

EMPIRICAL TESTS OF THE FACTOR ABUNDANCE THEORY: WHAT DO THEY TELL US?

Donald R. Davis
Harvard University and NBER

and

David E. Weinstein
Harvard University and University of Michigan

ESTIMATE, DON'T TEST?

In their masterful survey of recent empirical work on international microeconomics, Leamer and Levinsohn [1995] expound a distinctive view on the appropriate aims of empirical work in this area. In a phrase, they exhort researchers to "Estimate, don't test!" The foundation for this advice is a view that empirical work on trade has had very little influence on our beliefs about the causes of trade. And they believe that this is true because empirical researchers have taken the theory too seriously. In Leamer and Levinsohn's view, we should

Understand that theorems are neither true nor false. They are sometimes useful and sometimes misleading. If we approach a data base with the contrary attitude, hoping to determine the verity or falsity of a theorem, we may statistically "reject" the theory, but leave it completely unharmed nonetheless. After all, we already knew it wasn't literally true. [1995, 1341]¹

That the study of international trade has been dominated by theory is evident to anyone who turns the pages of the journals. Until recently, empirical work on international trade has been accorded little of the prestige bestowed on theoretical advances. And, no doubt, much of the existing empirical work fails to influence the profession strongly because it takes the theoretical caricature too literally.

Yet we believe that Leamer and Levinsohn understate the true influence exerted by the existing empirical work. And we believe that the anti-testing perspective is not the appropriate remedy. We will pursue these themes within the context of a review of selected work on the Heckscher-Ohlin theory of trade, and use the recent work of

Davis, Weinstein, et al. [1996] as an example of our own view of how to focus an empirical project in a way that will be persuasive.

HAS THE EMPIRICAL WORK BEEN INFLUENTIAL?

On the surface, the Heckscher-Ohlin theory of trade is the perfect example of a theory which has been statistically rejected time and again, but which is in Leamer and Levinsohn's phrase "completely unharmed." Leontief's [1953] findings, apparently at odds with the most basic predictions of the Heckscher-Ohlin theory, provoked a lot of worrying.

But Leontief's results did not prevent the Heckscher-Ohlin model from being the workhorse of international trade theory for the succeeding three decades. Economists had strong prior beliefs that endowments have to matter for production and trade patterns, and that the Heckscher-Ohlin model was the most sensible way of embodying these links. For this reason, Leontief's results became known as a paradox rather than a refutation.

Notwithstanding this, the contrary empirical results have had a decidedly negative impact on the influence of Heckscher-Ohlin within the profession. It is true that this remained muted during a period with no alternative theoretical framework. However the rise of the "new trade theory" in the 1980s was greatly smoothed by the apparent empirical shortcomings of Heckscher-Ohlin. This is true to the point where it now borders on the disreputable to believe that endowments matter much at all for intra-OECD trade.²

We want to argue that the disenchantment with Heckscher-Ohlin has proceeded too far. However this does not lead us to suggest that we dispense with testing of the theorems. Leamer and Levinsohn are surely correct when they assert that the theorems are neither correct nor incorrect. What, then, are we to expect empirical testing to do? Our view is that we want tests to narrow the bounds of beliefs that one can sensibly have about the world. Negative results do not render the theory useless. But they do suggest that when considering the data of interest, this particular constellation of assumptions may not be helpful in making sense of the world. Negative results are interesting insofar as they rule out otherwise appealing sets of assumptions. This also suggests a second approach to testing — to look for ways of weakening the strict assumptions of the theory in sensible directions in order to find a version that does in fact work. By successive approximation, we should learn which of the assumptions of the model are most crucial, and to which types of data sets one can sensibly apply the various versions of the theory.

In the sections that follow, we describe what we view as this process of successive approximation for the case of the Heckscher-Ohlin theory. We begin with a highly selective overview of previous empirical work, with an emphasis on the Heckscher-Ohlin-Vanek framework. In the subsequent section, we sketch some recent work of our own on the Heckscher-Ohlin-Vanek theory with decidedly more positive results. We conclude by noting some directions for additional empirical investigation.

PREVIOUS WORK: A STRING OF EMPIRICAL REJECTIONS

Leontief [1953] is the seminal empirical critique of Heckscher-Ohlin.³ He used data on input requirements and U.S. trade to measure capital to labor ratios in U.S. imports and exports separately. To universal surprise, widespread dismay, and scattered consternation, he showed that U.S. imports were more capital intensive than U.S. exports. This suggested that the United States is relatively labor abundant, a result ever after known as the "Leontief Paradox." Leamer [1980] showed, however, that Leontief applied a conceptually inappropriate test of the Heckscher-Ohlin hypothesis. When he re-examined the same data in a conceptually correct way, the paradox vanished. Nonetheless, this paradox refused to perish. Brecher and Choudhri [1982] pointed out that one (counterfactual) implication of Leamer's approach is that U.S. expenditure per worker would have to be lower than for the world as a whole. Stern and Maskus [1981] applied Leamer's [1980] approach to U.S. data for both 1958 and 1972, finding the Leontief paradox held in the former but not in the latter year. Extensive surveys of previous work on Heckscher-Ohlin can be found in Deardorff [1984] and Leamer and Levinsohn [1995].

Recent empirical work has focused on a higher dimensional version of the theory called Heckscher-Ohlin-Vanek. Rather than focusing on the pattern of trade in goods, it focuses on the net factor content of trade. Let c be an index for countries, and W stand for the world. Let V^c , s^c , and T^c be respectively the vector of endowments, the share of world income, and the net goods export vector of country c . Let B be a matrix of direct factor inputs and A be the input-output matrix. Then the fundamental Heckscher-Ohlin-Vanek relation can be expressed as:

$$(1) \quad B(I-A)^{-1}T^c = V^c - s^c V^W.$$

This theory has previously been examined by Maskus [1985] and Bowen, Leamer, and Sveikauskas [1987]. Equation (1) asserts a vector equality between what we may term the "measured" net factor content of trade (*lhs*) and the "predicted" net factor content of trade (*rhs*). If we assemble these for the full complement of countries, the theory asserts an equality between two matrices of measured and predicted net factor contents. Rows of the matrices correspond to various factors, while columns correspond to particular countries. Two basic tests have been applied to this. One is known as the "rank test." It looks across corresponding rows or columns of the two matrices to see if there is a high rank correlation. The other is known as the "sign test." It checks to see whether the factors that the theory predicts a country will export on net in fact are exported. This amounts to seeing whether corresponding cells of the two matrices have the same sign.

The results of Maskus [1985] and Bowen, Leamer, and Sveikauskas [1987] have severely undermined confidence in the robustness of the Heckscher-Ohlin framework. Maskus carried out both sign and rank tests on data for two time periods (1958 and 1972), and for three high quality factors (professional, unskilled labor, and capital). He reports results only for the United States, perhaps because the Leontief paradox

had focused on it. The sign test is correct for only one factor in 1958, but for all three in 1972. This might be seen to suggest that the Heckscher-Ohlin-Vanek relations fare well, at least in the latter period. However, the test lacks power. As Maskus notes, if we consider the alternative that the signs were determined randomly, we will have two or fewer sign failures out of six tries 34.4 percent of the time. Moreover, even if we limit ourselves to the 1972 data, under the same alternative, there will be no sign violations (as in his data) one in eight times. The results were, if anything, worse in the rank test. The direct measures of U.S. factor abundance relative to the rest of the world were stable, with physical capital most abundant, professionals second, and unskilled labor least abundant. However, the trade-imputed measures of factor abundance in 1958 suggested the United States was most abundant in unskilled labor, and least abundant in physical capital! The 1972 trade-imputed measure of factor abundance showed unskilled labor shifting dramatically to be least abundant, and reverses the relative abundance of physical capital and professionals. A repeat of the tests, restricted to OECD data, yielded no improvement. As Maskus notes, "paradoxical outcomes may be the rule rather than the exception" [1985, 209].

Bowen, Leamer, and Sveikauskas likewise report results widely viewed as undercutting Heckscher-Ohlin-Vanek. An important contribution was extending the test to a much broader set of countries (27) and factors (12). Thus, whereas the Maskus test was based on a matrix of only three cells for each time period, the Bowen, Leamer, and Sveikauskas matrix had 324 entries. Because of the greater dimensionality of the matrix, it became possible to conduct sign and rank tests not only for a single country across factors (as in Maskus [1985]), but also for a single factor across countries. The sign test was correct more than half of the time for eleven of the twelve factors, but was correct over 70 percent of the time for only four in twelve. The sign matches were correct more than half the time for 18 of 27 countries, but over 70 percent of the time for only 8 of the 27. Only 61 percent of the total sign matches were correct. They note that independence between the signs of corresponding entries can be rejected at the 95 percent level for only one factor in twelve, and for only 4 of the 27 countries. In effect, in determining which factors' services would on net be exported or imported, Heckscher-Ohlin did little better than a coin-flip.

The rank proposition fares no better. Bowen, Leamer, and Sveikauskas report both rank correlations and the proportion of correct rankings when entries are compared two at a time. A zero correlation is rejected for only four of the twelve factors and 8 of the 27 countries. Moreover, one factor and 5 countries have the wrong sign on the correlation. While the pairwise comparisons get over 50 percent correct rankings for 22 of the 27 countries, the same is true for only three of the twelve factors (all land variables). In sum, Bowen, Leamer, and Sveikauskas note that the sign and rank propositions yield the Heckscher-Ohlin-Vanek model relatively little support.

The pessimism regarding Heckscher-Ohlin was partly relieved by Trefler [1993], only to be revived by Trefler [1995]. They represent alternative approaches to resolving the problems identified by Bowen, Leamer, and Sveikauskas. The former follows up on Leontief's suggestion that the failure of Heckscher-Ohlin may be due to factor-based differences in efficiency. Trefler chooses the efficiency factors so that the

Heckscher-Ohlin-Vanek equations fit exactly. He then shows that the implied productivity differentials correlate nicely with evidence on cross-country differences in wages and rentals, suggesting a version of adjusted factor price equalization.

Trefler [1995] returned to the simple Heckscher-Ohlin-Vanek framework using the U.S. B matrix, B^{US} . In one exercise, he graphed the net factor trade residuals, $\epsilon_T = B^{US}(I - A)^{-1}T - (V^c - s^cV^w)$, against the predicted net factor trade, $V^c - s^cV^w$. Theory would predict that these should be centered around the line $\epsilon_T = 0$. Instead they closely followed the line $\epsilon_T = -(V^c - s^cV^w)$, or equivalently, $B^{US}(I - A)^{-1}T = 0$. This says that measured net factor trade is approximately zero, to which Trefler applied the colorful moniker "the case of the missing trade" [1995, 1032]. This is certainly a serious difficulty for an endowments-based model of trade. After considering a variety of alternatives, he identifies neutral technological differences and a home bias in consumption as the most likely culprits for the missing trade. However, by working with trade data — rather than production and absorption data — he can only test this indirectly. The technical differences are estimated rather than observed. Thus, while these papers provide some glimmerings of hope for Heckscher-Ohlin, they also serve to underscore the serious empirical shortcomings of the simple model.

Recently, a few authors have studied Heckscher-Ohlin using United States regional data, as in Horiba [1992], and Grimes and Prime [1993].⁴ The rationale for such studies is that the conventional assumptions of identical technologies and preferences, factor price equalization, and free trade are more likely to hold for regions of a single country than across different countries. However, an important flaw underlies both studies — the United States is treated as a closed economy, which is inherently anti-Heckscher-Ohlin. They assume that the factor content of regional consumption is proportional to the U.S. — rather than the world — endowment. If this were so, then the U.S. would have no reason to trade with the rest of the world. More importantly, if one ignores the rest of the world, the model does not have determinate predictions about the factor content of trade. The general model yields predictions for any region *vis a vis* the rest of the world, but not relative to any subset of regions or countries in the world. In addition, there are problems in interpreting the data. As Grimes and Prime [1993] note, their data treats U.S. regional exports to the rest of the world as though they are consumed at the port of exit, wherever they may come from, and wherever they are destined. They also adopt a curious measure of some of their endowments, one closer to a measure of output than input. For example, the measure of "agricultural endowment" is the "value of agricultural products sold by farms in millions of dollars" [*ibid.*, 26]. This comes close to reducing the production side of the Heckscher-Ohlin model to a tautology, and so testing only homothetic preferences. The authors deserve credit for attempting to implement a regional Heckscher-Ohlin model. However, data problems and methodological flaws make interpretation of their results difficult.

What has determined the influence wielded by the various papers surveyed above? It is surely not that they abandoned testing in favor of estimating a model to fit the data. Rather the influence of each has depended on the sensible and sensitive matching of theory and data to mark bounds of how well or poorly the data and theory match.

WHICH ASSUMPTIONS MATTER FOR THE HECKSCHER-OHLIN-VANEK THEORY?

In this section we will sketch some recent work of our own [Davis, Weinstein, et al., 1996] which likewise aims to determine which of the Heckscher-Ohlin-Vanek assumptions are crucial to get the theory to work. We make two important modifications from the simple Heckscher-Ohlin-Vanek framework. The first is to consider the possibility that the technology implemented is not identical everywhere. This leads us to focus on the implications of the model for regions of Japan for which the assumption of identical technologies is more plausible. In contrast to other regional models of Heckscher-Ohlin trade, we do this while continuing to embed our model in a theoretically appropriate way within the full world equilibrium. The second modification is to ignore trade data as such, and instead focus on the components — production and absorption. This allows us to test separately the Heckscher-Ohlin theories of the location of production and pattern of consumption. By examining them separately, we know the source of any deviations that arise when we assemble the pair to a full test of the trade theory.

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

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

We then assemble this information for a full test of the Heckscher-Ohlin theory of the net factor content of trade. Our earlier results showed that the theory could not account for the international pattern of production. Hence no point is served by looking at the implied net factor content of trade of the various countries, as positive results would have to be spurious. Instead, we focus on accounting for the net factor trade of the Japanese regions. Three approaches are developed. The first establishes a benchmark. It uses data on actual world factor endowments, implicitly assuming again that all countries use the same input coefficients. We show that the model performs poorly and replicate Treffer's "mystery of the the missing trade" [1995]. In the next two cases, we examine this using the endowments imputed to the world,

given their measured output, as if they had used the Japanese input coefficients. In the first of these, we assume that Japanese regional absorption is proportional to world net output. This model is a marked improvement over that based on measured world endowments, and largely eliminates the missing trade. In the last case, we consider this under the assumption that Japanese regional consumption is proportional to national consumption. This yields no further improvement, demonstrating that changing the consumption model is not necessary to better the model's fit.

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

In sum, we find that the Heckscher-Ohlin model under the conventional restrictive assumptions is a poor predictor of the international pattern of production, hence of net factor trade. However, this changes markedly when applied to predictions for regions of Japan. Given the long string of empirical failures of Heckscher-Ohlin, it is surprisingly successful as a theory of the location of production and the pattern of consumption, hence the net factor content of trade of these regions. We identify the assumption of universal factor price equalization as the crucial element to relax to improve the results on the Heckscher-Ohlin-Vanek theory.

CONCLUSION

Recently Leamer and Levinsohn have called for a re-orientation or empirical work in international microeconomics. It may be summarized in their injunction to "Estimate, don't test!" However a review of recent empirical work on the Heckscher-Ohlin-Vanek model suggests that the cumulation of tests is informative and influential, even where any one alone might not be. They give us successive approximations that allow us to see the circumstances under which a particular theory is useful. Thus we suggest instead that researchers both estimate and test. Rather than dispense with tests, we suggest a simple criterion: Do they narrow the range of sensible application of the theory? A test is useful when we can reply in the affirmative.

NOTES

1. Leamer and Levinsohn also provide a caution the other direction — not to treat the theory too casually. Leamer's [1980] own work on the conceptual problems underlying the Leontief paradox provides a striking object lesson in the dangers of turning to empirics without a clear tie to the theory.
2. The strong influence of the empirical results on economists' views of the world is exemplified by the interpretation of the results of Bowen, Leamer, and Sveikauskas [1987] (discussed below) by Krugman and Obstfeld [1994]. The latter say that the Bowen, Leamer, and Sveikauskas results confirm that "trade just does not run in the direction that the Heckscher-Ohlin theory predicts" [*ibid.*, 78]
3. It is worth noting that Leontief did not view his own work as contradicting the Heckscher-Ohlin model. Rather he saw it as evidence against the notion that the United States was capital abundant. Later authors, however, came to view his work as providing "the single biggest piece of evidence against the factor proportions theory" [Krugman and Obstfeld, 1994, 77].
4. An earlier set of papers sought to examine Heckscher-Ohlin with regional data, as in Horiba [1973]. However, these primarily relied on the method of Leontief [1953], which is vulnerable to the criticisms of Leamer [1980].

REFERENCES

- Bowen, H. P., Leamer, E. E. and Sveikauskas, L. Multifactor, Multicountry Tests of the Factor Abundance Theory. *American Economic Review*, December 1987, 791-809.
- Brecher, R. A., and Choudhri, E. The Leontief Paradox: Continued. *Journal of Political Economy*, August 1982, 820-23.
- Davis, D. R., Weinstein, D. E., Bradford, S. C., and Shimpo, K. The Heckscher-Ohlin-Vanek Model of Trade: Why Does it Fail? When Does it Work? NBER Working Paper #5625, June 1996.
- Deardorff. Testing Trade Theories and Predicting Trade Flows, in *Handbook of International Economics*, Volume 1, edited by R. Jones and P. Kenen. NY: Elsevier, 1984.
- Grimes, D. R. and Prime, P. B. A Regional Multifactor Test of the Heckscher-Ohlin-Vanek Theorem. Mimeo. University of Michigan and Kenesaw State College, 1993.
- Horiba, Y. On the Empirical Relevance of the Heckscher-Ohlin Trade Model. Mimeo. Tulane University, 1992.
- _____. Factor Proportions and the Structure of Inter-Regional Trade: The Case of Japan. *South-ern Economic Journal*, 1973.
- Krugman and Obstfeld. *International Economics: Theory and Policy*. NY: Harper Collins, 1994.
- Leamer, E. The Leontief Paradox Reconsidered. *Journal of Political Economy*, June 1980, 495-503.
- Leamer, E. and Levinsohn, J. International Trade Theory: The Evidence, in *Handbook of International Economics*, Volume 3, edited by G. Grossman and K. Rogoff. NY: Elsevier, 1995.
- Leontief, W. Domestic Production and Foreign Trade: The American Capital Position Re-Examined. *Proceeding of the American Philosophical Society*, 1953, 332-49.
- Maskus, K. V. A Test of the Heckscher-Ohlin-Vanek Theorem: The Leontief Commonplace. *Journal of International Economics*, November 1985, 201-212.
- Stern, R. M., and Maskus, K. E. Determinants of the Commodity Structure of U.S. Foreign Trade, 1958-1976. *Journal of International Economics*, May 1981, 11, 207-224.
- Trefler, D. International Factor Price Differences: Leontief Was Right! *Journal of Political Economy*, December 1993, 101, 961-87.
- Trefler, D. The Case of the Missing Trade and Other Mysteries. *American Economic Review*, December 1995, 1029-46.