National and International Returns to Scale in the Modern Theory of International Trade

Wilfred J. Ethier


Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28198206%2972%3A3%3C389%3ANAIRTS%3E2.0.CO%3B2-S


Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/aea.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
National and International Returns to Scale in the Modern Theory of International Trade

By Wilfred J. Ethier*

For over a quarter century, the Heckscher-Ohlin-Samuelson (H-O-S) trade model has thoroughly dominated work in the pure theory of international trade. Indeed this model is often identified as “the” modern theory of trade. But this dominance has always been rendered uneasy by a widespread suspicion that the salient facts of modern commerce are inconsistent with the theoretical structure. Two broad areas of suspicion may be identified.

The first, concerned with whether actual trade patterns and factor endowments are related as predicted by the theory, consists of the Leontief Paradox and the huge volume of work it stimulated. My interpretation of this literature is that the factor-endowments theory of trade fares reasonably well, but that its two-factor, two-commodity version is essentially an inadequate description of reality. I have on other occasions investigated the consequences of many goods and factors, and do not wish to do so now. This paper will accordingly confine itself to a factor-endowments model with two primary factors and two final goods.

The second (by no means unrelated) area of suspicion centers on the stylized fact that the largest and fastest growing component of world trade since World War II has been the exchange of manufactures between the industrialized economies. By contrast, the H-O-S model, and neoclassical theory generally, sees little basis for trade between similar economies. Two manifestations of this point may be found in the empirical literature. The first, of relevance to international monetary problems, concerns departures from purchasing-power parity among the industrial countries even for fairly disaggregated indices of traded goods (see Irving Kravis and Richard Lipsey). The second, a central concern of this paper, involves intraindustry trade.

Although two-way trade had long been observed, it first became of major concern when economists investigated the consequences of economic integration in Europe (see, for example, P. J. Verdoorn; Bela Balassa, 1966) and noted a tendency, not towards increased specialization, but rather for all countries to simultaneously increase exports of most categories of manufactures. Indeed in some cases specialization actually declined. Subsequent work divorced the phenomenon of intraindustry trade from economic integration, and exposed the pervasive expansion of such trade between all industrial countries and across most manufacturing sectors, irrespective of tariff barriers or their changes (Helmut Hesse; Herbert Grubel and Peter Lloyd, 1975; Emilio Pagoulatos and Robert Sorenson; and Richard Caves), furthermore this expansion does not appear to be a matter of transition between equilibria1 (Caves).

The most natural explanations of intraindustry trade, advanced by Gottfried Haberler (p. 34) long ago, are product heterogeneity within aggregates and border trade (and its seasonal analog). But accumulated empirical work (for example, Hesse; Grubel and Lloyd, 1975; and Caves) strongly suggests that these explanations are inadequate.2 Attention has

---

*Professor of economics, University of Pennsylvania. The research for this paper was supported by the National Science Foundation under grant no. SES-7925614. I have benefitted from seminars at the universities of Pennsylvania, Western Ontario, and Wisconsin, and Northwestern University. Richard Caves also supplied useful comments and suggestions.

1See Gary Hufbauer and John Chilas for a contrasting argument that intra-industry trade is due to the GATT method of tariff reduction and to transitory responses with quasi-fixed factors of production. Joseph M. Fnger also offers some contrasting evidence.

2Grubel and Lloyd (1975), for example, found significant Australian intraindustry trade even at the seven-digit level. I do not mean to suggest that such trade indicates a violation of commodity arbitrage and
accordingly shifted to product differentiation and economies of scale (in theory the two go together, as scale economies supply the limitation to the degree of differentiation).\textsuperscript{3}

International trade theory does contain a sizable literature on increasing returns to scale (see, for example, R. C. O. Matthews; James Meade; and Murray Kemp, 1969), but this literature would not appear to offer much that is relevant to the phenomenon at issue. Indeed it has not on balance had great influence on trade theory generally. This is no doubt largely because increasing returns establish a strong presumption for complete specialization, multiple equilibria, and indeterminacy, conclusions that both vitiate the standard comparative statics methodology of trade theory and that also bear questionable relevance to the facts of contemporary trade. Furthermore, this theory has generally assumed that the economies of scale were external to the firm and appropriate at an aggregate level, assumptions that preclude much relevance to intratrade. Empirical work has likewise encountered difficulty in coming to grips with scale economies, although many researchers have argued their importance (for example, Balassa, 1961, 1967; Donald J. Daly et al., Nicholas Owen, and Caves), and some have found relevance to intratrade.

Recently I pointed out (1979a) that scale economies resulting from an increased division of labor rather than from, say, an increased plant size, depended at an aggregate level upon the size of the world market rather than upon geographical concentration of the industry, as supposed in the traditional theory. Such "international" returns to scale were shown to be free of the presumption of indeterminancy and multiple equilibria\textsuperscript{4} characteristic of "national" returns to scale, to imply a theory of intratrade in intermediate goods, and generally to lead to conclusions much more in accord with the stylized facts of modern trade.

The purpose of this paper is to argue that, in the context of international returns, the factor-endowments theory is consistent with, and indeed helps to explain, the stylized facts discussed above. In order to do this I shall construct a simple but detailed model yielding international returns to scale, and employ this model to systematically explore the relations between such international returns, the traditional national returns to scale, and the modern (or factor endowments) theory of international trade. I have four broad messages. 1) International returns depend in an essential way on an interaction between the two types of scale economies, which is also an interaction of internal and external (to the firm) economies. 2) The basic theorems of the factor-endowments theory are essentially robust in the presence of such scale economies, in contrast to the conclusions of the traditional literature.\textsuperscript{5} Nevertheless there are some important modifications. 3) Intratrade, like interindustry trade, has a factor-endowments basis. However, such trade is basically \textit{complementary} to international factor mobility. 4) Although the existence of internal-scale economies and product differentiation are essential to the theory, the degree of such phenomena need not be an essential determinant of the degree of intratrade.

Recently Paul Krugman (1979, 1981); Avinash Dixit and Viktor Norman; Elhanan Helpman; and Colin Lawrence and Pablo Spiller have extended to an international context the Avinash Dixit-Joseph Stiglitz formalization of Chamberlinian monopolistic

\textsuperscript{4}International scale economies actually "bury" these problems rather than eliminate them, as will become clear below.

\textsuperscript{5}In this respect the present paper complements recent work establishing the general validity of the basic H-O-S propositions. See Kemp (1976); Ronald Jones and José Scheinkman; my 1974, 1979b articles; Winston Chang, et al.
competition; and Kelvin Lancaster has also applied to international trade his product-characteristics approach to consumption theory. These authors are concerned with the same stylized facts that my earlier paper (1979a) had been, and they reach some broadly analogous conclusions, although from entirely different theoretical starting points. The present paper is not concerned with the differentiated consumer goods that are their subject. However I do hope to treat the differentiated producer goods central to my own theory in such a way as to bring out the parallels between the approaches of these other authors and my earlier work on intra-industry trade (I cannot resist the temptation to point out that producers' goods are in fact much more prominent in trade than are consumers' goods).

I. National and International Scale Economies

A. Production

Capital and labor combine to produce wheat ($W$) and manufactures ($M$). Wheat is produced subject to constant returns to scale via a smooth production function of the familiar sort. Manufactures, on the other hand, are potentially subject to increasing returns to scale, and I suppose that the production function for $M$ is separable in the sense that I can write $M = km$, where $k$ is an index of scale economies and $m$ an index of the scale of operations; $m$ can be thought of as produced via a familiar smooth production function. Thus the endowment of capital and labor determines a transformation curve:

$$W = T(m).$$

The possibility of returns to scale in manufacturing arises, first, from exploitation of the division of labor, as in the hoary examples of Adam Smith's pin factory and the Swiss watch industry.Finished manufactures are costlessly assembled from intermediate manufactured components.$^6$ The number $n$ of components that are actually produced will be endogenously determined, and, I assume, only a fraction of a large number of potentially producible intermediate goods.$^7$ I am not interested in issues which depend upon distinctions between potential intermediate goods, so I assume that they are all producible from capital and labor via identical production functions, and that all produced components contribute in totally symmetric fashion to the finished manufactures. Under these assumptions, all components which are actually produced will be produced in equal amounts, and I denote the output of each such component by $x$, so that the total number of produced components of all types is $nx$. Assume that the output of finished manufactures is given by

$$M = n^{a-1}(nx)$$

for some parameter $a > 1$. It will be convenient to adopt the following specific form of the production function (2), where $x$ denotes the quantity of the $i$th component and $\beta$ is a parameter, $1 > \beta > 0$.

$$M = n^a \left[ \sum_{i=1}^{n} \left( \frac{x_i^\beta}{n} \right) \right]^{1/\beta}.$$

With all the $x_i$ equal to a common $x$, as will be the case in equilibrium, (2') reduces to (2). A higher value of $\beta$ indicates that components can be more easily substituted for each other in the assembly of finished manufactures. Thus lower values of $\beta$ correspond to greater "product differentiation" within the manufacturing sector. (Note the analogy to the ways in which Dixit-Stiglitz; Krugman, 1979, 1981; Dixit-Norman; and Helpman measure the utility obtained from a bundle of differentiated consumer goods.)

The gains from an increased division of labor would mandate the production of an infinitesimal amount of an infinite number of separate components if that were possible.

$^6$One could instead interpret the intermediate goods as successive stages, Austrian fashion, or alternatively, allow some of the components to be assembly services, so that assembly is not costless. Either interpretation would leave the balance of this paper unscathed.

$^7$I ignore the difficulty of interpreting nonintegral values of $n$. 
I assume that indivisibilities in the production of components prevents this (the division of labor is limited by the extent of the market). If the scale variable \( m \) is interpreted as an index of the number of bundles of factors devoted to manufacturing production, suppose that the number of such bundles required to produce \( x \) units of any component is \( ax + b \), for some \( a, b > 0 \). Then

\[
m = n(ax + b).
\]

The technology of the model is formally summarized by equations (1), (2) (or (2')), and (3), with the usual H-O-S model behind the transformation curve \( T(m) \).

Note that there are two distinct sources of increasing returns to scale. The individual component production functions, \( ax + b \), display what Balassa (1967, ch. 5) refers to as “economies of scale in the traditional sense,” and which I term “national” returns. These involve considerations of minimal plant size and the like, and they require total production \( x \) to be geographically concentrated. I assume that these economies are internalized by firms, and I shall examine equilibria in which the total output of each produced component is provided by a single firm in a single location. (I shall sometimes refer casually to \( b \) in equation (3) as “fixed costs,” but I will not assume that \( b \) is variable in the long run.)

The finished-manufactures production function (2), or (2'), displays constant returns to scale for a given value of \( n \). But an expansion of the manufacturing sector arising from an increased number of components (a rise in \( n \) with constant \( x \)) displays increasing returns, since \( M \) rises in greater proportion than \( nx \). These economies reflect not an increased plant size but rather a greater division of labor; they are what Balassa (1967, ch. 5) refers to as “horizontal specialization” or “vertical specialization,” and were the subject of my earlier paper (1979a), where they were called “international” returns to scale. Economies of this sort depend upon the size of the market for finished manufactures, and they do not require that all manufacturing output be concentrated at a single place. I assume that these economies are external to the individual firm. Components are assembled into finished manufactures by many competitive firms, each of which takes \( n \) as a parameter and consequently views itself as subject to constant returns to scale.

**B. Autarkic Equilibrium**

Consider a closed economy with the above technology. An individual producer of finished manufactures uses components, subject to (2'), with \( n \) as a parameter. If \( q_0 \) and \( q \) denote the prices, in terms of wheat, of some pair of produced components with outputs \( x_0 \) and \( x \), then cost minimization by producers of \( M \), subject to (2'), requires

\[
x_0 = x(q/q_0)^{1/(1-\beta)}.
\]

If \( n \) is sufficiently large so that the producer of each component acts as though his behavior does not influence that of other component producers, condition (4) is the demand curve faced by the producer of \( x_0 \), for given \( q \) and \( x \). This curve has an elasticity of \( 1/(1 - \beta) \). The component-producer purchases the services of primary factors in competitive markets and therefore has a cost function given by \(-T'(m)[ax_0 + b] \), where \( T'(m) \) is exogenous to the individual firm. This firm will therefore equate marginal revenue and marginal cost, and maximize its profit, by charging the price

\[
q_0 = -T'(m)a/\beta.
\]

Because of the symmetry assumption, \( q_0 \) is the price of each component that is actually produced. The profit of each component-producing firm is \( q_0x_0 + T'(m)[ax_0 + b] \). These profits will be driven to zero in equilibrium by the entry and exit of firms, that is, by variations in \( n \), as each component is produced by only one firm. Thus, from equation (5),

\[
x_0 = b\beta/a(1 - \beta).
\]

---

8 This assumption was implicit in my earlier paper (1979a) and explicit in Dixit-Stiglitz and in Krugman (1979).
Finally, the number $n$ of components is given by equation (3), for any given $m$ and for $x = x_0$.

$$n = (1 - \beta) m / b. \tag{7}$$

Equations (5), (6), and (7) now imply the value of $k$ for $M = km$:

$$k = \left( [(1 - \beta)/b]^{a - 1} \beta / \alpha \right) m^{a - 1}. \tag{8}$$

The relative supply price $P_S$ of $M$ in terms of wheat is given by $P_SM = q_0nx$, or $P_S n^a x = q_0 nx$, or $P_S = n^{1-a} q_0$. Thus substitution yields the supply curve of $M$:

$$P_S = -T'(m)/k. \tag{9}$$

The supply curve is illustrated in Figure 1. The term $M_0$ denotes the value of $M$ when $T = 0$, and the curve $P_S$ shows for each value of $M$ the minimum price at which that quantity would be supplied (i.e., expression (9)). Note from (6) that $x$ is independent of $m$. Thus a reallocation of resources from $W$ to $M$ involves an expansion of $M$ production by an increase in $n$. The manufacturing sector thus displays increasing returns to scale, allowing a negatively sloped supply curve, as illustrated.

The picture of autarkic equilibrium is completed by a description of the demand for final goods. I assume that a constant fraction $\gamma$ of income is always spent on manufactures. Each output combination $M$ and $W$ determines a demand price $P_D$ that will clear commodity markets: $P_D M = \gamma [W + P_D M]$. Thus the demand curve can be written as

$$P_D = \left[ \gamma / (1 - \gamma) \right] T(m) / km. \tag{10}$$

The demand curve has a negative slope, as illustrated in Figure 1, and it is easy to show that equations (9) and (10) imply that the demand curve intersects the supply curve from above. Thus in autarky the economy possesses a unique equilibrium which features production of both final goods.

C. International Equilibrium

Consider next free international trade between two economies—of the sort just described—identical in all respects other than factor endowments. Let $m$ and $m^*$ denote the scale of manufacturing operations at home and abroad (variables pertaining to the foreign country will be distinguished by an asterisk). In free-trade equilibrium the total output of each component will be concentrated in a single country. Thus, if $m$ and $m^*$ are both positive, the two countries produce distinct collections of components. Equation (4) continues to represent the demand curve faced by each component producer (in either country), and the same argument as before establishes that (6) gives the output of each component actually produced somewhere. Then, if $n_H$ and $n_F$ denote the number of components produced at home and abroad, respectively, it follows as before that $n_H = (1 - \beta) m / b$ and $n_F = (1 - \beta) m^* / b$ so that $n = n_H + n_F$ is given by

$$n = (1 - \beta) (m + m^*) / b, \tag{11}$$

and so, from (2), the world output of finished manufactures is

$$M + M^* = \left( \frac{\beta}{\alpha} \right) \times \left[ [(1 - \beta)/b]^{a - 1} (m + m^*) \right]^a. \tag{12}$$
What determines \( m, m^* \) and the relative prices? To answer this question I use the allocation-curve technique\(^9\) developed in my 1979a and forthcoming articles. For any \( m \) and \( m^* \), the world demand price of finished manufactures in terms of wheat is given by

\[
P_D = \frac{\gamma}{1-\gamma} \frac{T(m) + S(m^*)}{M + M^*}
\]

\[
= \frac{\gamma}{1-\gamma} \frac{a}{b} \left( \frac{b}{1-\beta} \right)^{a-1} \frac{T(m) + S(m^*)}{(m + m^*)^a}
\]

where \( S(m^*) = W^* \) denotes the foreign transformation curve. The home supply price \( P^H_S \) must be given by

\[
P_S^H = - \left[ (1-\beta)(m + m^*)/b \right]^{1-a} T'(m)a/\beta.
\]

Home-country equilibrium requires \( P_D = P_S^H \), or, from equations (13) and (14),

\[
\gamma[T(m) + S(m^*)]
\]

\[
+ (1-\gamma)(m + m^*)T'(m) = 0.
\]

Equation (15) defines the "home allocation curve"; the collection of \( m \) and \( m^* \) for which the home country is in equilibrium in the international economy. This relation is depicted as the \( HH' \) curve in each panel of Figure 2. It is easily shown that the curve has a negative slope, that \( P_D < P_S^H \) above the curve, and that \( P_D > P_S^H \) below. In Figure 2, \( m_0 \) and \( m^*_0 \) denote the manufacturing scales if the home and foreign economies, respectively, specialize to manufactures, so the world is confined to the rectangle \( 0m_0Em^*_0 \) in each panel, and that part of the graph of (15) lying outside this rectangle is irrelevant (it is easy to see that the graph of (15) cannot be completely excluded from the rectangle). If the home economy specializes, equilibrium is consistent with an excess of the relative demand price of the produced good over its relative supply price. Thus the home allocation curve also contains any part of \( 0m^*_0 \) that lies above the intersection of \( H'H \) with the vertical axis, and also any part of \( m_0E \) below its intersection with \( H'H \).

Similarly, the foreign allocation curve, showing the \( m \) and \( m^* \) for which \( P_D = P_S^H \), the foreign relative supply price, is given by

\[
\gamma[T(m) + S(m^*)]
\]

\[
+ (1-\gamma)(m + m^*)S'(m^*) = 0.
\]

The graph of (16) is shown as \( FF' \) in each panel of Figure (2). As illustrated, \( FF' \) must decline less steeply than \( H'H \), reflecting the fact that each country's supply price is relatively more sensitive than the world demand price to that country's allocation of resources compared to the foreign allocation. Any part of \( m^*_0E \) lying below \( FF' \) and any part of \( 0m_0 \) above \( FF' \) also constitute part of the foreign allocation curve.

International equilibrium requires each country to be individually in equilibrium in the world economy and is therefore determined by the intersection of the two allocation curves. Figure 2 depicts the five qualitatively distinct possibilities; equilibrium is shown by points \( F, H', D, H, \) and \( F' \) in panels (a)–(e), respectively (in addition there are the two boundary cases, featuring complete specialization in both countries, between cases (a) and (b) and between (d) and (e)). Evidently each case yields a unique international equilibrium.

The equilibrium values of \( m \) and \( m^* \) determine the relative price of finished manufactures in terms of wheat from equation (13) (or (14)). Then \( n \) is given by (11) and \( x \) by (6). The price of each component is as indicated by (5), or, if the home economy specializes to wheat, by the foreign analog.

Note that equation (12) can also be written \( M + M^* = k(m + m^*) \) where \( k = (\beta/a) \cdot [(1-\beta)/b]^{a-1}(m + m^*)^{a-1} \), thereby emphasizing that the extent of external scale economies depends on the size of the world market alone. If \( M \) and \( M^* \) denote the quantities of finished manufactures to which the respective countries obtain effective title as a result of their manufacturing activity (i.e., the

\(^9\) The remainder of this section is rather terse as it closely follows my 1979a article, to which the reader is referred for more detail.
Figure 2
most finished manufactures each country could consume without offering any wheat in exchange), then,

\[(17) \quad M = km; \quad M^* = km^*.\]

Note that \(M\) and \(M^*\) need not equal the quantities of finished manufactures actually assembled in each country. With assembly costless it can be divided in any way between the two countries. (If one of the components were an assembly process, \(M + M^*\) would be all assembled in the country where that process happened to be located.)

II. The Modern Theory

Armed with a complete description of international equilibrium, this section inquires into the fate of the four central propositions supplied by the H-O-S model.\(^{10}\) Henceforth manufacturing is assumed relatively capital intensive and wheat relatively labor intensive. Since inherent properties of the two factors do not figure in this paper, this pattern of factor intensities is in a formal sense only a definition. But the definition is not arbitrary: one intuitively associates a greater division of labor with significant use of capital. This could be introduced by letting manufacturing production explicitly take time, adding a positive interest rate, and allowing components to be producers’ goods generally. It is straightforward to extend the main results of this paper to such a context by using the techniques in my article (1979b), provided attention is restricted to steady states.

A. Factor-Price Equalization

Suppose that both countries operate both sectors, so that equilibrium is as in Figure 2(c). Then equations (15) and (16) hold, implying that \(T'(m) = S'(m^*)\). Likewise, \(T'\) and \(S'\) are necessarily unequal if one or both countries specializes completely.\(^{11}\) Now, with the assumption of identical technology, equality of \(T'\) and \(S'\) implies that domestically produced components sell at the same price as all foreign-produced components, from (14) and its foreign counterpart. Since \(T\) and \(S\) are generated by conventional H-O-S technologies, the standard factor-price equalization argument can now be employed to yield

PROPOSITION 1: Factor-Price Equalization. If both countries operate both sectors, and are not separated by a factor-intensity reversal, free trade implies factor-price equalization. A separating reversal, or (incipient) specialization in either country, precludes such equalization.

B. The Rybczynski Theorem

Note, first, that if domestic factor endowments are altered and \(T'(m)\) is kept unchanged, the standard Rybczynski argument applies exactly: an increase in capital will cause a more than proportionate rise in \(m\) and a fall in \(W\).

The next step is to relate the change in \(m\) to the production of individual components and finished manufactures. Expression (5) implies that constancy of \(T'(m)\) is equivalent to constancy of the relative prices of components in terms of wheat. Then (6) implies that endowment changes cause no changes in the outputs of those components that are produced both before and after, while (7) reveals that the number of produced components varies in proportion to \(m\). Thus:

PROPOSITION 2: Rybczynski. At constant relative component prices, an increase in the capital stock will absolutely reduce the production of wheat, have no effect on the outputs of all components initially produced, and increase the number of produced components in greater proportion than the rise in the capital stock itself.

Several features of this Rybczynski theorem deserve emphasis. First, despite the pervasive presence of economies of scale, the standard results are completely preserved at the interindustry level: changes in endow-

\(^{10}\)For a discussion of the basic H-O-S propositions see, for example, Caves and Ronald Jones.

\(^{11}\)Except, of course, for the borderline case of “incipient diversification” in one country.
ments produce magnified changes in the outputs of wheat and aggregate manufactures. Because all produced components sell at the same price and have equal outputs, aggregate component production can be measured by nx which changes proportionally to n since the adjustment of manufacturing output occurs through variations in the number of components rather than in the outputs of individual components.

Second, at a disaggregated level, endowment changes produce magnified changes in the outputs of wheat and of those components which change status (and thereby experience infinite proportional output changes). The (zero) proportional output changes of the other components will be weighted averages of proportional endowment changes if and only if the latter are in opposite directions.

Note, third, that the present version of the Rybczynski theorem places no restrictions on the relative price of finished manufactures. Nor does it make any claims about the effective output M of these final goods. This is because international economies of scale influence the production of final manufactures independently of the internal, national, economies associated with the production of individual components. The extent of the former depends upon patterns of production in both countries. Thus, in order to describe the effects of endowment changes in one country on that country’s (effective) output of finished manufactures, the behavior of the rest of the world must be considered.

With relative component prices held fixed, m* will be unchanged. Now equation (17) implies that the proportional change in M equals that in m plus that in k. With m* constant, the proportional change in k equals that in m multiplied by \((\alpha - 1)m/(m + m^*)\). This is straightforward: an increase in m takes the form of an increased number of components and thereby raises M both directly and indirectly through the scale economies made possible by the increased division of labor. The latter scale effect depends upon the rate \(\alpha - 1\) at which scale economies are realized, plus the extent to which an increase in m constitutes an increase in world manufacturing \(m + m^*\). In any event, M changes in the same direction as m and in even greater proportion, so that the effect of economies of scale is to simply accentuate the standard results.

With constant component prices, changes in the scale of manufacturing activity must change the relative price of finished manufactures. Equation (14) reveals that this price falls in the proportion \((\alpha - 1)m/(m + m^*)\hat{n}\), where I use a circumflex to denote proportional change. This is the same scale effect as above, entering in for the same reason. The relative price of finished manufactures falls in the proportion that k rises; thus the value of finished manufactures, in terms of either wheat or components, rises in the same proportion as aggregate component output. This is a reflection of the external character of the relevant scale economies.

**PROPOSITION 3:** At constant relative component prices, an increase in capital increases the output of finished manufactures relative to that of components (and therefore to the capital stock), and also reduces the price of finished manufactures, so that the value of the output of finished manufactures rises in the same proportion as that of aggregate components.

It is natural to wonder how the Rybczynski theorem is affected when the relative price of finished manufactures in terms of wheat is held constant. The exercise is not in fact very interesting because of the international interdependence. For example, equation (14) reveals that any increase in m brought about by endowment changes must be accompanied by an increase in T sufficient to leave \((m + m^*)^{1-\alpha}T\) unchanged. Thus the relative price of components in terms of wheat rises, and the factor prices and techniques must change as well. Furthermore, if the rest of the world is not specialized to wheat, the equilibrium condition of equal component prices requires foreign changes in factor prices and techniques.

**C. The Stolper-Samuelson Theorem**

Expression (5) implies that the wheat price of components is linked to factor prices in essentially the same way as are commodity...
prices in the standard H-O-S model. Then the usual argument establishes the following:

\[ \hat{P}_C = \theta_{km} \hat{r} + \theta_{Lm} \hat{w}; \]

\[ \hat{P}_W = \theta_{kw} \hat{r} + \theta_{LW} \hat{w}. \]

In these equations \( P_C, P_W, r, \) and \( w \) denote the nominal prices of industrial components, wheat, capital and labor, respectively (so that \( q = P_C/P_W \)). The \( \theta_{km} \) and \( \theta_{LW} \) denote capital's and labor's distributive shares in the wheat industry and analogously for \( \theta_{km} \) and \( \theta_{Lm} \) in the components industry. The following is immediate from (18) and (19).

**PROPOSITION 4:** An increase in the price of wheat relative to the price of all components will raise the wage relative to the prices of both wheat and components and will reduce the rent relative to the prices of both wheat and components.

Industrial components are intermediate goods, and only wheat and finished manufactures are consumed. Thus Proposition 4 says nothing about real factor rewards and so fails to come to grips with the essence of the Stoepfer-Samuelson Theorem. The price \( P_M \) of finished manufactures is related to that of components by \( P_M = n^{1-a}P_C, \) so that

\[ \hat{P}_M = \hat{P}_C (1 - (\alpha - 1)\hat{m}). \]

Now \( \hat{m} = \hat{m}[(m/(m + m^*)], \) and \( \hat{m} \) can be related to \( \hat{P}_M - \hat{P}_W \) by differentiation of (14) to obtain

\[ (\hat{P}_M - \hat{P}_W) = \left[ m \frac{T''}{T'} - (\alpha - 1) \frac{m}{m + m^*} \right] \hat{m}. \]

Substitution into (20) gives

\[ (\hat{P}_M - \hat{P}_C) = Z(\hat{P}_M - \hat{P}_W) \]

where the elasticity of the intraindustry price structure \( P_M/P_C \) with respect to the interindustry price structure \( P_M/P_W \) is \( Z = 1/(1 - [mT''/T'][(\alpha - 1)m/(m + m^*)]). \)

Note first that, although \( Z \) may be either sign, it can never fall between zero and unity. Thus variations in the intra-industry price structure are never damped versions of variations in the interindustry structure. This reflects the role of scale economies. A change in relative prices involves a reallocation of resources. If resources move into manufacturing, that is, if \( \hat{m} > 0, \) this will be associated with an increase in \( P_C/P_W \) to a degree dependent upon the curvature of the transformation curve. Since the cost of components contributes towards the cost of manufactures, this will also tend to be associated with a rise in \( P_M/P_W. \) This effect is simply the familiar H-O-S phenomenon, and I accordingly call it the intersectoral price effect of \( \hat{m}; \) it is reflected in the first term in the brackets in (21).

The variation in \( m \) will involve a variation in \( n \) rather than \( x, \) so individual firms will not experience internal scale economies. But the expansion of world manufacturing will induce external scale economies, and this scale effect will tend to reduce the cost of finished manufactures and will therefore work against the intersectoral effect. This scale effect, reflected in the second term in brackets in (21), is the same elasticity of \( k \) with respect to \( m \) encountered above. If the intersectoral effect dominates, \( P_M/P_W \) will therefore change in the same direction as, but to a lesser extent than, \( P_C/P_W, \) whereas a dominant scale effect\(^{12} \) means that a rise in \( P_C/P_W \) actually reduces \( P_M/P_W. \)

**PROPOSITION 5:** If the scale effect dominates the intersectoral effect \((Z > 0), \) changes in the intrasectoral price structure \( P_M/P_C \) are magnifications of changes in the intersectoral structure \( P_M/P_W. \) If the intersectoral effect dominates the scale effect \((Z < 0), \) the two relative prices always change in opposite directions.

We are now in a position to relate factor rewards to final-goods prices. Substituting equation (22) into (18) and (19) gives

\[ \hat{P}_M = \theta_{kw} \left( \frac{(\theta_{km}/\theta_{kw}) - Z}{1 - Z} \right) \hat{r} + \theta_{LW} \left( \frac{(\theta_{km}/\theta_{LW}) - Z}{1 - Z} \right) \hat{w}. \]

\(^{12}\) These conclusions can perhaps be made clearer by rewriting equation (22) as \((\hat{P}_C - \hat{P}_W) = (1 - Z)(\hat{P}_M - \hat{P}_W)\) and noting that \( Z \) cannot fall between zero and unity.
\[ \hat{P}_W = \theta_{kw} \hat{r} + \theta_{lw} \hat{w}. \]

Note first that (19) is not affected, so that the proportional change in the price of wheat is still a weighted average of those in the two factor rewards. Second, the coefficients of \( \hat{w} \) and \( \hat{r} \) in (23) sum to unity, as in (18). With labor designated intensive to wheat, \( \theta_{lm} < \theta_{lw} \), so that the exclusion of \( Z \) from between zero and unity renders the coefficients of \( \hat{w} \) in (23) positive. Thus that of \( \hat{r} \) is less than unity. This is all that can be said without restrictions on technology.

If \( Z < 0 \), the coefficient of \( \hat{r} \) in (23) is positive, so that \( \hat{P}_M \) is a weighted average of \( \hat{r} \) and \( \hat{w} \). Furthermore, the coefficient of \( \hat{r} \) in (23) exceeds that of \( \hat{r} \) in (19), and the usual Stolper-Samuelson argument follows. If, on the other hand, \( Z > 0 \), then in fact \( Z > 1 \), so the coefficient of \( \hat{r} \) in (23) is less than that in (19) and the standard results can not possibly hold.

**PROPOSITION 6:** Stolper-Samuelson. An increase in the price of manufactures relative to wheat raises the rent and lowers the wage relative to both final-goods prices, if and only if the intersectoral effect dominates the scale effect.

The reason for this result is straightforward. A dominant intersectoral effect ensures that a change in \( P_M/P_W \) produces a larger change, in the same direction, in \( P_C/P_W \), to which Proposition 4 can be applied to yield the standard result. A dominant scale effect precludes this result by causing \( P_C/P_W \) to change in the opposite direction. Note that Propositions 5 and 6 together give a very simple test for the validity of the Stolper-Samuelson theorem: The intrasectoral price structure \( P_M/P_C \) and the intersectoral structure \( P_M/P_W \) should always change in opposite directions.

If \( Z \) exceeds unity it may or may not also exceed \( \theta_{km}/\theta_{kw} \). If the scale effect is sufficiently dominant so that \( Z \) does exceed this latter term, the coefficient of \( \hat{r} \) in equation (23) is again positive and \( \hat{P}_M \) is a weighted average of \( \hat{w} \) and \( \hat{r} \). But the coefficient of \( \hat{r} \) in (23) is less than that of \( \hat{r} \) in (19). Thus the standard Stolper-Samuelson argument applies, but with the role of the factor-intensity pattern reversed, so that an “anti-Stolper-Samuelson” result follows.

**PROPOSITION 7:** An increase in the price of manufactures relative to wheat raises the wage and reduces the rent relative to both final-good prices, if and only if the scale effect dominates the intersectoral effect sufficiently so that \( Z > \theta_{km}/\theta_{kw} \).

Finally, if \( Z \) lies between unity and \( \theta_{km}/\theta_{kw} \), the coefficient of \( \hat{r} \) in (23) is negative, and that of \( \hat{w} \) in (23) exceeds that of \( \hat{w} \) in (19). Thus any change in \( P_M/P_W \) causes \( w \) to rise relative to one final-good price and fall relative to the other, but the rental still behaves as above.

**PROPOSITION 8:** The relative price of manufactures and the real income of capital vary in the same or opposite directions according as the intersectoral and intrasectoral price structures vary in identical or opposite directions.

Note that a large scale effect tends to destroy the duality between the Stolper-Samuelson and Rybczynski theorems, since it works to reverse the former while leaving the latter unscathed.\(^{13}\)

### D. The Heckscher-Ohlin Theorem

Suppose that both countries produce both goods in international equilibrium, as in Figure 2(c). Then \( T'(m) = S'(m^*) \), and Proposition 3 implies that the output ratio \( M/W \) is higher in the country endowed with the higher capital-labor ratio. Since the two countries consume the two final goods in equal proportions, each country is a net exporter of the sector intensive to its relatively abundant factor.

Next suppose that at least one country specializes completely. As illustrated in Figure 2, the home allocation curve is steeper than the foreign and, in this case, must lie

\(^{13}\)See Krugman (1981) for an interesting alternative treatment of income distribution in a model featuring relative factor-endowment differences and intraindustry trade in differentiated consumer goods.
either wholly outside or wholly inside the foreign. In the former case (Figure 2, panels (d) and (e)) the home country is a wheat importer and in the latter case (Figure 2, panels (a) and (b)) a wheat exporter. What distinguishes the two cases?

At point $F'$, $m$ and $m^*$ satisfy equation (16). In the home country, from equations (13) and (14),

$$P_D - P_S^H = \frac{a}{\beta} \left( \frac{b}{1-\beta} \right)^{\alpha-1} (m + m^*)^{1-\alpha} \times \left[ \frac{\gamma}{1-\gamma} \frac{T + S}{m + m^*} + T' \right].$$

Substitution of (16) into this expression reveals that, at $F'$, $P_D - P_S^H$ has the same sign as $[T'(m) - S'(m^*)]$. Now, if $F'$ falls in the interval $0m_0$, $m^* = 0$ and $0 < m < m_0$, whereas $0 < m^* < m_0^*$ if $F'$ lies in the interval $m_0E$. In either case, $T' - S'$ is necessarily negative if the foreign country is relatively capital abundant, provided that the two countries are not separated by a factor-intensity reversal. Then $P_D < P_S^H$ at $F$, which implies that $F'$ lies above the home allocation curve. A similar argument applied to $F$ reveals that the home allocation curve necessarily lies above the foreign when the home country is relatively capital abundant. Panels (a) and (b) of Figure accordingly illustrate home relative labor abundance, and panels (d) and (e) illustrate home capital abundance.

PROPOSITION 9: Heckscher-Ohlin. In international equilibrium, each country necessarily exports the good intensive in its relatively abundant factor, if the two countries are not separated by a factor-intensity reversal.

The quantity version of the Heckscher-Ohlin theorem remains intact. Clearly the price version cannot similarly escape unscathed: the scale effect can alter the relation between commodity and factor prices. Also, Section I showed that relative autarkic commodity prices depend upon national size as well as relative endowments. Since only the latter determine trade patterns, these patterns need not reflect autarkic price differences.

But in spite of all this, the price version does have a role to play, a role that depends on a distinction between intraindustry trade and interindustry trade. Basically, the pattern of free interindustry trade (the net exchange of manufactures for wheat) will correspond to what the pattern of relative factor prices in the two countries would be in the absence of such trade but with free intraindustry trade (the exchange of components). The latter in effect isolates the influence of internationally decreasing costs.

To see this, consider a quasi-autarkic equilibrium in which the two countries freely exchange industrial components, but in which there is no trade in wheat. That is, the home economy consumes $W = T(m)$ and $M = km$, where $k = \beta[(m + m^*)(1 - \beta)/b]^{\alpha-1}/a$, and analogously abroad. Such an equilibrium is as described in Section I.B. for each country, except that the two economies are linked by a common value of $k$ dependent upon $m + m^*$. The equilibrium values of $m$ and $m^*$ are the solution to

$$\gamma T(m) - (1 - \gamma) T'(m)(m + m^*) = 0$$
$$\gamma S(m^*) + (1 - \gamma) S'(m^*)(m + m^*) = 0.$$
rately predicted by comparative quasi-autarkic relative commodity prices.

Since in quasi autarky the two countries have distinct relative prices but common values of \( n \) and \( k \), the logic of Proposition 6 can be employed to compare the two national equilibria in the absence of a scale effect. This gives the present form of the price version of the Heckscher-Ohlin Theorem.

**PROPOSITION 10: Quasi-Autarky Theorem.** In the absence of a separating factor-intensity reversal, each country in free trade is a net exporter of the sector relatively intensive in that factor with the lower relative quasi-autarkic price in that country.

Note that the quasi-autarky theorem applies to a world such as that envisioned by Gary Hufbauer and John Chilas, who stress that trade between industrial countries is largely intratrade and involves much less interindustry specialization than does interregional trade within the United States. They see this as due in large part to the pattern of tariff reductions made since the war.

### III. The Factor-Endowments Basis of Intratrade

The previous sections have established the continued relevance to international trade of the basic Heckscher-Ohlin idea that trade is a substitute for international factor mobility. The present environment, for example, does no violence to the factor-price-equality theorem.

The purpose of this section is to establish that intratrade trade likewise has a factor-endowments basis, but that such trade is complementary to international factor mobility. Thus a similarity of factor endowments between nations tends to promote such trade as it limits the scope for interindustrial exchange. This property is emerging as a central feature of models with intratrade trade. The result was first deduced in my article (1979a) and subsequently appeared in the quite distinct work based on differentiated consumer goods.

### A. The Complementarity Theorem

Recall the assumption that finished manufactures are costlessly assembled from bundles of all components. The above analysis established that if both countries produce manufactures, they specialize to distinct collections of components.

It will prove convenient, and consistent with both the balance of this paper and most empirical work, to measure intratrade trade as what it would be if finished manufactures were assumed to be costlessly assembled where consumed, and if no component entered trade more than once, so that international trade in manufactures therefore consisted entirely of the shipment of components from their country of manufacture to where they are combined with other components and consumed. Then the home country’s import \( M_C \) and export \( X_C \) of manufactures (components) must be

\[
M_C = n_F x, \quad X_C = n_H x (1 - g);
\]

where \( g \) equals domestic national income as a fraction of world income (i.e., \( (PM + W)/(P(M + M^*) + W + W^*) \)). I shall use the relative index \(^{14}\) employed by Hesse, Grubel and Lloyd, Caves and others: \( \rho = 1 - |X_C - M_C|/(X_C + M_C) \). Then substitution yields

\[
\rho = \begin{cases} 
2 gn_F / [(1 - g)n_H + gn_F] & \text{if } n_H > gn, \\
2 (1 - g)n_H / [(1 - g)n_H + gn_F] & \text{if } n_H < gn.
\end{cases}
\]

Higher values of \( \rho \) indicate relatively more intratrade trade; \( \rho = 1 \) if all manufacturing trade is intratrade and \( \rho = 0 \) if it is all interindustry.

The complementarity between intratrade trade and factor movements should now be apparent. Let \( h \) and \( h^* \) denote the capital-labor endowment ratios at home and abroad, and designate the home country as the capital abundant one, so that \( h > h^* \). Suppose that the two countries freely trade, with neither country specialized and with

\(^{14}\text{Alternative measures are discussed in Grubel and Lloyd (1971).}\)
factor prices equalized. Suppose now that the factors are slightly "traded" between the two countries so as to reduce \( h - h^* \) while leaving each country's income unchanged at the unchanged factor prices and commodity prices. Then \( n_H \) falls, \( n_F \) rises by the same amount, and \( g \) is unchanged. Now, since the home economy is relatively capital abundant, its share of the world output of components \( (n_H/n) \) must exceed its share of world income, \( g \). Thus \( \rho \) equals the top expression in (24), which directly reveals that a relative-endowment-equalizing trade of factors must raise \( \rho \).

**PROPOSITION 11: Complementarity Theorem.** If both countries initially produce both goods, and if there are no separating factor-intensity reversals, a small relative-endowment-equalizing trade of primary factors will increase \( \rho \).

The basic complementarity property becomes most apparent upon the comparison of extreme cases. If the two countries' endowments differ sufficiently so that one country specializes to wheat there is no intraindustry trade and all trade is inter-industrial. If, on the other hand, \( K = K^* \) and \( L = L^* \), there is no basis at all for interindustry trade (each country will be self-sufficient in wheat), but intraindustry trade will be maximized since the two countries will produce distinct collections of an equal number of components.

**B. Technological Structure and Intratrade Trade**

While the Complementarity Theorem establishes that intratrade trade is sensitive to factor endowments, it is nonetheless clear that the existence of such trade is due to the assumptions about the technology of manufacturing production. Attention thus naturally focuses on the sensitivity of my measure of intratrade trade to the technological parameters: \( a, b, \) and \( \beta \).

Note, first, that these parameters do not influence the basic allocation of resources which, by Proposition 9, is determined as described by the modern theory of international trade. The intersection of the allocation curves determines \( m \) and \( m^* \), and these curves are invariant with respect to all three parameters. Attention therefore focuses on equations (6) and (7).

Consider the relative measure \( \rho \). Changes in the technological parameters will not change \( g \), because they produce offsetting changes on the volume and the relative price of finished manufactures, from (12) and (14). Substitution of (7) into the top line of (24) gives \( \rho = 2gm^*/[(1-g)m + gm^*] \). This immediately yields

**PROPOSITION 12: The relative index of intratrade trade is invariant with respect to the degree of product differentiation and the levels of fixed and marginal costs.**

An increase in marginal costs, by reducing \( x \) with \( n_H \) and \( n_F \) constant, lowers both types of trade in proportion. An increase in product differentiation also reduces interindustry trade pari-passu with intratrade trade. This is surprising: with the existence of intratrade trade dependent upon product differentiation, one might expect more differentiation to increase such trade. But the explanation is simple. From (7) the number of components does rise, but (6) reveals that the output of each falls in even greater proportion, because of fixed costs.

In sum, the technological parameters play a knife-edge role: their existence is crucial for the present theory, but changes in their values have few effects, or sometimes counterintuitive effects, upon intratrade trade. Empirical investigations have so far produced mixed results (see Pagoulatos and Sorenson; Caves; and references cited therein).

**C. Multilateral Trade**

A brief consideration of this paper's implications for multilateral trade further illustrates the above discussion and also brings out some features that can not arise at all in the two-country framework, but that reflect the stylized facts of modern trade.

The model used thus far is retained, except that many countries of the sort described
above are allowed. Suppose that no factor-intensity reversal is displayed by the common technology. Let $h_m$ and $h_w$ denote the capital-labor ratios in the two sectors in countries which diversify production at the existing prices. Denote by $h'$ the capital labor ratio having the property that any country with an endowment $h > h'$ necessarily imports wheat, and any country with an endowment $h < h'$ necessarily exports wheat.\(^{15}\) Assume $h_m > h' > h_w$. The various possibilities an individual country could experience are indicated in Table 1.

The concerns of the Complementarity Theorem are best brought out in Cases II, III, and IV where similarities of endowments (to each other and to the world as a whole) foster intraindustry trade while limiting interindustry trade. Cases I and V, on the other hand, illustrate behavior that can not arise in a two-country context. Countries in Case I engage in extensive intraindustry trade with the rest of the world (including other Case I countries), by virtue of their specialization to manufacturing, despite the fact that their endowments are quite different from that of the rest of the world as a whole. Countries in Case V, because they specialize to wheat, engage in no intraindustry trade at all with each other (or with anyone else), no matter how closely the factor endowments of these countries resemble each other. The relevance of all this to the stylized facts of contemporary trade should be clear.

\(^{15}\)If all countries diversify in production, $h'$ equals the world capital-labor ratio.

### Table 1—Possibilities in Multilateral Trade

<table>
<thead>
<tr>
<th>Case</th>
<th>Endowment</th>
<th>Production</th>
<th>Intraindustry Trade</th>
<th>Interindustry Trade</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$h \geq h_m$</td>
<td>specialized to $m$</td>
<td>with all other I, II, III, IV countries</td>
<td>imports $W$ from IV and V</td>
</tr>
<tr>
<td>II</td>
<td>$h_m &gt; h &gt; h'$</td>
<td>diversified</td>
<td>with all other I, II, III, IV countries</td>
<td>imports some $W$ from IV and V</td>
</tr>
<tr>
<td>III</td>
<td>$h = h'$</td>
<td>diversified</td>
<td>with all other I, II, III, IV countries</td>
<td>none</td>
</tr>
<tr>
<td>IV</td>
<td>$h' &gt; h &gt; h_w$</td>
<td>diversified</td>
<td>with all other I, II, III, IV countries</td>
<td>exports some $W$ to I and II</td>
</tr>
<tr>
<td>V</td>
<td>$h_w \geq h$</td>
<td>specialized to $W$</td>
<td>none</td>
<td>exports $W$ to I and II</td>
</tr>
</tbody>
</table>

### IV. Conclusion

This paper has developed a simple model of the interaction of national scale economies—internal to individual firms—with international returns to scale—external to firms—and with the modern, factor-endowments, theory of international trade. The result furnishes a detailed microeconomic backdrop to my earlier paper (1979a). In addition two conclusions emerge.

First, as formalized in the Complementarity Theorem, intraindustry trade in manufactures is complementary to international factor movements. Although the existence of such trade depends upon product differentiation and scale economies, these features play a knife-edge role and thereby leave the determination of the level of intraindustry trade largely to relative factor endowments. Second, the basic propositions of the modern theory of international trade remain, on the whole, essentially valid in the presence of scale economies, although some significant modifications do arise.

The second conclusion contrasts rather strongly with the traditional increasing-returns literature—thoroughly preoccupied with national returns to scale. The present treatment also gives national returns a prominent place. The indeterminacy and multiple equilibria characteristic of the standard analysis are still present (because national returns are still present), in terms of the location of production of individual components. But the sharp difference in my conclusions follows from the fact that, when national and international economies are
allowed to interact, disturbances to equilibrium typically take the form of changes in the number of production units rather than in their size, so that the concerns of the traditional theory do not arise. Of course it is possible to chip away at this conclusion by relaxing some of my assumptions, especially those that components are symmetric and that internal scale economies arise solely from fixed costs. These assumptions reflect in simple form my views of what is relatively important, but they cannot be literally accurate.

My earlier paper (1979a) argued that international increasing returns to scale are significant in the modern world economy. The present paper suggests that the conclusions of the earlier need not be altered even if national scale economies are also widespread and important. The resulting theory appears empirically relevant, with respect to both the stylized facts cited at the beginning of this article and also to recent studies—see Caves, and Rudolf Loertscher and Frank Wolter.

REFERENCES


Daly, Donald J., Key, B. A., and Spence, E. J., Scale and Specialization in Canadian Manufacturing, Staff Study No. 21, Economic Council of Canada, 1968.


Helpman, Elhanan, “International Trade in the Presence of Product Differentiation, Economies of Scale and Monopolistic Competition,” Journal of International Economics,


