NBER WORKING PAPER SERIES

A CENTURY OF MISSING TRADE?

Antoni Estevadeordal
Alan M. Taylor

Working Paper 8301
http://www.nber.org/papers/w8301

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2001

This paper is forthcoming in *American Economic Review*. For superb research assistance we are indebted to Brian Frantz. The paper has been improved by the suggestions of two anonymous referees and the editor. Some results from this paper were presented at the conference "Bertil Ohlin: A Centennial Celebration 1899–1999," held October 14–16, 1999, at the Stockholm School of Economics. For helpful comments we thank Robert Feenstra, the Ohlin conference participants, especially our discussant Rikard Forslid, as well as workshop participants at the Inter-American Development Bank, the University of California at Davis, the University of Barcelona, and the National Bureau of Economic Research. All errors are ours. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

A Century of Missing Trade?
Antoni Estevadeordal and Alan M. Taylor
NBER Working Paper No. 8301
May 2001
JEL No. F2, F11, F14, N70

ABSTRACT

In contemporary data, the measured factor content of trade is far smaller than its predicted magnitude in the pure Heckscher-Ohlin-Vanek framework, the so-called "missing trade" mystery. We wonder if this problem has been there from the beginning: that is, we ask if the Heckscher-Ohlin theory was so much at odds with reality at its time of conception. We apply contemporary tests to historical data, focusing on the major trading zone that inspired the factor abundance theory, the Old and New Worlds of the pre-1914 "Greater Atlantic" economy. This places our analysis in a very different context than contemporary studies: an era with lower trade barriers, higher transport costs, a more skewed global distribution of the relevant factors (especially land), and comparably large productivity divergence. These conditions might seem more favorable to the theory, but the results are still very poor.

Antoni Estevadeordal
Inter-American Development Bank
1300 New York Avenue, NW
Washington, DC 20577
Tel: 202-623-2614
Fax: 202-623-3030
Email: antonie@iadb.org

Alan M. Taylor
Department of Economics
University of California
One Shields Avenue
Davis, CA 95616-8578
Tel: 530-752-1572
Fax: 530-752-9382
Email: amtaylor@ucdavis.edu

Factor Abundance Theory in a Historical Context

Heckscher-Ohlin theory has been confounded by empirical contradictions dating back to the paradox unearthed by Wassily Leontief.[1] Following the notation and methodology of Jaroslav Vanek (1968), scholars have focused both on goods trade and its relationship to factor abundance (Edward E. Leamer 1980), as well as the factor content of trade itself (Harry P. Bowen, Leamer, and Leo Sveikauskas 1987). To try to deal with the Leontief paradox in the Heckscher-Ohlin-Vanek (HOV) theory one can allow for differences in cross-country productivities (Daniel Trefler 1993). To get an even closer fit, other modifications have been suggested by Trefler (1995) and Donald R. Davis and David E. Weinstein (1999) such as home bias in consumption, non-traded goods, and models without factor price equalization. Still, there remains a huge gap between theory and reality. Trefler (1995) coined the term "missing trade" to depict the extent to which measured trade is still negligible compared to the prediction of the pure theory.

Were they here to comment, Heckscher and Ohlin might not condone the use of their theory in today's very different global economic environment, but the fathers of trade theory might be impressed by the technical apparatus we have developed to evaluate their ideas. It is fair to say that their original exposition (Eli F. Heckscher and Bertil Gotthard Ohlin 1991) lacks for solid empirical evidence. They drew heavily on introspection and casual empiricism about trade between, and factor intensities in, the New and Old Worlds—for example, as in Ohlin's discussion of trends in wages and land rents in Europe and Australia (pp. 91–92). By the standards of modern empirical analysis we might consider this kind of evidence anecdotal—but can we do any better? For a fairer test of their own creative efforts, would not Heckscher and Ohlin wish for us to

[1] Leontief (1953a) shocked everyone when he computed a U.S. input-output table for 1947 and discovered that the seemingly capital-abundant and labor-scarce United States was actually engaging in net labor export via trade, with a capital-labor ratio in imports 60 percent higher than exports.

take our considerably refined empirical skills back in time and at least test the model in the historical context for which it was first designed? This is our goal. By bringing to the discussion new data from a different economic and political era, we can gain a new perspective.

In many ways, the pre-1914 period of economic liberalization and shrinking transport costs offers a superb historical laboratory for trade theory, and, by extension for the study of the political economy of a globalizing world. A strand of the economic history literature has already found strong support in that era for several features of standard theory, including predictions of factor price convergence and the pattern of goods trade.[2] In their book *Globalization and History* (1999), Kevin H. O'Rourke and Jeffrey G. Williamson blend these findings into a broader study of goods and factor market integration, its causes and impacts, in the period 1870 to 1914. Political scientists are also rightly fascinated by this period, for related reasons. Path-breaking studies such as Ronald Rogowski's *Commerce and Coalitions* (1989) introduced a new approach to comparative political economy informed by trade theory's penetrating insights on comparative advantage, factor endowments, factor rewards, and political cleavages between interest groups. In this broad multidisciplinary literature the Heckscher-Ohlin model stands as a central explanatory device, so it is essential that we be sure of its accuracy and usefulness. Until now we have not seen any direct tests of its validity in the manner of factor content studies: our paper fills that gap.

Several features of the pre-1914 era make it a potentially better laboratory for testing pure trade theory compared to today. We know, first, that there were much lower trade barriers then than now, and this could be why the theory fails in the present where

[2] On factor price equalization see Kevin H. O'Rourke, Alan M. Taylor, and Jeffrey G. Williamson (1996) and O'Rourke and Williamson (1994). On goods trade and factor endowments see Antoni Estevadeordal (1993).

tariffs and quotas lead to "too little" trade.[3] Second, in the last century, certain endowments were very skewed in their distribution, most famously the agricultural land that differentiated the endowments of the New World from the Old. Today, in contrast, many of the countries studied have much more similar endowment patterns: there may be enough variation to reveal missing trade, but the omission of economies with radically different endowments is a weakness.[4] Third, we note that there were considerable divergences in productivity across countries circa 1913, just as there are today. Over the course of the twentieth century we have seen dramatic productivity convergence only within a narrow club of countries—mostly the OECD.[5] By doing our tests circa 1913, we are in no way making the problem simpler for ourselves by avoiding an essential ingredient in the "missing trade" puzzle: the possibility of international productivity differences.[6]

---

[3] On trade barriers then and now see, for example, World Bank (1991). Tariff levels were in the single- or low-double-digit range before 1914, and much larger in the postwar period, especially in developing countries. Quotas were virtually nil before 1914, and considerable in the late postwar period. Supporting this idea, recent work by Dalia Hakura (1996) suggests that the contemporary missing trade problem has not been as serious in a set of trading countries with lower trade barriers, namely the European Union.

[4] For example, Davis and Weinstein (1999) find inevitably that OECD countries are clustered together with similar capital-labor ratios, a feature arising from those countries' similar levels of development and endowments. Their Rest-of-the-World data point lies far away from the OECD group, but this gives a great deal of leverage to one point, so much so that it is thought prudent to exclude it from the tests as a sensitivity check. And in terms of data quality, the Rest-of-the-World point uses less consistent data, making for a more fragile procedure.

[5] The first studies of long-run convergence (Moses Abramovitz 1986; William Baumol 1986) used the 16-country data of Angus Maddison (1982). Baumol was the first to note the postwar failure of unconditional convergence in wider samples that included less-developed countries. The origin of this failure was first identified by Steven Dowrick and Duc-Tho Nguyen (1989); they found conditional convergence controlling for investment and population growth.

[6] However, as Elhanan Helpman (1998) notes, productivity differences create another puzzle: namely, where do these differences in productivity originate? In historical work, a disturbing answer was brought to the fore by the controversial work of Gregory Clark (1985). He found no compelling economic reason why one New England cotton textile operative performed as much work as 1.5 British, 2.3 German and nearly 6 Greek, Japanese, Indian or Chinese workers. After controlling for capital intensities, breakdowns, human capital, learning, and other effects, Clark was forced to admit the possibility of a purely cultural origin of the differences.

Tests

Consider the standard Heckscher-Ohlin theory, in a world of C countries, I industries, and F factors. Let the net output in country c be $\mathbf{X}^c$ ($I \times 1$). The factor content of $\mathbf{X}^c$ is $\mathbf{BX}^c$, where $\mathbf{B}$ is a matrix ($F \times I$) of factor content coefficients. We make the standard assumptions that each country is within the cone of diversification, factor price equalization holds, and $\mathbf{B}$ is uniform across countries. Full employment implies that $\mathbf{BX}^c = \mathbf{V}^c$, where $\mathbf{V}^c$ ($F \times 1$) is the factor endowment of country c. With the standard assumption of uniform homothetic preferences, consumption $\mathbf{C}^c$ ($I \times 1$) in country c equals the country share of world expenditure (assumed equal to world output in this study) s^c times world consumption $\mathbf{C}^W$. By world market clearing, $\mathbf{C}^W = \mathbf{X}^W = \Sigma_c \mathbf{X}^c$. Hence, $\mathbf{C}^c = s^c \mathbf{X}^W$, and the net goods trade $\mathbf{T}^c$ of country c equals $\mathbf{T}^c = \mathbf{X}^c - \mathbf{C}^c = \mathbf{X}^c - s^c \mathbf{X}^W$. If we denote world factor endowment by $\mathbf{V}^W = \mathbf{BX}^W$, then

$$\mathbf{BT}^c = \mathbf{V}^c - s^c\, \mathbf{V}^W. \qquad (1)$$

Here, the left-hand side vector is measured factor content of trade (denoted MFCT_f) and the right-hand side is predicted factor content of trade (denoted PFCT_f). Four tests of equation (1) have been deployed in the contemporary literature, usually one factor at a time and using the set of countries c as the sample:

- The *sign test* focuses on whether the direction (i.e., sign) of MFCT_f matches that of PFCT_f. The results are displayed in terms of the fraction of correct predictions.
- The *variance ratio test* (VR) asks on whether the variance of MFCT_f is as large as PFCT_f. If the theory were a perfect fit, the ratio of the variances would be unity.
- The *slope test* calculates the slope coefficient from a regression of MFCT_f on PFCT_f. Again, if the theory were a perfect fit the slope would be unity.
- The *t-test* reports the t-statistic for the slope test where the null is a zero slope. This test can detect a positive and significant relationship of endowments to trade, though it need not be one-for-one.

What support for pure Heckscher-Ohlin theory have such tests yielded when applied to contemporary data? Precious little. As measured by the sign tests, the successful prediction rate just for the direction of factor trade is very poor, usually less than 50 percent—said to be "no better than a coin flip" (Trefler 1995). The "missing trade" problem in modern data squashes the variance ratio to less than 5 percent (0.03 in Trefler 1995; 0.0005 in Davis and Weinstein 1999), and the slope estimates turn out small, sometimes negative—albeit insignificant—values (Xavier Gabaix 1997; Davis and Weinstein 1999). We shall now try to see how well the theory fitted the facts in its own time.[7]

Data, Coding, and Aggregation

Data on net trade for the period circa 1913 were collected by Estevadeordal (1993) for $C = 18$ countries: Argentina, Australia, Austria-Hungary, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, and United States. The sources used were official national reports of trade statistics, originating from such agencies as the Board of Trade (U.K.) or the Department of Commerce (U.S). The principal problem in ensuring consistency across countries was to set up a universal classification scheme for industries, since, prior to World War Two, no standards had been developed and each country used its own classification. The solution was to laboriously construct country-specific concordances that would map each country's sectors into selected sectors of the Standard International Trade Classification (SITC, Revised 1961) at the two-digit level. In this way $\mathbf{T}^c$ was constructed for $I = 55$ sectors expressed in U.S. dollars at market exchange rates.[8]

---

[7] We are limited by data constraints from testing more embellished models such as Davis and Weinstein (1999) in a historical setting such as this, an issue we take up in the conclusion.

[8] For all the data described in this section, figures were collected for the year closest to 1913. Exchange rates were taken from international compendia of exchange rates, where available, or from national sources.

National product estimates were taken from Brian R. Mitchell (1980, 1983) and Maddison (1995) and expressed in U.S. dollars at market exchange rates, providing the basis for expenditure shares s^c.[9]

Endowment data $\mathbf{V}^c$ for $C = 18$ countries were collected by Estevadeordal (1993) for $F = 5$ types of factor: capital stock, skilled and unskilled labor force, agricultural land, and mineral resources (the latter proxied by annual production data). We made new capital stock estimates for 1913 using a perpetual-inventory method applied to pre-1913 annual investment rates and real outputs (Maurice Obstfeld and Taylor 2002). The results gave capital-output ratios for the terminal year 1913, and multiplying by national products yielded capital stocks in U.S. dollars. The labor force figures originate from Maddison (1982) and Mitchell (1980, 1983), using interpolation between census years. The data on agricultural land in hectares is largely from the League of Nations (1927). Mineral resources were estimated based on the value of annual production of coal and petroleum plus twelve other minerals and ores; quantities were drawn principally from Richard Pennefather Rothwell (various issues) and Mitchell (1980, 1983); prices from Neal Potter and Francis Christy (1962).[10]

We also need a factor use matrix $\mathbf{B}$, which depends on the direct factor use matrix $\mathbf{B}^d$ and the input-output matrix $\mathbf{A}$. Calculating $\mathbf{B} = \mathbf{B}^d(\mathbf{I} - \mathbf{A})^{-1}$ is straightforward if data on technology can be found to construct $\mathbf{B}^d$ and $\mathbf{A}$. In the pure version of the theory and empirics it is assumed that $\mathbf{B}$ is constant across countries. The objective can then be

---

[9] We do not calculate consumption or expenditure shares directly, but rather assume they are equal to income or output shares. That is we set $s^c = \text{GDP}^c/\text{GDP}^w$, and not, following the trade-balance correction of Trefler (1995) as $s^c = \text{C}^c/\text{C}^w$. This correction makes no material difference to our results.

[10] The twelve ores are bauxite; copper; iron; lead; manganese; nickel; phosphate; potash; pyrites; sulphur; tin; and zinc. Some data was also drawn from national sources of mineral production for various countries.

easily met if we can construct **B** for just one country, and, like Trefler (1993, 1995) we pick the United States as the source of the **B** data.[11]

Construction of a historically useful direct factor use matrix $\mathbf{B}^d$ for the U.S. is possible using the study of Mary Locke Eysenbach (1976).[12] She used the BLS-Leontief 1947 input-output table as the basis of her 165-industry classification scheme. Her capital and labor coefficients came from the census of 1899, and her natural resource coefficients, via Vanek (1963), from the 1947 input-output table. Already, the composite nature of her sources alerts us to the fact that her estimated $\mathbf{B}^d$ is not built from a consistent database at one point in time, and this drawback should be kept in mind. However, with this matrix available, it was straightforward to construct a concordance mapping the 165 industries into the aggregated classification based on $I = 55$ industries codes of the SITC scheme.

Another inconsistency problem with $\mathbf{B}^d$ arises because the categories (and the figures for the U.S.) do not exactly match the endowment data $\mathbf{V}^c$. This is not a problem for all inputs. Total labor is a commensurate count measure in both cases. For capital input, Eysenbach has a single stock measure expressed in U.S. dollars that we take as corresponding to our factor endowment definition of capital, up to a deflator. She measures nonrenewable resource inputs in the same units (dollars) as our endowment measure of mineral resources, up to a deflator. However, her renewable resources measure is not the same (neither in definition nor in units) as our endowment category of agricultural land. Here there is likely to be an insurmountable discrepancy between the

[11] A less restrictive but very data-intensive formulation would examine allow **B** to vary across countries. See Davis and Weinstein (1999), who were fortunate to find this information easily to hand in a consistent form in the OECD input-output database. Unfortunately, we have no consistent source of input-output tables circa 1913.

[12] Gavin Wright (1990) used Eysenbach's data in his study of resource abundance and U.S. industrial success from 1879 to 1940. The methodology in Wright's study followed in the Heckscher-Ohlin-Vanek tradition and examined the renewable and nonrenewable resource factor contents of exports and imports at the benchmark dates.

measurement concepts of the two sides. This will invalidate some of our tests: consistent units are needed for a meaningful benchmark of unity in the slope coefficient and variance ratio tests. However, we can still deploy the sign test and *t*-test to see if the directions of factor trade accord with theory.[13]

Our final data collection task was to find a suitable input-output matrix **A**. Leontief's 1947 input-output table is famous but the date is too late for our purpose; luckily, he also constructed a less well-known input-output table for the closer date of 1919 that we can employ here (Leontief 1953b). However, the 1919 input-output table was built around a smaller classification scheme of only 41 industries, so yet another concordance problem had to be solved in order to usefully align this dataset with the 55-sector SITC classification scheme used in all the previous calculations. Considering the extent of the overlap and consistency between these two classifications, we developed an

[13] A summary of units is as follows. Labor is measured as the size of the workforce, and skilled and unskilled are based on Eysenbach's census data on the left, and on literacy rates across countries on the right. (We do not use Eysenbach's skilled and unskilled labor disaggregation, because it is based on wages and occupational classifications, whereas the only cross-country **V** data we have partitions skilled and unskilled labor by literacy.) Capital is measured in dollars, and uses Eysenbach's U.S. census measure of factor use on the left-had side and a perpetual-inventory estimate of capital stocks on the right. Nonrenewable resources are in dollar terms from Eysenbach's census data on the left, and from Estevadeordal's minerals measure on the right. Renewable resources are not commensurate on the left and right of Equation (1), being based on Eysenbach's dollar-value census data on the left and Estevadeordal's measure of agricultural land areas on the right. Eysenbach's data does introduce deflation problems. Labor and capital in **B** are per 1919 dollar, but we found the differences with 1913 dollars were small enough not to change the results. Likewise, she measures nonrenewable resources in 1947 dollars per dollar of 1947 output, and here deflation must account for any changes in the relative price of nonrenewable resources from 1913 to 1947. However, our preliminary investigation showed the price trend for such resources in the United States to be almost exactly that of the GDP deflator, so again we made no price correction. When all is said and done, then, how consistent are the two sides of the equation? One test is to ask if $(\mathbf{B}^{US}\,\mathbf{X}^{US})_f / \mathbf{V}^{US}{}_f = 1$ for each supposedly commensurate factor *f*, since we have **X** data for the U.S. from Leontief's input-output table. For labor and capital we find the ratio is about 0.5 to 0.6, for nonrenewable resources about 1.8. We deem this a remarkably good fit given the disparate sources of data (from 1919 input output tables to 1947 resource coefficients to 1913 endowments) and the substantial manipulation of the data via a three-way concordance mapping.

Table 1
Industry Classification Descriptions

	Categories from Leontief's 1919 U.S. Input-Output Table
1	Agriculture
2	Flour & grist mill products; Bread & bakery products
3	Sugar, glucose & starch
4	Liquors & beverages
5	Tobacco manufactures
6	Slaughtering & meat packing
7	Butter, cheese, etc.
8	Other food industries; Canning & preserving
9	Iron mining; Non-Iron metal mining
10	Blast furnaces; Steel works & rolling mills
11	Other iron, steel & electric manufactures
12	Automobiles; Transportation
13	Smelting & refining; Brass, bronze, copper, etc. manufactures
14	Non-metal minerals
15	Petroleum & natural gas; Refined petroleum; Manufactured gas
16	Coal; Coke
17	Chemicals
18	Lumber & timber products; Other wood products
19	Paper & wood pulp; Other paper products
20	Yarn & cloth; Other textile producs
21	Clothing
22	Leather tanning; Other leather products
23	Leather shoes
24	Rubber manufactures
25	Industries, nes; Printing & publishing

Note: For Leontief-category numeric equivalents see Leontief (1953b).

aggregation scheme of $I = 25$ industries shown in Table 1 via a new set of concordances.[14]

Of course, having to shift to such a high level of aggregation is always regrettable, but particularly so here. We know that the use of output-weighted data (necessary to preserve full-employment conditions) can cause downward bias in MFCT, as Feenstra and Hanson (2000) demonstrated. Some industries within an aggregate sector may intensively use an abundant factor and tend to export more, but an output-weighted aggregation will not capture this, since it will not preserve the value of the factor content of trade. Sadly, given the limitations of the historical data, we cannot avoid this problem.

[14] A copy of the data and a data appendix are available from the authors upon request.

Table 2
Tests of Measured versus Predicted Factor Content of Trade

Productivity correction		(a) None				(b) GDP per capita			
Test		Sign	*t*	VR	Slope	Sign	*t*	VR	Slope
Capital	K	0.50	1.4	0.01	0.03	0.72	2.4	0.01	0.06
Labor	L	0.44	-1.1	0.00	-0.02	0.44	0.2	0.18	0.02
Resources-Renewable	Rr	0.67	2.6	—	—	0.83	3.0	—	—
Resources-Nonrenewable	Rn	0.78	2.4	0.38	0.35	0.78	3.6	0.51	0.52
Pooled	K, L, Rn	0.57	2.7	0.21	0.17	0.65	4.9	0.39	0.37

Notes: Sign = sign test; *t* = *t* -test; VR = variance ratio test; Slope = slope test.
Results with Trefler weights. For a description of tests and commensurability problems that affect Rr (slope and VR tests) see text.
Sources: See text and authors' appendix.

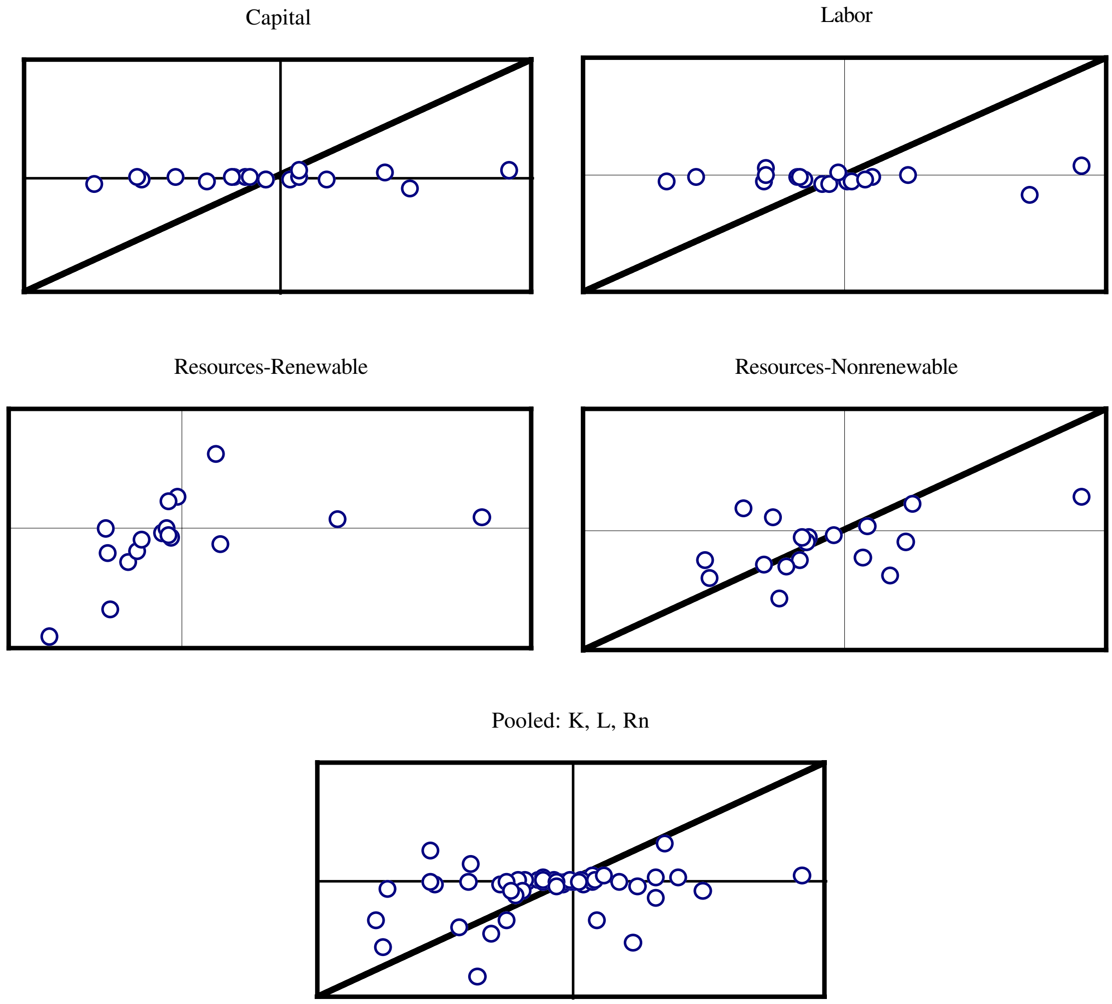
Results

The results of applying the four tests (sign, *t*, variance ratio, and slope) to 18 countries are shown in Table 2, Panel (a). We examine four individual factor types plus a set of pooled factor types. In cases where the factors are pooled, we need to worry about the commensurability not only on each side of the equation, but also from one type of factor to the next. Units of, say, labor and capital, will never be commensurate in a physical sense, so econometric adjustments are needed to permit valid estimation, specifically to ensure homoskedasticity. Following Trefler (1995), we weight each observation by $\omega_{fc} = 1/(\sigma_f s_c^{1/2})$ where the σ_f are the standard deviations of the pure HOV error term $MFCT_{fc} - PFCT_{fc}$ for each factor *f*, and where s_c is an adjustment for country size.[15]

The results are quite discouraging. For capital and labor all the tests offer almost no support for the theory. The sign test reveals a predictive power no better than a coin flip. The *t*-tests are insignificant and often of the wrong sign. The variance ratio and slope tests confirm that the fit is very poor, the slope is almost a horizontal line, and overall the model can explain less than 1 percent of the overall variance of the dependent variable. In other words, the model is performing in its historical setting just as badly as it performs

[15] This is our preferred specification, but alternative weighting schemes produce similar results.

Figure 1
Measured versus Predicted Factor Content of Trade



Notes: MFCT on vertical axis, PFCT on horizontal. Trefler weights, no productivity correction. See text and Table 2. Units on each axis are non-commensurate for renewable resources, hence 45-degree line is omitted.

with contemporary data, and it shows measures of fit as appallingly low as those from the present that we discussed earlier.

So far so bad, but our hopes pick up a little bit when resources are considered. For renewable resources, the non-commensurability problem confines us to the sign test and the *t*-test, but the results are more favorable. The sign test rises to 67 percent and the slope is significant and positive. For non-renewable resources, we can run the full battery of tests, and we find the best fit of all. The sign test shows that we get the direction of trade right for this factor in almost 4 out of every 5 cases, the *t*-ratio is a respectable 2.4, the variance ratio is 0.38 and the slope is 0.35. Finally, what the regressions are telling us can also be shown graphically, and Figure 1 depicts the scatter plots for the five cases in Panel (a). The poor fit for labor and capital is immediately apparent given the diffuse cloud of dots seen in each case. For resources, the basis for a tighter fit is also clearly visible, and the pooling is a mélange of the two.

Such results, though disappointing, are not too surprising given the equally weak findings of the recent literature using the basic, unadorned specification of the HOV hypothesis. To address these problems, various enhancements of the basic specification have been proposed. These looser specifications appeal to theory as a basis for adding additional parameters that allow for a better fit: for example, adjustments for factor productivity differences and home bias in consumption. We now apply each of these refinements to the historical data.

Factor Productivity Adjustment

Could the poor results be simply a manifestation of the Leontief problem? That is, could we be measuring factor endowments incorrectly in raw units instead of in effective units? Trefler (1993) showed that a way to correct for this problem is to rescale the endowment vector V_{fc} by some measure of relative productivity. If such a productivity correction δ_c is common to all factors in one country, then we would arrive at a productivity-corrected endowment vector of the form $\tilde{V}_{fc} = \delta_c V_{fc}$, and the analysis can then proceed as before. As a proxy for δ_c, we use the relative GDP per capita, following Trefler.[16]

The results are shown in Table 2, Panel (b). By our reading, productivity adjustments do help the model fit better, by reducing "missing trade," confirming the

[16] In each case we set U.S. equal to 1. GDP per capita measures were taken from Maddison (1991). Similar results were obtained using relative real wages from Williamson (1995) as a proxy for δ_c.

findings on contemporary data. The sign test rises well above the coin-flip level for capital, and improves for both types of resources. For non-renewable resources the slope rises to 0.52 and the variance ratio to 0.51. However, the joy is short-lived, since slopes and variance ratios are still demoralizingly low for both capital and labor, although the ratio of 0.18 for labor is still quite high by the standards of contemporary tests (where the ratio is normally is less than 0.1). Again, the pooled tests come out somewhere in between the good results for resources and the poor results for labor and capital, as expected.

Are we justified in using GDP per capita as a productivity proxy? If this were an imperfect measure of overall factor productivity, either due to measurement error, market failures, or deviations from pure Hicks-neutral technological shifts, then our results might be polluted. One way around this is to "let the data speak" by estimating the implied technology shift parameters, rather than imposing them. In this method, the parameters δ_c are chosen to maximize the fit of HOV equation, subject to the normalization that $\delta_{US} = 1$.[17] Accordingly, we estimate the corresponding variant of equation (1):

$$\mathbf{BT}^c = \tilde{\mathbf{V}}^c - s^c \tilde{\mathbf{V}}^W = \delta^c \mathbf{V}^c - s^c \sum_{c'} \delta^{c'} \mathbf{V}^{c'} . \qquad (1')$$

Tests are based on a regression of measured (left-hand side) versus predicted (fitted values on right-hand side). This method cannot be attempted on a single factor type since it would exhaust all degrees of freedom. We have to pool, so we must use Trefler weighting and omit renewable resources. The results appear in column (1) of Table 3. Here the findings are somewhat encouraging. We use up 17 out of 54 degrees of freedom (17 parameters, 54 observations), but to good effect. Over all factors, the sign test shows successful predictions in more than 4 out of 5 cases, the *t*-ratio is very significant, the

[17] Obviously, in a less interesting exercise, we can scale each factor and country, with a free choice of δ_{fc}, then we can obtain a perfect fit in the HOV model by using all degrees of freedom. (Trefler 1993; 1995).

Table 3
Productivity and Home Bias Parameters

		(1)	(2)	(3)
Factors		{K,L,Rn}	{K,L,Rn}	{K,L,Rn}
Productivity correction		Implied Hicks-Neutral	None	Imposed GDP per capita
Home Bias		No	Yes	Yes
sign		0.83	0.74	0.65
t		7.8	9.2	9.4
VR		0.54	1.55	1.49
slope		0.55	1.00	1.00
VR	K	0.05	0.07	0.14
	L	0.01	0.10	0.38
	Rn	0.64	1.83	1.49
Coefficient		δ	$1-\alpha^*$	$1-\alpha^*$
Argentina		1.03 (0.318)	0.40 (0.292)	0.43 (0.244)
Australia		0.87 (0.179)	-0.10 (0.394)	0.35 (0.267)
Austria		0.34 (0.108)	1.31 (0.336)	1.18 (0.228)
Belgium		0.63 (0.155)	1.48 (0.599)	1.21 (0.410)
Canada		0.58 (0.195)	1.59 (0.233)	1.74 (0.510)
Denmark		0.74 (0.362)	-0.11 (0.601)	0.03 (0.516)
Finland		0.59 (0.394)	0.64 (0.563)	0.65 (0.474)
France		0.72 (0.095)	0.56 (0.169)	0.48 (0.181)
Germany		0.63 (0.068)	0.65 (0.327)	0.85 (0.185)
Italy		0.64 (0.110)	0.75 (0.158)	0.69 (0.139)
Netherlands		0.73 (0.246)	-0.64 (0.375)	-0.91 (0.372)
Norway		0.19 (0.179)	1.22 (0.446)	1.21 (0.522)
Portugal		0.28 (0.153)	0.81 (0.570)	0.66 (0.544)
Spain		0.81 (0.151)	1.06 (0.208)	0.93 (0.227)
Sweden		0.80 (0.260)	0.58 (0.697)	0.94 (0.451)
Switzerland		0.55 (0.221)	0.37 (0.385)	0.49 (0.323)
United Kingdom		1.26 (0.119)	1.06 (0.132)	0.91 (0.165)
United States		1.00 (0.000)	0.68 (0.069)	0.53 (0.087)
Restrictions	$\delta=1$	0.00	—	—
(p-values)	Σ**BT**=0	—	0.07	0.00
Correlation of δ with:				
GDP per capita		0.62	—	—

Notes: Standard errors in parantheses. See text for explanation of tests and restrictions.

variance ratio is 0.54 and the slope of 0.55 is not to be sneezed at. We also reject the null restriction $\delta_c = \delta_{US} = 1$. However, a break down of the variance ratio test by factor shows that, again, the good fit for resources remains the driving force for these results (VR = 0.64). Missing trade is still overwhelming for labor and capital (VR < 0.05).

Do the implied δ_c make sense? In a sharp theoretical insight that illuminated some confusion in the debate Gabaix (1997) warned of the pitfalls of comparing the implied δ_c

to *seemingly-independent* measures of productivity, such as GDP per worker, and using these results for *inference about the fit of the HOV theory*. For example, in the case of labor, the measures are not independent at all. In fact, even if there is complete missing trade, the productivity correction turns out to be exactly a weighted total productivity measure. We can report that our δ_c look quite reasonable upon inspection, and they have a correlation with GDP per capita of about 0.7. But (to repeat) this says nothing at all about the fit of the HOV theory: only missing trade tests can do that. Still, the correlation is not irrelevant: it should reassure us that any better fit was achieved without the data having to be manipulated through implausible productivity corrections.[18]

We should not overlook the main points here. The results look "relatively" good. For comparison, the overall variance ratio was about 0.21 (or 0.39) in Table 2 without (respectively, with) the imposed productivity correction. And there is some improvement over the results on contemporary data: according to Trefler and Susan Chu Zhu (2000, Table 1), the highest all-factor VR of 0.33 was found by Trefler (1995), with all other examples yielding an overall VR less than 0.09. Hence, our historical study provides stronger support for the idea that the Leontief hypothesis will play an important part in reconciling the HOV theory to the data. But the results are not perfect, and we know that the fit for individual factors like capital and labor is poor, whilst the fit for resources is much, much better.

Home Bias in Consumption

A second extension to the basic model, also due to Trefler (1995), allows for home bias in consumption. This extension admits Armington preferences where country *c* consumption is now a weighted combination of home goods and foreign goods,

[18] Another check involves simple inspection of the implied δ_c, and an appeal to introspection to determine whether they look reasonable. They do, but with some exceptions. It might be questioned, for example, whether Norway really languished with a productivity of merely 10 percent of the U.S. level.

$\mathbf{C}^c = s^c[\alpha_c \mathbf{X}^c + \alpha_c^*(\mathbf{X}^W - \mathbf{X}^c)]$, with $\alpha_c > 1$ and $\alpha_c^* < 1$. National budget balance requires $\alpha_c s^c + \alpha_c^*(1 - s^c) = 1$, so one can eliminate the α_c. Here, the estimating equation (1) becomes

$$\mathbf{BT}^c = \alpha_c^*(\mathbf{V}^c - s^c\,\mathbf{V}^W). \tag{1″}$$

One way to estimate the α_c^* is to impose "world" market clearing, F restrictions of the form $\sum_c \mathbf{BT}^c = \sum_c \alpha_c^*\,(\mathbf{V}^c - s^c\,\mathbf{V}^W) = \mathbf{0}$.[19] Another is to let the α_c^* be free: since 18 countries do not comprise the entire world, most likely $\sum_c \mathbf{BT}^c = \mathbf{0}$ fails to hold.[20] We follow Trefler and impose restrictions but the results are not much different when the paramaters are free. In the case $(1 - \alpha_c^*) = 0$, we have no home bias and revert to the standard theory. In the case $(1 - \alpha_c^*) = 1$ we have complete home bias. Values of $(1 - \alpha_c^*)$ between zero and one are assumed to correspond to a varying degree of home bias.

Already, we can intuitively see what is going to happen. Suppose the home bias were constant across countries, with $\alpha_c^* = \alpha^*$. Clearly, a regression based on (1″) will then set α^* equal to the slope from (1) and the fit will improve dramatically. A corollary of the OLS algebra is that the implied slope of measured (left-hand side) versus predicted (fitted values on right-hand side) will be unity by construction! This eliminates the slope test as a meaningful criterion. However, the other tests are still good—for example, the variance ratio is not necessarily equal to unity in these regressions, nor is the R^2.

The results of applying these tests are shown in columns (2) and (3) of Table 3. Again, if we want to allow for country-specific parameters we can only gain sufficient degrees of freedom by pooling. We indicate whether an (imposed) productivity correction

[19] For simplicity, we omit the (empirically less-relevant) adjustments for trade imbalances discussed by Trefler (1995). Equation (1″) is estimated by OLS with the F linear restrictions imposed.

[20] We do indeed find that in all cases the null restriction $\Sigma_c\,\mathbf{BT}^c = \mathbf{0}$ is rejected, perhaps because of significant trade with countries outside the sample.

is included (2) or not (3).[21] Compared to Table 2, the sign test improves in column (2) a little (0.74 versus 0.50), but deteriorates in column (3) (0.65 versus 0.72), suggesting that once a productivity correction is used, the home bias effects are of little help. The variance ratio in both cases rises far above one, perhaps disconcertingly too far in the direction of excessive (not missing) trade. But of greater concern are the implied home bias parameters themselves. Judging whether these are reasonable is again a matter for our introspection, but many of these estimates are beyond the bounds of what theory permits. It is not clear what is implied by a value of $(1 - \alpha_c^*)$ that is outside the interval [0,1]. Although the standard errors on the are large—one cannot reject the hypothesis that all the coefficients lie in the unit interval—the point estimates remain disturbing given their extremely wide range. Thus, we react to the home-bias extension rather pessimistically. Like Trefler (1995), we find some strange implied values for the $(1 - \alpha_c^*)$ coefficients that make little sense in theory. Undoubtedly some kind of home-bias effect will be a necessary part of a complete trade theory; but, given that the results here are not markedly better as judged by fit, we are skeptical as to the usefulness and relevance of the home-bias correction for our sample, at least in this form.

Conclusion

We have shown how it is possible to implement a test circa 1913 of the HOV theory. For labor and capital the fit of the model is close to nonexistent. For resources, there is evidence that the model fits really quite well. Introducing extensions to the model, the fit is much improved by a Leontief-style productivity correction, but the results of a home bias correction appear, in many cases, quite implausible.

If we stand back for a moment, we can ask, following Trefler and Zhu (2000), whether our empirical approach constitutes a "cure" or merely a "diagnosis" of the failure of the HOV model in its original historical context. We take away an overall impression

[21] Though feasible, there would be few degrees of freedom left if we estimated *both* a productivity correction (17 parameters) and a home bias correction (18 parameters). We did try such regressions. They were not edifying. Among other problems, they implied *negative* productivity parameters, another reason for us to suspect the home bias approach employed here.

that the modified HOV model is still dysfunctional in its own time, but not as pathologically so as it is today. Knowing that skewed resource endowments were key motivations for Heckscher and Ohlin, adherents of their approach might take some comfort in the fact that missing trade appears less severe in the case of resources. But with respect to the full set of factors, the simple factor-content approach seems to work as well in its own time as it does today—that is, not very well at all.

Our historical tests therefore bring us to the same pessimistic point that the contemporary literature arrived at in 1995—the year Trefler announced the mystery of the "missing trade." For devotees of the Heckscher-Ohlin model, missing trade turns out to be just as big a mystery a century ago. It could yet be seen as an *even bigger* mystery given the historical and institutional circumstances of that time.

References

Abramovitz, Moses, "Resource and Output Trends in the United States Since 1870," *American Economic Review*, 1956, 46, 5–23.

Baumol, William, "Productivity Growth, Convergence and Welfare: What the Long-Run Data Show," *American Economic Review*, 1986, 76, 1072–85.

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

Clark, Gregory, "Why Isn't the Whole World Developed? Lessons from the Cotton Mills," *Journal of Economic History*, 1987, 47, 141–73.

Davis, Donald R., and David Weinstein, "An Account of Global Factor Trade," Photocopy, Harvard University, 1999.

Dowrick, Steven, and Duc-Tho Nguyen, "OECD Comparative Economic Growth 1950–85: Catch-Up and Convergence," *American Economic Review*, 1989, 79, 1010–30.

Estevadeordal, Antoni, "Historical Essays on Comparative Advantage, 1913–1938," Ph.D. dissertation, Harvard University, 1993,

Eysenbach, Mary Locke, *American Manufactured Exports, 1879–1914: A Study of Growth and Comparative Advantage*, New York: Arno Press, 1976.

Feenstra, Robert C., and Gordon H. Hanson, "Aggregation Bias in the Factor Content of Trade: Evidence from U.S. Manufacturing," *American Economic Review*, 2000, 90, 155–60.

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

Hakura, Dalia, "A Test of the General Validity of the Heckscher-Ohlin Theorem for Trade in the European Community," Photocopy, International Monetary Fund, 1996.

Heckscher, Eli F., and Bertil Gotthard Ohlin, *Heckscher-Ohlin Trade Theory*, Cambridge, Mass.: MIT Press, 1991. Translated, edited, and introduced by Harry Flam and M. June Flanders.

Helpman, Elhanan, "The Structure of Foreign Trade," National Bureau of Economic Research, Working Paper Series No. 6752, October 1998.

League of Nations, *Population and Natural Resources*, Geneva: League of Nations, 1927.

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

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

Leontief, Wassily W., *The Structure of the American Economy 1919–1939*, Oxford: Oxford University Press, 1953.

Maddison, Angus, *Phases of Capitalist Development*, Oxford: Oxford University Press, 1982.

Maddison, Angus, *Dynamic Forces in Capitalist Development: A Long-Run Comparative View*, Oxford: Oxford University Press, 1991.

Mitchell, Brian R., *European Historical Statistics, 1750–1975*, 2nd ed., New York: Facts on File, 1980.

Mitchell, Brian R., *International Historical Statistics: The Americas and Australasia*, Detroit: Gale Research, 1983.

O'Rourke, Kevin H., Alan M. Taylor, and Jeffrey G. Williamson, "Factor Price Convergence in the Late Nineteenth Century," *International Economic Review*, 1996, 37, 499–530.

O'Rourke, Kevin H., and Jeffrey G. Williamson, "Late-Nineteenth Century Anglo-American Factor-Price Convergence: Were Heckscher and Ohlin Right?" *Journal of Economic History*, 1994, 54, 892–916.

O'Rourke, Kevin H., and Jeffrey G. Williamson, *Globalization and History: The Evolution of a Nineteenth-Century Atlantic Economy*, Cambridge: MIT Press, 1999.

Obstfeld, Maurice, and Alan M. Taylor, *Global Capital Markets: Integration, Crisis, and Growth*, Japan-U.S. Center Sanwa Monographs on International Financial Markets, Cambridge: Cambridge University Press, 2002. In progress.

Potter, Neal, and Francis T. Christy, *Trends in Natural Resource Commodities: Statistics of Prices, Output, Consumption, Foreign Trade, and Employment in the United States, 1870–1957*, Baltimore: Published for Resources for the Future by Johns Hopkins Press, 1962.

Rogowski, Ronald, *Commerce and Coalitions: How Trade Affects Domestic Political Alignments*, Princeton, N.J.: Princeton University Press, 1989.

Rothwell, Richard Pennefather, *The Mineral Industry, its Statistics, Technology, and Trade*, New York: Scientific Publishing Company, various issues.

Trefler, Daniel, "International Factor Price Differences: Leontief Was Right!" *Journal of Political Economy*, 1993, 101, 961–87.

Trefler, Daniel, "The Case of the Missing Trade and Other Mysteries," *American Economic Review*, 1995, 85, 1029–1046.

Trefler, Daniel, and Susan Chun Zhu, "Beyond the Algebra of Explanation: HOV for the Technology Age," *American Economic Review*, 2000, 90, 145–49.

Vanek, Jaroslav, *The Natural Resource Content of United States Foreign Trade 1870–1955*, Cambridge, Mass.: MIT Press, 1963.

Vanek, Jaroslav, "Factor Proportions Theory: The *N*-Factor Case," *Kyklos*, 1968, 21, 749–55.

Williamson, Jeffrey G., "The Evolution of Global Labor Markets since 1830: Background Evidence and Hypotheses," *Explorations in Economic History*, 1995, 32, 141–96.

World Bank, *World Development Report*, New York: Oxford University Press, 1991.

Wright, Gavin, "The Origins of American Industrial Success, 1879–1940," *American Economic Review*, 1990, 80, 651–68.