

MichU
DeptE
ResSIE
D
#212

RESEARCH SEMINAR IN INTERNATIONAL ECONOMICS

Department of Economics
The University of Michigan
Ann Arbor, Michigan 48109-1220

SEMINAR DISCUSSION PAPER NO. 212

How Robust Are Sign and Rank Order Tests of the Heckscher-Ohlin-Vanek Theorem?

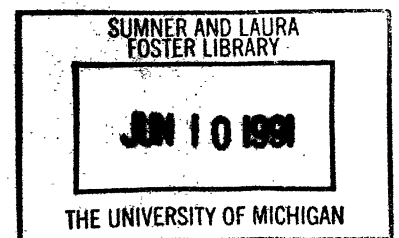
by

Wilhelm K. Kohler*

Revised Version

Published in *Oxford Economic Papers*, January 1991.

*The University of Innsbruck. Visiting Scholar, The University of Michigan, Ann Arbor, 1987-88.



HOW ROBUST ARE SIGN AND RANK ORDER TESTS OF THE HECKSCHER–OHLIN–VANEK THEOREM?

By WILHELM KOHLER*

1. Introduction

DURING the past two decades, the *Heckscher–Ohlin (HO) theorem* linking the structure of a country's foreign trade to its relative factor abundance has been among the most frequently tested propositions of economic theory. We are now confronted with a massive amount of empirical evidence, but this evidence is, overall, inconclusive. While it is more or less universally accepted that the theory in its strict form is refuted by empirical evidence, the situation is less clear-cut if we look at the somewhat less demanding rank order or sign propositions.

One of the key problems in trying to reach unanimous conclusions from the literature lies with the multiplicity of test designs employed. Thus, some of the tests have relied on inappropriate generalizations of the intuition drawn from the two-dimensional version of the HO theorem to the higher dimensional case. This holds true both for the approach adopted in Leontief's (1953) pioneering study (see Leamer, 1980) and for many of the regression studies in commodity or country space (see Deardorff, 1984, and Kohler, 1988a). Moreover, many of these studies are what Maskus (1985) has called 'incomplete' tests of the HO theory because they do not involve independent observations on all three key variables of the theory: factor intensities, factor endowments, and commodity trade.¹

Even 'complete' tests rigorously derived from generalized versions of the theory have frequently been inconclusive, because they have concentrated on different versions of the HO theorem, thus often arriving at conflicting test results. The generalized version of the factor proportions theory used in these studies is the so-called *Heckscher–Ohlin–Vanek (HOV) model* (Vanek, 1968), and it has become standard practice to test both a *rank order* and a *sign hypothesis* derived from this model (see, for instance, Harkness, 1983, Maskus, 1985, and Bowen *et al.*, 1987).

While it has been observed and is, therefore, commonly accepted that sign

* A large part of the research presented here was done while I was visiting scholar at the Department of Economics, University of Michigan, Ann Arbor. I wish to thank Alan Deardorff and Robert Stern for detailed and very helpful comments on an earlier paper focusing on the ideas of chapter 6 of the present one. Comments by two anonymous referees are also gratefully acknowledged. Remaining errors are, of course, my own responsibility. Thanks are also due to the Fonds zur Förderung der wissenschaftlichen Forschung, Vienna, for the financial support received through an Erwin-Schrödinger fellowship.

¹ That any appropriate test design necessarily involves independent observations of these three key variables was first pointed out by Leamer and Bowen (1981).

and rank order tests may produce contrasting results for identical data sets, it appears to have gone unnoticed that even different ways to perform a sign or a rank order test on the basis of the HOV model may lead to different outcomes, if the data do not strictly satisfy the HOV equations. Such a non-robustness would cast serious doubt on the common practice of drawing conclusions from tests of a single rank order and sign hypothesis, more or less arbitrarily chosen among the several hypotheses suggested by the HOV model.

The present paper will concentrate on four different rank order and two different sign propositions of the HOV model, which have, at various times, all been the subject of empirical analysis in the literature, and it will establish conditions, under which a given data set will support one (rank order or sign) hypothesis while at the same time rejecting another. This will shed some light on the circumstances influencing the likelihood of such non-robust test results.

The paper will also examine the empirical relevance of this phenomenon of non-robustness using the multi-country data set published in Bowen *et al.* (1987). Furthermore, it will be shown that the apparent conflict between two recent re-examinations of the Leontief paradox by Brecher and Choudhri (1982a) and Casas and Kwan Choi (1985) must be attributed to this non-robustness of sign tests of the HOV theorem.

2. The model

Travis (1964) and Vanek (1968) have shown that the standard assumptions of the HO theory imply the following vector equation, which relates the factor content of a country j 's net exports of factor services to the difference between that country's resource endowment and that of the whole world in the following way:

$$At_j = f_j - \alpha_j f_w, \text{ where} \quad (1)$$

- A is a $(k \times n)$ matrix of direct plus indirect input coefficients of primary factors (there are k factors and n goods);
- t_j is country j 's vector of net commodity exports;
- f_j is country j 's vector of factor endowments;
- f_w is the world endowment vector; and
- α_j is the ratio between country j 's domestic absorption and world income: $\alpha_j = (y_j - b_j)/y_w$, where y indicates GDP and b indicates the trade balance.

Unlike the HO theorem of the two-dimensional world and the generalized version of it introduced by Deardorff (1982), the above equation also assumes free trade and incomplete specialization inasmuch as these are necessary for international equalization of factor prices. We may note also that initial versions of this simple model have been set up in value rather than quantity terms.²

² See Vanek (1968), Bertrand (1972), and Horiba (1971).

A value terms formulation of (1) can be written as

$$\Theta Pt_j = Wf_j - \alpha_j Wf_w, \text{ where} \quad (2)$$

Θ denotes the $(k \times n)$ matrix of factor shares, and W and P , respectively, indicate diagonal matrices of commodity and factor prices, as prevailing in free trade equilibrium.³ Such a formulation is slightly more useful than equation (1), because empirical tests almost always use observations of values, at least for trade patterns.

In what follows, subscripts i and h will be used to indicate two arbitrary factors of production. Thus, the i th row of equation (2) can be written as

$$T_{ji} = E_{ji} - \alpha_j E_{wi}, \quad (3)$$

or, equivalently, as

$$\frac{T_{ji}}{E_{ji}} = 1 - \alpha_j \frac{E_{wi}}{E_{ji}}. \quad (4)$$

It is perhaps fair to say that equation (3) in its strict form has almost always been regarded as an 'a priori incredible' model (see, for instance, Leamer, 1984, p. 45). And for this reason, it has become standard practice in empirical testing to maintain hypotheses that represent 'weakened' forms of equation (3). These are sign and rank order hypotheses. They may be called 'weaker' than the above equations, because any data strictly satisfying the above equations will necessarily also support any sign or rank order hypothesis while the reverse is not true.⁴ Thus, calling sign or rank order propositions 'weakened' forms of equation (3) should not be taken to indicate that their violation is only weak evidence against HOV. The contrary is true: the 'weaker' (in the sense above) the proposition violated the stronger is the evidence against the strict version of the theory.

3. Rank order hypotheses

There are two straightforward rank order hypotheses suggested by equations (3) and (4).

First, equation (4) proposes that ranking all factors in terms of country j 's proportional net factor exports (T_{ji}/E_{ji}) is equivalent to ranking them in terms of country j 's share in world endowments (E_{ji}/E_{wi}) .⁵ This hypothesis, henceforth called RH-I, is the one tested by Williams (1970), Harkness (1983),

³ Note that Θ is parametric under Cobb–Douglas technologies, in which case the value terms formulation of the HOV model also holds under non-equalized factor prices, as long as the 'law of one price' holds for all commodities. In this case, equation (2) becomes: $\Theta Pt_j = W_j f_j - \alpha_j \sum_{i=1}^m W_i f_i$. For other treatments on non-equalized factor prices within the HOV model, see Bertrand (1972), and Brecher and Choudhri (1982b).

⁴ Maskus (1985, p. 206) uses the terms 'weak' versus 'strong' prediction to indicate this difference.

⁵ $(T_{ji}/E_{ji}) > (T_{jh}/E_{jh})$ if and only if $-E_{wi}/E_{ji} > -E_{wh}/E_{jh}$ or, equivalently, if $E_{ji}/E_{wi} > E_{jh}/E_{wh}$.

and Maskus (1985). In a somewhat modified form it has been tested by Leamer (1980), Stern and Maskus (1981), and Sveikauskas (1983).⁶

A second rank order hypothesis directly following from equation (3) would state that ranking factors in terms of absolute net factor exports is equivalent to ranking them in terms of the difference between country j 's endowment and world endowment scaled down by α_j , country j 's share in world absorption. This will henceforth be called RH-II.

Some writers, like Casas and Kwan Choi (1985) and Bowen *et al.* (1987), have argued that it is intuitively more appealing to use income ratios instead of expenditure ratios for the definition of factor abundance. This can be done by adjusting the factor content part of the equation instead of the endowment part for the aggregate trade imbalance. Thus, under balanced trade

$$T_{ji}^* = E_{ji} - \frac{y_j}{y_w} E_{wi}, \quad (5)$$

where a * superscript indicates hypothetical balanced trade variables. Equation (5) assumes that the same prices would clear world markets if trade were balanced as in the actual case of unbalanced trade. But that is exactly what the assumption of identical and homothetic preferences implies. Hence, equation (5) relates 'hypothetical balanced trade net factor exports' to factor abundance. From equation (4), T_{ji}^* is related to the actual factor content of trade as follows:

$$T_{ji}^* = T_{ji} - \frac{b_j}{y_w} E_{wi}, \quad (6)$$

and if the right-hand sides of both (5) and (6) are observable, they can be compared with one another, and this leads to a third rank order hypothesis, henceforth called RH-III. RH-III is really very similar to RH-II, the only difference being that adjustment for the aggregate trade imbalance is being made on the trade side instead of the endowment side of the equation. Notice that there is no such modified version of the first rank order hypothesis.

Bowen *et al.* (1987) use yet another version of the rank order hypothesis, which differs from RH-III by a multiplicative transformation: every element of both rank orders is divided by the respective world endowment and then divided by y_j/y_w . The resulting measure of factor abundance is $[(E_{ji}/E_{wi})/(y_j/y_w) - 1]$, and the corresponding rank order hypothesis, henceforth called RH-IV, states that ranking factors in terms of $(E_{ji}/E_{wi})/(y_j/y_w)$ is equivalent to ranking them in terms of $(T_{ji}^*/E_{wi})/(y_j/y_w)$, where T_{ji}^* is defined as in (6) above.

The crucial point to be made here is that, given the data do not strictly satisfy equation (2), choosing among these four rank order hypotheses may well be decisive for the outcome of the test.

⁶ The modification in these latter three studies is that net factor exports are related to the factor content of domestic consumption instead of domestic production in order to form the factor content ranking of factors. In doing so they make use of the fact that consumption is the difference between production and trade.

Consider, for instance, RH-I and RH-II. To see that some given data set may, indeed, yield different test results for these two hypotheses, it is sufficient to realize that corresponding entries of the rank orders to be compared are multiplied by a number that varies from entry to entry. Starting with R-I, we obtain R-II by multiplying every element of both rank orders by the respective domestic endowment. And if the elements of one rank order are sufficiently close to each other, and the domestic endowments are sufficiently dissimilar for different factors, this will change the rank order on one side while at the same time preserving it on the other. The following analysis will establish the exact conditions, under which this will happen.

Thus, suppose RH-I is satisfied while RH-II is not. This can be written as

$$\frac{T_{ji}}{E_{ji}} > \frac{T_{jh}}{E_{jh}}, \quad (7)$$

$$\frac{1 - \alpha_j \frac{E_{wi}}{E_{ji}}}{1 - \alpha_j \frac{E_{wh}}{E_{jh}}} > 1, \quad (8)$$

$$T_{ji} > T_{jh}, \quad (9)$$

$$\frac{1 - \alpha_j \frac{E_{wi}}{E_{ji}}}{1 - \alpha_j \frac{E_{wh}}{E_{jh}}} < \frac{E_{jh}}{E_{ji}}. \quad (10)$$

Inequalities (7) and (8) state that the RH-I is supported, whereas inequalities (9) and (10) state that RH-II is violated. In (8) and (10) it has been assumed that $(1 - \alpha_j(E_{wh}/E_{jh})) > 0$. If we also assume $T_{ji} > 0$, we can summarize the above four inequalities as follows:

$$\frac{T_{jh}}{T_{ji}} < 1 < \frac{E_{jh}}{E_{ji}} < \frac{E_{wh}}{E_{wi}}, \quad (11)$$

which is perfectly possible. But E_{jh}/E_{ji} not only has to be greater than one, it also has to be greater than $(1 - \alpha_j(E_{wi}/E_{ji}))/ (1 - \alpha_j(E_{wh}/E_{jh}))$, while at the same time being smaller than E_{wh}/E_{wi} .

Thus, $(1 - \alpha_j(E_{wi}/E_{ji}))/ (1 - \alpha_j(E_{wh}/E_{jh}))$ and E_{wh}/E_{wi} are the lower and upper bounds, respectively, of an interval, in which E_{jh}/E_{ji} has to lie for RH-II to be violated while at the same time RH-I is supported. To be sure, if the data satisfy the strict equality in equation (2) above, E_{jh}/E_{ji} will always lie outside this interval, and any rank order test will support the model. But given some

deviation from this equality in the data at hand, E_{jh}/E_{ji} may well come to lie inside this interval, and the outcome of a rank order test will then depend upon the hypothesis chosen.

If $(1 - \alpha_j(E_{wh}/E_{jh}))$ is negative, inequalities (8) and (10) are reversed, and the equivalent of condition (11) is

$$\frac{T_{jh}}{T_{ji}} < \frac{E_{jh}}{E_{ji}} < \frac{E_{wh}}{E_{wi}} < 1, \quad (12)$$

which is also possible. The above-mentioned interval for E_{jh}/E_{ji} remains the same, but it now lies below unity. For negative $(1 - \alpha_j(E_{wh}/E_{jh}))$, inequality (8) implies a violation of RH-I and inequality (10) implies that RH-II is supported. This again leads to condition (11). Conversely, for positive $(1 - \alpha_j(E_{wh}/E_{jh}))$, such a violation would occur for opposite inequalities in (8) and (10), leading again to condition (12).

There are two cases, in which RH-I and RH-II will necessarily be either jointly violated or jointly supported for the two factors i and h . The first is $E_{ji} = E_{jh}$, and the second is given by opposite signs of the two net factor exports plus opposite signs of $(1 - \alpha_j(E_{wi}/E_{ji}))$ and $(1 - \alpha_j(E_{wh}/E_{jh}))$.

Turning now to RH-II and RH-III, we first note that they differ from one another by an additive transformation, and different test outcomes are caused by circumstances similar to the ones discussed above. Thus, for instance, R-II will be violated and RH-III supported if the following conditions hold in addition to (9) and (10) above:

$$T_{ji} - \frac{b_j}{y_w} E_{wi} > T_{jh} - \frac{b_j}{y_w} E_{wh}, \quad (13)$$

$$E_{ji} - \frac{y_j}{y_w} E_{wi} > E_{jh} - \frac{y_j}{y_w} E_{wh}. \quad (14)$$

This can be summarized as

$$\frac{b_j}{y_w} E_{wi} - \frac{b_j}{y_w} E_{wh} < (E_{ji} - \alpha_j E_{wi}) - (E_{jh} - \alpha_j E_{wh}) < 0 < T_{ji} - T_{jh}. \quad (15)$$

In this case conflicting test results are excluded if $E_{wi} = E_{wh}$ and, of course, if $b_j = 0$, that is if country j 's trade is balanced. In the latter case R-II and R-III coincide. Conversely, a sufficiently large trade imbalance and a sufficiently large difference between world endowment with the two factors i and h cause differing results for tests of RH-II and RH-III.⁷

⁷ The trade balance has also been shown by Aw (1983) to be important for the question of whether or not comparing the factor intensity of exports to that of imports as in Leontief (1953) constitutes a valid test of the Heckscher-Ohlin theorem in the two factor case. By way of contrast, all tests discussed here are equally valid irrespective of the dimensions involved and whatever the value of the trade balance. Instead, the present issue is that a high trade balance will increase the likelihood that different, but equally valid, tests may lead to conflicting results.

Finally, we may consider the relationship between RH-III and RH-IV. But given that they differ from one another by two successive multiplicative transformations, as shown above, we can invoke arguments similar to the ones pertaining to RH-I and RH-II, and hence a more detailed analysis can be omitted here. In particular, for the relationship between RH-III and RH-IV, dissimilarities across factors in world endowments and dissimilarities across countries in their share in world income play a role similar to that of domestic endowments in the comparison, discussed in detail above, between RH-I and RH-II.

All the above rank order hypotheses have been formulated for comparisons across factors for a given country j . This is the case that one usually observes in the literature. But they can, of course, just as well be formulated for comparisons across countries for a given factor. Both procedures are equally valid rank order tests of the HOV model, and will, accordingly, both be covered in the empirical analysis below.

4. Sign hypotheses

As mentioned in the introduction, it has become standard practice to also perform sign tests on the basis of the HOV model. Clearly, there is a corresponding sign test to every one of the different rank order tests discussed above. If we indicate these sign tests as SH-I through SH-IV, it is immediately obvious that SH-I and SH-II will always be simultaneously rejected or supported. The same holds true for SH-III and SH-IV. But choosing between SH-II and SH-III may be decisive for the outcome of the test. Thus, for instance, SH-III will be rejected and SH-II supported if

$$T_{ji} > 0, \quad (16)$$

$$E_{ji} - \alpha_j E_{wi} > 0, \quad (17)$$

$$T_{ji} - \frac{b_j}{y_w} E_{wi} > 0, \quad (18)$$

$$E_{ji} - \frac{y_j}{y_w} E_{wi} < 0. \quad (19)$$

(16) and (17) state that SH-II is supported, whereas (18) and (19) state that SH-III is rejected. Taken together, the above four inequalities imply

$$0 < \frac{y_j}{y_w} E_{wi} - E_{ji} < \frac{b_j}{y_w} E_{wi} < T_{ji}. \quad (20)$$

Thus, if b_j/y_w and/or E_{wi} are sufficiently large for $(b_j/y_w) E_{wi}$ to be greater than $(y_j/y_w) E_{wi} - E_{ji}$, while at the same time being less than T_{ji} , SH-III will be rejected and SH-II will be supported. As with RH-II and RH-III above, we observe a decisive role of the trade balance.

Again, if the data satisfy equation (2) exactly, the situation just discussed will never be observed, but in actual practice there will always be some deviations from this exact equality, and this may result in conflicting test results.

5. Empirical analysis

Having established conditions for non-robustness of rank order and sign tests of the HOV theorem, a natural question to ask now is whether or not these conditions are likely to be met by a typical data set in the actual practice of empirical work. One of the most comprehensive data sets ever compiled in the spirit of the Heckscher–Ohlin theory has been published in Bowen *et al.* (1987), henceforth BLS, and I have used this data set for a thorough empirical analysis of the robustness problem discussed above. The data cover the 1967 factor content of net exports, 1966 factor endowments, 1966 GNP, and the trade balance of 1966 for twenty-seven countries and a total of twelve factors. As in BLS, world factor endowments were calculated as the sum of the twenty-seven country endowments. Similarly, world GNP was approximated by summing up the twenty-seven country GNPs.

I have tested all four rank propositions, both across factors and across countries, and the two sign propositions SH-II and SH-III. The focus of the analysis was not so much whether or not this data set support the HOV model, but rather whether or not, or to what extent these tests do or do not exhibit robustness in the sense discussed above.

The results are summarized in Table 1, which gives the percentages of turnovers in test outcomes observed for different combinations of hypotheses. It is quite clear from Table 1 that the phenomenon of non-robustness as discussed above is of considerable relevance in the present data set. The *factor rankings* for individual countries appear to be much more sensitive to the particular hypothesis chosen than the *country rankings* for individual factors.

TABLE 1
Total Number* (in %) of Changes in Test Results between Different Hypotheses

		RH-II	RH-III	RH-IV	S-III
RH-I	Factor rankings	42.09	48.60	43.76	
	Country rankings	27.27	25.98	8.64	
RH-II	Factor rankings		34.52	43.76	
	Country rankings		23.18	24.13	
RH-III	Factor rankings			27.64	
	Country rankings			26.66	
SH-II					46.30

* The total number of two-by-two factor rankings is $66 \times 27 = 1,782$, and the total number of two-by-two country rankings is $351 \times 12 = 4,212$. The total number of sign comparisons is $27 \times 12 = 324$.

There are sixty-six two-by-two factor rankings for every country. This gives a total number of 1,782 factor rankings, of which more than 42% change from being correct under RH-I to being incorrect under RH-II, or vice versa. Changes of similar frequency occur between RH-I and RH-III, RH-I and RH-IV, and RH-II and RH-IV, whereas the number of changes is somewhat lower, but still considerable, for RH-II and RH-III, and RH-III and RH-IV.

The country rankings for individual factors are, overall, less sensitive. Of the 4,212 possible two-by-two rankings, between 23 and 27% change from being correct under one to being incorrect under another hypothesis, or vice versa, with a low of 8.64% for RH-I and RH-IV. The number of inversions is again considerable between the two sign hypotheses SH-II and SH-III: 46.3%.

While Table 1 clearly establishes the empirical relevance of the phenomenon of non-robustness of rank order and sign tests of the HOV theorem, this does not necessarily show up in summary statistics across factors or countries. These are presented in Tables 2 and 3, which correspond to Tables 3 and

TABLE 2
Different Rank Order and Sign Tests, Country by Country for all Twelve Factors*

Country	RH-I		RH-II		RH-III		RH-IV		SH-II	SH-III
	%	τ	%	τ	%	τ	%	τ		
Argentina	54.55	0.09	50.00	0.00	37.88	-0.24	57.58	0.15	66.67	33.33
Australia	40.91	-0.18	39.39	-0.21	39.39	-0.21	40.91	-0.18	41.67	33.33
Austria	65.15	0.30	34.85	-0.30	50.00	0.00	46.97	-0.06	25.00	66.67
Belg.-Lux.	68.18	0.36	78.79	0.58	54.55	0.09	65.15	0.30	75.00	58.33
Brazil	84.85	0.70	74.24	0.48	46.97	-0.06	84.85	0.70	41.67	16.67
Canada	54.55	0.09	66.67	0.33	68.18	0.36	60.61	0.21	75.00	75.00
Denmark	31.82	-0.36	31.82	-0.36	30.30	-0.39	28.79	-0.42	33.33	50.00
Finland	60.61	0.21	50.00	0.00	66.67	0.33	57.58	0.15	33.33	75.00
France	75.76	0.52	62.12	0.24	62.12	0.24	72.73	0.45	66.67	41.67
Germany	72.73	0.45	89.39	0.79	75.76	0.52	77.27	0.55	91.67	66.67
Greece	83.33	0.67	72.73	0.45	93.94	0.88	75.76	0.52	50.00	91.67
Hong-Kong	96.97	0.94	84.85	0.70	96.97	0.94	92.42	0.85	50.00	100.00
Ireland	75.76	0.52	74.24	0.48	93.94	0.88	69.70	0.39	50.00	91.67
Italy	65.15	0.30	81.82	0.64	81.82	0.64	66.67	0.33	58.33	66.67
Japan	72.73	0.45	63.64	0.27	63.64	0.27	69.70	0.39	83.33	66.67
Korea	96.97	0.94	46.97	-0.06	62.12	0.24	66.67	0.33	16.67	75.00
Mexico	89.39	0.79	74.24	0.48	90.91	0.82	83.33	0.67	41.67	91.67
Netherlands	37.88	-0.24	28.79	-0.42	33.33	-0.33	34.85	-0.30	33.33	58.33
Norway	42.42	-0.15	36.36	-0.27	31.82	-0.36	37.88	-0.24	58.33	25.00
Philippines	87.88	0.76	65.15	0.30	66.67	0.33	80.30	0.61	33.33	50.00
Portugal	65.15	0.30	40.91	-0.18	59.09	0.18	54.55	0.09	25.00	66.67
Spain	75.76	0.52	56.06	0.12	53.03	0.06	56.06	0.12	33.33	66.67
Sweden	65.15	0.30	50.00	0.00	59.09	0.18	56.06	0.12	66.67	41.67
Switzerland	78.79	0.58	71.21	0.42	65.15	0.30	69.70	0.39	75.00	66.67
UK	84.85	0.70	69.70	0.39	90.91	0.82	74.24	0.48	50.00	91.67
USA	45.45	-0.09	39.39	-0.21	56.06	0.12	54.55	0.09	25.00	50.00
Yugoslavia	63.64	0.27	21.21	-0.58	50.00	0.00	50.00	0.00	16.67	83.33

* Per cent is the percentage of correct rankings in the total number of sixty-six two-by-two rankings of factors and τ is Kendall's coefficient of rank correlation (significant at the 5% level if boldfaced).

TABLE 3
Different Rank Order and Sign Tests, Factor by Factor for all Twenty-seven countries*

Factor	RH-I		RH-II		RH-III		RH-IV		SH-II	SH-III
	%	τ	%	τ	%	τ	%	τ		
Capital	73.50	0.47	55.84	0.12	49.00	-0.02	57.26	0.15	59.26	55.56
Total labour	69.23	0.38	51.28	0.03	50.71	0.01	59.26	0.19	44.44	66.67
Agric. labour	56.13	0.12	53.85	0.08	51.85	0.04	60.11	0.20	55.56	62.96
Cler. labour	77.49	0.55	49.86	-0.00	62.11	0.24	57.55	0.15	29.63	74.07
Prof. labour	80.91	0.62	55.56	0.11	48.15	-0.04	56.13	0.12	14.81	77.78
Man. labour	78.06	0.56	46.15	-0.08	22.22	-0.56	37.32	-0.25	62.96	22.22
Prod. labour	63.82	0.28	52.42	0.05	61.25	0.23	67.24	0.34	29.63	70.37
Sales labour	74.36	0.49	50.14	0.00	53.28	0.07	61.25	0.23	37.04	66.67
Service labour	73.22	0.46	45.58	-0.09	49.86	-0.00	64.10	0.28	40.74	66.67
Arable land	67.52	0.35	79.20	0.58	78.92	0.58	76.64	0.53	70.37	70.37
Forest	69.23	0.38	76.92	0.54	76.07	0.52	67.81	0.36	70.37	70.37
Pasture	51.28	0.03	68.38	0.37	66.95	0.34	59.83	0.20	70.37	51.85

* Per cent is the percentage of correct rankings in the total number of 351 two-by-two rankings of countries and τ is Kendall's coefficient of rank correlation. Boldfaced τ -values indicate statistical significance at the 5% level.

2, respectively, in BLS.⁸ Columns 1 through 8 of Table 2 give the percentages of correct factor rankings, two at a time, for every country under the four rank order hypotheses, and the corresponding Kendall rank correlation coefficients. Columns 9 and 10 show the percentages of matching signs for the two sign predictions SH-II and SH-III. Table 3 presents analogous summary statistics for individual factors.

It is evident that the non-robustness established by Table 1 does not affect these summary statistics equally for all countries or factors. Thus, some of the countries, like Australia, Denmark, the Netherlands, and Norway show very similar results for all four rank order hypotheses, whereas different hypotheses yield drastically different summary statistics for Austria, Brazil, Korea, and the UK. Among the twelve factors, capital and the different categories of labour appear to be much more affected in the summary rank order statistics by the non-robustness established above than do total labour and the three categories

⁸ Column 8 and 7 of the present Tables 2 and 3 should correspond to columns 2 and 3, respectively, of Tables 3 and 2 in BLS. However, I was not able to exactly reproduce all of their results. With only one minor exception (arable land) the rank correlation coefficients for individual factors in column 8 of Table 3 above are equal to those reported by BLS in column 2 of their Table 2, but the same does not hold true for the percentages of correct rankings (column 7 in Table 3 above and column 3 in Table 2 of BLS). The percentages of sign matches in these two tables are the same except for capital and clerical labour. An attempt to reproduce their Table 1 revealed differences for clerical labour in all countries, and capital in France and Germany, and arable land in Italy.

An exact correspondence between columns 2 and 3 of their Table 3 and our columns 8 and 7 of Table 2 was not to be expected since their calculations excluded total labour, whereas the above figures include total labour. The differences, however, are almost negligible in most cases.

of land. The two sign predictions in general show significantly more turnovers than any two rank order predictions.

It might be argued that the differences shown by Tables 2 and 3 between the four different rank order hypotheses are, overall, surprisingly low, given the non-robustness established by Table 1. One might even be tempted to conclude that this non-robustness does not pose any serious problem for reaching a general conclusion from the summary statistics. The general picture portrayed by columns 1 and 2, it might be argued, is not really drastically different from that portrayed by columns 3 and 4, or 5 and 6, or 7 and 8. But this is a rather delicate issue, because Tables 2 and 3 may mask some of the existing non-robustness. This is because the percentages given in Table 1 represent switches in both directions, i.e. from correct rankings and signs to incorrect ones and vice versa. Thus, it would be possible, in principle, to observe a 100% turnover of results between two different hypotheses, while at the same time consistently observing a 50% sign or rank order match for both hypotheses. Concluding a 50% support of the theory in such a situation would appear very odd indeed, since this support would originate from two completely disjoint subsets of the sample. This suggests that the appropriate way to examine the robustness of rank order and sign tests of the HOV model is the one underlying Table 1 above. Without such an examination, any test result must remain seriously in doubt.

6. Re-examinations of the Leontief paradox

Another example of a non-robust sign test can be found in the famous Leontief paradox.⁹ This has become evident in a recent debate, in which various authors have re-examined Leontief's original data through the lens of the HOV-model, and in which contradictory test results have emerged. It is very important for a correct interpretation of this debate to realize that what we have here is nothing but an example of non-robustness in the sense discussed above. Since this has not generally been acknowledged, it seems worthwhile to be shown in some detail.

The debate was sparked off by Leamer (1980), who could show that, contrary to the long tradition of interpreting Leontief's findings as paradoxical, the original data do, in fact, support RH-I. It must be emphasized that this is *not* a question of robustness in the sense discussed above. It is simply a matter of using a generally valid rank order test, such as RH-I, as opposed to Leontief's original procedure which is not generally correct.¹⁰ However, Brecher and Choudhri (1982a) have subsequently pointed out that Leontief's data violate SH-II for labour, and they conclude that this constitutes a modified Leontief paradox. This has, in turn, been questioned by Casas and Kwan Choi (1985), who argue that the Brecher-Choudhri version of the paradox is nothing but

an artefact of the aggregate trade surplus of the 1947 US economy, and that adjusting for this surplus removes the paradox. It is at this stage that we observe the non-robustness, because it can be shown that the only difference between Casas and Kwan Choi and Brecher and Choudhri is that the former choose SH-III instead of SH-II. And from the above analysis it is then no longer surprising that they obtain a different test result.

Casas and Kwan Choi's starting-point is equation (5) above, which links 'hypothetical balanced trade net factor exports' to their preferred measure of factor endowment. Their presumption is that this endowment measure was clearly negative for the 1947 US economy, and thus T_{ji}^* should also be negative. Since they do not observe E_{wi} , they cannot use equation (6) for an independent observation of T_{ji}^* . But identical and homothetic tastes, a standard assumption of the HOV model, imply that the factor content of domestic absorption is $(y_j - b_j/y_w)E_{wi}$, and substituting for E_{wi} in equation (6) then gives

$$T_{ji}^* = E_{ji} - \frac{y_j}{y_j - b_j} C_{ji}, \quad (21)$$

where C_{ji} is the factor content of domestic absorption. This is the equation used by Casas and Kwan Choi to infer 'hypothetical balanced trade net factor exports', whereby C_{ji} is taken from $C_{ji} = E_{ji} - T_{ji}$. But this, together with $C_{ji} = (y_j - b_j/y_w)E_{wi}$, gives equation (6) above, which shows that Casas and Kwan Choi, in effect, test SH-III.

It is essential for a correct interpretation of this debate to realize that SH-II and SH-III are equally valid tests of the HOV model. In particular, SH-II is by no means less satisfactory than SH-III in its adjustment for an aggregate trade imbalance. This was already pointed out above. Hence, rather than concluding a definite resolution of the modified Leontief paradox from the result obtained by Casas and Kwan Choi, we realize that what we observe here is another example of non-robustness of sign tests of the HOV model.

7. Conclusion

There is a sizable body of literature in which Vanek's (1968) generalized version of the factor proportions theory has been subject to empirical testing. A common problem faced by much of this literature was that the model in its strict form was regarded as being 'clearly incredible' without any empirical observation. And the general strategy that has emerged from this problem was to derive 'weakened' propositions that somehow maintain the essence of the theory without being 'a priori incredible'. These 'weakened' propositions were either rank order or sign propositions regarding the factor content of net exports and various measures of factor abundance. The specific formulation of the rank order or sign hypothesis was, in general, made dependent on which specific measure of factor abundance one considered to be most appealing. In any case, this question was generally regarded as being of secondary importance.

⁹ The material of this section draws on Kohler (1988b).

¹⁰ In addition to Leamer (1980), see also Aw (1983).

The present paper has shown that this practice is seriously flawed by the fact that different formulations of the rank order or sign proposition may lead to different test results for a given data set, if these data do not satisfy the strict version of the model. Such tests of the HOV model may thus show a considerable amount of non-robustness against seemingly harmless transformations of the rank order or sign hypothesis. This was shown both analytically and empirically, whereby use was made of the comprehensive multi-country data set published by Bowen *et al.* (1987). This paper focused on four different rank order and two different sign hypotheses, and in some cases as many as 40% of the two-by-two rankings or sign structures turned out to change from being correct under one to being incorrect under another hypothesis, or vice versa.

There are two conclusions that one can draw from these results. First, given the non-robustness established above, the practice of using intuition to choose one among the various possible specific formulations of the rank order or sign hypothesis appears highly problematical. If one wishes to perform such rank order or sign tests at all, one should at least also examine the robustness of these tests for the given data set, such as illustrated in Table 1 above. Drawing conclusions from any one rank order or sign test without addressing the issue of robustness may involve a significant amount of hidden arbitrariness.

Secondly, the results of the present paper suggest a wholly different approach to testing the factor proportions theory. Rather than trying to see whether or not the data support a more or less arbitrarily 'weakened' form of the 'incredible' model, tests of the theory should be aimed at trying to see just how well, or how badly, the data support the model in its strict form. Moreover, the phenomenon of non-robustness established in this paper offers additional reinforcement for the strategy, followed by Bowen *et al.* (1987), of embedding the Heckscher–Ohlin factor endowment proposition in a more general model allowing specific deviations from the factor endowment explanation of trade, and testing the Heckscher–Ohlin proposition against specific alternatives in a regression framework.

University of Innsbruck, Austria, and University of Konstanz, W. Germany

REFERENCES

- AW, B. Y. (1983). 'Trade Imbalance and the Leontief Paradox', *Weltwirtschaftliches Archiv*, **119**, 734–8.
- BERTRAND, T. J. (1972). 'An Extension of the N-Factor Case of the Factor Proportions Theory', *Kyklos*, **25**, 592–6.
- BOWEN, H. P., LEAMER, E. E., and SVEIKAUSKAS, L. (1987). 'Multicountry, Multifactor Tests of the Factor Abundance Theory', *American Economic Review*, **77**, 791–809.
- BRECHER, R. A. and CHOUDHRI, E. U. (1982a). 'The Leontief Paradox, Continued', *Journal of Political Economy*, **90**, 820–3.
- and — (1982b). 'The Factor Content of International Trade without Factor-Price Equalization', *Journal of International Economics*, **12**, 277–83.
- CASAS, F. R. and KWAN CHOI, E. (1985). 'The Leontief Paradox: Continued or Resolved?', *Journal of Political Economy*, **93**, 610–15.
- DEARDORFF, A. V. (1982). 'The General Validity of the Heckscher–Ohlin Theorem', *American Economic Review*, **72**, 683–94.
- (1984). 'Testing Trade Theories and Predicting Trade Flows'. In R. W. Jones, and P. B. Kenen (eds.), *Handbook of International Economics*, vol. 1, pp. 467–517. Amsterdam: North-Holland.
- GIFT, R. and MARXSEN, W. (1984). 'Aggregation and the Factorial Content of Trade', *Journal of Political Economy*, **92**, 979–84.
- HARKNESS, J. (1983). 'The Factor-Proportions Model with Many Nations, Goods and Factors: Theory and Evidence', *Review of Economics and Statistics*, **65**, 298–305.
- HORIBA, Y. (1971). 'A Note on the Factor Proportions Theory in the N-Factor Case', *Kyklos*, **24**, 339–43.
- KOHLER, W. K. (1988a). 'Modeling Heckscher–Ohlin Comparative Advantage in Regression Equations: A Critical Survey', *Empirica—Austrian Economic Papers*, **15**, 263–93.
- (1988b). *A Further Note on the Leontief Paradox*. Research Seminar in International Economics, University of Michigan, Seminar Discussion Paper No. 212.
- LEAMER, E. E. (1980). 'The Leontief Paradox Reconsidered', *Journal of Political Economy*, **88**, 495–503.
- (1984). *Sources of International Comparative Advantage. Theory and Evidence*. Cambridge, Mass.: MIT Press.
- BOWEN, H. P. (1981). 'Cross Section Tests of the Heckscher–Ohlin Theorem: Comment', *American Economic Review*, **71**, 1040–43.
- LEONTIEF, W. (1953). 'Domestic Production and Foreign Trade: The American Capital Position Re-examined', *Economia Internazionale*, **7**, 9–38.
- MASKUS, K. E. (1985). 'A Test of the Heckscher–Ohlin–Vanek Theorem: The Leontief Commonplace', *Journal of International Economics*, **19**, 201–12.
- STERN, R. M. and MASKUS, K. E. (1981). 'Determinants of the Structure of U.S. Foreign Trade (1958–76)', *ibid.* **11**, 207–24.
- SVEIKAUSKAS, L. (1983). 'Science and Technology in the United States Foreign Trade', *Economic Journal*, **93**, 542–54.
- TRAVIS, W. P. (1964). *The Theory of Trade and Protection*. Cambridge, Mass.: Harvard University Press.
- VANEK, J. (1968). 'The Factor Proportions Theory: The N-Factor Case', *Kyklos*, **21**, 749–55.
- WILLIAMS, J. R. (1970). 'The Resource Content in International Trade', *Canadian Journal of Economics*, **3**, 111–22.

