and Slaughter (2001), Beaulieu (2002), and Beaulieu and C. Magee (2004) do find statistical support for the Stolper-Samuelson proposition in their analyses of political attitudes and behavior with regard to protectionist legislation, their finding that an individual's industry of employment influences political attitudes about protectionism indicates the need to recognize the existence of labor immobility within countries for a better understanding of the forces affecting relative factor prices. Bernard, Redding, and Schott (2005) and Bernard et al. (2005) point to the relative immobility of productive factors within a country as one of the reasons for their findings of significant differences in relative factor price across regions within the United Kingdom and the United States.

6 Conclusions and Related Research Topics

6.1 Successfully Modifying the HO Model

Two notable revolutions in real-side international economics have occurred over the 75 years covered by this review: one dealing with trade theory and the other with the empirical testing of propositions derived from trade theory. On the theoretical side, there was an abandonment of the classical "real cost" approach that combined positive and normative economics in an unsatisfactory manner. It was replaced with a neoclassical framework that utilized the concepts of opportunity cost and reciprocal demand to explain the determination of countries' outputs and levels of trade. The integration of the factor-proportions approach of Heckscher and Ohlin into this framework then permitted trade economists to formulate a general equilibrium theory that tightly tied together product prices and factor prices. The normative side of international economics was separated from the positive side and treated as a part of welfare economics. As Krugman (1994) points out in his comments at the conference celebrating the fiftieth anniversary of the Stolper-Samuelson theorem, these developments stimulated the

flourishing of real-side trade theory into a formal, model-oriented field.

The second revolution is the explosion of empirical studies aimed at rigorously testing theoretical trade models and, in particular, tying empirical tests closely to the appropriate theory. Leamer has long been a strong advocate of this point (see Leamer 1984, ch. 2, and Leamer and Levinsohn 1995). His discovery (Leamer 1980) that Leontief (1953) had improperly tested the Heckscher-Ohlin proposition is a classic example of its importance, and the point is now routinely made by empirical trade economists.

Real-side international economists have clearly made enormous progress in better understanding the interrelationships among the forces shaping trading patterns by emphasizing the dual importance of formulating tightly reasoned theoretical models and of rigorously testing the hypotheses derived from these models. However, an unintended consequence of this more formalistic approach—until recent years—seems to have been an overfascination with the elegance of the HOS and HOV models at the cost of minimizing the role of forces other than relative factor endowments in determining trading patterns. The typical real-side graduate text would generally begin with an exposition of Ricardo's one-factor model with fixed labor coefficients that vary between countries, but this is treated largely as a convenient introduction to the variable-proportions HOS model with its assumption of identical technology among countries. After thoroughly discussing in several chapters the basic theorems of this model and the complications that arise with more than two goods and two factors (i.e., the HOV model), the text would then typically devote only a few chapters to other less rigorous theoretical trade models that introduced such features as increasing

returns, differentiated products and monopolistic competition, oligopolistic markets, and endogenous changes in technology, factor endowments, and tastes. Given the elegant HOV framework, most trade economists seemed to behave as if there no longer was a strong reason to develop new, rigorous alternative theoretical trade models.

The inconsistency of the empirical results of tests of the Heckscher-Ohlin proposition over a number of decades following Leontief's 1953 test on 1947 US trade data also did little to encourage trade economists to modify the HOV model. As discussed in chapter 3, Leamer (1980) was able to show that Leontief's so-called paradox in 1947 US trade is reversed if the proper adjustment to the US trade balance is made. This adjustment fails, however, to reverse Baldwin's 1971 finding of a paradoxical result for 1962 US trade data nor Brecher and Choudhri's (1982a) finding of another seemingly paradoxical relationship with Leontief's 1947 results. Another example of inconsistent test results is, first, the claim by Stern and Maskus (1981) that the Leontief paradox disappears in 1972 US trade data and, then, Maskus's (1985) demonstration that this conclusion is erroneous. The introduction by BLS (1987) of a definition of factor abundance that differs from the one used by Leamer (1980) as well as by BLS (1986) in their working paper and that gives conflicting results about the existence of the paradox in 1967 US trade data also added to the inconsistency of the empirical findings on the HO proposition.

Trade economists seemed to fully appreciate the necessity of modifying the basic HOV model only after the findings that emerged from testing this model rigorously using carefully collected data on trade, factor input requirements, and factor endowments covering a large number of countries and several productive factors. These findings clearly

demonstrated how badly the model preformed. In particular, the findings by BLS (1987) that the signs of the two sides of the HOV equation match only a bit better than 50 percent of the time and by Trefler (1995) that the right-hand side of this equation grossly overpredicts the actual amounts of countries' trade forced trade economists to seek modifications of the HOV model. Interestingly some of those who have made the greatest contributions in modifying the Heckscher-Ohlin model tend to downplay the importance of their modifications. Davis and Weinstein (2001a, p. 1445), for example, remark: "Yet it is startling that such a plausible and simple set of departures from the conventional model allows us to so accurately match the international data."[1] The number of modifications may be small, but by dropping such assumptions as identical technologies across countries and factor price equalization, the extent to which the new models differ from the workhorse HOV model in an analytical sense is very significant.

The main goal of those who have modified the basic HOV model has been to account for the large differences between the volume of countries' factor trade predicted by the HOV model and the actual trade measured by the investigators. In other words, it has been to account for what Trefler has called "the mystery of the missing trade." In undertaking this task, researchers have not relied on well-formulated alternative trade theories but rather on general notions of what might account for the differences between predicted and measured trade. For example, the introduction of neutral technology differences among countries is not motivated by a well-formulated theory that predicts such differences but rather by the fairly obvious existence of a technology gap between richer and poorer countries. The assumption of neutral technology differences among countries is a parsimonious means of dealing with the issue of technology gaps. Thus recent empirical trade economists have backed away somewhat from the earlier dictum that empirical investigations must be tightly tied to trade theory.

The key conclusion emerging from the various modifications to the basic HOV model is that relative factor endowments do matter for understanding patterns of factor trade embodied in goods. Differences in technology among countries generally play a more important role in accounting for differences in countries' trading patterns in productive factors, but once these plus the existence of endowment differences sufficiently large to leave countries in different cones of diversification are taken into consideration, there is significant statistical support for the Heckscher-Ohlin proposition.

As noted in the preceding chapter, efforts to identify Stolper-Samuelson and Rybcyznski effects empirically have been less extensive and generally less rigorous than those aimed at testing the Heckscher-Ohlin proposition. However, investigators have, for example, found strong evidence that in addition to labor-saving technical progress and changes in relative factor supplies, increased import competition has contributed to increased wage inequality in the United States. There has been less success in directly identifying Rybczynski effects.

6.2 Further Broadening of the Analytical Framework

6.2.1 Endogenous Technical Change

Although efforts to narrow the empirical gap between the predicted and measured factor content of trade by modifying the HOV model have been successful, the theoretical

foundations for introducing these modifications have not been well explained. For example, given the importance of differences in levels of technology across countries in accounting for the differences between the predicted and measured factor content of trade, trade economists need to broaden their theoretical framework to present causal explanations of why these nonrandom country differences in technology among countries arise and persist. Of course, endogenous growth theory is an enormously complex field, and trade economists can hardly be expected to do much more than adapt the contributions from other fields to their particular models. There are, however, endogenous growth models that focus on some of the same key causal variables as trade economists do, in particular, on relative factor endowments. This presents opportunities for linking differences in relative factor endowments and technological progress.

Grossman and Helpman (1991, ch. 5), for example, formulate a model in which the supply of skilled labor and the rate of technological changes are endogenous. Skilled workers differ from unskilled workers in that they have devoted a certain amount of time to additional schooling. These two productive factors produce differentiated manufactured goods or work in an industrial research lab that produces innovations in the form of new products or quality improvements. Skilled labor is also assumed to be employed relatively more intensively in the industrial research lab than elsewhere in the economy. To illustrate the interrelationships in their model, the authors suppose that there is an increase in the productiveness of time spent in school as the consequence, for example, of an increase in public investment in education capital. Entrants into the labor force who previously were indifferent to becoming unskilled or

skilled workers suddenly prefer to become skilled workers, and the supply of skilled labor increases relative to unskilled workers. This lowers the relative wages of skilled workers, thereby promoting an expansion of research activities and a rise in the rate of technological progress. Thus, to quote Grossman and Helpman: "A country that has a greater steady-state supply of the factor most essential for industrial research will allocate more resources to R&D in equilibrium and will experience faster innovation and growth as a consequence" (Grossman and Helpman 1991, p. 140).

In explicitly introducing trade among countries, Grossman and Helpman also present a variety of endogenous growth models that highlight the interrelationships between trading patterns, innovation, and growth. They show, for example, how technological knowledge can be exchanged through trading transactions and results in increased competition among world innovations that leads to a reduction in the duplication of efforts and an increase in the aggregate productivity of R&D activities. But they also explore alternative trading scenarios in which economic integration can reduce the growth rate of a country. Thus, given the significant role that recent empirical work has shown differences in technology levels to play in accounting for differences in trade patterns, it is important that trade economists integrate models of endogenous technical change into their basic factor proportions framework.

6.2.2 *Foreign Direct Investment*

A key insight of Ohlin that forms the modern basis for analyzing international trade with more than two factors and two goods, namely the HOV model, is that trade in goods

can usefully be viewed as trade in the services of productive factors. With this perspective the question naturally arises as to what economic forces determine whether factors of production move across borders embodied in goods and services or whether the factors themselves move across borders.

Investigations of this question with regard to the international flow of goods versus capital have a long and rich history in international economics. Mundell (1957) is a well-known study that utilizes the standard two-country, two-good, two-factor (capital and labor) HOS model to study this question. Beginning with a free-trade equilibrium such as described in chapter 2 where factor prices are equalized, Mundell drops the assumption of capital immobility between the two countries and also assumes that the relatively labor-abundant country imposes a specific import duty on its imports of the capital-intensive good from the capital-abundant country. Since this increases the return to capital in the labor-abundant country, capital will move from the capital-abundant country to the labor-abundant country. As Mundell demonstrates, this continues until the capital/labor endowment ratios are the same in both countries and there no longer is any trade between the countries. Product prices and factor prices return to the same levels that existed in the initial free trade situation with capital immobility. Thus the flow of capital between the two countries serves as a substitute for trade in goods.

Later authors (e.g., Helpman 1984b; Markusen 1984) focus their modeling efforts on foreign direct investment undertaken by multinational firms. Helpman (1984b) analyzes the emergence of vertical multinationals in which a firm's headquarters are located in one country and its production activities in another. There are two sectors and two factors

(skilled and unskilled labor) in his model.[2] One sector produces a differentiated product with a constant elasticity-of-substitution production function, while the second sector produces a homogeneous product. A variety of the differentiated good requires a certain amount of skilled labor to cover its fixed headquarters costs and utilizes only unskilled labor in its variable costs. Headquarter services do not necessarily have to be supplied in the same country in which the variable-cost activities are performed. However, production of a unit of the homogeneous good requires both skilled and unskilled labor that cannot be divided between countries. The differentiated product is assumed to be more skilled labor-intensive in its production than the homogeneous good.

First, suppose that the endowments of skilled and unskilled labor for two countries lie in the same cone of factor diversification so that the returns to skilled and unskilled labor are the same in both countries. With the same factor prices and technologies in both countries, there is no incentive for producers of the differentiated good to split their headquarters and variable-cost activities between the two countries. However, now suppose that factor endowments differ sufficiently between the two countries that factor prices would not be equalized if headquarters and variable-cost activities for the differentiated good had to take place in the same country. In these circumstances the relative price of skilled labor will be lower in the capital abundant country. Consequently producers of the differentiated good in the capital-abundant country have an incentive to shift their variable-cost activities for this good to the labor-abundant country, and producers of the differentiated good in the labor-abundant country have an incentive to shift their headquarter activities to the capital-abundant country. The

final outcome will be the equalization of factor prices with the capital-abundant country specializing on headquarters services for the differentiated product. Thus the vertical multinationals model associates differences in direct foreign investment among countries with differences in relative factor endowments among the countries.

Horizontal multinationals are firms that have facilities for producing the same product in multiple countries and selling their outputs in local markets. The intuition behind formal horizontal multinational models such as Markusen (1984) and the various extensions elaborated in Markusen (2002) are quite straightforward.[3] The existence of firm-level scale economies are the driving force for direct investment. These arise from such activities as R&D expenditures undertaken in a headquarters location. The improved product or production processes from these expenditures often involve a "public goods" aspect such that the innovation can be used in additional plants without reducing the marginal product of the innovation in existing plants. Thus two-plant firms, for example, have fixed costs that are less than double those of a single-plant firm. These relationships imply that given the existence of transportation costs for shipping the product between countries together with plant-level and firm-level scale economies, horizontal foreign investment will take place between similar countries. However, in contrast to the vertical foreign investment model, the existence of very different relative factor endowments can have the effect of discouraging horizontal multinationals.

6.2.3 *Heterogeneous Firms*

Heterogeneity among firms within an industry in the sense that they produce different varieties of a differentiated product using the same technology has already been discussed in section 2.4. However, another type of firm heterogeneity that has received considerable attention in recent years is variations in productivity levels among firms within an industry. Interest in this subject has been stimulated in part by the recent willingness of governments to make firm-level data that they collect available to researchers.[4] Trade economists have been particularly interested in determining whether there are significant differences in the economic characteristics of firms engaged in foreign activities such as exporting and foreign direct investment from firms that just serve the domestic market.

The findings from these studies have been somewhat surprising and have stimulated modeling efforts to explain them. Only a relatively small proportion of firms in an industry export, and these tend to be both more productive and larger than those that only serve the domestic market. There are large sunk costs in exporting in addition to those related to undertaking production domestically.[5] Among researchers who developed theoretical models to explain these relationships are Melitz (2003), Bernard et al. (2003), and Helpman, Melitz, and Yeaple (2004). The following discussion summarizes Helpman's (2006) description of Melitz's (2003) model.[6]

Assume a continuum of firms in a differentiated-product sector that each face the same elasticity of demand for varieties sold in the domestic market. The firms differ in terms of their labor productivity (the only factor of production), however. Firms also only discover their productivity level, which is randomly drawn from a given productivity-distribution function, after they have entered the market by incurring a given fixed cost and face given variable production costs per unit of output. For firms with low

productivity levels, the profit-maximizing price to charge and the resulting sales will not be sufficient of cover variable production costs plus their fixed costs. They will choose not to produce. But firms lucky enough to have drawn a high productivity level will undertake production and earn a net profit. Moreover the larger a firm's productivity level, the greater will be its sales and size.

Now assume that there is also an export market in which firms can sell their varieties and that the elasticity of demand (though not necessarily the level of demand) is the same in this market as in the domestic market. Trading costs to export markets (e.g., transportation costs and tariffs) are assumed to take the melting iceberg form; that is, more than one unit of every variety must be shipped for one unit to arrive. In addition there are fixed costs in each export market.

The productivity level of some firms that can profitably produce for the domestic market will be high enough so that they can also cover the fixed and variable costs involved in exporting. Thus these firms can increase their overall profits by exporting. Consequently there are three category of firms: (1) those with the lowest levels of productivity that close down because they lose money from domestic sales as well as exporting, (2) those with intermediate levels of productivity that can make a profit by just selling in the domestic market, and (3) those with the highest productivity that sell their varieties both domestically and in export markets. The latter group will also be the largest because of their greater sales.

As Helpman points out, the description of Melitz's model presented thus far is a static version, However, Melitz specifies his model in dynamic terms in which there is a constant

probability of the death of every firm, regardless of its productivity level. Steady state equilibrium is characterized by a constant number of firms where the inflow of new firms equals the outflow brought about by the death of firms. This framework permits Melitz to shed meaningful insights on the entrance, exit, and turnover of firms in response to exogenous economic shocks.

Consider, for example, the effects of a proportional reduction of tariffs in all countries. This reduces the costs of exporting and thus both lowers the productivity level at which firms can profitably export and raises the profits of firms already exporting. However, the resulting increase in the demand of labor raises labor costs and forces some of the least productive firms that are only selling in the domestic market to exit the industry.[7] Overall turnover in the sector increases but also so does average productivity.

Helpman, Melitz, and Yeaple (2004) extend the model to analyze horizontal foreign direct investment. In this situation there is a "proximity-concentration" trade-off (see Brainard 1997) between saving on trading costs but incurring increased fixed costs. They show that only the most productive firms will tend to establish overseas production facilities rather than export from domestic production facilities.

A deficiency of the heterogeneous firms models is that a firm's productivity level is just a matter of chance. However, just as researchers discovered that whether a firm exports or just serves the domestic market is not a random event, so too may it be that a firm's productivity level is not just a matter of chance. It may be, for example, that the initially most productive firms are formed and managed by individuals with exceptional managerial and technical skills

and knowledge relative to their competitors. Research is needed to better understand both the demand and relative supply forces determining the performance capabilities of different firms.

6.2.4 *Other Research Areas*

Further theoretical and empirical research on the three topics described above seem particularly important for improving our understanding of the economic influences affecting factor and commodity patterns of international trade, but there are a number of other economic conditions whose relative importance in determining the nature of trade are in need of clarification and further testing. The role of scale economies in shaping trading patterns is a good example. As pointed out in chapter 1, Ohlin sometimes seemed to treat these as being of equal importance in shaping trading patterns as relative factor endowments. Antweller and Trefler (2002) and Davis and Weinstein (2003), whose research on this subject was reviewed in chapter 4, find empirical support for the importance of scale economies in influencing trade and production patterns, but they do not integrate these findings into their general modified HOV models developed in Trefler (1995) and Davis and Weinstein (2001a). Thus we do not have a good understanding of the importance of scale economies compared to other factors influencing production and trading patterns such as relative factor endowments, cross-country differences in technology, and differences in preferences. Research devoted to this goal is very much needed.

A better understanding of the role of demand differences among countries in determining trading patterns and why these differences arise are other subjects deserving of further

research. The three best-known, multi-country and multi-factor empirical studies that modify the HO model, namely those by BLS (1987), Trefler (1995), and Davis and Weinstein (2001a), utilize quite different methodologies in analyzing the demand side of trade. As these authors note, all these approaches have their drawbacks, and the authors all encourage further research in this area.

The existence of differentiated products needs to be integrated much more tightly into the modified HO models. As trade theorists such as Krugman (1979, 1980) emphasize, there are usually a number of varieties of a particular manufactured good that are similar but not identical to each other in the many ways described by Edward Chamberlin (1933) in his pioneering book on monopolistic competition. It is interesting that Ohlin (1933, p. 96) pointed out that most manufactured goods were differentiated and that this accounted in part for intra-industry trade.

As pointed out in chapter 4, Schott (2003a,b) finds heterogeneity of output within industries across countries and strong support for the notion that a country's product mix varies with relative factor endowments. Factor-proportions specialization taking place *within* products means that countries may produce in different cones of factor diversification even though they produce goods that are classified as being in the same industry. Additional research in this area seems promising in further developing an alternative to the HOV model that is supported by empirical analysis but is not overly complex theoretically. Integrating the impact of traded intermediates into international models without factor price equalization is still another topic on which both additional theoretical and empirical research is needed, as is the effects of nontraded goods and services in influencing trading patterns.

There have been a number of studies with the objective of determining the restrictiveness of world trade barriers and the benefits from moving toward free trade.[8] However, there have been fewer analyses aimed at assessing the impact of actual trade barriers on the HO theorems. As discussed in section 4.5.2, Staiger et al. (1987) found that their simulation of free trade conditions in the world did not change the ordering of net exports by the United States and Japan to the rest of the world of ten productive factors from the order that existed under actual trading conditions with trade barriers. In view of both much better estimates of existing trade barriers and wider availability of input–output relationships for countries, it would be very interesting to replicate and extend their calculations.

The relative relationship among factor prices is another HO relationship that theoretically can be significantly affected by trade barriers. Section 5.1.2 reviews part of the extensive debate on the effects of changes in trade patterns and changes in relative wages. The most widely accepted view seems to be that changes in technology have contributed most to the widening gap between the wages of skilled and unskilled labor but that changes in trade have also been a contributing factor. There is a need, however, for rigorous studies detailing just what part of the trade changes are attributable to changes in protection and, in turn, what the effect of these changes have been on relative wages.

There are a variety of other topics on which further research can improve and expand our understanding of the determinants of factor and commodity patterns of international trade. But the ones discussed in this and previous chapters seem to be among the most promising. Research on these topics needs to involve both theoretical and empirical analyses. As Leamer and Levinsohn (1995) emphasize

in their *Handbook* article, trade theorists have traditionally been little influenced by the work of trade empiricists. The dominance on the theoretical side for many years of the unmodified HOV model, despite growing empirical evidence of its inadequacy as a stand-alone explanation of trading patterns, is a manifestation of this separation of theoretical and empirical analyses in the trade field. Fortunately, much of the recent research reviewed in this monograph indicates that this separation of theory and empirics is changing.