. "The Factor Proportions Theory: The *N*-Factor Case." *Kyklos,* October 1968, *21*(4), pp. 749–56.