5

The Factor Content of Trade

Donald R. Davis and David E. Weinstein

Chapter Outline

Study of the factor content of trade has become a laboratory to test our ideas about how the key elements of endowments, production, absorption, and trade fit together within a general equilibrium framework. Already a great deal of progress has been made in fitting these pieces together. Nevertheless, the existing research raises a great many questions that should help to focus empirical research in the coming years. Among the more pressing issues is a deeper consideration of the role of intermediates, the role of aggregation biases, and of differences in patterns of absorption. This work should provide a more substantial foundation for future policy work developed within a factor content framework.

1 Introduction

The concept of the factor content of trade originates with Vanek (1968). The original formulation is based on a simple model of international factor price equalization, or what is more precisely termed the "integrated equilibrium" (Helpman and Krugman (1985)). Under conditions of competition in goods and factor markets, free international arbitrage, common constant returns to scale technologies, and adequate restrictions on the distribution of world endowments, both goods and factor prices will be equalized internationally. Under these conditions, a good will embody fixed amounts of the services of the productive factors, independently of where it is produced. Trade then can be conceived of in two ways. The first is as the overt exchange of goods that traditional theory addresses. Vanek's contribution was to recognize that we could equally think of trade as the international exchange of the services of factors embodied in those goods. Vanek's formulation of the problem allowed an extension of the logic of the Heckscher–Ohlin theory to settings in which the pattern of trade may be indeterminate but in which the net factor content of trade may nonetheless be determinate. Expression of the theory in this form also highlights the deep logic of the Heckscher–Ohlin theory in its focus on the relative availability of factors. When we move beyond a fully integrated world economy, as we do at length below, we will have to take care in defining the factor content of trade appropriate to each setting.

A reasonable first question is why we should care about the factor content of trade. We think there are two good reasons. The first is that the study of the factor content of trade is a laboratory for general equilibrium. A first statement of general equilibrium is that the elements of the system should "hang together." In the case of international trade, the elements of interest are the technologies, productive endowments, outputs, and demands of all countries in the world. Study of the factor content of trade then becomes a first test of the reasonableness of our assumptions about how these elements interact. If our theories perform poorly in matching measured and predicted factor contents of trade, then this may point in directions in which our theories need to be modified for understanding the world. If our theories perform well, then this indicates that the relevant framework may be a reasonable representation of the world and so also a reasonable framework for policy studies.

Indeed, the second important reason for considering studies of the factor content of trade is precisely that they may one day prove helpful in addressing policy questions of the impact of openness on national income levels and distribution. There already is a substantial applied literature mapping measures of the factor content of trade into impacts on domestic relative wages for the US and other members of Organization for Economic Cooperation and Development (OECD). Likewise, there has developed a theoretical literature seeking to establish the conditions under which such a mapping makes sense. We believe that these literatures have been very important in clarifying the issues that need to be addressed in future work. However, we also believe that the results of the empirical studies must be treated with caution, since the theoretical frameworks in which such calculations have been shown to make sense bear little resemblance to the frameworks preferred in the studies of the factor content of trade. We conclude that a major area for future work is taking the empirical frameworks favored by the studies of factor content and working out within them the consequences of international integration on incomes and inequality.

The past 15 years have seen wide swings in trade economists' views of models of the factor content of trade. Early studies, such as Bowen, Leamer, and Sveikauskas (1987, hereafter BLS) and Maskus (1985), seemed very damning. It was not so easy to see that this represented only a phase in the development of the literature. More recent studies, such as Trefler (1995) performed the signal service of identifying anomalies in the data which further research could aim to understand. The most recent studies, such as Davis and Weinstein (2001a), have been much more positive for amended versions of the theory.

We do not at all want to suggest that all issues about the factor content of trade are settled. Future work needs to gather better and more extensive data sets, to consider more carefully the role of traded intermediates, cross-country differences in demand, the role of trade costs, and so on. But the progress made in the past 15 years surely holds promise that this will continue to be a fertile area for research.

2 THEORY

2.1 The Simple Heckscher–Ohlin–Vanek Model

We begin with the standard model of Heckscher–Ohlin–Vanek (HOV). Let there be G goods, each produced under perfect competition with constant returns to scale. Let there be F primary factors of production with factor markets competitive. Let technologies for all goods and quality of all factors be common for all countries of the world. Let there be at least as many goods as factors, i.e., $G \geq F$. Assume that trade between countries is free, so that goods' prices are equalized. Assume that the distribution of world endowments among countries satisfies the requirements to replicate the integrated equilibrium.[1] Then factor prices will be equalized (FPE), and for all countries $c \in C$, there is a common technology matrix:

$$\mathbf{B} = \mathbf{B}^c \tag{5.1}$$

Columns of this matrix represent input coefficients for a given good. Rows represent the input coefficients for a given factor across all goods.

For country c, let the output vector be given by $\mathbf{X}^c$ and the primary input vector be given by $\mathbf{V}^c$. Then under the maintained assumption that the technology matrix $\mathbf{B}$ is common to all countries, full employment of resources implies that:

$$\mathbf{B}\mathbf{X}^c = \mathbf{V}^c \tag{5.2}$$

Demand is assumed to be identical across countries and homothetic. Let $\mathbf{D}^c$ be country c's vector of final goods demand, $\mathbf{X}^W$ be the world output vector, and s^c be country c's share of world spending. Then, with free trade equalizing goods prices,

$$\mathbf{D}^c = s^c \mathbf{X}^W \tag{5.3}$$

Equation (5.3) provides a statement about demand for goods. By pre-multiplying by the common technology matrix $\mathbf{B}$, we can convert this to a statement about the factor content of consumption. First, we note that under the hypothesis of a common $\mathbf{B}$,

$$\mathbf{B}\mathbf{X}^W = \mathbf{V}^W \tag{5.4}$$

Then it follows that:

$$\mathbf{BD}^c = s^c\mathbf{V}^W \tag{5.5}$$

Finally, noting that the net trade vector is $\mathbf{T}^c = \mathbf{X}^c - \mathbf{D}^c$, we can difference equations (5.2) and (5.5) to arrive at the statement of the simple HOV model:

$$\mathbf{BT}^c = \mathbf{V}^c - s^c\mathbf{V}^W \tag{5.6}$$

For future reference, it is convenient to call $\mathbf{BT}^c$ the *measured* factor content of trade (MFCT) and $\mathbf{V}^c - s^c\mathbf{V}^W$ the *predicted* factor content of trade (PFCT). The first depends on trade flows weighted by a technology matrix. The latter is based on endowments relative to average endowments for a country of that size in the world.

It is straightforward to incorporate non-traded goods into the model with FPE.[2] Let $\mathbf{V}^{cN}$ be country c's devotion of primary factors to non-traded production. Then the residual endowments available for production of exportables are $\mathbf{V}^{cT} \equiv \mathbf{V}^c - \mathbf{V}^{cN}$. Similarly, for the world as a whole, the endowments devoted to production of tradables are $\mathbf{V}^{WT} = \Sigma_c\mathbf{V}^{cT}$. The predicted factor content of trade will then be the difference between the residual factors available for production of exportables and the factor content of consumption of tradables, or $\mathbf{BT}^c = \mathbf{V}^{cT} - s^c\mathbf{V}^{WT}$. However, with FPE and free trade, it is also true that $\mathbf{V}^{cN} = s^c\mathbf{V}^{WN}$. If we note this and add $\mathbf{V}^{cN} - s^c\mathbf{V}^{WN} = 0$ to the right-hand side of $\mathbf{BT}^c = \mathbf{V}^{cT} - s^c\mathbf{V}^{WT}$, we see that this returns us to equation (5.6). That is, so long as there is FPE, the presence of non-traded goods affects the predicted factor content of trade not at all, and empirical researchers are free to ignore them in spite of the fact that non-traded sectors are in practice very large. As we will see below, this changes importantly when FPE no longer holds.

2.2 Incorporating Intermediates into the HOV Model

When output of goods requires inputs both of primary factors and other goods as intermediates, we need to amend the foregoing. The matrix $\mathbf{B}$ should now be interpreted as the matrix of primary, or direct, factor inputs. In addition, there is an input–output matrix $\mathbf{A}$, with dimension $G \times G$. Each element in the input–output matrix is the unit input requirement of one good in the production of another good, where it is important to remember that industries may use their own output as an input. A row of the input–output matrix indicates the unit input requirement of a given good in the production of all other goods (e.g., how much steel is used in the production of a unit of trucks, planes, etc.). A column, then, indicates how a given good uses all other goods (e.g., how much steel, trucks, etc. are used in the production of a unit of planes). With the presence of intermediate usage, we have to distinguish between gross output, which we now call $\mathbf{X}^c$ and net output available for final demand, which we denote by $\mathbf{Y}^c$. Let $\mathbf{I}$ be the $G \times G$ identity matrix. Then the relation between net and gross output is given simply by:

$$\mathbf{Y}^c = (\mathbf{I} - \mathbf{A})X^c \tag{5.7}$$

Equations (5.2) and (5.4) continue to hold as factor market clearing conditions even in the presence of intermediates. Of course, only net output is available for final consumption, so (5.3) must be amended to:

$$\mathbf{D}^c = s^c \mathbf{Y}^W \tag{5.8}$$

Assuming $(\mathbf{I} - \mathbf{A})$ is invertible, we can define a new matrix of total, or direct plus indirect, factor inputs, given by:

$$\overline{\mathbf{B}} = \mathbf{B}(\mathbf{I} - \mathbf{A})^{-1} \tag{5.9}$$

With a little algebra, this allows for a statement of HOV in the presence of intermediates

$$\overline{\mathbf{B}}\mathbf{T}^c = \mathbf{V}^c - s^c \mathbf{V}^W \tag{5.10}$$

2.3 First Tests of the HOV Model

Equation (5.10) is based on observable variables, so can provide a test of the HOV model. A welcome feature for empirical implementation is that it can be implemented even if the researcher has data on only a subset of the primary factor inputs. Let $\overline{\mathbf{B}}^f$ be the fth row of the technology matrix, V^{fc} be country c's endowment of factor f, and V^{fW} be the corresponding sum for the world. Then we can imagine constructing matrices of HOV predictions with dimension $F \times C$ equal to the number of factors times the number of countries on each side of the equation. The typical elements of such matrices will be of the form:

$$\overline{\mathbf{B}}^f \mathbf{T}^c = \mathbf{V}^{fc} - s^c \mathbf{V}^{fW} \tag{5.11}$$

Various tests could be applied to this. Since corresponding elements of the matrices are supposed to be equal, one would at least like them to have the same sign. That is, one would like countries to be measured to export the services of factors that the theory identifies as abundant there. This may be called a sign test. An alternative is to note that corresponding columns and rows of the matrices are supposed to be equal, so one can test the weakened hypothesis that they will have a high rank correlation. For columns, this implies that, holding fixed a country, measured net exports by factor correspond to the abundance across these factors indicated by endowments relative to typical country endowments. For a row, this implies that, holding fixed the factor, measured net factor service exports across countries correspond to those predicted based on national and world endowments.

One strong requirement for implementing (5.10) is that it requires data on world factor endowments. This motivates an alternative that is frequently employed. Divide both sides of the equation by the income share and you get:

$$\frac{\bar{\mathbf{B}}\mathbf{T}^c}{s^c} = \frac{\mathbf{V}^c}{s^c} - \mathbf{V}^W \tag{5.12}$$

Do this for two countries, c and c', take the difference and multiply through by the income share of the first, and you obtain:

$$\bar{\mathbf{B}}\left(\mathbf{T}^c - \frac{s^c}{s^{c'}}\mathbf{T}^{c'}\right) = \mathbf{V}^c - \frac{s^c}{s^{c'}}\mathbf{V}^{c'} \tag{5.13}$$

As noted, this has the strong advantage in empirical implementation that one needs data only on the countries for which the bilateral comparisons are to be made. However, there is also a disadvantage. Suppose that the true model, instead of being (5.10) is instead:

$$\bar{\mathbf{B}}\mathbf{T}^c = \mathbf{V}^c - s^c\mathbf{V}^W + s^c\mathbf{\Lambda}^{c,c'} \tag{5.14}$$

where the last term is any systematic deviation from HOV that is proportional to country size but otherwise common among the included countries.[3] If we apply the bilateral difference approach to (5.14), we again end up with (5.13). That is, the approach of (5.13) will not detect systematic and potentially large deviations from HOV that are of the form indicated by the last term in (5.14).

2.4 Adjusted Factor Price Equalization

The requirement in (5.1) that input requirements are identical everywhere is stringent, but can be relaxed. Suppose that across countries, there are differences in factor quality of a pure factor augmenting nature. In such a case, we need to distinguish between natural and efficiency units of factors. Suppose that there exists a common technology matrix $\bar{\mathbf{B}}$ and, for each factor and country, an adjustment scalar π^{fc} that satisfies:

$$\bar{\mathbf{B}}^{fc} = \pi^{fc}\bar{\mathbf{B}}^f \tag{5.15}$$

Then we can say that the endowments of country c expressed in efficiency units are:

$$\mathbf{V}^{cE} = \frac{1}{\pi^{fc}}\mathbf{V}^c \tag{5.16}$$

If all the other requirements of the HOV world are satisfied, this leads to a restatement of the factor content of trade in terms of efficiency units of factors:

$$\bar{\mathbf{B}}\mathbf{T}^c = \mathbf{V}^{cE} - s^c\sum_{c'}\mathbf{V}^{c'E} \tag{5.17}$$

If country 0 is the base that defines $\overline{\mathbf{B}} \equiv \overline{\mathbf{B}}^0$, and if the return for factor f in country 0 is w^{f0}, then for any country c

$$w^{fc} = \frac{w^{f0}}{\pi^{fc}} \tag{5.18}$$

A special case of adjusted FPE is when for a fixed c, $\pi^{fc} = \pi^{fc}$ for all factors. This accommodates a world with scalar country differences in total factor productivity, which here will be interpreted just as common differences in factor quality.

2.5 Failure of Factor Price Equalization

Under a variety of conditions, factor prices may fail to equalize, even in efficiency terms. There are a variety of ways of dealing with this failure of FPE while continuing to work with predictions about factor contents of trade.

2.5.1 An FPE club

One approach is to ask how HOV should be modified if a subset of the world shares FPE, but not necessarily the whole world. Suppose that there is a set of regions, $r \in R$, that shares FPE even though this may not hold for countries $c \notin R$. Call R the FPE club. Then club members share a common technology matrix $\overline{\mathbf{B}}^R$. Hence for each $r \in R$ we continue to have

$$\mathbf{B}^R \mathbf{X}^r = \mathbf{V}^r \tag{5.19}$$

If trade continues to equalize goods prices and the other standard HOV demand assumptions hold, then:

$$\mathbf{D}^r = s^r \mathbf{Y}^W \tag{5.20}$$

We can always pre-multiply this by $\overline{\mathbf{B}}^R$, the difference being that absent world FPE, it is no longer the case that $\overline{\mathbf{B}}^R \mathbf{Y}^W$ will equal $\mathbf{V}^W$. Instead, the corresponding equation is that for members of the FPE club, and a common technology matrix $\overline{\mathbf{B}}^R$,

$$\overline{\mathbf{B}}^R \mathbf{T}^r = \mathbf{V}^r - s^r \overline{\mathbf{B}}^R \mathbf{Y}^W \tag{5.21}$$

This restriction can be examined for all members of the FPE club.

2.5.2 Understanding the breakdown of FPE

A huge advantage of the last approach is that we can apply it to members of the FPE club without taking a stand on why FPE is breaking down. The big drawback of this approach is that we care why FPE breaks down. If FPE fails in a systematic way, then we can identify hypotheses that allow us to test an appropriately modified version of HOV.

If FPE breaks down, then this should show up in an examination of the technology matrices of the respective countries. Hence, a first approach is to return to equation (5.1) and check if there is a common $\mathbf{B} = \mathbf{B}^c$ for all c. FPE would imply that these technology matrices are equal. Unfortunately, a finding that they are not equal need not invalidate FPE if there are more goods than factors and our industrial categories, which define columns of the **B** matrix, are themselves composed of products of heterogeneous factor content.

To make this point clearly, it is best to depart from generality and think about a world of two factors, say of capital and labor. Let there be two countries, with the home country more capital abundant than the foreign. Suppose that there is a Dornbusch–Fischer–Samuelson (1980) continuum of goods with varying capital intensities. Assume as well that the distribution of world endowments is consistent with re-creating the integrated equilibrium. Then the HOV predictions will hold exactly, even though the pattern of trade in goods may not be fully determined because the number of goods exceeds the number of factors. Now assume that there are strictly positive costs of trade, although we can think of them as vanishingly small. In this case, the pattern of trade becomes determinate. The home country concentrates its exports among the most capital-intensive goods, and vice versa for the foreign country. Goods of intermediate factor intensity are non-traded. Formally, the exports of the countries are disjoint. If, however, our system of classifying goods into industries aggregates into the same industries goods of very different factor content produced in the two countries, then we will expect to find in the data that country capital abundance is correlated with industry capital intensity.[4] However, this would only be true for traded goods. With the vanishingly small trade costs doing little to disturb FPE, we should not find this correlation among non-traded goods. A second issue is that the average input coefficients that we calculate in the data are a weighted average of the goods that we actually export and our non-traded goods. In this simple framework, goods of intermediate factor intensity fall into the non-traded sector, so we tend to underestimate the true factor content of trade. Suppose that we could correct for this problem and define a technology matrix $\overline{\mathbf{B}}^{cDFS}$ that reflects the actual factor intensity of production of c's exports, E^c. Let the imports of c from a country c' be $\mathbf{M}^{cc'}$ (and for simplicity, let country c's demand for its own output be denoted $\mathbf{M}^{cc}$), then the appropriate HOV equation is:

$$\overline{\mathbf{B}}^{cDFS}\mathbf{Y}^c - \left[\overline{\mathbf{B}}^{cDFS}\mathbf{D}^{cc} + \sum_{c'}\overline{\mathbf{B}}^{c'DFS}\mathbf{M}^{cc'}\right] = \mathbf{V}^c - s^c\mathbf{V}^W \tag{5.22}$$

The foregoing has allowed for differences in technology matrices for tradables by country, even though "approximate" FPE holds. The reason that the matrices differed for tradables, but not for non-tradables is that the former reflected heterogeneity of goods in an industry in spite of the approximate FPE, while this implied common input coefficients for the non-traded goods where homogeneity is more plausible.

An alternative that can be investigated is a breakdown of FPE, so that there is a systematic correlation between country capital abundance and industry input usage not only in tradables (where this now suggests specialization) but also in non-tradables (where this suggests factor substitution). Interestingly, the key to distinguishing this from the former case is to look for this systematic correlation in the input coefficients in the non-traded sector.

If this indicates a breakdown of FPE, there is yet another adjustment that needs to be made. Capital abundant countries with high wages will use more capital per worker in non-traded sectors, implying that standard measures of excess factor supplies overstate how much of the abundant factor is available for production of exportables, hence tend to predict too high a volume of factor trade. Since, in practice, non-traded sectors are large, these adjustments to the theoretical model may matter quite a lot. Let the appropriate matrix for country *c* in this case be $\mathbf{B}^{cH}$ (after Helpman, who suggested such an approach). Let a superscript *T* indicate traded output and an *N* indicate endowments dedicated to the production of non-traded goods. Then the appropriate measure for this amended HOV model is:

$$\overline{\mathbf{B}}^{cH}\mathbf{Y}^{cT} - \sum_{c'}\overline{\mathbf{B}}^{c'H}\mathbf{M}^{cc'} = [\mathbf{V}^c - s^c\mathbf{V}^W] - [\mathbf{V}^{cN} - s^c\mathbf{V}^{WN}] \qquad (5.23)$$

Here the measured factor flows of trade are measured using the *producer's* technology. The predicted factor flows are also adjusted for the fact that countries abundant in a factor tend to use that factor more intensively in non-traded production, so have less available for production of exportables than indicated in the standard HOV equation.

2.5.3 Cross flows of factor services in a many-cone world

When endowment differences lead to a breakdown of FPE, then a set of equilibrium factor prices and the associated goods that can be competitively produced at these factor prices define a "cone" in factor space. Countries whose endowments lie within the same cone share FPE, while those that lie in different cones do not.

It is simplest to think about this in the case of two factors, say capital and labor. Consider a many-cone world. For simplicity, ignore non-traded goods, assume that each cone contains just one country and that the number of goods produced is sufficiently large that boundary goods produced by countries in different cones can be safely ignored. Consider the case of a country of intermediate capital abundance. Such a country will find that it trades both with countries that are more capital abundant than itself and also countries where the reverse is true. Importantly, the country should be a net *importer* of capital services from the countries more abundant than itself, and *simultaneously* a net *exporter* of capital services to those countries less abundant in capital. This suggests an important caution on any implicit welfare conclusions based on the magnitude of a country's total net factor trade services. The point is that even if the country is close to a zero net trader in the services of capital and labor when considering its trade with all countries, it could nonetheless be enjoying significant gains from factor service trade by being able to trade with countries both more and less capital abundant than itself. Let $C^+(c)$ denote the set of countries more capital abundant than *c*, and $C^-(c)$ the set of countries less capital abundant than *c*. Let $\mathbf{E}^{cc'}$ be exports from *c* to *c′* and $\mathbf{M}^{cc'}$ be imports to *c* from *c′*. Then, for example, the factor content of trade of *c* with those countries more capital abundant than itself is:

$$\overline{\mathbf{B}}^{c} \sum_{c' \in C^{+}(c)} \mathbf{E}^{cc'} - \sum_{c' \in C^{+}(c)} \overline{\mathbf{B}}^{c'} \mathbf{M}^{cc'} = \sum_{c' \in C^{+}(c)} s^{c'} \mathbf{V}^{c} - s^{c} \sum_{c' \in C^{+}(c)} \mathbf{V}^{c'} \tag{5.24}$$

Naturally a similar condition could be written down for trade with those countries less abundant in capital than country c. Moreover, under the conditions stated, which imply full specialization in tradables, such factor content predictions can be written down bilaterally.

2.6 The Factor Content of Trade with Traded Intermediates

If intermediates are traded and all countries use identical techniques, the standard HOV equations can be implemented because goods will always embody the same amount of factors regardless of where they are produced. The mathematics becomes significantly more complicated in the case where FPE fails and intermediates are traded. The reason is that all exports and imports embody a combination of domestic and foreign factors. Hence, the factor content of any country's trade is going to depend on all of the input–output relationships and technological coefficients in all countries.

Surprisingly, Trefler (1996) claims to have modeled a world with traded intermediates without using any information about the requisite input–output matrices. The starting point of his work is the principle that HOV equations must *always* be of the form $\mathbf{V}^c - s^c\mathbf{V}^W = \mathbf{F}^c$, where the definition of $\mathbf{F}^c$ varies with the model. Note that the fixed point of this approach, $V^c - s^c\mathbf{V}^W$, will only under restrictive circumstances be the predicted factor content of trade. That is, from the start, he abandons the idea of developing counterparts to predicted and measured factor contents of trade.

By using data identities, market clearing, and a stronger than usual assumption on demand – that *bilateral* consumption patterns are proportional to world income shares – Trefler (1996) derives a relation $\mathbf{V}^c - s^c\mathbf{V}^W = \mathbf{F}^c$. Since Trefler *defines* $\mathbf{V}^c - s^c\mathbf{V}^W$ to be the factor content of trade, it tautologically follows that $\mathbf{F}^c$ is likewise the factor content of trade, even though *neither term actually is the net exports of factors embodied in trade*. This is an important point that can be lost to the reader. Since Trefler is only interested in the aggregate restrictions on factor usage rather than tracking factor service flows, he is able to bypass using a complete set of input–output matrices. This enables him to show that $\mathbf{F}^c$ can be decomposed as follows:

$$\mathbf{V}^c - s^c\mathbf{V}^W = \mathbf{F}^c_{Cons} + \mathbf{F}^c_{Inter} + \sum_{c'} \mathbf{B}^{tc'}(\mathbf{D}^{cc'} - s^c\mathbf{D}^{Wc'}) \tag{5.25}$$

where $\mathbf{B}^{tc'}$ is some transform of the standard technology matrix for country c'. The first two terms on the right-hand side might seem to be – but are *not* – the factor content of trade in final goods and intermediates respectively. As noted, while they might look like the factor content of intermediate and final goods trade, Trefler (1996) and close inspection makes clear this is not the case. The final term of $\mathbf{F}^c$ is a term that would equal zero if the strong assumption that *bilateral* consumption patterns are proportional to world income shares were exactly correct. What he has derived is a relationship between endowments, non-standard technology matrices, and trade, but it is hard to see what the economic meaning of these terms may be. This suggests that understanding how to incorporate traded intermediates into factor content studies remains an important area for future research.

2.6.1 Cost restrictions on factor content absent FPE

An alternative approach to testing factor content predictions in the absence of FPE relies on the properties of cost minimization. Define the factor content of imports to c from c' as

$$\mathbf{M}_V^{cc'} \equiv \overline{\mathbf{B}}^{c'II}\mathbf{M}^{cc'} \tag{5.26}$$

Letting the vector of factor prices in c be denoted $\mathbf{w}^c$, Helpman (1984) shows that a restriction on costs implies that:

$$(\mathbf{w}^c - \mathbf{w}^{c'})'\mathbf{M}_V^{cc'} \leq 0 \tag{5.27}$$

In simple terms, this restriction just says that if instead of importing the factor services we actually import, we had instead hired these same factors in our local market c at prices $\mathbf{w}^c$ to produce these goods ourselves, the cost would have been at least weakly greater than our import bill (equal to the cost of producing that import vector in the foreign country). Obviously, a correlative restriction to (5.27) can be placed on the costs of factor service imports to c'. We can also look at the difference between two such equations for c and c' to arrive at:

$$(\mathbf{w}^c - \mathbf{w}^{c'})'(\mathbf{M}_V^{cc'} - \mathbf{M}_V^{c'c}) \geq 0 \tag{5.28}$$

On average, country c imports from c' the services of those factors that are relatively costlier in c than c'.

The restrictions in (5.27) and (5.28) can in principle be taken to data, since they involve observable post-trade factor prices and measurable factor service flows. However, there are two difficulties in implementing these. The first is that, in contrast to the other approaches to HOV derived above, it is crucial to have information about *all* factors of production. A second difficulty is that one must be able to measure with confidence the factor returns in each country, including the rental on capital.

3 Data Issues

Empirical analysis requires the researcher to confront a spectrum of questions. What data is required to test the theory? When alternative measures are available, how do we choose among them? When alternative sources of data exist and they do not provide identically the same values, how do we choose among them? When some analytic elements must be constructed from more primitive data, how do we choose the method for such construction? Is there reason to believe that measurement error in the relevant data is large? Is there a way to minimize the impact of such measurement error on tests of the theory?

3.1 What is the Data?

The data required depend on the variant of the HOV theory to be tested and whether we also want to test some of the subsidiary hypotheses. Tests of HOV always require some measure of a technology matrix. Standard HOV theory requires a matrix of total (direct plus indirect) factor requirements. The next choice is how many technology matrices one wants to work with. Most of the literature has worked with a single technology matrix – typically that of the US, although occasionally also of Japan. When these papers (e.g., BLS, Trefler, 1993, 1995) have contemplated technological differences, these have been treated as parametric deviations from the US technology matrix. An alternative approach is to use distinct technology matrices for the countries in the study, as in Davis and Weinstein (2001a) or Hakura (2001).

Tests of HOV always require some measure of endowments. Standard HOV theory requires endowments of the entire world, although in practice researchers have worked with endowments for the largest set of countries they can obtain. When the tests are in the bilateral difference format of equation (5.13), endowments are only required for the countries considered. In standard tests of HOV, tests are factor by factor, so that omission of some factors from the test does not affect results for included factors. By contrast, tests of cost restrictions when FPE fails require data for all factors. An important question in practice is whether one wants to use coarse or fine definitions of the factors themselves. Alternative implementations of HOV have used two, three, up to as many as twelve factors of production. At times the choice is mandated by data availability or compatibility. At other times, there is a real choice. For example, do we want to characterize labor endowments by occupational category: managerial, production, service, sales, etc.? Or do we want to see it as stocks of high-skilled and low-skilled labor? The former has the advantage of providing a more detailed division of the labor types. However, the latter is probably closer to our idea of factors that can flow across occupational categories. Researchers also must address the issue of the compatibility of the data. Is a skilled worker in the US the same factor as a skilled worker in Cambodia? Are there quality differences? How shall we sum across countries to obtain some measure of world endowments?

The measurement of trade flows is relatively more straightforward. Standard HOV theory requires for each country only its net trade vector with the rest of the world. Some variants of HOV require that exports and imports be measured as gross flows, since they may have different factor content under what is nominally the same industry. Further refinements may require that this be refined to examine bilateral trade flows.

Implementations that examine the production side of HOV typically require some measure of output. Depending on the question at issue, this may require gross or net output.[5] The difference, of course, also requires an input–output matrix. Here there are choices about whether one computes this with absorption of domestic intermediates separated from those that are imported from various sources.

Standard HOV requires some measure of a country's share of world absorption. Often this is taken as the country's GDP share. Sometimes this is adjusted for trade imbalances, although these have typically had scant impact on the HOV results. The demand side of HOV also places restrictions on the pattern of absorption that can be examined more directly. Unfortunately, absorption is often not directly observed, and so inferences can only be made indirectly from factor service flow calculations.

Implementation of standard HOV does not require data on factor prices. However, some tests of HOV (e.g., Trefler, 1993) have relied on factor price data in a subsidiary manner to confirm the plausibility of parameters calculated from the data. When FPE fails, implementation of tests of cost restrictions relies importantly on high quality factor price data for *all* factors.

3.2 Data Quality and Compatibility

Trade is the difference between output and absorption. The factor content of trade is the difference between endowments (the factor content of production) and the factor content of absorption. From this perspective, trade in goods or factor services can be thought of as a residual that is frequently an order of magnitude smaller than output or absorption. This fact brings to the fore the issues of data quality and compatibility.

We have already talked about the data inputs required to test factor content theories. The researcher is then faced with the question of which data source to rely on for measures of the relevant variables. An unfortunate fact is that measures of the same variable for the same country and time period, but drawn from alternative sources, frequently differ – and the differences need not be small! There are many reasons for this, potentially including different definitions, different choices about which exchange rates are used to convert figures, different methodologies for constructing key variables, and so on.

Indeed, at times these problems may be sufficiently large that it might be impossible to observe relationships based on net factor flows even were the relationship to exist. A case in point is the study of BLS. BLS had stressed the importance of the fact that their tests use three independent sources of information on endowments, technology, and trade. To their credit, BLS report in their footnote 14 a check on measurement error in these data. If you pre-multiply the US gross output vector by the US technology vector, this should deliver as an identity the US endowments employed. Since BLS constructed the endowment data separately from the US technology matrix, this identity does not hold in their data, and frequently departs quite sharply from equivalence. For example, the imputed endowment of capital based on the technology matrix and the output levels exceeds the endowment of capital by more than 100 percent. Smaller, but substantial errors exist for other factors. Since one way of interpreting their tests is as a check of whether the entire world uses US technology, it should be more than a little troubling that in their data, even the US does not use US technology.

Such concerns lead us to believe that a great deal of attention must be paid to consistency in the construction of the data. National capital stocks cannot be constructed independently of the way that capital is constructed when the technology matrix is put together. Definitions for other variables likewise must be consistent. This also suggests the value of using one data source as an ultimate authority for a given project. When there are discrepancies between alternate sources, resolve them based on this authority and do re-scaling of supplementary data as needed. Naturally, this requires care in the selection of the highest quality database as the authority. But it at least allows the theory some chance to escape being swamped simply by inconsistent definitions of the same variables.

4 Factor Content: What are the Tests?

In the absence of a clear alternative framework relating endowments and trade, it is not possible to test HOV against a well-specified null hypothesis. As a result, researchers typically run horse races between various versions of the model, e.g., HOV with neutral technical shifts, home bias, etc. The statistical framework is typically Bayesian and simply asks which version of HOV is best supported by the data.

A major problem with such statistical tests of HOV is that one of the models must be deemed "best" even if it has little explanatory power. As a result, researchers also rely on goodness of fit criteria as a means of "testing" HOV. Typically researchers have focused on five such criteria. The first two measures are non-parametric. *Sign tests* ask whether the sign of MFCT is the same as that of PFCT. These tests identify what share of the data would lie in quadrants one and three if one plotted MFCT against PFCT. A strength of this test is that large outliers are unlikely to affect the results. The major weakness, of course, is that countries with small PFCT may have many sign errors without it indicating a major problem for the theory. Rank tests put a little more structure on the data by asking whether countries that are predicted to be large exporters (importers) of a factor are measured to do so. A problem in these tests arises when there are a large number of countries that have similar PFCTs.

A second major class of tests is the regression tests. Here three tests are standard. The first two arise from the slope and R^2 of a regression of MFCT on PFCT. In addition to the slope and R^2 tests, Trefler has utilized the missing (factor service) trade test (MT) which is defined as the variance of the MFCT, σ_M^2, divided by the variance of PFCT, σ_P^2.

How these tests are related is best understood by thinking about how each statistic is calculated. The formulas for the slope coefficient and R^2 in a linear regression are:

$$\hat{\alpha} = \frac{\sigma_{MP}}{\sigma_P^2} \quad \text{and} \quad R^2 = \frac{\sigma_{MP}^2}{\sigma_P^2 \sigma_M^2}$$

where $\hat{\alpha}$ is the slope coefficient of a regression of MFCT on PFCT. A little algebra shows that the three tests are related:

$$MT = \frac{\hat{\alpha}^2}{R^2}$$

In other words, any two tests are sufficient for identifying the outcome of the third.

4.1 First Generation Studies: A String of Empirical Rejections

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

The first real tests of HOV were conducted by Maskus (1985) and BLS (1987). The analytic foundation is given by equation (5.11). Their results severely undermined confidence in the robustness of the Heckscher–Ohlin framework. Maskus (1985) carried out both sign and rank tests on data for two time periods (1958 and 1972), and for three high quality factors (professional, unskilled labor, and capital). He reports results only for the US, perhaps because the Leontief paradox had focused on it. The sign test is correct for only one factor in 1958, but for all three in 1972. This might be seen to suggest that the Heckscher–Ohlin–Vanek relations fare well, at least in the latter period. However, the test lacks power. As Maskus notes, if we consider the alternative that the signs were determined randomly, we will have two or fewer sign failures out of six tries 34.4 percent of the time. Moreover, even if we limit ourselves to the 1972 data, under the same alternative, there will be no sign violations (as in his data) one in eight times. The results were, if anything, worse in the rank test. The direct measures of US factor abundance relative to the rest of the world were stable, with physical capital most abundant, professionals second, and unskilled labor least abundant. However, the trade-imputed measures of factor abundance in 1958 suggested the US was most abundant in unskilled labor, and least abundant in physical capital! The 1972 trade-imputed measure of factor abundance showed unskilled labor shifting dramatically to be least abundant, and reverses the relative abundance of physical capital and professionals. A repeat of the tests, restricted to OECD data, yielded no improvement. As Maskus noted, "paradoxical outcomes may be the rule rather than the exception."

BLS likewise report results widely viewed as undercutting Heckscher–Ohlin–Vanek. An important contribution was extending the test to a much broader set of countries (27) and factors (12). Thus, whereas the Maskus test was based on a matrix of only three cells for each time period, the BLS matrix had 324 entries. Because of the greater dimensionality of the matrix, it became possible to conduct sign and rank tests not only for a single country across factors (as in Maskus (1985)), but also for a single factor across countries. The sign test was correct more than half of the time for 11 of the 12 factors, but was correct over 70 percent of the time for only 4 in 12. The sign matches were correct more than half the time for 18 of 27 countries, but over 70 percent of the time for only 8 of the 27. Only 61 percent of the total sign matches were correct. They note that independence between the signs of corresponding entries can be rejected at the 95 percent level for only one factor in twelve, and for only four of the 27 countries. In effect, in determining which factors' services would on net be exported or imported, Heckscher–Ohlin did little better than a coin-flip.

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

It is hard to overstate the impact that the Leontief and BLS studies had on the profession. Krugman and Obstfeld (1994, p. 78) summarized the thinking at the time in their textbook writing, "trade just does not run the direction that the Heckscher–Ohlin theory predicts." The problem was that there was no alternative. Ricardian and scale economies models were useful at explaining many problems, but it was hard to imagine that educational levels in the US had no impact on the US industrial mix. In large measure because we had to have some theory about these linkages, empirical researchers continued to search for ways of reconciling the theory with the data. In order to do this, they adopted two main approaches. The first was to see if simple amendments to the theory would yield new insights, and the second was to test the theory with better data.

The pessimism regarding Heckscher–Ohlin was partly relieved by Trefler (1993), only to be revived by Trefler (1995). One of his key insights was that it was not enough to simply say HOV fails without understanding why it fails. Trefler's two papers represent alternative approaches to resolving the problems identified by BLS. The former follows up on Leontief's suggestion that the failure of Heckscher–Ohlin may be due to factor-based differences in efficiency. Trefler chooses the efficiency factors so that the Heckscher–Ohlin–Vanek equations fit exactly. He then shows that the implied productivity differentials correlate nicely with evidence on cross-country differences in wages and rentals, suggesting a version of adjusted factor price equalization. Trefler (1995) returned to the simple Heckscher–Ohlin–Vanek framework. We had learned from Staiger (1988) that there were systematic ways in which factor content predictions missed the mark; Trefler went on to show us what those systematic problems were. In one exercise, he graphed the net factor trade residuals, $\boldsymbol{\varepsilon_T} = \mathbf{B}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{T} - (\mathbf{V} - \mathbf{sV}^W)$, against the predicted net factor trade, $\mathbf{V} - \mathbf{sV}^W$. Theory would predict that these should be centered around the line $\boldsymbol{\varepsilon_T} = 0$. Instead they closely followed the line $\boldsymbol{\varepsilon_T} = -(\mathbf{V} - \mathbf{sV}^W)$, or equivalently, $\mathbf{MFCT} = \mathbf{B}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{T} = 0$. This says that measured net factor trade is approximately zero, to which he applied the colorful moniker "the case of the missing trade."

An important insight into the work of Trefler was identified by Gabaix (1997). Gabaix tried to understand why the results looked so good in Trefler (1993) and so bad in Trefler (1995). What Gabaix realized was that these two sets of results were linked. In the first paper, Trefler calculated productivity parameters, $\boldsymbol{\pi}^{fc}$, that solved the following problem:

$$F^{fc} = \pi^{fc} V^{fc} - s^c \sum_c \pi^{fc} V^{fc}$$

The second paper had demonstrated that the LHS of this equation was very close to zero. If we set it as exactly zero, then

$$0 = \pi^{fc} V^{fc} - s^c \sum_c \pi^{fc} V^{fc}$$

After a little algebra and remembering that s_c is the share of country c's GDP in the world, it is possible to show that

$$\frac{\pi^{fc}}{\pi^{fc'}} = \frac{Y^c / V^{fc}}{Y^{c'} / V^{fc'}} \left[\left(\frac{\tilde{V}^W - \tilde{V}^c}{\tilde{V}^W - \tilde{V}^{c'}} \right) \left\{ \frac{(1 - s^{c'})}{(1 - s^c)} \right\} \right]$$

where Y^c is country c's GDP. If both c and c' are sufficiently small relative to the world, then the term in brackets on the right converges to unity. In that case, the calculated relative productivities are:

$$\frac{\pi^{fc}}{\pi^{fc'}} = \frac{Y^c / V^{fc}}{Y^{c'} / V^{fc'}}$$

In short, the productivity parameters would simply be GDP per factor. Hence, as long as wages are correlated with GDP per capita, it will not be surprising that the measured productivity parameters would be correlated with wages.

To drive home the point that missing trade was responsible for Trefler's 1993 results, Gabaix did an experiment in which he began with the hypothesis that the measured factor content of trade is minus the HOV prediction, i.e. –MFCT = PFCT.

Using this equation, he shows that the calibrated π^{fc} differ little from those of Trefler and relative π^{fc} correlate with relative wages nearly as well as in Trefler (1993). What this makes clear is that evidence that the calibrated relative π^{fc} correlate well with relative wages could not be used as evidence in favor of HOV.

These were not problems that Trefler could have foreseen when he wrote the original paper. However, it underscored Trefler's contention that understanding the mystery of the missing trade would be critical to understanding what was wrong with HOV. Indeed, once you understand that the MFCT is essentially zero, much of the HOV econometrics becomes quite simple.

Consider, for example, Trefler's preferred specification involving an Armington home bias. His estimating equation is:

$$F^{fc} = V^{fc} - s^{c}\left[(1-\alpha_c^*)\frac{Y^W}{Y^c}V^{fc} + \alpha_c^* V^{fW}\right] + \mu^{fc}$$

If we make the assumption that trade balances are a small share of GNP and hence $s^c \approx Y^c/Y^W$, then this equation collapses to:

$$F^{fc} = \alpha_c^*(V^{fc} - s^c V^{fW}) + \mu^{fc}$$

or

$$\text{MFCT}^{fc} = \alpha_c^* \,\text{PFCT}^{fc} + \mu^{fc}$$

We already know from the first part of Trefler's paper that the LHS of this equation is close to zero so we may not be surprised to find evidence that α_c^* is much smaller than unity.

This may be evidence in favor of an Armington home bias, but it could be the result of any other process that results in little measured net factor trade. For example Conway (2001) argues that MFCT will be small if there is little factor mobility since trade will not move factors away from their autarky allocations. Using Trefler's data he estimates

$$F^{fc} = \frac{1}{1+\gamma^f}[V^{fc} - (\beta+1)s^c V^{fW}] + \varepsilon^{fc}$$

where γ^f and β are parameters to be estimated. He finds γ^f to be significantly greater than zero while β is indistinguishable from zero. In other words, for the case where β is zero, we can rewrite the specification as:

$$\text{MFCT}^{fc} = \alpha^f \text{PFCT}^{fc} + \mu^{fc}$$

Fundamentally, the difference between the two papers is in how they shrink PFCT to match MFCT – Trefler does it by country, and Conway by factor. In both cases $\hat{\alpha}$ will be significantly less than unity as long as MFCT is small.

The relationship between the various tests that we derived earlier gives us an insight into why these specifications succeed at eliminating the mystery of the missing trade. Recall that these authors declare victory over the missing trade when $\sigma_M^2 \sigma_{P'}^2 \geq 1$ where P' can now be defined as $\hat{\alpha}$ PFCT. Recalling our earlier discussion of regression tests, we know that this condition can be rewritten as $\sigma_M^2 \, (\hat{\alpha}^2 \sigma_P^2) \geq 1$ or MT/$\hat{\alpha}^2$, which just equals $1/R^2$. This implies that the missing trade statistic is bounded below by 1 and above by infinity in this type of specification. Hence any specification that can be written as $\text{MFCT}^{fc} = \alpha \text{PFCT}^{fc} + \mu^{fc}$ is guaranteed to deliver a missing trade statistic above one and so appear to solve the mystery of the missing trade. This "solution" is illusory and provides no information about the economics underlying missing trade. Oddly enough, as the fit deteriorates, missing trade will shift toward excess trade. This may help explain why in the preferred specification of Trefler (1995), the missing trade statistic is much larger than unity even though Gabaix finds that PFCT has almost no explanatory power.[6]

Is this all that is going on? First, we have already noted that such specifications are mathematically guaranteed to "solve" the mystery of the missing trade. The only remaining question is what they tell us about factor service flows. Gabaix (1997) noticed that tests of equation (5.6) can fail miserably even for preferred econometric specifications. In neither Trefler (1995) nor Conway (2001) do the authors take the estimated parameters and go back to the original puzzle to see if they successfully reconcile MFCT and PFCT. When Gabaix does this using Trefler's data, he finds that the amended model does little to reconcile predicted and measured factor trade.[7]

4.2 Putting the Pieces Together

As researchers puzzled over why HOV performed so badly, they began to ask which parts of the theory were causing the problems. As we have already noted, trade theory necessarily contains a theory of production and a theory of absorption. Davis, Weinstein, Bradford, and Shimpo (1997, hereafter DWBS) were the first to recognize that this naturally suggests that tests of HOV can be broken up into tests of production and absorption models. This enables one to test the theories directly on the relevant data rather than trying to infer parameters about demand and production from the factor content of trade.

Aware of many of the problems that had plagued testing of HOV on international data, DWBS developed a new approach to testing HOV. Several elements of that approach are worth noting. First, it examines the production and absorption sides separately. Prior HOV tests working with *trade* data could make inferences about the source of difficulties, but could not examine them directly. Second, we sought to bridge our own and prior work by starting with a strict HOV model and relaxing assumptions one at a time. This allowed us to identify which assumptions seemed to be crucial in driving the results. Third, we developed an approach that allowed us to make HOV predictions when only a subset of the world shared FPE. This draws on the analytics embodied in equation (5.21). In our case the relevant "FPE club" was a set of ten regions of Japan. Finally, we worked a data set in which identities held. The results, in contrast to prior work, were very positive for HOV.

The step-by-step approach in DWBS allowed us to see which elements of the theory were causing problems. We first considered 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. This indicated that although the theory was a powerful means of talking about production within an FPE club, it performed poorly as a description of international production patterns.

We then turned to the Heckscher–Ohlin theory of the pattern of consumption. We examined this first by considering Japanese regional absorption, which the theory suggests should be proportional to world net output. The Heckscher–Ohlin model of proportional absorption does surprisingly well under this assumption. Indeed while Trefler (1995) was forced to estimate home bias parameters from factor content data, we could examine the question on the actual consumption data. What we found was that the assumption that Japanese consumption differed from that in the rest of the world did no better than the standard prediction of homotheticity. 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 assembled 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. 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.

In sum, DWBS found 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.

DWBS was clearly only a stepping-stone in the understanding of how to implement HOV. While world trade and endowments were critical elements of the tests, there was a serious question why the international production model fared so poorly. Without answering that question, it would be impossible to understand how HOV worked internationally.

A paper closely related to the work of DWBS is that of Hakura (2001). Where DWBS focused on asking, if you assume that all countries use a common technology matrix, how badly does the model perform, Hakura approached the question from the opposite perspective. Taking direct measures of technology matrices for four OECD countries, she asked how much improvement we might attain if we got the technology side of the model to fit perfectly. Her answer – quite a lot! One can note some drawbacks of this approach. Because the technology matrices she works with fit as a matter of construction, her empirical exercises cannot "test" any of the hypotheses underlying the model of production. Rather, it must ask, when the production model fits as an identity, are the assumptions about international demand patterns sufficiently incorrect as to throw off the basic HOV predictions (in the bilateral difference form)? The answer is no. This left open the question of why technology matrices differ and in particular whether these differences can be systematically related to fundamental characteristics of the countries' trading system.

The starting point for Davis and Weinstein (2001a) was the realization that the existing literature had one major drawback. The hypothesized amendments concern technology and absorption. Yet, with the exception of Hakura, the empirical tests draw on only a single direct observation on technology (typically that of the US) and no observations whatsoever concerning absorption. Moreover, they aimed to understand whether these differences could be related to systematic differences among the included countries on the basis of theory.

The 1995 publication of the OECD's input–output database dramatically improved our ability to test trade theory. Prior to that, researchers had no access to large numbers of compatible IO tables. Its publication enabled us to construct technology matrices for ten rich OECD countries as well as for a composite rest of the world (ROW). The data cover both manufacturing and non-manufacturing with two factors of production, capital and labor.

An examination of the technology matrices allows testing of the nature of differences in techniques across countries. These differences in techniques correspond to a variety of economic hypotheses that can be related to observed characteristics of the countries. These allow us to make inferences not only about whether efficiency differences exist across countries but whether these efficiency differences are sufficient to capture the cross-country differences or whether one needs to take specific account of the failure of the world to replicate an integrated equilibrium. Using parameter estimates obtained from analyzing the technology matrices, one can then take the fitted technology matrices and apply them to the trade data to see which, if any, of the hypotheses may help to resolve the mystery of the missing trade.

Having gone this far purely from examining the technology matrices, one can take the further step of asking how much additional gain would come from a model that more accurately predicts the volume of trade than the frictionless model traditionally used. That is, how much of the missing net factor trade is due to the low volume of product trade? Here we estimate a gravity model and use the fitted values, in addition to our preferred model of production, to predict the factor content of trade.

Our estimation strongly rejects the traditional assumption of identical technologies, even for the ten rich OECD countries. Allowing for Hicks-neutral productivity differences greatly improves the fit of the production model, but surprisingly does very little to eliminate the mystery of the missing trade. A hypothesis that industry input usage is correlated with country factor abundance, which would not hold in conventional HOV models, is strongly confirmed in the data. If this held only in tradable sectors, then it would be possible that this correlation reflects only aggregation. But it holds about as strongly in non-tradable sectors as well, which indicates a breakdown in FPE and hence a departure from the integrated equilibrium. Once this departure from FPE is recognized, it is crucial to re-examine the treatment of non-traded goods within the predictions. In the conventional model in which all countries use the same techniques of production and preferences are identical and homothetic, the factor content of trade is invariant to the presence of non-traded goods. However this is not true when FPE breaks down. In this case, capital abundant countries use more capital per worker in non-traded sectors, which leaves the residual available for production of tradables diminished and so lowers the predicted factor content of trade. Allowing for the fact, very evident in the production data, that industry input usage is correlated with country capital abundance dramatically improves the performance of the model. The major previous research efforts had left measured factor trade as a minuscule proportion of predicted factor trade. In this last exercise, predicted factor trade is approximately 60 percent of predicted net factor trade. If one goes further to incorporate the fact that the volume of trade is smaller than predicted by the frictionless model, then measured factor trade rises to roughly 80 percent of that predicted.

In short, a few simple modifications provide a dramatically improved ability of the model to match the data. These modifications include cross-country Hicks-neutral efficiency differences; a breakdown of FPE with the consequence that industry input usage is correlated with country factor abundance; a recognition that the breakdown of FPE has important consequences for factor usage in non-tradables; and the fact that trade volumes are smaller than predicted by the frictionless model. Suitably modified, HOV works well.

4.3 An Integrated World or Not?

The preferred specification of Davis and Weinstein (2001a) is a multi-cone Heckscher–Ohlin model, so one in which the world fails to operate as an integrated equilibrium. In Davis and Weinstein (2001b), we show that thinking about the world in this explicitly non-integrated framework provides new insights about the nature of world trade. It yields strong restrictions that are counter-intuitive from the standpoint of an analysis based on the assumption of FPE, but that are nonetheless strongly endorsed by the data. *Inter alia*, the theoretical and empirical analysis allow us to gain a deeper understanding of the "mystery of the missing trade," the nature of intra-industry trade, and the role that net factor trade plays in the world at large, as well as specifically within the OECD.

The analytic model is based on Helpman (1984), and features many goods, factors, countries, and production cones. The principal results we derive are as follows. Using countries' actual technology matrices, we show that true net factor trade is much larger than that reported by previous studies. Moreover, the net factor trade looms quite large when scaled by resources employed in tradable sectors. In contrast to results from integrated equilibrium theory, our model predicts that the typical country will be a net exporter of the services of a factor to the set of countries less abundant in that factor, and a net importer of services of the *same* factor from the remaining countries. This prediction is strongly confirmed in the data.

We are able to decompose the true factor content of trade into the conventional measure plus three sources of error. We show that the traditional measure of the factor content of trade is much smaller than, and essentially uncorrelated with, the true measure of net factor trade. This is an important reason for the "mystery of the missing trade." It is important to realize, however, that this is a misnomer. What Trefler identified was missing *factor service* trade not necessarily missing trade *per se*. In other words, measured factor service trade could be small without trade volumes being small. One of the sources of error concerns intra-industry trade. While it has become a convention in the integrated equilibrium analytic literature to define intra-industry trade as the exchange of goods of similar factor content, this is not what the data reveal. We verify directly in the data that intra-industry trade between countries consists of the exchange of goods that differ systematically in their factor content, and that these differences reflect endowment differences. Indeed, the data show that for the typical country, approximately 40 percent of total *net* factor trade is accomplished via *intra*-industry exchange of commodities. For several rich countries, including the US, over two-thirds of the net factor trade is accomplished via intra-industry trade. Finally, our results demonstrate that trade *among* the rich countries of the OECD composes an important share of net factor trade for many of these countries.

5 Further Work

5.1 Incorporating Intermediates

Trefler and Zhu (2000) have criticized Davis and Weinstein (2001a) for the treatment of traded intermediates. Their paper builds on Trefler (1996) and seeks to implement the relationship described in equation (5.25). Trefler and Zhu (2000) use data on four countries (the US, Belgium, France, Germany and the Netherlands) and proceed in three stages. First, they demonstrate that the variance of $\mathbf{V}^c - s^c\mathbf{V}^W$ is large relative to the variance of the sum of the first two terms of $\mathbf{F}^c$ (i.e., $\mathbf{F}^c_{Cons}$ + $\mathbf{F}^c_{Inter}$). Second, they show that assuming no trade in intermediates can, under some circumstances, improve the results. Finally, they show that even if they use the right level of intermediates trade and get the consumption side of the model right, the variance of $\mathbf{V}^c - s^c\mathbf{V}^W$ is substantially larger than the variance of $\mathbf{F}^c$.

It is very hard to know what to conclude from the exercise. The theory is based on Trefler (1996), and as we noted in the theoretical section, equation (5.25) has no economic implications beyond being an equation that *must* hold if some data identities, market clearing conditions, and a stronger than usual assumptions on demand also hold. Consider what this implies for the first test. When Trefler and Zhu (2000) constrain the term, $\sum_{c'} \mathbf{B}^{tc'}(\mathbf{D}^{cc'} - s^c\mathbf{D}^{Wc'})$, to be zero, it is not clear what it means to have one side of equation (5.25) with larger variance than the other since neither side has economic content. In particular, since the expression $\mathbf{F}^c_{Cons} + \mathbf{F}^c_{Inter}$ is *not* the net exports of factor services embodied in final and intermediates goods trade, what does it mean to say that the variance of $\mathbf{V}^c - s^c\mathbf{V}^W$ is bigger than that of $\mathbf{F}^c_{Cons} + \mathbf{F}^c_{Inter}$? This looks similar to a missing trade test *notationally*, but the meaning is completely different.

Things become more confusing when they try to incorporate traded intermediates. Their theory requires them to have information on the volume of final and intermediate goods exports and imports by country, but they lack the data. Hence their critique of Davis and Weinstein (2001a) is not based on *actual* flows of final and intermediate goods, but on what would happen if these flows were at some hypothetical level. So how do they calculate the level of final goods trade? They assume that the share of final goods in total trade equals the share of final demand divided by the sum of gross output and total imports. This makes little sense. Suppose the world is perfectly specialized and only final goods are produced. In this case, a country with a value of output (so total spending) equal to one dollar will consume one dollar's worth of final goods drawn from a variety of countries. The ratio of final absorption to the sum of output and imports will be less than unity even though there are no intermediates! By using this ratio as their indicator of the final goods share of trade, Trefler and Zhu systematically overstate the importance of traded intermediates. The reason why this matters is that by using the wrong levels of intermediates, equation (5.25) can fail even if the model is correct.

Finally, the last test of Trefler and Zhu is particularly puzzling. We have already argued that equation (5.25) must hold as an identity if some basic relationships are true. What they find is that $\mathbf{V}^c - s^c\mathbf{V}^W$ does not equal $\mathbf{F}^c$. How can this be? The relationship will only be violated if a full employment or market clearing condition fails or if their estimates of the level of final and intermediate goods trade are inaccurate. All of these constitute implementation or definitional problems and not real tests of the theory.

Hence the Trefler–Zhu exercise examines the relative variances of terms that have no real economic content and are equal only under highly restrictive circumstances. We don't think that much can be learned from the exercise. This notwithstanding, we do believe that the broader point of Trefler and Zhu (2000), that research ultimately needs to give a more complete account of the role of intermediates in factor service trade is correct. But we believe it needs to be done according to a theory designed to track measured and predicted factor contents as developed earlier.

5.2 Other Extensions

Feenstra and Hanson (2000) test the implication of Davis and Weinstein (2001a) that factor content of exports differs systematically from domestic production. They find strong evidence that it does. In particular, they find that the factor content of US trade rises in skill intensity as they use increasingly disaggregated data. Feenstra and Hanson go on to argue that this bias may help explain why Davis and Weinstein did not fully eliminate the missing trade phenomenon. While much work still needs to be done before this type of aggregation bias can be implemented into a full HOV model, we think that the results are important and encouraging.

Another area that cries out for more research is the demand side of the model. Clearly there is a lot less trade than one would expect in a frictionless but specialized world. This causes us to overestimate factor service flows. Understanding what is driving this puzzle seems to be a very important question for understanding HOV. Unfortunately, our models of absorption and trade in a world with frictions are still not well developed; however we feel that this is also an area that may be important to explore.

6 Conclusions

Study of the factor content of trade has become a laboratory to test our ideas about how the key elements of endowments, production, absorption, and trade fit together within a general equilibrium framework. Already a great deal of progress has been made in fitting these pieces together. Nevertheless, the existing research raises a great many questions that should help to focus empirical research in the coming years. Among the more pressing issues is a deeper consideration of the role of intermediates, the role of aggregation biases, and of differences in patterns of absorption. This work should provide a more substantial foundation for future policy work developed within a factor content framework.

Acknowledgment

The National Science Foundation provided support for the research reported in this chapter.

Notes

1 See Helpman and Krugman (1985).
2 For the conditions under which FPE will hold in the presence of non-traded goods, consult Helpman and Krugman (1985).
3 One of many possible examples is if the included countries share a demand structure that is identical to each other but systematically different from that of the rest of the world.
4 Note that this is very much at odds with the standard relation of Rybczynski, where capital abundance of a country affects output composition but not capital intensity by industry.
5 One should note that net output is conceptually quite distinct from value added. Net output of a good in a country equals its total output of that good less the output of that good used up as intermediates in that country. Note that it is possible for net output to be negative, for example when a country does not produce a good but does use imports as intermediates. Value added in an industry is the value of all output in an industry less the value of *all* intermediate inputs used to produce that output. Note that in a long-run equilibrium value added cannot be negative.

6 Similarly in Conway (2001), one derives estimates of the missing trade statistic of 50 or more in the preferred specifications.

7 It is also worth noting the role played by neutral technical differences. In Trefler (1995) the technical shift terms enter the analysis two ways: first as estimated parameters and second as data. When the technology parameters are estimated, the Schwarz criterion rejects them, but when the parameters are assumed to be proportional to per capita income, the Schwarz criterion accepts them. This result arises from the fact that the estimated parameters are similar to per capita income, and while the Schwarz criterion contains penalties for additional parameters it has no penalty for data transformations. Technically speaking, the results for the technology and consumption model cannot be compared with the other results of the paper because the underlying data are different.

References

Bowen, Harry P., Edward E. Leamer, and Leo Sveikauskas 1987: Multicountry, multifactor tests of the factor abundance theory. *American Economic Review*, 77, 791–809.

Brecher, Richard A. and Ehsan U. Choudhri 1982: The Leontief paradox, continued. *Journal of Political Economy*, 90, 820–3.

Conway, Patrick J. 2001: The mystery of the missing trade: Comment. *American Economic Review*, 92, 1, 394–404.

Davis, Donald R. and David E. Weinstein 1998: An account of global factor trade. NBER working paper no. 6785.

Davis, Donald R. and David E. Weinstein 2000: International trade as an "integrated equilibrium": New perspectives. *American Economic Review*, 90, 150–4.

Davis, Donald R. and David E. Weinstein 2001a: An account of global factor trade. *American Economic Review*, 91, 1423–54.

Davis, Donald R. and David E. Weinstein 2001b: Do factor endowments matter for North–North trade? NBER working paper no. 8516.

Davis, Donald R., David E. Weinstein, Scott Bradford, and Kazushige Shimpo 1997: Using international and Japanese regional data to determine when the factor abundance theory of trade works. *American Economic Review*, 87, 421–46.

Deardorff, Alan V. 1984: Testing trade theories and predicting trade flows. *Handbook of International Economics*. Amsterdam: New Holland, 467–517.

Dornbusch, Rudiger, Stanley Fischer, and Paul A. Samuelson 1977: Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods. *American Economic Review*, 67, 823–39.

Feenstra, Robert C. and Gordon H. Hanson 2000: Aggregation bias in the factor content of trade: Evidence from U.S. manufacturing. *American Economic Review*, 90, 155–60.

Gabaix, Xavier 1997: The factor content of trade: A rejection of the Heckscher–Ohlin–Leontief hypothesis. Harvard University, mimeo.

Grimes, Donald R. and Penelope B. Prime 1993: A regional multifactor test of the Heckscher–Ohlin–Vanek theorem. University of Michigan and Kenesaw State College, mimeo.

Hakura, Dalia S. 2001: Why does HOV fail? The role of technological differences within the EC. *Journal of International Economics*, 54, 361–82.

Helpman, Elhanan 1984: The factor content of foreign trade. *Economic Journal*, 94, 84–94.

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

Horiba, Yutaka 1997: On the empirical content of the factor-contents theory of trade: A regional test. Osaka Economic Papers no. 47, 1–11.

Krugman, Paul R. and Maurice Obstfeld 1994: *International Economics: Theory and Policy*. Boston: Addison Wesley.

Leamer, Edward E. 1980: The Leontief paradox, reconsidered. *Journal of Political Economy*, 88, 495–503.

Leamer, Edward E. and James Levinsohn 1995: International trade theory: The evidence. *Handbook of International Economics*, 3, 1339–94.

Leontief, Wassily 1953: Domestic production and foreign trade: The American capital position re-examined. *Proceedings of the American Philosophical Society*, 97, 332–49.

Maskus, Keith E. 1985: A test of the Heckscher–Ohlin–Vanek theorem: The Leontief commonplace. *Journal of International Economics*, 19, 201–12.

Staiger, Robert W. 1988: A specification test of the Heckscher–Ohlin theory. *Journal of International Economics*, 25, 129–41.

Staiger, Robert W., Alan V. Deardorff, and Robert M. Stern 1987: An evaluation of factor endowments and protection as determinants of Japanese and American foreign trade. *Canadian Journal of Economics*, 20, 449–63.

Stern, Robert M. and Keith E. Maskus 1981: Determinants of the structure of U.S. foreign trade, 1958–76. *Journal of International Economics*, 11, 207–24.

Trefler, Daniel 1993: International factor price differences: Leontief was right! *Journal of Political Economy*, 101, 961–87.

Trefler, Daniel 1995: The case of the missing trade and other mysteries. *American Economic Review*, 85, 1029–46.

Trefler, Daniel 1996: The structure of factor content predictions. University of Toronto, mimeo.

Trefler, Daniel and Susan Chun Zhu 2000: Beyond the algebra of explanation: HOV for the technology age. *American Economic Review*, 90, 145–9.

Vanek, Jaroslav 1968: The factor proportions theory: The n-factor case. *Kyklos*, 21, 749–56.