PUTTING THE “NEW” INTO NEW TRADE THEORY:
PAUL KRUGMAN’S NOBEL MEMORIAL PRIZE IN ECONOMICS

J. Peter Neary

Number 423
March 2009

Manor Road Building, Oxford OX1 3UQ
PUTTING THE “NEW” INTO NEW TRADE THEOREY: PAUL KRUGMAN’S NOBEL MEMORIAL PRIZE IN ECONOMICS *,†

J. Peter Neary

University of Oxford and CEPR

March 17, 2009

Abstract

This paper reviews the scientific contributions of Paul Krugman to the study of international trade, on the occasion of his receipt of the 2008 Nobel Memorial Prize in Economics. A simplified exposition is presented of some of his principal findings, including: the effects of trade on firm scale and product diversity in a general model of monopolistic competition; the integration of monopolistic competition with factor endowments theory; the implications of transport costs, including home-market effects and the possibility of agglomeration in models of economic geography; and the positive and normative consequences of oligopolistic trade.

Keywords: economic geography; imperfect competition; intra-industry trade; monopolistic competition; oligopoly and trade; product differentiation.

JEL Classification: F12

*Address for Correspondence: Department of Economics, University of Oxford, Manor Road Building, Oxford, OX1 3UQ, U.K.; tel.: +44 (0)1865-271085; e-mail: peter.neary@economics.ox.ac.uk.
1 Introduction

The award of the 2008 Nobel Memorial Prize in Economics to Paul Krugman is a fitting recognition of his achievement in reorienting the field of international trade. Throughout the 1950s, 1960s and 1970s, most research in trade pursued research programmes implicitly defined by the work of two earlier winners of the Prize, James Meade and Bertil Ohlin. They set a rich agenda, in normative and positive analysis respectively, which I will discuss further below. However, it exhibited one key feature: it was carried out almost exclusively under the assumption of perfect competition. Starting in the late 1970s, the agenda was rewritten by Krugman and others. By the early 1980s much of the exciting new research in the field was being carried out in models of imperfect competition, and the literature was rich with references to “new trade theory”. Nearly thirty years later, the novelty has worn off but the change in orientation has remained. The fruits of applying models from industrial organization to positive and normative issues in open economies are an integral part of modern trade theory.

In this paper I attempt to summarize and put in context Paul Krugman’s key scientific contributions, and give some simple diagrammatic expositions of his main results. For the most part I concentrate on the areas for which he was awarded the Nobel Memorial Prize: in the words of the Prize Committee, “for his analysis of trade patterns and location of economic activity.”1 Section 2 outlines his model of intra-industry trade, first presented in a simple but fully-specified general-equilibrium framework in [11]. Section 3 sketches how this model was integrated by Krugman and others with the Heckscher-Ohlin model of trade based on international differences in factor endowments. Section 4 turns to his treatment of transport costs, in generating both home-market effects on trade and a propensity for agglomeration of economic activity, the central prediction of his model of economic geography (see [24]). Section 5 reviews his contributions to the theory of trade under oligopoly, both its positive and normative aspects, while Section 6 discusses briefly some of his contributions to other topics in trade and to other fields of economics, as well

1For more detailed overviews, see Helpman and Krugman [1], [2], Fujita, Krugman and Venables [7], and Grossman and Rogoff (1994), especially the chapters by Brander and Krugman himself. My exposition draws in part on Neary (2001 and 2003).
as his more recent role of public intellectual.

2 A New Theory of Trade


2.1 Precursors: Empirical Anomalies and Theoretical Tools

Ever since the publication of David Ricardo’s Principles of Political Economy and Taxation in 1817, a cornerstone of international trade theory has been the theory of comparative advantage: when trade is free, countries will and (from their own perspectives) should specialize in the production of those goods which they produce relatively more efficiently. Ricardo himself took international productivity differences as given, whereas in the early twentieth century Eli Heckscher and Bertil Ohlin highlighted international differences in factor endowments as a major source of comparative advantage. This view, which received its definitive formalization in the work of Samuelson (1949, 1953) and Jones (1965), came to dominate the field from the mid twentieth century onwards. Indeed, it continues to play an important role in discussions of trade, as can be seen from the opening chapters of any international economics textbook (including Krugman and Obstfeld [6]).

However, it became increasingly clear in the 1970s that not all features of international trade were well explained by factor endowments or any other comparative-advantage-based theory. Such theories predicted that trade should involve the exchange of different goods and should be greater the more countries differed in their relative production possibilities. By contrast, most of the enormous growth in trade in the decades after the Second World War came in relatively similar goods (manufactures) between relatively similar countries (developed ones). Moreover, this feature was robust to the level of
disaggregation: empirical work by Grubel and Lloyd (1975) and others showed that, no matter how finely industries were defined, a high proportion of trade took place within industries rather than between them.

The increased recognition of such *intra-industry* as opposed to *inter-industry* trade was not just of interest from the perspective of positive science. Comparative advantage theories predicted that adjusting to trade liberalization would be disruptive, as the process of specialization encouraged factors of production to move out of sectors in which each country had a comparative disadvantage and into its expanding sectors. Though classroom expositions of this process often suggested a smooth movement along a production possibility frontier, it was clear that factor specificities and factor-price rigidities could make for painful adjustment in the short run. However, while this was undoubtedly true of many real-world episodes of adjustment to trade liberalization, it seemed inadequate as a description of the major such episode in the immediate post-war era: the integration of the relative similar economies of Western Europe to form the European Economic Community (the forerunner of the European Union). Studies by Balassa (1967) and others showed that this process proceeded with surprisingly little costs of adjustment. Once again, specialization appeared to be intra-industry rather than inter-industry.

Given these empirical findings, the stage was set for a new theoretical direction. However, a prerequisite for this was an analytic framework which would allow for both increasing returns to scale and differentiated products. Edward Chamberlin’s model of monopolistic competition had in principle provided such a framework in 1933. But its reliance on verbal and geometric reasoning made it hard to embed in small-scale models of general equilibrium; so much so that a thirty-four-years-on celebration of Chamberlin’s contribution bemoaned its relative lack of influence on economics in general, and in particular the fact that “the theory of monopolistic competition has had virtually no impact on the theory of international trade” (Johnson, 1967, p. 203). What was needed was a tractable specification of preferences and costs which would make it possible to apply Chamberlin’s insights at the global level, and the technical tools for doing just that were first assembled by Dixit and Stiglitz (1977). On the preferences side, they assumed
that utility \( u \) was a symmetric and additively separable function of the consumption levels \( x_i \) of a large but variable number \( N \) of distinct goods:

\[
\begin{align*}
  u &= \sum_{i=1}^{N} v(x_i) \\
  v'(x_i) &> 0, \quad v''(x_i) < 0
\end{align*}
\] (1)

On the cost side, Dixit and Stiglitz made it intellectually respectable to replace Chamberlin’s intricate U-shaped average cost curves with a simple specification involving a fixed cost and a constant variable cost, both common to all firms, so total costs for a typical firm \( i \) are:

\[
C_i = (f + ay_i) w
\]

(2)

(where \( y_i \) is firm output and \( w \) is the wage rate). Dixit and Stiglitz themselves applied their framework to classic issues in industrial organization. But it contained all the ingredients needed to explain intra-industry trade: the utility function generates a demand for differentiated products, and under a mild restriction implies that consumers value diversity;\(^3\) the cost function implies that returns to scale are increasing, as higher output \( y_i \) spreads the fixed costs \( fw \) more thinly; and the assumption that each firm produces a single product means that industry adjustment alters the number of varieties \( N \) available to consumers. What was needed was a master-chef to take these ingredients and use them to construct a new theory of trade.

\(^2\)A similar formulation was adopted by Spence (1977), and other ways of specifying tastes for differentiated goods drawing on Hotelling, Lancaster and Salop among others were explored around the same time and continue to be used in many fields of economics. However, for reasons discussed in more detail in Neary (2003), the Dixit-Stiglitz specification was the one that lent itself most easily to embedding in general equilibrium.

\(^3\)In a symmetric equilibrium, utility (1) becomes \( u = Nv(x) \) and the budget constraint is \( Npx = I \). Holding income \( I \) and prices \( p \) fixed, the effect of an increase in diversity on utility (using “hats” to denote proportional changes) is \( \hat{u} = \hat{N} + \varepsilon_x \hat{x} = (1 - \varepsilon_x) \hat{N} \), where \( \varepsilon_x \equiv \frac{dx}{dx} \) is the elasticity of the sub-utility function. Hence, the consumer exhibits a preference for diversity provided there are diminishing returns to the consumption of each individual variety; i.e., provided \( \varepsilon_x \) is less than one. In the CES case, to be discussed further below, this condition is always met: \( v(x) = x^\theta \) implies that \( \varepsilon_x = \theta \), so \( \hat{u} = (1 - \theta) \hat{N} \), where \( \theta < 1 \).


2.2 Increasing Returns and Product Differentiation

Thus was the scene set for Krugman’s 1979 paper, which introduced probably the simplest possible fully-specified general equilibrium model in which intra-industry trade could be rigorously demonstrated. Suppose that there are \( k \) identical countries with \( n \) goods produced per country in equilibrium, so the total number of varieties available to consumers when trade is free is \( N = kn \). In each country there are \( L \) households, each of whom supplies a unit of labour (the only factor of production) and maximizes the utility function (1) facing given prices \( p_i \) for each good. This leads to individual demand curves in implicit form: \( v'(x_i) = \lambda p_i \). Here \( \lambda \) is the individual household’s marginal utility of income, which depends on their income and on the prices of other goods. However, provided the number of goods is large, each firm rationally takes \( \lambda \) as fixed: echoing Chamberlin, the demand curve a firm perceives for its own product depends on its price only. Of course, the total quantity demanded comes from all households in all countries, given by the market-clearing condition for the output of each firm:

\[
\text{Goods-Market Equilibrium: } y = kLx
\]

But with identical consumers, the perceived elasticity of demand facing firm \( i \), \( \frac{p_i \partial y_i}{y_i \partial p_i} \), depends only on the consumption of an individual household \( x_i \). Differentiating the demand function, this elasticity can be written as \( \sigma(x_i) \equiv -\frac{v'(x_i)}{v''(x_i)x_i} \), and, following Krugman, we assume that it is decreasing in consumption: \( \frac{d\sigma(x_i)}{dx_i} < 0 \). Higher consumption, or, equivalently, a lower price, makes households less responsive to price.

After these preliminaries, the equilibrium of the model can be illustrated in Figure 1. Each firm maximizes profits by setting its marginal revenue given \( \lambda \) equal to its marginal cost \( aw \). Writing the first-order condition in terms of the perceived elasticity of demand

---

4Following Krugman, the key geometric innovation is to link the two equilibrium conditions for the firm via the goods-market-clearing condition (3). Krugman illustrated this in \( \{p/w, x\} \) space rather than in \( \{p/w, y\} \) space as here. The latter yields a simpler link with the full-employment condition in the lower panel, and also facilitates comparison with the standard partial-equilibrium diagram due to Chamberlin, which is useful for understanding the incentives faced by individual firms: see Neary (2001).
Figure 1: General Monopolistically Competitive Equilibrium

\[ \frac{\partial \xi}{\partial \mathbf{x}} \text{ and dropping firm subscripts because of the symmetry assumption, gives:} \]

\[ \text{Profit Maximization (MR=MC): } \frac{p}{w} = \frac{\sigma(x)}{\sigma(x) - 1} a \]  

(4)

Recalling that \( \sigma(x) \) is decreasing in consumption, this implies that higher levels of
consumption allow firms to charge higher prices. Hence (4) is represented, for given values of \( k \) and \( L \), by the upward-sloping locus \( MR = MC \) in the upper panel of Figure 1. The other equilibrium condition in each sector is that profits are driven to zero by free entry and exit of firms, so price must equal average cost:

\[
\text{Free Entry (} p = AC) : \quad \frac{p}{w} = \frac{f}{y} + a
\]  

This implies the downward-sloping relationship between output and the price-wage ratio illustrated in the upper panel of Figure 1. Finally, the model is completed by the requirement that each country’s labour supply \( L \) must equal the demand from all domestic firms:

\[
\text{Labour-Market Equilibrium (} LME) : \quad L = n (f + a)\]

This equation implies a negative relationship between equilibrium firm size \( y \) and the number of firms \( n \), as illustrated in the lower panel of Figure 1. The full model then consists of the four equations (3), (4), (5), and (6), in four unknowns: \( p/w \), \( x \), \( y \) and \( n \). The upper panel illustrates the determination of the equilibrium values of \( p/w \) and \( y \), with \( x \) determined by (3) and \( n \) determined residually in the lower panel.

Now suppose that the world economy expands: the simplest example of this is an increase in \( k \), representing the addition of more identical countries. Inspecting the equations, the only direct effect of this is to disrupt goods-market equilibrium: world demand rises, so every firm must increase output by an equal proportionate amount if firms are to continue maximizing profits at the same prices. Thus the \( MR = MC \) curve shifts to the right as shown in Figure 1. But if prices did not change, as at point \( E \), firms would now be earning positive profits. Hence, prices must fall and the new equilibrium must be at point \( B \). Firms move down their average cost curves, producing more at lower costs, with the benefits passed on to consumers in the form of lower prices.\(^5\)

\(^5\)Totally differentiating (4) gives \( \hat{p} = \hat{\varepsilon}_\mu \hat{x} \), where \( \varepsilon_\mu \equiv -\frac{1}{\sigma(x)-1} \frac{\sigma(x)}{x} \frac{dx}{dx} > 0 \) is the elasticity of the mark-up with respect to consumption. Totally differentiating (5) gives \( \hat{p} = -\theta_f \hat{y} \), where \( \theta_f \equiv \frac{f}{f+ay} \) is the share of fixed costs in total costs. Finally, (3) implies that \( \hat{x} = \hat{y} - \hat{k} \). Solving gives: \( \hat{y} = \frac{\varepsilon_\mu}{\varepsilon_\mu + \theta_f} \hat{k} > 0 \), \( \hat{x} = -\frac{\theta_f}{\varepsilon_\mu + \theta_f} \hat{k} < 0 \), and \( \hat{p} = -\frac{\sigma(\hat{x})}{\sigma(\hat{x})-1} \frac{\hat{x}}{\hat{x}} \frac{dx}{dx} \hat{k} < 0 \), so firm output rises but the per capita consumption and price of each variety fall.
resource constraint, the lower panel of the figure shows that an increase in firm output can only come about if the number of domestic firms falls. However, the total number of varieties produced in the world rises, so consumers benefit from an increase in diversity as well as a fall in prices.6 Finally, because consumers demand all varieties, there is an increase in trade, all of which is intra-industry. Thus the model is consistent with the empirical evidence on intra-industry trade, and also shows that it leads to gains from trade: just as in Ricardo, there are precise predictions about both positive and normative aspects of trade, with the difference that the countries are identical so there is no role for comparative advantage.

2.3 Models of Infinite Variety

Krugman’s landmark 1979 paper was thus the first to present a coherent general-equilibrium analysis of the kind of trade that, in the words of the Nobel Memorial Prize Committee: “enables specialization and large-scale production, which results in lower prices and a greater diversity of commodities.” It was arguably also the last, since most of the subsequent literature has concentrated on a special case in which only the final prediction, a greater diversity of goods consumed, remains true. The problem which soon emerged is that, though the model is deceptively simple, for most purposes it is not nearly simple enough. In particular, the general Dixit-Stiglitz specification (1) does not yield closed-form solutions. As a result, Krugman himself (in [12] and many later papers) and most subsequent writers on monopolistic competition and trade have worked with a special case where the sub-utility function takes a constant-elasticity-of-substitution form:

\[
u = \left( \sum_{i=1}^{N} x_i^\theta \right)^\frac{1}{\theta} \quad 0 < \theta < 1, \quad \theta = \frac{\sigma - 1}{\sigma} \quad (7)\]

where the parameter \(\theta\) is a measure of substitutability, closely related to the elasticity of substitution \(\sigma\).

---

6The proportional change in the number of domestic firms (which is also the change in the number of domestically-produced varieties) is: \(\hat{n} = -\frac{\epsilon\mu (1-\theta) f}{\epsilon\mu + \theta f} \hat{k} < 0\); however, the change in the total number of active firms in the world is: \(\hat{N} = \hat{k} + \hat{n} = \frac{(1+\sigma)e^{\theta}}{\epsilon\mu + \theta f} \hat{k} > 0\).
The implications of this for the effects of an expansion in the world economy are easily seen. With $\sigma$ now a constant, the $MR = MC$ locus in Figure 1 is horizontal, and is unaffected by changes in the number of countries. Profit maximization fixes the price-cost margin as a function of the elasticity of substitution only, $\frac{p}{aw} = \frac{\sigma}{\sigma-1}$, so price is a constant mark-up over marginal cost: $p = \frac{\sigma}{\sigma-1}aw$. Free entry in turn pins down the size of each firm as a function only of the elasticity of substitution and the ratio of fixed to variable costs: from (4) and (5), $y = (\sigma - 1) \frac{L}{a}$. When the number of countries increases, the initial equilibrium in Figure 1 is now unaffected: there is no change at all in price-cost margins, scale of production, or firm numbers per country; only the destination of home output changes, with a larger share exported in exchange for more imports, leading to a greater range of varieties, and thus higher utility, for domestic consumers. The fact that even $n$ does not change is an artefact of the particular shock assumed; for example, a rise in the population of each country, $L$, shifts upwards the $LME$ curve in the lower panel of Figure 1, so inducing entry of more firms. However, it remains true that most shocks leave firm scale unaffected, with any necessary adjustment in domestic production coming via changes in the number of firms: the CES special case de-emphasizes the implications of increasing returns and concentrates attention on the range of varieties available to consumers. Notwithstanding this limitation, the CES specification (7) has proved enormously convenient, allowing a rich exploration of a wide range of topics, including important recent work on choice of organizational form by Antras (2003) and on heterogeneous firms by Melitz (2003). It also formed the basis for Krugman’s own work on the role of factor endowments and transport costs, to be considered in the next two sections.

---

7 These results can also be seen from the explicit solutions in footnotes 6 and 7, with $\varepsilon_\mu$, the elasticity of the mark-up with respect to consumption, set equal to zero.

8 One way to avoid the prediction that firm scale is fixed by tastes and technology is to assume that fixed and variable costs require different factors, so relative factor prices affect optimal firm size. This was pioneered by Lawrence and Spiller (1983) and has proved useful in the literature on economic geography: see Section 4.2 below.
3 Integrating Monopolistic Competition with Factor Endowments

Readers of Krugman’s landmark 1979 paper might have assumed that he came to bury comparative advantage as an explanation of trade patterns: he certainly did not praise it. But revolutions, especially intellectual ones, are more successful when they encompass the old view rather than totally supplanting it. So an important step in consolidating the monopolistically competitive approach to international trade was to show how it could be embedded in a Heckscher-Ohlin competitive trade model, a task which Krugman undertook in his 1981 *Journal of Political Economy* paper “Intraindustry Specialization and the Gains from Trade” [13]. Krugman was not the only economist to present a synthesis of this kind, nor was his model the most general. Similar work was carried out around the same time by Avinash Dixit and Victor Norman, Kelvin Lancaster, Elhanan Helpman, and Bill Ethier. But Krugman’s presentation in [13] was the most elegant and forceful, and his masterly synthesis with Helpman in their 1985 book [1] provided the definitive statement of the core of new trade theory. It did so by building on the core of old trade theory, which consisted of a suite of four theorems: the factor-price equalization theorem, showing how goods trade integrates factor markets internationally; the Rybczynski and Heckscher-Ohlin theorems, linking factor endowments to production and trade patterns respectively; and the Stolper-Samuelson theorem, relating changes in goods prices to changes in real factor rewards.

---

9. Norman set out a concise integration of CES preferences and the two-sector Heckscher-Ohlin model in an unpublished 1976 mimeo which formed the basis of the concluding sub-section of his path-breaking book with Dixit (1980). Lancaster (1980) and Helpman (1981) integrated the two-sector model with preference structures drawing on Hotelling and Vickrey, respectively, which assumed a population of symmetrically distributed heterogeneous consumers, each with an ideal variety which can be represented by a point on a line or a circle, and who incur psychic costs the further is the variety they actually consume from their ideal. This was in many respects a theoretically more satisfactory way of modelling aggregate demand for differentiated products than the Dixit-Stiglitz approach, but its complexity made it more difficult to apply to other problems. Finally, Ethier (1982a) presented an alternative synthesis of Heckscher-Ohlin and monopolistic competition, which reinterpreted the Dixit-Stiglitz CES specification of preferences as a production function with a large and variable number of differentiated inputs used to produce a single final consumption good. This allowed a formalization of Adam Smith’s view that the division of labour is limited by the extent of market, and in turn directly influenced the development of endogenous growth theory by Romer (1987).
### 3.1 Replicating the Heckscher-Ohlin Equilibrium

International equalization of the prices of immobile factors is hardly a persuasive description of reality, but it is a crucial benchmark in trade theory, since it makes it possible to decompose the problem of solving a full world general equilibrium. This turned out to be even more true when monopolistic competition was embedded in the standard two-factor, two-sector, two-country model. The first step in doing this, already taken by Dixit and Stiglitz, was to model consumers as choosing between two goods, one a homogeneous good ("Food") produced by a perfectly competitive sector, and the other a composite aggregate ("Manufactures") made up of the outputs of $N$ monopolistically competitive firms.\(^{10}\) The most tractable way to do this was to assume a nested utility function, with the upper tier Cobb-Douglas and the manufacturing sub-utility function in the lower tier taking a CES form just as in the previous section:

$$U = X_{F}^{1-\mu} u_{M}^{\mu} \quad u_{M} = \left( \sum_{i=1}^{N} x_{i}^{0} \right)^{\frac{1}{\theta}}$$  \(8\)

where $\mu$ is the fixed budget share of manufactures. This implies demand behaviour essentially identical to that already outlined, and so, for given costs, the equilibrium of the monopolistically competitive sector was unchanged.

How are costs determined? The beauty of the factor-price equalization theorem in perfect competition is that, with equal numbers of goods and factors, goods prices and technology alone determine factor prices and hence production costs. It turns out that, with CES preferences, the same holds under monopolistic competition. Consider the simplest case, where each country has exogenous endowments of two factors, labelled capital $K$ and labour $L$, whose prices $r$ and $w$ are determined on competitive factor markets. In the perfectly competitive food sector, free entry and exit ensures that the output price (which we can set equal to one for convenience) equals average cost $c_{F}$, and

\(^{10}\)In [13] Krugman assumed that both industries were monopolistically competitive, but considered only symmetric differences in factor endowments: i.e., he focused on points along the downward-sloping diagonal $AB$ (not drawn) in Figure 2. By contrast, Helpman and Krugman in [1] allowed for arbitrary differences in factor endowments, but for the most part focused on the case where only one industry is monopolistically competitive.
profit-maximizing firms choose techniques to minimize average cost given the prices of the factors of production:

\[ 1 = c^F(w, r) = a_{LF}w + a_{KF}r \] (9)

where \( a_{LF} \) is the cost-minimizing labour requirement per unit output in the food sector, etc. As for manufacturing, it is simplest to assume that fixed and variable costs use the same factor proportions, so total costs for firm \( i \) are given not by (2) but by:

\[ C_i = (f + a_{yi}) c^M(w, r) \quad c^M(w, r) = a_{LM}w + a_{KM}r \] (10)

Now the pricing condition for each monopolistically competitive firm is as before, except that marginal cost depends on the costs of both factors of production:

\[ p = \frac{\sigma}{\sigma - 1} (a_{LM}w + a_{KM}r) \] (11)

Apart from the constant \( \sigma \), equations (9) and (11) are identical to the corresponding equations in the perfectly competitive Heckscher-Ohlin model, and as there they form a sub-system which determines factor prices given the relative price of manufactures \( p \).

Of course, the goods price cannot be taken as exogenous: as we have seen, it too is endogenous in general equilibrium. However, with free trade in goods and identical homothetic tastes in both countries, there is an important set of country configurations within which the same goods and factor prices must obtain. In the words of Helpman and Krugman, this is the set of factor endowments which “replicate the integrated equilibrium”, meaning that they have the same prices as would obtain if the two countries were fully integrated, so factors were perfectly mobile internationally. Following Samuelson (1949, pp. 194-195) and Dixit and Norman, the trick is to take world endowments as fixed, and consider alternative ways of partitioning the world into countries, with any particular partition represented by a point in the world Edgeworth-Bowley endowment box as in Figure 2. Clearly any point along the diagonal \( OO^* \) satisfies the condition: total
demand and supply of each good from the world as a whole are unaffected by neutral reallocations of factors between countries, so the same prices obtain. More generally, the same holds for any allocation in which the factor proportions used in each sector are the same as those in the integrated equilibrium. If these factor proportions are represented (relative to the home-country origin $O$) by the rays $OD$ and $OF$, then the factor-price equalization or FPE set consists of the parallelogram $ODO^*F$.\footnote{Krugman in \cite{13} assumed that each factor is specific to one sector, in which case the FPE set coincides with the whole Edgeworth-Bowley box.}

It might be thought that this argument hinges on constant returns to scale, and so does not survive when one of the sectors is monopolistically competitive with increasing returns to scale. But CES preferences take care of that too, provided we keep in mind that factor proportions reflect the requirements for both fixed and variable costs. With two factors of production, the full employment condition (6) must be replaced by two conditions, one for each factor (with the home country’s aggregate endowments denoted by $L$ and $K$):

\begin{align}
L &= a_{LF}Y_F + a_{LM}(f + ay)n \\
K &= a_{KF}Y_F + a_{KM}(f + ay)n
\end{align}

\text{(12)}

Figure 2: The FPE Set and Inter-Industry Trade
The variable factor requirements $a_{LF}$ etc. in each sector are determined as long as factor prices remain equal to their integrated-equilibrium levels, given by (9) and (11); the fixed-cost factor requirement $f$ is fixed by assumption; and the output of each monopolistically competitive firm $y$ is determined by the free entry condition as we saw in Section 2. Hence these equations determine the total output of food $Y_F$ and the number of manufacturing firms $n$ as linear functions of home factor endowments. Since both countries have the same technology, corresponding equations hold for the foreign country, and so any allocation of factors within the FPE set generates the same total outputs of each good, which in turn implies that it is consistent with the same relative goods price.

### 3.2 Inter- and Intra-Industry Trade

While the world equilibria depicted in Figure 2 look identical to those in a perfectly competitive model, the implications for trade patterns are considerably richer. Two types of trade, driven by different determinants, now take place.

As far as net, or inter-industry trade is concerned, its direction and volume are given by the standard results of Heckscher-Ohlin theory. With homothetic tastes, the ratio of factors used to produce, and hence embodied in, the consumption bundle of each country equals the world ratio. Hence the world consumption point in factor space must lie along the diagonal $OO'$. If the relative endowment point also lies along the diagonal, for example at the mid-point $E$, it follows that net trade must be zero. Now consider a reallocation of the world’s resources from $E$ to $E'$. This represents a rise in the relative endowment of capital in the home country, which by the Rybczynski theorem implies a rise in the relative output of the capital-intensive good. With world endowments held fixed, the reverse happens in the other country. Hence, provided the equilibrium remains in the FPE set, the same world output of goods is produced, but its distribution across countries changes. Since consumption remains along the diagonal, the Heckscher-Ohlin prediction about trade patterns follows: the now more capital-abundant home country is a net exporter of the more capital-intensive good. Indeed, because of the convenient linear structure implied by factor-price equalization, loci of equal volumes of inter-industry trade

14
are given by the dashed parallel lines in Figure 2, with higher trade volumes corresponding to lines further from the diagonal.

Independently of this inter-industry trade, consumers' taste for diversity drives intra-industry trade in manufactures. Consider now Figure 3, which repeats the essential features of Figure 2. One important special case corresponds to the set of points along the diagonal \( OO^* \) in the world endowment box: at such points there is no inter-industry trade, and so no net trade in manufactures, as we have seen. However, there will nevertheless be intra-industry trade, as consumers at all locations demand both home and foreign varieties of manufactures. Moreover, it is straightforward to see that its volume will be greatest at the mid-point \( E \). Here the two countries are of equal size, and so each country exports exactly half the output of its manufacturing sector. For a movement away from \( E \) to a point on \( OO^* \) closer to \( O^* \), say \( E'' \), the home country produces more manufacturing varieties and so imports fewer from abroad. Hence the volume of intra-industry trade is lower.\(^\text{12}\)

\(^{12}\)Along the diagonal \( OO^* \) the Heckscher-Ohlin model effectively reduces to a one-factor model, identical to that in Section 2, with \( \frac{K}{N} = \frac{L}{L^*} = \frac{K}{K^*} \). The total value of intra-industry trade, \( V_{IIT} \), equal twice the value of the home country's imports \( M \), which in turn equals the share of its spending on imported manufactures \( \theta_{IIT} \) times its income \( I \). The share \( \theta_{IIT} \) equals the product of the budget share of manufactures \( \mu \) times the share of world varieties produced in the foreign country \( \frac{N^*}{N} \); the latter equals
Movements away from $E$ to off-diagonal points such as $E'$ also reduce the volume of intra-industry trade. The total number of manufacturing varieties produced in the world is constant everywhere in the FPE set. But at $E'$ there are fewer produced abroad because of the Rybczynski effect. Hence, the home country once again imports fewer and the volume of intra-industry trade is lower.\textsuperscript{13} Thus the volume of intra-industry trade falls for any movement away from the point $E$, and it is not difficult to show, following Helpman and Krugman, that loci of constant values of intra-industry trade have the shapes shown by the dashed lines.\textsuperscript{14} Figures 2 and 3 combined thus illustrate one of the central results of new trade theory: two types of trade coexist, with net or inter-industry trade driven by differences between countries in comparative advantage, here determined by relative factor endowments; while intra-industry trade is encouraged by similarities in country size.

Since the stimulus for monopolistically competitive models of trade had come from the observation of extensive intra-industry trade between similar countries, empirical work applying the new models soon followed.\textsuperscript{15} Helpman (1987) was a pioneering paper, which for the first time confronted the monopolistically competitive model with data, and showed that its main predictions were consistent with manufacturing trade between advanced economies. Subsequent work has explored the robustness of this finding. Helpman himself pointed out that the form of the trade volume equations implied by the monopolistically competitive model is identical to the “gravity” equation, which, as Anderson (1979) had shown, could be derived under perfectly competitive assumptions, assuming only that goods were country-specific. It was also shown by Hummels \[ \frac{L}{L^*} \text{ along } OO^*. \]

As for home income, it equals the sum of factor payments, which can be expressed in terms of the home capital-labour endowment ratio $k \equiv \frac{K}{L}$: $I = wL + rK = (w + r) L$. Combining: $V_{IIT} = 2M = 2 \mu \frac{w}{r} (w + r) LL^*$. With $w$ and $r$ fixed by factor price equalization, and $k$ fixed by the slope of $OO^*$, it follows, just as in Krugman [11], that the value of imports is proportional to the product of country labour forces $LL^*$, which is maximized at $E$.

\textsuperscript{13} For simplicity, assume that the slope of the line $EE'$ equals the ratio of factor prices. In that case, home income is constant: $dI = w dL + r dK = w dL + r (\frac{r}{w} dL) = 0$. Hence, from the last footnote, total home imports of manufactures definitely fall.

\textsuperscript{14} In general, the value of intra-industry trade equals twice home imports of manufactures: $V_{IIT} = 2M = 2 \mu \frac{w}{r} I$. At all points in the FPE set, home income $I = wL + rK$, which is linear in $L$ and $K$. As for the number of foreign varieties $n^*$, it is also linear in $L$ and $K$ from the foreign-country analogues of (12). Hence, any given value of $M$ implicitly defines an iso-IIT locus in the FPE set, giving $K$ as a quadratic function of $L$.

\textsuperscript{15} See Helpman (1999) and Feenstra (2003) for overviews.
and Levinsohn (1995) that such specifications appeared to work equally well for trade flows between non-OECD countries, which one would expect to be dominated more by comparative advantage considerations. On the other hand, Evenett and Keller (2002) showed that the gravity equation gives a better fit to data for countries with a higher share of intra-industry trade, thus providing strong circumstantial evidence in favour of the monopolistically competitive approach.

More recently, empirical work in this area has exploded further, stimulated by dramatic falls in computing costs and increases in data availability. In particular, a great deal of new data has become available at the level of individual firms, which has led to a switch in emphasis away from testing theories at industry level to estimating parameters at firm level. Much of this work is based on an extension of the monopolistic competition model to allow for heterogeneous firms due to Melitz (2003), and since this is based firmly on the work of Krugman, recent work continues the empirical tradition of monopolistically competitive models initiated by him.

### 3.3 Trade and Income Distribution

The final result of the Heckscher-Ohlin model, the Stolper-Samuelson theorem, predicted an unambiguous source of conflict in income distribution. Indeed, one of the key take-home messages of “old” trade theory was that, although there are always aggregate gains from trade, there are nevertheless always some losers. In particular, in the two-factor, two-good case, the zero-profit conditions (9) and (11) imply that either real wages or the real return to capital must fall in response to any shock.\(^\text{16}\)

However, Krugman in [13] pointed out that there is more to be said in the presence of differentiated products. Factor prices are determined by (9) and (11) as before, but these do not correspond to real factor rewards because the true cost of living with differentiated products is over-estimated by the common price of manufactures. This follows from the fact that, unlike the general utility function (1), the nested CES case (8) implies that

\(^{16}\)Jones (1965) called this a “magnification effect”: irrespective of the source of the shock, (9) and (11) imply that, following a rise in the relative prices of manufactures, \(\hat{p} > \hat{p}_F\), the changes in factor prices must satisfy the inequality \(\hat{r} > \hat{p} > \hat{p}_F > \hat{w}\), if and only if manufacturing is capital-intensive relative to food. Hence changes in factor prices are magnified relative to changes in goods prices.
tastes are homothetic, so there exists a unique true cost-of-living index, which can be written in a simple and explicit closed form:

\[ P = P_F^{1-\mu} P_M^\mu \]

Where

\[ P_M = \left( \sum_{i=1}^{N} \frac{p_{i}^{1-\sigma}}{\sigma} \right)^{1-\sigma} = N^{\frac{1}{1-\sigma}} p \]  

(13)

Like the utility function itself, the true cost-of-living index is a Cobb-Douglas aggregate of sub-indices for food and manufacturing, \( P_F \) and \( P_M \). The first expression for \( P_M \) corresponds to the general case where varieties differ in price, so the manufacturing sub-index is a CES function of all prices. The second shows that, in the symmetric case where all prices are the same, the manufacturing sub-index is less than the common price \( p \), and if the number of varieties is large it can be much less than \( p \).\(^{17}\)

The implications of this for changes in income distribution are immediate. Provided only that they have common tastes, the gains from increased diversity accrue to all factor-owners. So, while factor incomes measured in the usual way are pushed in opposite directions by the Stolper-Samuelson mechanism, the losing factor may be more than compensated by the fall in the true cost of living. As Krugman pointed out, this may explain the apparent lack of conflict in the adjustment to growing trade in post-war Western Europe suggested in the work of Balassa (1967) and others, already mentioned in Section 2.1. To the extent that most of the growth in trade was intra-industry, it required relatively little reallocation of factors; and, even if some changes in relative prices occurred, their effects may have been swamped by the benefits of the increased range of choice available to all consumers. Of course, the extent to which changes in product diversity affect living standards in practice is an empirical question, and detailed explorations have had to wait for the more recent availability of highly disaggregated data on household purchases and low-cost computing power to analyze them. A study of this kind by Broda and Weinstein (2006) shows that the number of imported varieties into the U.S. increased by a factor of three from 1972 to 2001, and estimates that this has raised the welfare of U.S. consumers by an amount equivalent to 2.6 percent of GDP. Though

\(^{17}\)The elasticity of substitution \( \sigma \) must be greater than one from the firm’s second-order condition, so \( N \) is raised to a negative power in (13), and hence \( N^{\frac{1}{1-\sigma}} \) is less than one.
modest relative to real growth, and though subject to the proviso that it takes CES preferences as a maintained hypothesis, this finding suggests the practical importance of gains from trade due to increased variety, as highlighted by Krugman.

4 Transport Costs and Economic Geography

The key to the elegant Helpman-Krugman synthesis of Heckscher-Ohlin and Dixit-Stiglitz was its focus on cases where both goods and factor prices were equalized internationally. However, Krugman also explored the case where transport costs cause international differences in goods prices. This was to prove perhaps the most innovative of his contributions, since it led to two new predictions: home-market effects on trade, and the possibility that manufacturing activity may agglomerate even when countries are ex ante identical.

4.1 The Home-Market Effect

The home-market effect was introduced in Krugman’s 1980 American Economic Review paper “Scale Economies, Product Differentiation, and the Pattern of Trade” [12], and, as there, we can return for this sub-section to the labour-only model of Section 2, though we retain from Section 3 the two-sector structure and the nested CES preferences given by (8).\footnote{The approach here follows Helpman and Krugman ([1], pp. 205-209) rather than Krugman ([12]) though the formal differences are slight.} We also assume as before that food is produced competitively and freely traded internationally: the latter is not at all realistic but it allows a clear focus on the effects of costly transportation of goods produced under monopolistic competition. It also simplifies the model a lot since it implies international factor price equalization: free trade in food equalizes both food prices and wages, so only manufacturing prices differ between countries.

To distinguish explicitly between the two countries, we denote foreign variables by an asterisk, with the total number of varieties in the world equal to: $N = n + n^*$. We continue to assume that all firms are identical, so “factory-gate” prices are the same at home and abroad: as before, $p = \frac{\sigma}{\sigma-1}aw$. The new feature is that international trade
incurs “iceberg” transport costs, meaning that for every \( \tau \) units of a manufacturing good shipped (where \( \tau \) is greater than one), only one unit arrives. This raises the price to consumers of an imported variety from \( p \) to \( \tau p \). The home cost-of-living sub-index for manufactures is still a CES aggregate of the prices of all varieties as in (13), but now imports are more expensive:

\[
P_M = \left[ \sum_{i=1}^{n} p_i^{1-\sigma} + \sum_{i=1}^{n^*} (\tau p_i)^{1-\sigma} \right]^{\frac{1}{1-\sigma}} = \left[ s_n + (1 - s_n) \tau^{1-\sigma} \right]^{\frac{1}{1-\sigma}} N^{\frac{1}{1-\sigma}} p \tag{14}
\]

Here \( s_n \equiv n/N \) is the home-country share of manufacturing in the world. Compared with the price index in free trade, (13), higher transport costs \( \tau \) raise the cost of living for a given distribution of manufacturing \( s_n \). More important is the effect of changes in \( s_n \) itself: a rise in \( s_n \) means that home consumers save on transport costs as they purchase more domestic varieties and fewer imported ones; hence the true cost of living falls. Since the opposite is true in the foreign country, there is a negative relationship between the relative cost of living at home and abroad, \( P/P^* \), and the home-country share of manufacturing \( s_n \):

\[
\frac{P}{P^*} = \left[ \frac{s_n + (1 - s_n) \tau^{1-\sigma}}{(1 - s_n) + s_n \tau^{1-\sigma}} \right]^{\mu} \tag{15}
\]

This is illustrated by the downward-sloping locus in panel (a) of Figure 4.

This “price-index effect” is the first building block of the home-market effect. The second is the observation that, with fixed factory-gate prices of manufactures, goods-market equilibrium imposes a negative relationship between the two countries’ relative price indices and relative incomes. The key to this is the demand behaviour implied by the utility function (8). From the perspective of a typical domestic firm, the demand it faces from both domestic and foreign consumers depends negatively on the price they pay, with an elasticity equal to \( \sigma \), and positively on the portion of income spent on manufacturing, \( \mu I \) at home and \( \mu I^* \) abroad, all expressed relative to the relevant manufacturing price.

\[\text{Recall from (13) that the overall price index at home is } P = P_F^{1-\mu} P_M^\mu \text{ and that food prices are equal internationally. Hence } \frac{P}{P^*} = \left( \frac{P_M}{P_M^*} \right)^{\mu} \]
Figure 4: The Price-Index Effect and the Home Market Effect

index:

\[ x = \left( \frac{p}{P_M} \right)^{-\sigma} \frac{\mu I}{P_M} \quad x^* = \left( \frac{\tau p}{P^*_M} \right)^{-\sigma} \frac{\mu I^*}{P^*_M} \]  

(16)

Note that the price paid by foreign consumers must be grossed up by transport costs. A final consideration is that shipments to the foreign market exceed foreign demand for imports \( x^* \) because of the wastage incurred in transit. Hence the market-clearing condition for a typical home-produced variety is that total output \( y \) must equal total demand \( X \) made up as follows:

\[ y = X = x + \tau x^* = \mu p^{-\sigma} \left[ IP_M^{\sigma-1} + \tau^{1-\sigma} I^* \left( P^*_M \right)^{\sigma-1} \right] \]  

(17)

From the firm’s partial-equilibrium perspective, the expression in square brackets is exogenous, so the demand it faces is a simple iso-elastic function of the price it charges, with incomes, price indices, and transport costs given. However, in general equilibrium, the output and factory-gate price of each variety are fixed and the same in each country as we have seen, so incomes and price indices must adjust to ensure that goods markets clear. Combining (17) with the corresponding equation for foreign firms yields the
required negative relationship between relative price indices and relative incomes:

\[
P/P^* = \left( \frac{s_I}{1 - s_I} \right)^\tau^{1-\sigma}
\]

where \( s_I \equiv I / (I + I^*) \) is the home-country share of world income. This is illustrated by the downward-sloping locus in panel (b) of Figure 4.

The final step is to combine the two panels. Assume that the two countries are initially in equilibrium denoted by points A and B. Now assume some exogenous shock raises the relative income of the home country, as shown by the leftward arrow in panel (b). From panel (b), we can see that for goods markets to return to equilibrium, the home country whose income has risen must experience a fall in its relative price level. But in that case, from panel (a), it must also acquire a larger share of world manufacturing. Moreover, the curve in panel (a) is less elastic than that in panel (b): the relative price indices are more responsive to the home country’s share of income than to its share of manufacturing.\(^{20}\)

Hence the prediction that changes in the relative size of a country’s home market have a magnified effect on the relative size of its manufacturing sector, an idea which had been argued informally in the past but not demonstrated rigorously before. The consequences for trade patterns are immediate: the larger country demands more of both goods, and so it is a net exporter of manufactures and imports food in return.

Krugman himself noted that the home-market effect is not consistent with a perfectly competitive theory of trade: “the ... argument that countries will tend to export those kinds of products for which they have relatively large domestic demand ... is wholly dependent on increasing returns; in a world of diminishing returns, strong domestic demand for a good will tend to make it an import rather than an export” ([12], p. 955). This suggested a possible route to discriminating empirically between trade theories based on increasing returns and those based on constant or diminishing returns, an idea which was explored in a series of papers by Davis and Weinstein ((1999), (2003).) They found

\(^{20}\)Only when transport costs are infinite (so \( \tau^{1-\sigma} \) is zero) do (15) and (18) coincide: in autarky, the home share of world manufacturing must equal its share of world income. The two equations can be solved to give a linear relationship between manufacturing and income shares: \( s_n = (1 - \phi)^{-1} [(1 + \phi) s_I - \phi] \), where \( \phi \equiv \tau^{1-\sigma} \), as Helpman and Krugman ([1]) show.
persuasive evidence that increasing returns industries exhibit more pronounced home-market effects in comparisons across Japanese regions, but much weaker evidence using international data, from which they concluded that home-market effects were likely to be strongest in cases where technology and factor prices were similar.

4.2 Agglomeration and Economic Geography

The home-market effect takes incomes as exogenous, but, in the final section of his 1979 paper, Krugman sketched the idea that incomes might be endogenous because of international factor mobility. This idea was fully worked out in his most-cited paper, entitled “Increasing Returns and Economic Geography,” published in the *Journal of Political Economy* in 1991 ([24]).

The central idea in this model is that increasing returns coupled with factor mobility lead to centrifugal pressures, which may render unstable an initial symmetric equilibrium. The intuition for why centrifugal forces are likely to be strong is easy to formulate, and as Krugman noted in [4], had been put forward in non-technical accounts in the past. In the terminology of Hirschman (1958), an influx of mobile factors into one country leads to a “backward linkage” as the demand for immobile factors rises, so raising income; and to a “forward linkage” as the induced fall in the price level raises the real return to mobile factors and so encourages more in-migration. However, formal modelling was needed to pin down these arguments, and to establish how they might sometimes be counter-balanced by centripetal forces, since clearly not all economic activity is agglomerated.

Krugman himself examined this model under the same assumptions about technology made so far, though closed-form solutions are not possible in that case. The results are easier to derive, and to illustrate diagrammatically, if we follow Forslid and Ottaviano (2003) and assume that production takes a special non-homothetic form. Variable costs require unskilled labour, which is also used in food production. Though immobile internationally, its wage $w$ is the same at home and abroad as before. Fixed costs, by contrast, require the services of skilled workers, one per firm, who are internationally mobile but in fixed supply to the world as a whole: they can be thought of as entrepreneurs, whence
the label “Footloose Entrepreneur” for this variant of the model. Hence total costs for a home firm are given not by (10) but by:

\[ C_i = r + awy_i \]  

(19)

where \( r \) is the home return to entrepreneurship, which need not equal the foreign return \( r^* \). Applying the results from Section 2.3, output is proportional to the return to entrepreneurship, \( y = (\sigma - 1) \frac{r}{aw} \), and the equilibrium price is fixed, \( p = \frac{\sigma}{\sigma - 1} aw \). It is these features which make this version of the model particularly simple to use.

To illustrate the possibility of agglomeration, assume that the two countries are initially identical, so there exists a symmetric and fully diversified equilibrium, with exactly half the world’s entrepreneurs located in each country. This is illustrated by points \( A \) and \( B \) in panels (a) and (b) of Figure 5, which repeat the essential features of Figure 4. Now assume that the location of entrepreneurs is given in the short run, but that over
time they move between countries in response to international differences in their real reward:

$$\frac{ds_n}{dt} = \psi\left(\frac{r}{P/P^*}\right), \quad \psi(1) = 0, \quad \psi' > 0$$  \hspace{1cm} (20)

This tells us more about the initial equilibrium. Recall that the curve in panel (b) linking relative price levels and home's income share was derived by assuming that firm output levels were equal at home and abroad. With international factor mobility, this is no longer guaranteed in general, though it must hold in a symmetric equilibrium, as at $B$. However, because of the special cost assumptions, equality of outputs, $y = y^*$, must also imply equality of nominal returns to entrepreneurship, $r = r^*$. Hence, to ensure that the expression in brackets in (20) equals one, the price levels in the two countries must also be equal at the initial equilibrium points $A$ and $B$.

Now we need to test the initial equilibrium for stability. Consider what happens if (perhaps because of a random shock) the share of entrepreneurs in the home country rises by a small amount. In particular, we want to establish if this increases the real reward differential, in which case (20) implies that more entrepreneurs are encouraged to move into the home country, so initiating a cumulative process which leads the world economy further away from the initial symmetric equilibrium.

It turns out that the initial rise in the home share of entrepreneurs has three distinct effects. The first works through the price-index effect: the influx of entrepreneurs and hence of firms lowers the relative price level at home, leading to a new equilibrium at point $A'$ in panel (a). If relative incomes did not change, then as the leftward arrow labelled “1” in Figure 5 indicates, this would imply that the world economy was now at point $B'$ in panel (b). Because this is below the equal-reward equilibrium locus, it represents a situation where output per firm and so the reward to entrepreneurship is lower at home than abroad. Hence, this effect is a centripetal or stabilizing one: the price level effect of agglomeration tends to be self-correcting, since it is better for firms not to locate too close to their competitors.

Potentially offsetting this are the other, centrifugal, impacts of the initial in-migration. The second effect arises because entrepreneurs bring their demand for goods with them.
when they move, and so they raise the relative national income of the home country. This is shown by the relationship between home’s share of world income and its share of world entrepreneurs in panel (c). (Panel (d) simply maps $s_n$ through the 45° line.) This direct income effect works against stability: it tends to raise firm output at home relative to abroad, and so to raise the relative rental, encouraging more migration. It is illustrated in the diagram by the clockwise arrows labelled “2” in panels (c) and (d), which tend to push the new temporary equilibrium to the right of $B'$. The diagram illustrates the very special case where the first and second effects exactly offset each other, so their net effect is to restore the economy to a new point $B''$ on the equilibrium locus.

This would be the end of the story if migration responded to international differences in nominal rewards, but as (20) specifies, rational migrants respond to relative real rewards. Hence there is a third effect of the initial movement of entrepreneurs, working like the first through the price index, but now tending to raise the real return to entrepreneurship in the home country, and so encouraging yet further in-migration. Though not illustrated directly in the diagram this third effect tends to push the equilibrium locus in panel (b) inwards, as denoted by the arrow labelled “3”, which expands the $y > y^*$ region corresponding to instability.

The net outcome of these three effects is indeterminate and depends on the underlying parameters. For example, higher transport costs protect local production in the foreign country: the price-index effect is larger and so entrepreneurs who move to the home country are more likely to face an incentive to return. By contrast, a higher preference for manufacturing $\mu$ raises both centrifugal tendencies, increasing the positive impact of migration both on local demand and on the real return to entrepreneurship. If these latter effects dominate, then the initial diversified equilibrium is unstable and manufacturing agglomerates in the home country, which becomes the industrial “core” of the world economy, leaving the foreign country to languish as a deindustrialized “periphery”. Note in particular that these relative roles could just as easily have been reversed: which country acquires the world’s manufacturing depends on the direction of the initial small movement of entrepreneurs away from the diversified equilibrium, almost as imperceptible
as the flapping of a butterfly’s wings.\textsuperscript{21}

Krugman extended the basic insight of a potentially unstable diversified equilibrium in two directions. First was to show that the two specialized or core-periphery equilibria are also potentially unstable, but that there is a range of the exogenous parameters in which all three equilibria are stable. When the parameter in question is the level of transport costs, this result raises the possibility of all kinds of interesting dynamic phenomena, which exhibit “hysteresis” or path-dependence, so history matters for the outcome. For example, the diversified equilibrium may be stable in response to small shocks but not to large ones, and a reduction in transport costs may lead to deindustrialization of the periphery, which a return to protection is unable to reverse. The second direction was to consider intermediate inputs as an alternative driver of agglomeration than international factor mobility. This extension was pioneered by Tony Venables (1996) and presented in an elegant two-country model by Krugman and Venables ([26]). It turns out that this model has very similar properties to the mobile-factor ones, and is more plausible when applied to international rather than interregional contexts.

Krugman’s work on economic geography, synthesized in his book with Masahisa Fujita and Tony Venables ([7]), has spawned a huge literature, especially in Europe. Much of this has been theoretical, exploring the robustness of the model’s predictions to changes in assumptions as well as considering its policy implications.\textsuperscript{22} There have also been some empirical applications, and it seems likely that the model will continue to inspire applied work. From a large literature, two papers are of particular interest because they attempt to test explicitly a key prediction of Krugman’s model, that large but temporary shocks can have hysteretic effects, moving the pattern of industrial location from one equilibrium to another. Davis and Weinstein (2002) fail to find evidence for this in a study of the effects of the Allied bombing campaign on Germany during World War II. By contrast, Redding and Sturm (2008) show that it is consistent with the location pattern of German airports: this changed dramatically following the post-war partition.

\textsuperscript{21}Matsuyama (1995) surveys other examples in economics where monopolistic competition generates what Gunnar Myrdal called “circular and cumulative causation,” and what would today be called the “butterfly effect,” a metaphor for the sensitive dependence on initial conditions of a dynamical system. \textsuperscript{22}Baldwin et al. (2003) review both.
of Germany, but failed to return to its earlier equilibrium following German reunification in 1991. More broadly, Krugman’s work has also served to bring the fields of regional and urban economics closer to international trade. Even though many of the approaches used in these fields are very different from monopolistic competition à la Dixit-Stiglitz, a major part of Krugman’s legacy is his role in reviving interest in location theory and showing that the tools of mainstream applied theory can be applied to it.

5 Oligopoly and Trade

The Nobel Committee’s citation of Krugman’s work illustrates the empirical motivation for new trade theory with the example of Sweden exporting Volvo cars to Germany while Germany exports BMW cars to Sweden. But, especially recalling that Volvo is now owned by Ford, this example is hard to reconcile with the atomistic firms of monopolistic competition, which never earn profits in equilibrium, take the demand functions they face as given, and do not interact strategically with their competitors. This reflects the fact that there were always two strands to new trade theory, each building on a different tradition in industrial organization. On the one hand, the monopolistically competitive models discussed in previous sections made it possible, in Krugman’s words, “to get the issue of market structure out of the way as simply as possible” ([20], p. 1179), in order to allow a clear focus on increasing returns and especially product differentiation as drivers of intra-industry trade. On the other hand, a different strand of work applied oligopolistic models of small-group competition to trade issues, in the process bringing market structure and strategic interaction to the foreground. Largely through Krugman’s influence, the monopolistically competitive framework has come to dominate trade theory today almost as much as did the perfectly competitive Heckscher-Ohlin theory in the 1960s. Yet, remarkably, Krugman has also made essential contributions to the second oligopolistic approach, exploring both its positive implications for trade patterns and its

23Combes, Mayer and Thisse (2008) provide a recent overview.
24A third strand, pursued in particular by Ethier (1979, 1982b) as well as by Krugman himself (see Section 6 below), returned to an earlier tradition of modelling increasing returns while avoiding issues of market structure altogether, by assuming that perfectly competitive firms enjoy external economies of scale.
normative implications for trade policy.

5.1 Reciprocal Dumping

The simplest oligopoly model in a trade context is a Cournot duopoly where two firms produce identical products, one located in the home country and the other in the foreign. Brander (1981) pioneered the analysis of this model, assuming that international transport costs are zero, and adding the key assumption that national markets are segmented. This implied both a new explanation of trade and a new justification for it. From a positive perspective, each firm has an incentive to sell in the other’s market. Hence the model predicted two-way trade or “cross-hauling” of identical products, even in the perfectly symmetric case, where both firms and markets were identical. Ships passing each other in the night, carrying identical goods in opposite directions, were remarkable enough in themselves. What was particularly interesting was the normative implication of this seemingly pointless trade. Moving from autarky to free trade in effect means moving from monopoly to duopoly, and the resulting fall in price and increase in sales raises consumer surplus by more than enough to offset the loss in profits by the home firm and so to raise domestic welfare overall. Brander’s model thus provided a particularly clear illustration of the potential importance of the competition effect of trade in disciplining domestic producers and lowering prices for consumers.

However, the case of zero transport costs was implausibly stark in this context. Krugman joined forces with Brander in [14] to show that transport costs leave the basic results unchanged and also enrich the model considerably. First, assuming the two markets are identical, the home firm charges the same price in each. But this requires it to absorb part of the transport cost in its export price. As a result, its profit margin is lower on its export sales, implying that it is “dumping” in the foreign market. Since the foreign firm is doing exactly the same in the home market, the Cournot equilibrium with transport costs exhibits “reciprocal dumping” by both firms.

The second implication of this model is that welfare is related in a complicated way to the level of transport costs. The home country’s welfare equals the sum of consumer
Figure 6: Components of Home Welfare

surplus $B$ and the home firm’s profits in its home and foreign markets $\pi$ and $\pi^*$. The effects of changes in transport costs on these three components are shown in Figure 6. Consumers always gain from any additional competition, so consumer surplus increases monotonically as transport costs fall from the prohibitive level ($t = \bar{t}$) to zero. However, the effect on the home firm’s profits are more complicated, because a fall in transport costs has opposite effects on home and foreign sales and hence on home and foreign profits: it exposes the firm to greater competition in its home market, while expanding its opportunities in the export market. As a result, total profits do not fall monotonically with transport costs. This is most easily seen by considering the two extremes of free

25When the countries are symmetric, this is exactly equivalent to the welfare decomposition in Brander and Krugman, who measure world welfare as twice the sum of domestic consumer surplus and total profits by both firms in the home market. Switching from wasteful transport costs to tariff revenue costlessly reimbursed to consumers would not affect the argument in qualitative terms, since the change in tariff revenue is zero in the neighborhood of autarky.
trade \((t = 0)\) and autarky \((t = \bar{t})\). Starting first from free trade, where \(t\) is zero, exports are harmed more by an increase in own costs than home sales are helped by an equal rise in rival’s costs; hence total sales and profits fall for a small increase in \(t\) at free trade. But now consider starting instead from autarky, where \(t\) is initially at the prohibitive level \(\bar{t}\). A reduction in \(t\) raises profits on export sales and lowers profits at home. However, if the reduction is small, total profits must fall. Exports are initially zero, so the rise in sales has a negligible effect on profits in the export market; whereas home sales are initially maximal, so the fall in sales has a first-order effect on home-market profits. Hence, overall profits fall for a small reduction in \(t\) at autarky. Since total profits fall for a small movement away from either autarky or free trade, their relationship to trade costs must be U-shaped. Finally, consider the sum of consumer surplus and profits for a small fall in \(t\) starting in autarky. Consumer surplus rises because the price falls, but profits on home sales fall both because the price falls and because sales are reduced. The price effects cancel, so the total fall in profits outweighs the rise in consumer surplus.

Thus home welfare (the sum of profits and consumer surplus) is also a U-shaped function of the tariff. Note finally that, since the source of this paradoxical effect of trade costs on welfare is the response of profits on home sales, it cannot arise when all profits are competed away by free entry, as Brander and Krugman showed. Hence oligopoly with restricted entry is essential for this perverse effect of trade liberalization on welfare in a symmetric world, just as it is for the prediction of two-way trade in identical goods.

### 5.2 Strategic Trade Policy

The reciprocal dumping theory of trade in identical goods is an important adjunct to other theories, since it highlights the implications for trade of strategic interaction and persistent profits. However, it has never attained the same status as the theories of comparative advantage or product differentiation, perhaps because literal two-way trade in identical goods is empirically implausible, perhaps because it was not embedded in

\[26\text{With consumer surplus denoted by } B = u(X) - pX, \text{ where total home consumption } X \text{ equals sales by the home firm } x \text{ plus imports } y, \text{ the change in consumer surplus is } dB = -Xdp. \text{ As for profits on home sales, } \pi = (p - c)x, \text{ the change in this is } d\pi = (p - c)dx + xdp. \text{ In the neighborhood of autarky, imports } y \text{ are zero, so } x = X \text{ and } dB + d\pi = (p - c)dx \text{ which is negative as the tariff falls.}\]
general equilibrium. The same is not true of the application of oligopolistic models to policy questions, the sub-field which became known as “strategic trade policy”.

Notable early papers in this field by Spencer and Brander (1983) and Brander and Spencer (1985) showed that in oligopolistic markets there is an economic case for helping home firms to compete against foreign rivals. Provided the home government can credibly commit in advance to subsidize or tax the home firm, then it can influence the outcome of the game between firms. In particular, if firms play a Cournot game, a subsidy to exports induces a domestic firm to play more aggressively, raising both its profits and its market share at the expense of foreign competitors. Krugman ([15]) identified a different mechanism whereby assisting home firms could help them in export markets. If marginal costs fall with total output (unlike the simple Dixit-Stiglitz case considered so far), then supporting a firm in its home market reduces its costs of producing for export. In Krugman’s memorable phrase, import protection serves as export promotion in this case. At the same time, Krugman was careful to point out that welfare need not increase, for the kind of second-best reasons familiar from the extensive literature on trade and welfare initiated by James Meade: the gains from lower costs and higher export sales could be offset by the direct cost of the subsidies. To explore the practical importance of these issues, Krugman with Richard Baldwin in [19] developed and calibrated a simulation model of the market for random access memory (RAM) computer chips: at first sight a plausible example of the case for import promotion, given the apparent success of the Japanese government in using a protected home market to allow domestic firms move down their cost curves and subsequently compete successfully against established rivals in export markets. This study was notable for taking account of the dynamic nature of competition in markets with rapidly changing technology, but it concluded that there was no convincing evidence that the Japanese government’s policies had raised welfare.

This theme, that strategic trade policy is theoretically exciting but of limited relevance to real-world policy-making, was taken up by Krugman in his 1987 paper “Is Free

---

27 A rare empirical application is by Friberg and Ganslandt (2006). They estimate the reciprocal dumping model using data from the market for bottled water in Sweden, and conclude that such apparently wasteful two-way trade does not reduce welfare.
Trade Passé?” ([17]). The answer he gave was a qualified but firm “No”. In part this reflected theoretical qualifications to the argument for subsidizing exporting firms. Eaton and Grossman (1986) showed that the result was not robust to the assumptions made about firm behaviour: if firms competed by choosing prices rather than quantities, then intervention was still called for, but it should take the form of a tax on a domestic firm to prevent it over-producing. In a different vein, Dixit and Grossman (1986) pointed out that the argument for subsidization neglected general-equilibrium feedbacks: if applied to a broad swathe of industries, a subsidy policy would drive up factor prices, which could nullify its effect in helping the targeted firms. In part it reflected pragmatic considerations of political economy: the gains from intervention were likely to be captured by lobby groups or competed away by foreign retaliation. In this respect, the 1980s turned out to be a re-run of the 1890s. In the late nineteenth century, the optimal tariff argument had provided a logically watertight argument in favour of restricting trade, but its relevance to real-world policy-making was always open to question. Edgeworth famously suggested that it should be labelled “poison” lest politicians misuse it. In the same way, the insights of strategic trade policy continue to inform debates on industrial and R&D policy, but they have only slightly dented economists’ enthusiasm for what Krugman summarized as “a sadder but wiser argument for free trade as a rule of thumb in a world whose politics are as imperfect as its markets.”

6 Other Contributions

With close to a hundred papers in professional journals and many more in collected works and other volumes, it would be impossible to do justice to all of Krugman’s scholarly contributions in a single article.\textsuperscript{28} Moreover, he has written on a wide range of topics, always combining one of the best prose styles in the profession with an ability to construct elegant, insightful and useful models. Here I mention briefly only a few other highlights.

Though best-known for his work on imperfect competition and trade, Krugman has

\textsuperscript{28}As of January 2009, JSTOR lists ninety-two journal articles by him, ISI lists eighty-five, but these are only the tip of an iceberg.
also written many papers on trade using perfectly competitive models. Some of these have dealt with topics that were non-traditional when he wrote, though they have since come into the mainstream, such as technology and hysteresis. In an early paper ([9]), he constructed an ingenious model of “North-South” trade, in which the ranges of goods produced in both the high-wage North and the low-wage South continually expand, the former through innovation, the latter through technology transfer. Another important paper ([18]) showed that dynamic increasing returns driven by learning could generate hysteresis in industrial structure. In particular, this model suggested that an over-valued real exchange rate, whether due to monetary contraction or the “Dutch Disease” consequences of a resource boom, could leave permanent scars on the real economy. Hysteresis also featured in [22]: published in the same year as his economic geography paper, this paper featured explicit dynamics though a more reduced-form specification to show how history and expectations interact to determine the long-run equilibrium.

Krugman has also written on more traditional topics in trade, showing the power of neoclassical theory to illuminate current controversies. In some cases he has argued for the importance of factor endowments, as in his advocacy in [25] of the view put forward by Young (1995) that factor accumulation and reallocation rather than productivity growth, or “perspiration rather than inspiration,” explain the spectacular growth of East Asian economies in recent years. In others he has argued against, as in his disavowal in [27] of the importance of Stolper-Samuelson effects as an explanation for the rise in the relative wages of skilled workers in recent decades.

Apart from his contributions to trade, Krugman’s best-known work is in international macroeconomics. An early and highly influential paper [10] adapted a model due to Salant and Henderson (1978) to the study of currency crises and stimulated a large literature. This paper showed how fixed exchange-rate targets were vulnerable to speculative pressure. Reasoning by analogy might suggest that the same would be true of so-called “target zones”, where exchange rates are allowed to fluctuate within a band, and intervention only occurs when the exchange rate approaches the limits of the band. However, Krugman showed in [23] that this was not true if the target zone was credible:
on the contrary, in that case it made exchange rates less rather than more volatile. As for exchange rate volatility itself, an underlying explanation of this was proposed by both Krugman with Richard Baldwin in [16] and [21] as well as by Dixit (1989): in the face of both uncertainty and adjustment costs, agents do not have an incentive to respond to small exchange-rate movements. Hence, production and trade exhibit hysteresis, varying rather little most of the time, and requiring large exchange-rate movements to shift them from one quasi-steady state to another.

Finally, while this is not the place to attempt an assessment of Krugman’s non-scientific writings, no discussion of his work could fail to mention his transition from Academic Superstar to Public Intellectual. Through his extensive writings, including a regular column for the *New York Times*, monographs and textbooks at every level, and books on economics and current affairs for the general public (ironically, the latter are known as “trade books” in publishing circles), he has probably done more than any other writer to explain economic principles to a wide audience. Nor has he shied away from controversy, with the target of his attacks shifting from old trade theory in the 1980s, to “pop internationalists” and other “peddlers of prosperity” in the 1990s (see [3] and [5]), to the politics and economics of President George W. Bush in the twenty-first century (see [8]). Throughout, he has been wide-ranging, stimulating, and provocative: floating like a butterfly, pollinating like a bee, and, sometimes, stinging like a wasp. With these targets no longer in his sights, his many readers can look forward with relish to where he will next focus his energies, now that (as of February 2009) liberalism is back in fashion in the United States but, for other reasons, the world seems a more dangerous place.

7 Conclusion

A former secretary of the physics Nobel Prize committee is on record as quoting an unnamed economist to the effect that economics does not deserve any more Nobel Prizes: “All the mighty firs have fallen. Now there are only bushes left.” (Nasar, 1998, p. 368) Perhaps he was cross at Krugman’s much more apt remark that “economics is harder
than physics; luckily it is not quite as hard as sociology” ([3], p. xi). At the very least, the anonymous quotation risks holding economists to a higher level of achievement than is the norm in chemistry, medicine or physics. It is true that the Dixit-Stiglitz specification of preferences was an essential precondition for Krugman’s work on intra-industry trade and economic geography. But standing on the shoulders of giants does not by itself diminish the importance of a contribution. (The phrase was, after all, used by Newton to describe his life’s work: Nobel-worthy if anyone’s was.) Taking new theoretical tools and adapting them to explain empirical anomalies is the essence of scientific progress, and Krugman’s achievement has been to do just that in international trade. As noted earlier, he has also made important contributions to other fields, especially to international macroeconomics, and *Nullum quod tetiget non ornavit* - Everything he has touched he has adorned. But putting the “new” into “new trade theory” remains his major professional achievement. The tractable and flexible models of increasing returns and product differentiation which he helped develop and for which he campaigned tirelessly are now central to every trade economist’s tool-kit. As well as providing novel insights in themselves, they have proved consistent with the Heckscher-Ohlin theory which predated them, and have lent themselves to further empirical and theoretical elaboration.
Books by Paul Krugman Cited in the Text


Articles by Paul Krugman Cited in the Text


References


