


Provided by the author(s) and University College Dublin Library in accordance with publisher policies. Please cite the published version when available.

Title	Of hype and hyperbolas : introducing the new economic geography
Author(s)	Neary, J. Peter
Publication date	2000-12-11
Series	UCD Centre for Economic Research Working Paper Series; WP00/19
Publisher	University College Dublin. School of Economics
Link to online version	http://www.ucd.ie/economics/research/papers/2000/WP00.19.pdf
Item record/more information	http://hdl.handle.net/10197/1315

Downloaded 2017-11-19T12:44:07Z

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa) 

Some rights reserved. For more information, please see the item record link above.



**OF HYPE AND HYPERBOLAS:
INTRODUCING THE NEW ECONOMIC GEOGRAPHY**

J. Peter Neary¹
University College Dublin and CEPR

April 2000
(This revision December 11, 2000)

JEL: F12, R10

Keywords: agglomeration; Dixit-Stiglitz preferences; economic geography; increasing returns to scale; pecuniary externalities.

Address for Correspondence: Department of Economics, University College Dublin, Belfield, Dublin 4, Ireland; tel.: (+353) 1-706 8344; fax: (+353) 1-283 0068; e-mail: peter.neary@ucd.ie.

¹ Some of the ideas below have been presented at a seminar in Leuven and at conferences in Bergen, Geneva and Royaumont. I am very grateful to participants on these occasions and to Mary Amiti, Richard Baldwin, Steven Brakman, Frédéric Robert-Nicoud, Gianmarco Ottaviano, Jacques Thisse, Tony Venables, the editor and two anonymous referees for comments. This paper was written while I was visiting the Laboratoire d'Econometrie of the Ecole Polytechnique, Paris, and it forms part of the Globalisation Programme of the Centre for Economic Performance at LSE, funded by the UK ESRC.

**OF HYPE AND HYPERBOLAS:
INTRODUCING THE NEW ECONOMIC GEOGRAPHY**

Abstract

I take the publication of *The Spatial Economy* by Fujita, Krugman and Venables as an opportunity to expound and assess the "new economic geography". I use a simple diagram to show how the basic model works; give a self-contained algebraic derivation of the main results; note some interesting extensions; comment on some weaknesses of the model, which make it less suitable than might at first appear for discussing issues of industrial structure; try to place the field in relation to "old" economic geography and to the theory of international trade; and discuss its empirical and policy implications.

1. Introduction

"New economic geography" has come of age. Launched by Paul Krugman in a 1991 *JPE* paper; extended in a series of articles by Krugman, Masahisa Fujita, Tony Venables and a growing band of associates; recently institutionalised with the appearance of a new journal; it has now been comprehensively synthesised between hard covers.²

Such rapid progress would make anyone dizzy, and at times the authors risk getting carried away by their heady prose style. They find the predictions of one model so plausible that they call it "History of the World, Part I" (p. 253); they describe the pattern of world industrialisation implied by another as "a story of breathtaking scope" (p. 277); and on his website Krugman expresses the hope that economic geography will one day become as important a field as international trade. This sort of hype, even if tongue-in-cheek, is not to everyone's taste, especially when the results rely on special functional forms and all too often can only be derived by numerical methods. What next, the unconvinced reader may be tempted to ask? the tee-shirt? the movie?

In this paper I argue that, despite the hype, there is interesting work here which deserves to be better known. The key contribution of the new economic geography is a framework in which standard building blocks of mainstream economics (especially rational decision-making and simple general equilibrium models) are used to model the trade-off between dispersal and agglomeration, or centrifugal and centripetal forces. The approach thus gives a choice-theoretic basis for a "propensity to agglomerate": only a propensity, since agglomeration is a possible outcome but not an inevitable one.

I begin with exposition. In Section 2 I try to show that the basic new economic geography model is much simpler than most presentations suggest. I use a simple diagram to illustrate how the model works, and give a self-contained algebraic derivation of the main results (with technical details in appendices). But that in turn suggests that for some purposes the model may be too simple. In Section 4 I argue that it is and try to disentangle which lessons are wedded to the special functional forms typically used and which are of more

² Masahisa Fujita, Paul Krugman and Anthony J. Venables (1999): *The Spatial Economy: Cities, Regions and International Trade*, Cambridge Mass.: MIT Press, pp. xiii+367; henceforward "FKV". The first issue of the *Journal of Economic Geography*, published by Oxford University Press, and with an editorial board representing both "old" and "new" approaches, appeared in January 2001. Gianmarco Ottaviano and Diego Puga (1998), Masahisa Fujita and Jacques-François Thisse (2000) and Steven Brakman, Harry Garretsen and Charles van Marrewijk (2001) provide overviews of the field.

generality. In between, Section 3 highlights some of the novel insights which come from applying the model to topics in regional economics, urban economics and international trade. Finally, in Section 5 I discuss how much is new in the new economic geography, and in Section 6 I comment on policy implications and empirical tests.

2. The Core of the Core-Periphery Model

Like so much recent work in a variety of fields, the underlying model in most of the new economic geography is one of Chamberlinian monopolistic competition, parameterised using the symmetric constant-elasticity-of-substitution utility function of Avinash Dixit and Joseph Stiglitz (1977). As the authors disarmingly admit, the book "sometimes looks as if it should be entitled *Games You Can Play with CES Functions*"! Except that these are not everyone's idea of games. The authors have made a huge effort to expound the model clearly, but even they cannot prevent the calculations from degenerating at times into a near-impenetrable soup of CES algebra. All the more reason then to try and present the main ideas in a simple diagram. In this section, I first illustrate equilibrium in a closed economy, and then extend it to take account of transport costs, factor mobility and intermediate inputs.

2.1 Equilibrium à la Chamberlin-Dixit-Stiglitz

Start with a closed economy; or, more in the spirit of the book, a world in which all economic activity is concentrated at a single point. There are two sectors. Agriculture is perfectly competitive and produces a homogeneous good under constant returns to scale. Manufacturing is monopolistically competitive, and produces many varieties under increasing returns to scale. Each sector uses a single factor specific to it: farmers in agriculture, workers in manufacturing. Aggregate utility is a Cobb-Douglas function of agricultural output A and of a CES sub-utility function M derived from consuming manufactures:

$$U = M^\mu A^{1-\mu}, \quad M^{\frac{\sigma-1}{\sigma}} = \sum_i^n m_i^{\frac{\sigma-1}{\sigma}} \quad (1)$$

where m_i is the amount of each variety of the manufactured good demanded and σ (which must exceed one) is the elasticity of substitution between varieties.³ From (1), μ is the share

³ I follow FKV's notation for the most part, though with some simplifications. Here I assume a finite number of firms, which makes the equations a bit neater at the cost of some hand-waving: I need to assume that the number of firms is sufficiently large that the integer constraint is not binding. This seems reasonable, and is standard in the literature, though see some intriguing speculative comments on the role of integer constraints on page 223.

of nominal income Y spent on manufactures. This in turn is allocated between individual varieties according to demand functions which are log-linear in own price p_i and in total spending on manufactures μY , both deflated by a manufacturing price index P .⁴

$$m_i = \mu \left(\frac{p_i}{P} \right)^{-\sigma} \frac{Y}{P} \quad \text{where:} \quad P^{1-\sigma} = \sum_j^n p_j^{1-\sigma} \quad (2)$$

Turn next to manufacturing firms. Because the manufacturing sub-utility function embodies a preference for diversity, and there are increasing returns to scale, each firm produces a distinct variety. Hence the number of firms, n , is also the number of varieties consumed, and firm output q_i equals the demand for that variety m_i . The typical firm therefore faces the demand function (2). However, a key feature of the Dixit-Stiglitz approach is that firms ignore the effects of their actions on income Y and on the industry price index P . Hence the demand curve perceived by the typical firm is not (2) but rather:

$$q = \phi p^{-\sigma} \quad \text{where:} \quad \phi = \mu Y P^{\sigma-1} \quad (3)$$

with the intercept ϕ assumed to be taken as given by the firm. Since all firms are identical, the subscript i can be suppressed. Marginal revenue ($d(pq)/dq$) is then easily shown to be:

$$MR = \frac{\sigma-1}{\sigma} p \quad (4)$$

Equations (3) and (4) define two constant-elasticity curves in $\{p, q\}$ space, labelled D and MR respectively in Figure 1. (Ignore the numbered arrows in the figure for now.)

Figure 1 here

The cost structure of a typical firm is even simpler. Workers are hired at a given wage w , fixed costs are Fw and variable costs per unit output are cw . Hence total costs are:

$$C = (F + cq)w \quad (5)$$

This implies a horizontal marginal cost curve at the level cw , and an average cost curve which

⁴ P is the unit expenditure function dual to the sub-utility function M . Note that it is defined in terms of *negative* exponents (since σ must exceed one): a frequent source of confusion.

is a rectangular hyperbola with respect to the vertical axis and the marginal cost curve. These curves are also illustrated in Figure 1, labelled MC and AC respectively.

Equilibrium for the firm exhibits Chamberlin's tangency condition, as Figure 1 shows. Moreover, the special functional forms imply very simple expressions for equilibrium price and output. From profit maximisation, marginal revenue equals marginal cost:

$$\frac{\sigma-1}{\sigma}p = cw \quad (6)$$

So the price-marginal-cost mark-up depends only on σ ; while free entry drives profits to zero: $\pi = pq - (F+cq)w = 0$, which simplifies, using (6), to give the level of output:

$$q = (\sigma-1)\frac{F}{c} \quad (7)$$

So the equilibrium output of each firm is totally independent of developments outside the industry: it depends only on the cost parameters F and c and on the elasticity of substitution σ . Changes in any other parameters or variables (including income Y and the wage rate w) lead to adjustments in industry output via changes in the number of firms only.⁵

2.2 Transport Costs and the Home-Market Effect

So far, I have interpreted the model as a closed economy. However, everything said until now applies also to a multi-regional or multi-national economy, provided we make the strong assumptions (not shocking to trade theorists, though possibly so to others) of identical tastes, identical technology and no barriers to goods trade. Of these three assumptions the first and second are maintained throughout the book. The next step is to relax the third by introducing transport costs, modelled using the so-called "iceberg" assumption, that a constant fraction of output is lost in transportation. This seemingly innocuous change takes us from a world with equalised factor prices, whether under Ricardian (one-factor) assumptions as in Krugman (1979), or Heckscher-Ohlin (two-factor) assumptions as in Avinash Dixit and Victor Norman (1980, Section 9.3), and Elhanan Helpman and Paul Krugman (1985); and transports

⁵ This adjustment mechanism can be written $\dot{n} = f(\pi), f(0)=0, f' > 0$. To test for local stability, totally differentiate π in the neighbourhood of equilibrium: $d\pi = (pq/\sigma)\hat{q}$, where $\hat{q} \equiv dq/q$. Hence profits are monotonic in demand. When Y and w are given, q (and hence π) is decreasing in n from (3), so equilibrium is stable. Matters are different when Y and w are endogenous, as in Section 2.3.

us to a world with international differences in goods and factor prices as in Krugman (1980).

So, switch to two countries, labelled "1" and "2" respectively. (In the relevant sections of the book they are called "regions": more on this in Section 4.) Assume that the cost of transporting a unit of manufactures in either direction is $T-1$ times the producer price (which is also the price paid by consumers in the country of production); this amounts to assuming that, for every T units shipped, only one unit arrives. Agricultural goods by contrast incur no transport costs. (I return to this assumption in Section 4.) How much of the preceding model must be changed? Perhaps surprisingly, the answer, is that, apart from adding country subscripts to utility, wages and so on, only two equations need to be altered. First, the demand function (3) facing a typical firm located in country 1 must be extended to incorporate exports. A little manipulation gives:⁶

$$q_1 = \phi_1 p_1^{-\sigma} \quad \text{where:} \quad \phi_1 = \mu [Y_1 P_1^{\sigma-1} + Y_2 P_2^{\sigma-1} T^{1-\sigma}] \quad (8)$$

Total demand depends positively on the industry price index (P_1 and P_2) and on the level of manufacturing expenditure (μY_1 and μY_2) in both markets, and negatively on the level of trade costs T . Note however, that firms are oblivious to all these complications: the perceived demand function is identical to (3). Herein lies the simplicity of iceberg transport costs. CES preferences yield log-linear demand functions. Adding log-linear transport costs changes the level but not the elasticity of the perceived demand functions.

The second equation which must be changed is that for the price index. In a closed-economy equilibrium, with all varieties selling for the same price, (2) simplifies to $P^{1-\sigma} = n p^{1-\sigma}$. Matters are more complicated with two countries and transport costs, because the consumer prices of goods from different countries are not the same: the n_1 domestic varieties cost p_1 , while the n_2 imported varieties cost $p_2 T$. Hence the price index becomes:

⁶ Since each firm produces a distinct variety, the demand which a typical country-1 firm faces in market s is identical to (2): $q_{1s} = \mu (p_{1s}/P_s)^{-\sigma} Y_s/P_s$, though with the proviso that p_{1s} equals the price paid by consumers in market s . This matters for exports, since $p_{12} = p_1 T$ (where p_1 is the producer or mill price). In addition, this demand function gives the level of export sales, whereas we must gross up by transport costs to get export shipments: so total demand equals $q_1 = q_{11} + T q_{12}$. Combining these expressions gives (8).

$$P_1^{1-\sigma} = n_1 p_1^{1-\sigma} + n_2 (p_2 T)^{1-\sigma} \quad (9)$$

Since every firm sells in both markets, the price index is decreasing in the number of firms in both markets (because greater variety benefits consumers) and is increasing in trade costs.

Equations (8) and (9), with the corresponding equations for country 2, define a subsystem which, for given demands Y_1 and Y_2 and given wages w_1 and w_2 (so from (6) prices are also given), determines the size of the manufacturing sector in both countries, n_1 and n_2 , and the price indices P_1 and P_2 . Inspecting the equations shows they share an important feature: provided transport costs are strictly positive ($T > 1$), the left-hand sides are more responsive to home than to foreign variables. This asymmetry has two useful implications, which are most easily seen by linearising the model around the symmetric equilibrium (where $Y_1 = Y_2$, $n_1 = n_2$, etc.) and considering a small increase in the size of country 1 (so country subscripts can be dropped: $\hat{Y} = \hat{Y}_1 = -\hat{Y}_2$, $\hat{n} = \hat{n}_1 = -\hat{n}_2$, etc., where a circumflex denotes a proportional rate of change). Differentiating first equation (9) gives:

$$\hat{P} = -\frac{Z}{\sigma-1} \hat{n} + Z\hat{p}, \quad \text{where:} \quad Z \equiv \frac{1 - T^{1-\sigma}}{1 + T^{1-\sigma}} \quad (10)$$

Here Z is an index of transport costs, which ranges from zero to one. With wages and hence p fixed, equation (10) gives what FKV call the "price index effect": because imports incur transport costs but home-produced varieties do not, the cost of living is lower the larger the market. Next, differentiating (8) gives:

$$\hat{q} = Z [\hat{Y} + (\sigma-1)\hat{P}] - \sigma\hat{p} \quad (11)$$

If firm output and the price of each individual variety are fixed, an increase in demand can only be accommodated by a fall in the industry price index. From (10), this in turn requires an increase in the number of varieties: eliminating \hat{P} from (10) and (11) gives:

$$\hat{n} = \frac{1}{Z} \hat{Y} \quad (12)$$

This gives what Krugman (1980) calls the "home-market effect": the country with higher demand has (since $Z < 1$) a proportionately larger share of manufacturing.

2.3 Migration and Agglomeration

The home-market effect is of independent interest since it predicts that countries with a larger home-market tend to export manufactures.⁷ This can be given a core-periphery interpretation, as in Krugman and Venables (1990): larger countries have a disproportionately larger share of manufacturing. However, by itself the home-market effect is only a prelude to a complete theory of economic geography. It assumes rather than explains international differences in incomes, since as yet there is no force pushing towards agglomeration in the model. Two such forces are considered in the book, international labour mobility (following Krugman (1991)) and intermediate inputs (pioneered by Venables (1996) and developed by Krugman and Venables (1995)). In this sub-section, I concentrate on the former.

Assume therefore that the farmers who produce agricultural goods are country- as well as sector-specific, with equal numbers in each country. Hence (with no agricultural transport costs) the wage and the price of agricultural output can be set equal to unity everywhere. By contrast, the workers employed in manufacturing move internationally in response to differences in real wages (ω_s in country s , equal to nominal wages w_s deflated by the consumer price index P_s^H). Note that the relevant price index is the local one: workers consume only in the country where they are employed.

We are finally ready to address the central theoretical question in the new economic geography: when will the equilibrium exhibit diversification, with manufacturing produced in both countries, and when will it exhibit agglomeration in one country, or a "core-periphery" pattern? These are actually two distinct questions, since it turns out that equilibrium is not unique for all parameter values. To answer them, consider in turn the local stability of diversified and agglomerated equilibria. Assume that workers migrate rapidly between countries in response to differences in real wages and that manufacturing firms enter or leave the industry sluggishly in response to profits or losses.⁸

⁷ The effect has also been shown to hold in some though not all other models. See Head, Mayer and Ries (1999).

⁸ A digression on dynamics. Following Krugman (1991), FKV test equilibria for stability by asking how migration by a single worker affects relative real wages. If they end up higher in the host country, then further migration is encouraged and the initial equilibrium is unstable. This heuristic argument is equivalent to testing for local stability a dynamic adjustment mechanism, whereby migration is driven by the real-wage ratio. Two comments are in order. First, this adjustment mechanism assumes that firms are always in equilibrium;

Take first the case where the world economy is in a symmetric diversified equilibrium, and go through the following thought experiment. Assume that a single new manufacturing firm enters in country 1 (and, for symmetry, a single firm exits in country 2) and ask how this affects the incentives for further entry or exit. If profits in country 1 (which are initially zero) fall relative to country 2, the diversified equilibrium is stable: the new firm is encouraged to exit and the initial equilibrium is restored. However, if relative profits in country 1 rise, the initial equilibrium is unstable. More firms are encouraged to enter and the world economy moves towards an equilibrium with agglomeration: more than half of world manufacturing locates in country 1.

There are three effects of entry, indicated by the numbered arrows in Figure 1. The first arises from the price-index effect: an extra firm lowers the industry price index which reduces the demand facing each existing firm. In Figure 1, the demand and marginal revenue curves shift downwards, as indicated by the arrow numbered "1". This competition effect reduces profits and so encourages stability of the diversified equilibrium. By contrast, the two other effects raise profits and so encourage instability and a tendency towards agglomeration.

The second effect is a *demand* or *backward* linkage. An extra firm raises the demand for labour in country 1. This puts incipient upward pressure on local wages, which encourages foreign workers to migrate. The new workers in turn raise demand for local varieties. The demand and marginal revenue curves shift upwards in Figure 1, as indicated by the arrow numbered "2", and this tends to *raise* profitability.

It is easy to see which of these two effects dominates, if we assume provisionally that

i.e., that entry and exit of firms occurs infinitely faster than migration. Reversing these assumptions, as in Puga (1999), does not affect the stability conclusions. I follow Puga's approach here since it can be more easily related to Figure 1 and since (unlike FKV's approach) it applies with minimal modifications to the intermediate-goods model of Section 2.4. Second, both these dynamic adjustment mechanisms are myopic: migrating workers or newly-entering firms consider only current returns (relative real wages or relative profits) in making their decisions. FKV justify such non-rational expectations by appealing, rather unsatisfactorily, to the "replicator dynamics" of evolutionary game theory. (See pp. 7-8 and p. 77, footnote 2: "our model can, if one likes, be regarded as an evolutionary game."). I prefer a more robust defense. Just as in the comparison between Neary (1978) and Michael Mussa (1978), replacing *ad hoc* with forward-looking dynamics leads to greater elegance and a more satisfactory basis for normative analysis. However, it makes no difference to positive comparative-statics predictions. This has been demonstrated formally for the internationally-mobile-labour variant of the core-periphery model by Richard Baldwin (1999b).

nominal wages do not change. From (10) and (11), the stabilising competition effect depends on Z : when transport costs are higher, a reallocation of firms has a greater impact in lowering the price level and raising competition in the expanding country. As for the demand linkage, consider the effects of entry on country 1's labour market. Supply is L , and demand is $n(F+cq)$, so from (7) the labour-market equilibrium condition is:

$$L = n\sigma F \quad (13)$$

(Once again, country subscripts can be dropped since we are linearising around a symmetric equilibrium: whatever goes up in country 1 must come down in country 2.) If equilibrium is to be maintained, entry of new firms must be matched by an equi-proportionate increase in the labour force. To see how this affects demand, consider its effect on national income. With the number of farmers fixed and their wage constant by choice of numeraire, the proportional change in national income equals the changes in the manufacturing wage and labour force, times the share of the sector in GNP, μ :⁹

$$\hat{Y} = \mu (\hat{w} + \hat{L}) \quad (14)$$

Hence, with \hat{w} fixed, the magnitude of the demand linkage depends on μ . Returning to (11), the demand linkage dominates the competition effect at the initial wage, so entry by a new firm raises the demand facing all existing firms, if and only if μ is greater than Z . When this condition holds the initial diversified equilibrium is unstable: agglomeration is more likely when the share of manufacturing in national income is high and when transport costs are low.

However, this condition ($\mu > Z$) underestimates the pressure towards agglomeration, because migration is driven by differentials in *real* rather than nominal wages. As we have seen, entry by a new firm lowers the price index in country 1. But this induces a third effect, since it reduces the cost of living for workers and so tends to raise real wages in country 1. The resulting migration restores international equality of real wages, which means that the nominal wage must fall. The result is a *cost* or *forward* linkage which shifts the average and

⁹ Since the initial equilibrium is symmetric, there is no inter-industry trade. Hence the share of manufacturing in each country's GNP equals its share in consumption, μ . Note that this has nothing to do with choice of units. Following Krugman (1991), FKV choose units such that the world labour force equals μ and the world supply of farmers equals $1-\mu$. But this is not needed to prove any of the results which follow, and I find it distracting. I comment on other choices of units in Section 4.

marginal cost curves downwards in Figure 1, as indicated by the arrow numbered "3", and so *raises* profitability further.

Deriving a necessary and sufficient condition for stability of the diversified equilibrium requires combining the three effects explicitly, and is best left to the Appendix. However, it is fairly easy to see how stability is affected by changes in the three key parameters, T , μ and σ . Higher transport costs T always encourage stability. For sufficiently high T , imports are so expensive that home production is always profitable;¹⁰ while for sufficiently low T , diversification is always unstable since the countries are *ex ante* identical (neither has a comparative advantage). Somewhere in between is a threshold level of trade costs T^B , the "break" point, at which the diversified equilibrium is on the brink of instability. (Such a threshold must exist provided there is some incentive for agglomeration, which means, provided μ is strictly positive.) A convenient way to summarise the results is to consider how T^B varies with the other parameters. As shown in the Appendix:

Proposition 1: T^B is increasing in μ and decreasing in σ .

The threshold value T^B must be increasing in μ since both potentially destabilising effects are more important the greater is μ . First, the greater the share of manufactures in national income, the greater the demand linkage whereby entry by a new firm induces migration which raises demand and hence encourages further entry. (In older times this would have been called a "bootstraps" effect.) Hence the larger the range of trade costs at which diversification is not an equilibrium. Second, the greater the budget share of manufactures, the greater the cost linkage whereby entry by a new firm lowers the cost of living and so encourages further migration and entry. Finally, the threshold value of T is negatively related to σ , the elasticity of substitution in demand. Higher σ means that consumers view different varieties as closer substitutes. Such a reduced preference for diversity leads to an equilibrium with fewer

¹⁰ Actually, this is an assumption rather than an intrinsic property of the model. When Z equals one, the numerator of equation (22) in the Appendix equals $-(\sigma - I - \sigma\mu)(1 - \mu)$. Hence, if $\sigma - I < \sigma\mu$, the diversified equilibrium is unstable even when Z is one (transport costs are infinite, so the economies are closed). It makes sense to rule out this possibility, where the propensity to agglomerate is always one. So we assume that the preference for diversity and/or the budget share of manufactures are sufficiently low to ensure that $\sigma - I > \sigma\mu$. FKV call this the "no-black-hole" condition.

varieties and a higher output of each. (Recall equation (7).) As a result, both countries are more likely to hold on to some production at low trade costs.

Transport costs lower than the break level T^B are sufficient for the existence of an agglomerated equilibrium. Are they necessary? To show that the answer is "no", consider a different thought experiment. Assume that the world economy initially exhibits a core-periphery pattern and ask whether there are incentives for a firm to deviate from it. Assuming (arbitrarily) that manufacturing is agglomerated in country 1, can a firm earn positive profits by entering in country 2?

We can answer this question heuristically by referring again to Figure 1, this time with all three arrows *reversed*. Assume that the initial equilibrium illustrated is that in the core (country 1) and ask what are the incentives for a potential firm to enter in the periphery. The first incentive is that it faces less competition from country-1 firms in serving country-2 consumers. Hence the demand and marginal revenue curves are higher (represented by the arrow numbered "1" in reverse). On the other hand, a periphery firm has worse market access to consumers in the core: this demand linkage is represented by the arrow numbered "2" in reverse. Finally, because the periphery has fewer firms to begin with, the bulk of manufactures consumed there must be imported, incurring transport costs. The cost of living is therefore higher in the periphery, so, to induce workers to migrate, a new firm in the periphery must pay higher nominal wages. This cost linkage is represented by the arrow numbered "3" in reverse.

In the light of this intuitive discussion it is not surprising that there is once again a threshold level of trade costs, T^S , the "sustain" point, at which agglomeration ceases to be an equilibrium. This new threshold has the same relationship to the underlying parameters as T^B , and for similar reasons:

Proposition 2: T^S is increasing in μ and decreasing in σ .

Finally, a key result is that the two thresholds can be unambiguously ranked:

Proposition 3: T^S is higher than T^B provided only that μ is strictly positive.

There is always some range of trade costs at which half the world's manufacturing firms are

content to locate in each country but a single firm on its own has no incentive to exit in the core or enter in the periphery. The Appendix proves the result and shows that it is driven by the cost linkage: if workers cared only about nominal wages, perhaps because they consumed only in their home country, the break and sustain thresholds would coincide.¹¹

In the light of Propositions 1 to 3, Figure 2 illustrates how the number and types of equilibria vary with the level of trade costs T . (FKV call this diagram evocatively a "tomahawk bifurcation".) The vertical axis measures λ , the share of the world manufacturing labour force located in country 1; solid and dotted lines denote stable and unstable equilibria respectively. Recall that the countries are *ex ante* identical, so at every level of trade costs there exists a symmetric diversified equilibrium. However, from Proposition 1, it is unstable for trade costs below T^B . Similarly, from Proposition 2, the core-periphery equilibrium cannot be sustained for trade costs above T^S . At or below T^S , there are *two* stable core-periphery configurations: manufacturing may agglomerate in either country. Finally, from Proposition 3, there are three stable equilibria between T^B and T^S , one diversified and two agglomerated.¹² It is this gap between T^B and T^S which opens up the possibilities of multiple equilibria and hysteresis that are the distinctive predictions of the economic geography literature and which are highlighted in the book. Within the range where $T^B < T < T^S$, both agglomeration and diversification are possible equilibria, so history and policy have a potential role in influencing which equilibrium prevails.

Figure 2 here

Now we are ready to apply the model. For example, falls in trade costs from initially high levels will lead to catastrophic agglomeration once T^B is reached; subsequent reversals of the trend in trade costs (as long as they remain below T^S) will not restore the diversified equilibrium; and whichever country is lucky enough to first acquire the manufacturing agglomeration will hold on to its initial advantage indefinitely. Complicating the model by

¹¹ Hence international capital mobility, at least as normally modelled in trade theory, with capital reallocating internationally in response to differences in *nominal* rentals, would not give the same richness of results: T^B and T^S would coincide. See Baldwin (1999a).

¹² Could there be other stable equilibria? The inherent symmetry of the model makes it unlikely, but no analytic proof is provided in the book, so we have to rely on simulation results to be sure.

adding other exogenous variables clearly opens the door to many more such "stories of breathtaking scope", or, if you prefer, interesting comparative-static applications.

2.4 Intermediate Inputs and Agglomeration

We have seen that models which exhibit a propensity to agglomerate require increasing returns and transport costs. But we have also seen that these features are not enough: some mechanism which actually brings about agglomeration is also needed. So far, international labour mobility has been the mechanism assumed. An alternative route is to allow for inter-industry linkages. The potential importance of intermediate inputs in models of monopolistic competition, and the relative ease of incorporating them into the Dixit-Stiglitz model, have been clear since Wilfred Ethier (1982). It turns out that they provide another channel for agglomeration and lead to an analysis which is almost identical to that of labour mobility.

Assume that labour is now permanently country-specific, though it is used in both agriculture and industry. Manufacturing also uses an intermediate good, which is a CES aggregate of the output of all manufacturing firms (both domestic and foreign). For convenience, assume this aggregate has the same elasticity of substitution as the manufacturing sub-utility function M , so the price index for intermediates facing producers in country 1 is just P_I (given as before by (9)). Intermediates are then combined with labour to form a Cobb-Douglas composite input, with unit cost W_I and intermediate cost share α :

$$W_I = w_1^{1-\alpha} P_I^\alpha \quad (15)$$

Hence, production costs depend positively on the local price index P_I .

The second way in which the model is altered is that the demand for each variety comes not only from consumers but also from firms. So, for example, the demand from country-1 consumers μY_I which appears in every demand function must be replaced by total country-1 expenditure on manufactures E_I , given by:

$$E_I = \mu Y_I + \alpha n_I p_I q_I \quad (16)$$

Local consumers spend μY_I as before; in addition the n_I local firms spend a fraction α of their revenue $p_I q_I$ on intermediates.

It is now straightforward to write down the full model. The cost of the composite input W_I replaces the wage rate in the pricing equation (6); but otherwise the equations for

individual firms derived in Section 2.1 are unchanged. And total country s expenditure, E_s , replaces consumer expenditure μY_s in the demand function (8) (and in the corresponding demand function facing firms in country 2); but otherwise the equilibrium equations derived in Section 2.2 are unchanged. (In particular, the price index equation (9) still holds.) The model can then be solved in exactly the same way as in Section 2.3. Matters are particularly simple if we make the further assumption that the agricultural wage rate is always fixed. (The consequences of relaxing this are considered in Section 3.) This in turn fixes the manufacturing wage and gives the model a partial equilibrium flavour, but the pay-off is considerable. FKV show that the results are very similar to those of the model with internationally mobile labour. In particular, the expressions for the break and sustain levels of transport costs are identical to (23) and (29), except that α , the cost share of intermediate goods, replaces μ , the budget share of manufactures. This is because, as equations (15) and (16) show, it is now α which determines the magnitude of the cost and demand linkages. Hence, the analysis of the model (whether diagrammatically or algebraically) proceeds essentially as in the internationally mobile labour case.

3. Extending the Model

Almost every chapter of the book relies on the basic model just outlined. From among many interesting extensions, let me highlight three.

Generalising the model to many locations might seem impossible without resort to simulation. However, the authors succeed in deriving a local analytic result for a multi-location extension by borrowing a result from theoretical biology due to the mathematician Alan Turing. Assume a circular world in which a continuum of locations, each identical to one of the two countries in the last section, is spread evenly around the circumference. (FKV call this the "racetrack" economy.) Start from a uniform distribution of manufacturing (the "flat earth" equilibrium). Whether this symmetric equilibrium is broken depends on the stability of a dynamic labour-adjustment mechanism similar to that used to calculate the threshold T^B above. Turing's result can now be invoked to derive an analytic expression for the (algebraically) largest eigenvalue of the linearised dynamic system. If this is positive, the symmetric equilibrium will be broken and the economy will move towards a pattern of agglomeration. The agglomerated locations will be spread evenly around the circle in a pattern which can be represented by a sine curve, and the Turing result allows their frequency

to be calculated. It turns out that exactly the same forces which apply at the two-country level determine this frequency: agglomerated equilibria will be fewer and located more sparsely around the circle the lower are transport costs, the higher the share of spending on manufacturing, and the lower the elasticity of substitution in demand. This shows that the forces found to be crucial in the two-country case apply much more generally.

A second extension retains the assumption of a continuum of locations but assumes that they are spread along an infinite line rather than around the circumference of a circle. Now, population and hence economic activity are not spread evenly along the line. Rather the population is initially large enough to support only a single location for manufacturing (which naturally can be called a "city"). A final difference is that the only variable factor, labour, is equally suited to both manufacturing and agriculture. Hence both sectors must pay the same wage in equilibrium. All this makes the set-up identical to that commonly used in formal models of urban economics. Now add a manufacturing sector such as that of Section 2 potentially active at each point, and ask how the equilibrium pattern of manufacturing location changes as the population grows. It turns out that the details can only be solved by simulation, but they can be expressed in terms of a neat technical trick. This calculates a "market potential function" which gives as a function of distance from the city the *virtual* manufacturing wage at each location (i.e., the wage at which production there just breaks even) relative to the current equilibrium wage. This curve is always downward-sloping in the neighbourhood of an existing city: increasing returns plus transport costs mean that every city casts a "shadow" within which it is not profitable for other manufacturing to locate. But it need not be monotonically decreasing and may hit unity at other locations, in which case manufacturing (and so cities) will be established there. Simulations show that, as population grows, the initially monocentric city system gives rise to increasing numbers of cities which emerge to produce for the hinterland. Finally, the model is extended to allow for many industries, differing in terms of their substitution elasticities and transport costs, following Fujita, Krugman and Tomoya Mori (1999). Again, simulations show that the extended model implies a hierarchy of cities: a central city in which all industries are located, and an evolving series of subsidiary cities which attract "lower-order" industries (meaning those with higher transport costs and/or higher elasticities of substitution). This pattern of urban growth mimics real-world developments, and provides a theoretical rationale for a hierarchical or "central place" urban system of the kind described by location theorists such as Walter Christaller

(1933) and August Lösch (1940).

A third extension takes the intermediate goods variant of the basic model and extends it to full general equilibrium by allowing wages to be endogenously determined (for example, by diminishing returns to labour in agriculture, because land is in scarce supply). Analytic results for this case are not available, but simulations reveal a common feature, independent of the details of different models. As before, industry is dispersed and wages are equalised internationally at high transport costs; and falls in transport costs lead to agglomeration. The new feature is that agglomeration drives up wages in the core country, which introduces a disincentive for industry to locate there. As transport costs fall further, this disincentive comes to dominate, so industry once again becomes dispersed and (when transport costs reach zero) wages are equalised internationally. This suggests a general phenomenon: there is likely to be a U-shaped relationship between the level of manufacturing transport costs and the propensity to agglomerate. Clearly many factors other than rises in agricultural wages (such as housing as in Helpman (1997)) can cause diseconomies of agglomeration and so lead to dispersal of manufacturing and convergence of per capita incomes in what Krugman and Venables (1995) call the "globalisation" phase of the world economy.

Following Puga and Venables (1996), this model is then extended to allow for many countries and industries, and to consider the effects of exogenous growth (uniform increases in the efficiency of all factors) rather than exogenous falls in transport costs. Simulations of the resulting pattern of industrialisation show that manufacturing gradually spreads to other countries, with industries that are more labour-intensive and have weaker inter-industry linkages leading the way (since they benefit more from lower wages in countries which have not yet industrialised and benefit less from agglomeration in the initial core country).

4. Limitations of the Model

Models are unrealistic by definition. Simplifying assumptions are a virtue when the goal is to illustrate possibilities that previously could not be shown to hold. However, when it comes to applying a model to other questions, the assumptions may not be so helpful. In this section, I argue that the model used throughout the book has a number of special features which make it less suitable for addressing some issues. In particular, I want to single out the ways it treats increasing returns, firms' strategies, transport costs and, finally, space itself.

Increasing returns are intrinsic to the model of course and the cost function (5) is

probably the most natural way of specifying such a technology in a simple tractable way. However a quirk of the Dixit-Stiglitz model is that standard measures of the degree of returns to scale turn out to depend not on the cost parameters F and c but rather on the elasticity of substitution. Though σ starts as a taste parameter it ends up as an index of returns to scale. There is nothing wrong with this in principle, but it raises issues in interpreting the results and also highlights the shadowy role played by individual firms in the model.

Figure 3 here

Figure 3 illustrates how different values of σ affect the equilibrium. Consider first the case where σ is high. As an assumption this can be expressed in various equivalent ways: different varieties are closer substitutes; products are less differentiated; there is less preference for diversity. As a consequence, equilibrium is at a point such as A. From (7), firm output is relatively high, so fewer varieties are produced with higher output of each. From (6), price is close to marginal cost, so the demand and marginal revenue curves are flat and close together. By contrast, if σ is low, equilibrium is at a point such as B. Now there is a greater preference for diversity, so more varieties are produced with less output of each. To cover the fixed costs of shorter production runs, the price-marginal-cost margin must be high, with the demand and marginal revenue curves relatively convex and far apart.

Since the only underlying difference between the two cases is the value of σ , it can be interpreted as "an inverse index of equilibrium economies of scale" (Krugman (1991), p. 490). (In equilibrium, $\sigma/(\sigma-1)$ equals the ratio of the composite factor's marginal product to its average product, or one over the output elasticity of total costs.) But this terminology risks suggesting to the unwary that the theory of the firm embodied in the model is richer than in fact it is. It is not that industries with a higher value of σ have access to a different technology, rather that the pattern of demand encourages them to produce at a different scale. Better surely to say that returns to scale are more highly exploited at A than at B.

All of this is obscured by a particular normalisation used for much of the book. Units are chosen to set the fixed and marginal cost parameters F and c equal to μ/σ and $(\sigma-1)/\sigma$ respectively. This is harmless for local analysis, and may help in deriving analytical expressions for the break and sustain levels of transport costs, equations (23) and (29). (Though as I have shown in Section 2, it can be dispensed with.) But in non-local comparisons it seems both unnecessary and undesirable. Unnecessary since one of the points

of simulations is surely to avoid reducing the dimensionality of the parameter space. Undesirable because the effects of greater substitutability in demand (a rise in σ) cannot be distinguished from the effects of a higher ratio of variable to fixed costs. Changes in one are exactly offset by changes in the other such that (from (7)) both firm output q and the unit input requirement $F+cq$ are always equal to μ . This in turn can lead to questionable interpretations. For example, bakery shops are given as an illustration of a high- σ industry (p. 193). Figure 3 shows that this makes little sense in the absence of normalisation.

This discussion reflects a worrying fact: like the traditional competitive model, the Dixit-Stiglitz model has almost nothing to say about individual firms. Except for the fact that it incorporates increasing returns, the new economic geography has industrial organisation underpinnings which are very rudimentary. In particular, the assumption of free entry - a perfectly elastic supply of firms at all locations - allows almost no role for strategic interactions between firms.¹³ As a result, while costs may be fixed they are never sunk, so firms, industries and even cities are always free to move. Footloose cities seem particularly odd. They might make sense in a long-run model, but even then not in the context of a growing economy with given technology such as that explored in Part III of the book. The authors note (p. 213) that introducing urban infrastructure would avoid the problem but that begs the question of exactly how to model it other than in an *ad hoc* ways.

Completely footloose firms also sit oddly with the discussion of the emergence of industrial clusters or a "manufacturing base" in Chapter 16, since the logic of the model is that such clusters are themselves footloose. At a superficial level the model is of course prejudiced in favour of a manufacturing base. The prior existence of other firms at one location lowers the costs of entry for a new firm there. Yet this feature alone does not seem to capture many of the issues which arise in discussions of industrial policy. Hysteresis or lock-in arises only from pecuniary (market-mediated) externalities in the new economic geography, not from technological externalities or local public goods (such as R&D

¹³ There is only one mention of strategic behaviour in the book: a brief reference (on page 52) to Alasdair Smith and Anthony Venables (1988), who show that an increase in market size has a pro-competitive effect if firms adopt either Cournot or Bertrand conjectures, so enhancing the home-market effect. This is an interesting point, but like Pierre-Philippe Combes (1997), it allows only for *post-entry, one-shot*, strategic behaviour. When oligopolistic firm conjectures are combined with free entry and exit, it is the latter which drives most of the results.

spillovers). Further, the myopic Chamberlinian firms used throughout the book cannot engage in industrial strategies to shore up their position. They cannot make strategic commitments to create artificial barriers to entry, nor vertically integrate to internalise the externalities arising from the combination of intermediate inputs with increasing returns. And of course out-sourcing or cross-border horizontal mergers in response to changes in trade policy, technology or market size are not allowed. The absence of these considerations makes the model less relevant to current debates on industrial location than it at first appears.

Transport costs are intrinsic to the model: if they are low enough, there are no barriers to agglomeration (though congestion effects may make it unprofitable); if they are high enough, there can be no departure from diversification. Such elementary insights should be robust to alternative specifications of transportation technology. But as we have seen, the detailed results in the book rely on the assumption of "iceberg" transport costs, usually attributed to Paul Samuelson. This amounts to assuming that the technology used to transport a good is identical to that used to produce it. This works for icebergs; also for grain if the main cost of transporting it is the grain consumed on the way by the horses pulling the wagon. (FKV note (p. 59) that von Thünen gave this example in 1826, so anticipating Samuelson by well over a century.) But it is obviously a very special case. An older literature in trade theory tried to model the transportation sector explicitly in a Heckscher-Ohlin framework, treating it just like any other industry, with an endogenous capital-labour ratio. (See Rodney Falvey (1976) for example.) That approach never led to simple or easily summarisable results, and is now largely forgotten. Perhaps it is due a revival. Even more important, in a field which emphasises the importance of increasing returns, would be to allow for fixed costs in transportation. Of all industries, it seems to be characterised by very high ratios of fixed to variable costs, which is presumably why it is typically either publicly owned or highly concentrated. This is particularly important if transport costs are interpreted broadly to include the communications and other costs associated with trade, which are likely to exhibit network externalities. (See Richard Harris (1995).) Overall, the models in this book make interesting predictions about the effects of exogenous changes in per-unit transport costs. However, they are not well suited to examining positive questions such as the links between transportation, technology and agglomeration (will what Danny Quah (1997) has called the increasing "weightlessness" of GNP facilitate backyard capitalism?); or normative questions such as the optimal level of provision of transport infrastructure.

A final point about transport costs is that, as noted earlier, the basic model assumes that they apply only to manufactures, and that agricultural goods (which really means goods produced under constant returns) are transported freely. Donald Davis (1998) criticises this on two counts. He makes the empirical point that real-world transport costs appear to be at least as high for the latter, and he shows theoretically (in the model of Section 2.2 above) that this neutralises the home-market effect. Chapter 7 of the book derives similar results (without referring to Davis) but gives them a totally different spin. Whereas Davis concludes that there is "no compelling argument [...] that market size will matter for industrial structure", FKV note that a reduction of agricultural transport costs may trigger agglomeration. So, relatively low agricultural transport costs are either a necessary and implausible condition for agglomeration, or a source of yet more "stories of breathtaking scope": take your pick.

Geography is about space, and as the authors note (p. 325), a strength of the book is that it integrates the continuous approach to space of location theory with the discrete approach of trade theory. However, they never mention an offsetting drawback: space is almost always one-dimensional, whether along a line or (as in the "racetrack" economy) around a circle. (The sole exception is an interesting digression in Chapter 13 showing how forks in a valley are natural locations for agglomeration). Perhaps it will prove possible to extend the Dixit-Stiglitz approach to a two-dimensional plain, though the complexities involved (such as the overlapping circular "shadows" cast by competing cities) suggest that the results may be long on trigonometry and short on elegance. Even if this can be done, the worry will remain that omitting strategic considerations misses much of importance. Strategic location issues in two dimensions are unresolved, but no references are given to the extensive literature on this topic. (Fujita and Thisse (1996) is a recent overview.)

Apart from dimensionality, it is also necessary to specify the spatial unit of analysis. The book deals in turn with regions, cities and countries, but there is nothing intrinsic to the models which conclusively identifies these units. In particular, FKV assert, with little discussion, that the distinction between regions and countries coincides with the difference between factor mobility and intermediate inputs as the trigger for agglomeration. Ultimately, this is an empirical judgement and it is easy to think of counter-examples. (Labour mobility from some small to large European countries is considerable; and surely Silicon Valley, if it fits the new economic geography paradigm at all, reflects a regional concentration of intermediate supply rather than of migration-induced final demand for computer hardware and

software?) More generally, the new economic geography, with its emphasis on pecuniary externalities, faces competition in the market for ideas from alternative theories at both ends of the spatial spectrum. At the local level, pure externalities such as knowledge spillovers may well prove to be the determining force behind most agglomerations. Krugman (1991) could dismiss technological externalities as a satisfactory basis for modelling spatial agglomeration with the quip "How far does a technological spillover spill?" But subsequent empirical work by David Audretsch and Maryann Feldman (1996) among others has shown that their geographic reach can be measured, and that they are typically very localised. As for the national level, it is not obvious that agglomeration effects are nearly so crucial there. It is possible for example to give a factor-endowments-based explanation for the pattern of East Asian development, which is at least as plausible as Puga and Venables's agglomeration story. (See Jaime Ventura (1997).) And, as I note in Neary (1999), calibration studies tend to find that the extent of intermediate input usage is not sufficient to yield strong agglomeration effects at the national level. Only further theoretical and empirical work will tell at what level of spatial analysis, if any, the forces emphasised in this book are the dominant influences on agglomeration and economic development.

5. Old Versus New Economic Geography?

"With friends like this, who needs enemies?" the authors may wonder if they have read this far. But in fact the new economic geography has real enemies out there; or, more prosaically, reputable academic critics who question the whole research program rather than just the details of its implementation. In particular, economic geographers of a more traditional kind, typically based in geography rather than economics departments, have not taken kindly to the self-proclaimed arrival of a new paradigm. Ron Martin (1999) summarises their objections and concludes that what he calls the 'new geographical economics' "represents a case of mistaken identity: it is not that new, and it most certainly is not geography." No doubt this hostility is partly induced by Krugman's prose style: guaranteed to influence people but not necessarily to win friends, not at least from the ranks of those wedded to different approaches. But there are genuine issues at stake. Is this just a turf war, or have economists taken a bridge too far?

Martin's central objection to the new economic geography is that its mathematical formalism leads it to neglect "real places". Economic geographers of whom he approves

avoid abstract theory, and focus instead on the contingency and particularity of real communities going about "the ordinary business of life". This methodological position reflects the "movement away from logical positivism" that took place in geography in the late 1970s. Though not averse to theory in general, even to economic theory provided it is "heterodox", economic geographers prefer to build explanations "'from below', often relying upon close dialogue with individual agents and organisations, and linking this 'local' knowledge with wider, larger stylised facts and conceptual frameworks."

Economists call these case studies, and regularly bemoan the fact that not enough get done. But the key issue is whether such studies can ever be "theory-free". Most economists would argue not, and would follow FKV in preferring an explicit model to inform (though not blindly predetermine) a case study rather than letting the facts speak for themselves. It is ironic that Krugman should be the target of such criticism, since he has shown unusual sympathy with heterodox traditions. In the same book (Krugman 1995) where he takes "old" economic geography to task, he also argues that increased formalism in economics as a whole held back progress in development and spatial economics, since pre-Dixit-Stiglitz no clean way of modelling imperfect competition was available. This may seem like trying to have it both ways, but it reflects the attitude of most working economists to what FKV call (p. 79) "aggressively unrealistic" models. They are, as the Roman poet Propertius said of his lover, difficult to live with, impossible to live without.

As for the "new" in "new economic geography", FKV make explicit their intellectual debt to location theory and regional science.¹⁴ But they also make a persuasive case that they have extended their insights significantly by giving them behavioral underpinnings. It is one thing to make assertions about processes of cumulative causation leading to agglomeration, or to describe urban systems as central-place hierarchies. It is quite another to explain such patterns as the outcome of decentralised self-interested behaviour.

Given mainstream economists' impatience with methodology, it is a fair bet that debates between old and new economic geographers will have little influence on the direction of future economic research. The genie is out of the bottle, and FKV are sure to inspire

¹⁴ Mention could also be made of Hesham Abdel-Rahman (1988), who used the Dixit-Stiglitz specification to explain agglomeration at the city level, and with Ping Wang (1995) developed a model of a core-periphery urban system.

further theoretical developments.¹⁵ Such work is important, but probably hard empirical work is even more so if non-believers are to be persuaded that the new approach represents a promising way forward. And, on a more hopeful note, perhaps there is potential common ground in the process of empirical implementation, scope for integrating the theoretical insights of the new economic geography with the detailed contingent knowledge of the old.

6. From Theory to ... What?: Empirical and Policy Applications

FKV do not present any empirical work of their own.¹⁶ Nor do they discuss the policy implications of the new economic geography at length. But it is in these areas that the ultimate usefulness of the field will most likely be determined.

As far as empirics are concerned, the increasing availability of large data sets has encouraged a recent explosion of studies showing the relevance of location to many economic phenomena.¹⁷ Among many examples, see the work of Audretsch and Feldman on local R&D spillovers mentioned above; John Gallup and Jeffrey Sachs (1998) who show the influence on living standards of geographical features such as latitude and length of coastline; and the revival of interest in the "gravity equation" which emphasises the role of distance in explaining trade patterns. (See Alan Deardorff (1998) for a discussion of its theoretical underpinnings and further references.) While these studies have developed independently of

¹⁵ Interesting recent contributions not already mentioned include Gianmarco Ottaviano, Takatoshi Tabuchi and Jacques-François Thisse (1999), who develop an agglomeration model with linear rather than Dixit-Stiglitz preferences: they find that the basic insights remain intact, though incorporating transport costs enriches the model's outcomes; and Baldwin (1999a) and Philippe Martin and Gianmarco Ottaviano (1999), who embed economic geography in explicit growth models.

¹⁶ The only empirical chapter in the book is a digression on the sizes of cities and the difficulties of accounting for the empirical regularity known as "Zipf's Law" or the "rank-size rule": to a remarkably close approximation, data on city size tend to show that the r 'th largest city has $1/r$ of the total population. This pattern is not exhibited by the simulations of the authors' model discussed in Section 3 above, although Xavier Gabaix (1999) has shown that it can be explained by assuming that both the growth rate and the variance of the growth rate of a city's population are independent of city size.

¹⁷ This is not to mention the many empirical studies of regional growth and convergence. At the risk of descending into semantics, I am inclined to agree with Martin (1999) that these do not constitute economic geography proper (either old or new) since for the most part they merely take regions as units of analysis without specifically considering issues of relative location (such as transport costs, market access, etc.).

the new economic geography, there is also a small but growing body of empirical work which builds directly on it. Most of the papers are unpublished, so a full evaluation is premature. Nevertheless some patterns can be identified.

One group of papers takes seriously the injunction of Edward Leamer and James Levinsohn (1995): "estimate do not test". Stephen Redding and Anthony Venables (2000) use cross-section data on 103 countries to estimate the intermediate-goods model of Section 2.4, with incomes and firm numbers exogenous. They use the estimated coefficients from a gravity equation for trade flows to construct measures of demand and cost linkages;¹⁸ these in turn are used to explain cross-country variations in incomes and in manufacturing wages and prices. Remarkably, they find that demand and cost linkages alone explain up to 70% of the cross-country variation in per capita income and up to 50% of the variation in manufacturing wages. Gordon Hanson (1998) uses Helpman's (1997) variant of the mobile-labour model of Section 2.3 to estimate the determinants of ten-year changes in wages from 1970 to 1990 in all 3,075 counties of the continental United States. Including distance-weighted measures of housing stock and wages in adjacent states, as the theory suggests, improves the fit over an *ad hoc* specification (which only includes distance-weighted incomes). The estimated equation explains over 20% of the variation in the dependent variable and yields parameter estimates which are reasonably consistent with theory. Finally, whereas the other two papers use distance to proxy for transport costs, Pierre-Philippe Combes and Miren Lafourcade (2000) use a very detailed data set on actual transport costs in mainland France. They estimate a specification based on a model similar to FKV's to explain employment patterns in 64 sectors (including services as well as manufacturing) across 341 employment areas and find that, when allowance is made for intermediate inputs, transport costs have a significant impact on specialisation patterns in nearly all sectors.

These papers show the usefulness of the economic geography model in guiding empirical work. However, they do not test the model against plausible alternatives, and the specifications they estimate are not necessarily inconsistent with other perspectives. A different set of papers attempts to discriminate between economic geography and alternative theories. Sukkoo Kim (1995) uses a long data set (from 1860 to 1987) on U.S. manufacturing employment to measure trends in regional specialisation and industrial localisation. He argues

¹⁸ This empirical technique, discussed further below, seems to originate with Leamer (1997).

that the evidence is inconsistent with an important role for external economies, since industry localisation patterns are negatively correlated with characteristics associated with external economies. In particular, high-tech sectors are less localised. Kim does not discuss the distinction between technological and pecuniary externalities, but his findings are also inconsistent with the latter provided we interpret high-tech sectors in the economic geography framework as sectors with higher shares of intermediate inputs and higher "equilibrium economies of scale", which the model predicts should be more agglomerated. By contrast, Kim finds that average plant size (his proxy for scale economies) explains industry localisation patterns over time (recall that the new economic geography model relies on increasing returns but has nothing to say about firm or plant size); while raw material intensity (his proxy for Heckscher-Ohlin endowment effects) explains localisation patterns across industries.

More encouraging results for the economic geography model come from a series of papers by Donald Davis and David Weinstein. They focus on the home-market effect, which as we have seen is a key building block of the model, though not sufficient for it. They note that this effect does not arise in models where trade is determined by comparative advantage. In such models, a positive shock to home demand for a good only affects home production if it raises price; but the price increase calls forth extra imports and so dampens the effect of demand on home production. By contrast, equation (12) shows that in a monopolistically competitive model with transport costs, a positive shock to home demand has a magnified effect on home production. This difference suggests a test which can discriminate between the two sets of models. In a first application to OECD data, Davis and Weinstein (1996), they found no evidence for home-market effects: above-average home demand for a sector was typically associated with a less-than-proportional excess of local production above the world average. However, in subsequent work, Davis and Weinstein (1998), they interpret "home" demand more broadly: as in Redding and Venables, they use the estimated coefficients from a gravity equation for trade flows to measure *effective* home demand (including demand from trading partners, inversely weighted by distance). This time the home-market effect shows up as important for between a half and two thirds of manufacturing output (with comparative advantage determining trade patterns for the remainder). These results complement their finding in Davis and Weinstein (1999) that home-market effects are important in explaining the structure of production across Japanese regions.

To summarise, the overall thrust of these papers is support for Krugman (1980) from Davis and Weinstein, but rejection of Krugman (1991) by Kim. But much more work needs to be done. In particular, I am not aware of any cross-industry tests of the model's predictions about the effects of the "three T's": tastes, technology and transport costs (all potentially measurable variables). As we saw in Section 2, agglomeration is more likely in industries with higher demand (both final and intermediate), greater product diversity and lower transport costs. Further work is needed to refine these implications and to devise plausible tests of them against real rather than straw alternatives. If the theory fails to find support from this process of empirical testing, it will not fade away but it will be much less persuasive. Of course, falsifiability at least in principle is a strength rather than a weakness of the model. The ultimate defense of the new economic geography is one that FKV themselves might prefer not to make: it could be wrong.

With the empirical verdict not yet in, is it too early to draw policy conclusions? The field's potential to throw light on policy is undoubtedly part of its appeal. This is especially true in Europe, where the development of the theory has coincided with debates on the effects of moves to deepen European Union, through greater economic and monetary integration, and to widen it, through enlargement. Fears that these developments may de-industrialise the periphery have so far proved groundless: in the immediate aftermath of the January 1999 launch of European Monetary Union, the fastest-growing countries in the European Union have been small (the Netherlands), peripheral (Spain) or both (Finland and Ireland). Nevertheless, it is clear that Europe is at a different stage of economic-geographical evolution from the U.S. Kim (1995) shows that U.S. inter-regional specialisation reached its peak during the inter-war years but (as of 1987) has fallen continuously and substantially since. By contrast, Karen Helene Midelfart-Knarvik et al. (2000) show that international specialisation in the EU increased from 1980-83 to 1994-97. In the light of this relevance to topical issues, it is hardly surprising that the new economic geography has stimulated more academic interest in Europe than in the U.S., as a comparison of the recent CEPR and NBER trade programmes shows. Moreover, the approach has already been used fruitfully as a background to policy discussions, as in Pontus Braunerhjelm et al. (2000). But using the models to evaluate policy options explicitly is surely premature.

The key problem is that the policy implications of the basic core-periphery model are just too stark to be true. The model turns Sartre's "Hell is other people" on its head:

agglomeration is unambiguously good for you. Because the cost of living is lower in the core, it is always better to live there than in the periphery, with the level of utility in a diversified economy lying in between. Even when congestion costs are added to the model, utility is always at least as great in the core as in the periphery and is usually higher. Faced with multiple equilibria which have a clear welfare ranking, it is tempting to suggest a role for government in "picking equilibria". This in turn may encourage a new sub-field of "strategic location policy", perhaps drawing on fifteen years' work on strategic trade policy, which, as James Brander (1995) and Peter Neary and Dermot Leahy (2000) argue, has produced much interesting theory but no simple robust rules to guide policy making. All these are temptations to be resisted, since they take too literally the neat structure of the model, and ignore the econometric difficulties in estimating the non-linear, non-monotonic relations it predicts. In any case, greater industrial specialisation, even if it takes the form of more agglomeration, is not the same as de-industrialisation once we allow for many sectors. And, of course, the neglect of strategic issues which I discussed in Section 4 is particularly relevant in this context. No harm then that FKV are mostly neutral on the applicability of the models to policy (except for some tentative interventionist thoughts in the chapter on industrial clustering).

6. Conclusion

In this paper, I have taken the appearance of FKV as an opportunity to review the "new" economic geography. It is not the only approach to location and agglomeration which economists have taken. Many authors such as Brian Arthur (1986) and Robert Lucas (1988) have theorised about the role of regions and cities in economic development. But no other body of work does quite the same thing as the new economic geography: explain agglomeration in a theoretical framework which is tractable, has solid micro foundations, and makes testable empirical predictions. So, to paraphrase Robert Solow (1962), everyone should read this book, or at least encourage their students to do so!

Remember though that Solow's remark was made about the two-sector growth model, as emblematic of the 1960s as mini-skirts or the Beatles, though not as long-lasting. Will the new economic geography prove more durable? I suspect that it will, though maybe not as a distinct field. Instead, I am tempted to suggest that it will survive as "merely" another simple general equilibrium model, supplementing the trade theorist's tool-kit; to quote Solow

again, another "general equilibrium model of matchbox size" (since even a continuum of identical matchboxes arranged symmetrically around a circle is, well, just a matchbox).¹⁹ Saying this risks sounding disparaging (and falls short of the authors' ambitions). But it is high praise in my view.

The new economic geography is firmly in the tradition of trade theory in at least two respects. First, it follows Ohlin in viewing interregional and international trade as different manifestations of the same influences. National boundaries are often coterminous with some economic forces, such as public policies, language and institutions. But most influences on trade and industrial patterns operate similarly in large regions and small countries, and so the study of them should use the same analytic tools. Ohlin's view was not ignored - there are many examples of trade models applied to regional issues - but as long as trade theory remained tied to the competitive paradigm its relevance was questionable. (Though as Krugman (1999) notes, Ohlin himself gave an important role to increasing returns as a determinant of trade patterns.) In stressing the relevance to regional issues of models derived from trade theory, Krugman has not so much created a new sub-field as extended the applicability of an old one.

Second, both the strengths and weaknesses of the new economic geography reflect the fact that it follows the tradition of focusing on a single cause of trade and location: technology in Ricardo, factor endowments in Heckscher-Ohlin, monopolistic competition in Krugman (1979), or oligopolistic competition in Brander (1981). Models such as these, which strip away the superfluous to focus on a single feature, are essential for understanding the world. Adding pecuniary externalities to the list and working out their implications is a significant achievement. But no one monocausal model can hope to capture the complexities of any applied problem, certainly not a model where space is one-dimensional, firms are identical and infinitesimal, and every function in sight is either constant-elasticity or a rectangular hyperbola. So, hold on the tee-shirt, skip the movie, but do read this book, possibly the best on interregional and international trade and location since Ohlin.

¹⁹ Those who remember the lessons of Ronald Jones (1965) will have noted how the exposition in Section 2 is simplified by using standard techniques for analyzing simple general equilibrium models: logarithmic differentiation, a focus on factor-market equilibrium conditions, etc.

Appendix: Proofs of Propositions²⁰

Proof of Proposition 1: Totally differentiating the model at the symmetric equilibrium gives six equations in eight variables, \hat{q} , \hat{Y} , \hat{P} , \hat{p} , \hat{n} , \hat{w} , \hat{L} and $\hat{\omega}$. Equation (10) gives the change in the home price index, and equation (11) gives the change in demand facing an individual home firm. Firms always maximise profits, so from (6) prices and wages move together:

$$\hat{p} = \hat{w} \quad (17)$$

The full-employment condition (13) implies that the labour force and the number of firms in the home country must always move together:

$$\hat{L} = \hat{n} \quad (18)$$

From the definition of the real wage ω :

$$\hat{\omega} = \hat{w} - \mu\hat{P} \quad (19)$$

Finally, equation (14) gives the change in GNP.

As noted in footnote 7, two alternative approaches can be taken to testing whether the symmetric equilibrium is stable. FKV assume that firms are always in equilibrium whereas migration takes time. In my notation, they set \hat{q} equal to zero (since output is always at its equilibrium value, given by (7)), take \hat{L} as exogenous, and solve for $\hat{\omega}/\hat{L}$. Here, following Puga (1999), I assume instead that workers are always in equilibrium whereas entry and exit take time. Hence, I set $\hat{\omega}$ equal to zero (migration ensures that the real wage is the same in both countries) take \hat{n} as exogenous, and solve for \hat{q}/\hat{n} . (From footnote 5, changes in profits are monotonic in changes in q .) Both approaches yield the same stability condition.

To solve for \hat{q}/\hat{n} , first use (10) and (14) to eliminate \hat{P} and \hat{Y} from (11), and then use (17) and (18) to replace \hat{p} and \hat{L} by \hat{w} and \hat{n} respectively. This yields:

$$\hat{q} = (\mu - Z)Z\hat{n} - [\sigma - \{\mu + (\sigma - 1)Z\}Z]\hat{w} \quad (20)$$

The first term on the right-hand side gives the first result quoted in the text: with equal nominal wages in the two countries ($\hat{w}=0$), equilibrium is stable if and only if μ is less than

²⁰ Propositions 1 and 2 first appear in Krugman and Venables (1995) and Krugman (1991) respectively; Propositions 3 and 4 appear to be new.

Z. Next, to impose equality of *real* wages, I need to solve for the change in the price index from (10), setting $\hat{p}=\hat{w}=\mu\hat{P}$ (using (19) with $\hat{\omega}$ equal to zero):

$$\hat{P} = - \frac{Z}{(\sigma - 1)(1 - \mu Z)} \hat{n} \quad (21)$$

Note in passing that the coefficient of \hat{w} in (20) is negative and that \hat{P} itself is negative from (21). (This is just the price index effect with real rather than nominal wages fixed.) These results confirm the assertion made in the text: the condition $\mu > Z$ underestimates the pressure toward agglomeration, since the symmetric equilibrium can be unstable even when the condition is not met. To find the full stability condition, substitute from (21) into (20) (once again setting \hat{w} equal to $\mu\hat{P}$):

$$\hat{q} = \frac{(2\sigma - 1)\mu - \{\sigma(1 + \mu^2) - 1\}Z}{(\sigma - 1)(1 - \mu Z)} Z\hat{n} \quad (22)$$

The denominator is positive since $\sigma > 1 > \mu, Z$. Hence the threshold value of Z is that which sets the numerator equal to zero. Using the definition of Z from (10), we can solve in turn for the threshold value of T :

$$T^{\mathbf{B}}_{+ -}(\mu, \sigma) = \left[\frac{\{\sigma(1 + \mu) - 1\}(1 + \mu)}{\{\sigma(1 - \mu) - 1\}(1 - \mu)} \right]^{\frac{1}{\sigma - 1}}. \quad (23)$$

The proposition follows by differentiating this expression with respect to μ and σ .

Proof of Proposition 2: Consider the initial agglomerated equilibrium. Since μ is the fraction of world income spent on manufactures, national income in country 1, Y_1 , equals its agricultural output X_A plus world manufacturing output $\mu(Y_1 + Y_2)$. By contrast, country 2 has no manufacturing, so its income Y_2 is simply X_A . Combining these equations gives:

$$Y_2 = \frac{1 - \mu}{1 + \mu} Y_1 \quad (24)$$

Income is lower in the periphery because there are no manufacturing workers there. Next, from equation (9) for the manufacturing price index, with n_2 initially zero:

$$P_2 = P_1 T \quad (25)$$

Manufacturing prices are higher in the periphery because all manufactures must be imported.

Now, consider whether there are incentives for firms to exit in the core and enter in the periphery. This will be profitable if a typical firm can sell more in the periphery than in the core. Assume to begin with that migration equalises nominal wages. Then relative demands depend on the ratio of the demand intercepts. That in the core is given by (8), which simplifies, using (24) and (25), to:

$$\phi_1 = \frac{2\mu}{1+\mu} Y_1 P_1^{\sigma-1} \quad (26)$$

As for the demand intercept in the periphery, it is given by an equation symmetric to (8):

$$\phi_2 = \mu [Y_1 P_1^{\sigma-1} T^{1-\sigma} + Y_2 P_2^{\sigma-1}] \quad (27)$$

Using (24), (25) and (26), potential demand in the periphery relative to the core is:

$$\phi_2 = H \phi_1 \quad \text{where:} \quad H \equiv \frac{1+\mu}{2} T^{1-\sigma} + \frac{1-\mu}{2} T^{\sigma-1} \quad (28)$$

Because of transport costs, a periphery firm is at a disadvantage ($T^{1-\sigma} < 1$) in serving the larger market, but has an advantage ($T^{\sigma-1} > 1$) in serving the smaller local market. The balance of advantage is ambiguous and depends on whether H is greater or less than one.

Finally, just as in the earlier discussion of the stability of the diversified equilibrium, migration equalises *real* not nominal wages. This places potential periphery firms at a further disadvantage, since they must charge higher prices to cover the higher nominal wages needed to compensate workers for the higher cost of living: $p_2/p_1 = w_2/w_1 = \omega_2 P_2^\mu / \omega_1 P_1^\mu = T^\mu$ (from (25)). Hence, the full expression for relative demands, which must equal one when deviation from the agglomerated equilibrium is just profitable, is:

$$\frac{q_2}{q_1} = \frac{\phi_2 p_2^{-\sigma}}{\phi_1 p_1^{-\sigma}} = H \left(\frac{p_2}{p_1} \right)^{-\sigma} = H T^{-\sigma\mu} = 1 \quad (29)$$

Since $T^{-\sigma\mu}$ is less than one, the level of trade costs sufficient to equalise nominal wages (i.e., just sufficient to set H equal to one) over-estimates the incentives to deviate from the

agglomerated equilibrium: at that level, real wages in the periphery are still lower than in the core. To equalise real wages requires a higher level of trade costs, defined by setting the left-hand side of (29) equal to one.

Equation (29) is trivially satisfied for $T=1$: with no transport costs, firms are indifferent between locations. At $T=1$ the left-hand side is decreasing in T , so q_2/q_1 must fall below one, but it must rise above it for high T (provided the "no-black-hole" condition is met). The second value of T at which (29) is satisfied defines a new threshold, T^S , at which agglomeration is just on the margin of being an equilibrium. The proposition is proved by totally differentiating (29) and evaluating in the neighbourhood of T^S .

*Proof of Proposition 3:*²¹ Direct comparison of the two thresholds T^B and T^S is not promising, since (29) only defines the latter implicitly. Instead, consider the left-hand side of (29), which gives the incentive to deviate from the agglomerated equilibrium, and evaluate it at the break threshold T^B . This defines a function h of μ :

$$h(\mu) \equiv H[T^B(\mu)] \cdot [T^B(\mu)]^{-\sigma\mu} \quad (30)$$

By inspection, $h(0)=1$: if there is no demand for manufactures, there is no propensity to agglomerate at any level of transport costs, so the break and sustain thresholds coincide. The proposition is proved by showing that, if there is any finite demand for manufactures (i.e., for any value of μ in the interior of the admissible region: $\mu \in (0, (\sigma-1)/\sigma)$), there is no incentive to deviate from the agglomerated equilibrium (i.e., $h(\mu) < 1$), when $T=T^B$.

Begin with the first derivative of $\ln h$ (writing H^B as shorthand for $H[T^B(\mu)]$):

$$\frac{d\ln h}{d\mu} = \frac{d\ln H^B}{d\mu} - \sigma \ln T^B - \sigma\mu \frac{d\ln T^B}{d\mu} \quad (31)$$

This is zero at $\mu=0$ (obvious for the second and third terms; refer to (34) below for the first). To eliminate the term in $\ln T^B$, proceed to the second derivative:

²¹ Many thanks to Frédéric Robert-Nicoud for spotting an error in my first proof and for pointing me in the direction of this one.

$$\frac{d^2 \ln h}{d\mu^2} = \frac{d^2 \ln H^B}{d\mu^2} - 2\sigma \frac{d \ln T^B}{d\mu} - \sigma\mu \frac{d^2 \ln T^B}{d\mu^2} \quad (32)$$

Evaluating this requires explicit expressions for $\ln T^B$ and its derivatives:

$$\begin{aligned} (\sigma-1) \ln T^B &= \ln \alpha - \ln \beta + \ln(1+\mu) - \ln(1-\mu) \geq 0 \\ (\sigma-1) \frac{d \ln T^B}{d\mu} &= \frac{\sigma}{\alpha} + \frac{\sigma}{\beta} + \frac{1}{1+\mu} + \frac{1}{1-\mu} > 0 \\ (\sigma-1) \frac{d^2 \ln T^B}{d\mu^2} &= -\frac{\sigma^2}{\alpha^2} + \frac{\sigma^2}{\beta^2} - \frac{1}{(1+\mu)^2} + \frac{1}{(1-\mu)^2} > 0 \end{aligned} \quad (33)$$

where $\alpha \equiv \sigma - 1 + \sigma\mu > 0$ and $\beta \equiv \sigma - 1 - \sigma\mu > 0$ (the latter is the no-black-hole condition). I also need explicit expressions for $\ln H^B$ and its derivatives:

$$\begin{aligned} \ln H^B &= \ln \eta - \ln \alpha - \ln \beta \geq 0 \\ \frac{d \ln H^B}{d\mu} &= \frac{2\sigma(3\sigma-2)\mu}{\eta} - \frac{\sigma}{\alpha} + \frac{\sigma}{\beta} \geq 0 \\ \frac{d^2 \ln H^B}{d\mu^2} &= \frac{2\sigma(3\sigma-2)}{\eta^2} [(\sigma-1)^2 - \sigma(3\sigma-2)\mu^2] + \frac{\sigma^2}{\alpha^2} + \frac{\sigma^2}{\beta^2} \end{aligned} \quad (34)$$

where: $\eta \equiv (\sigma-1)^2 + \sigma(3\sigma-2)\mu^2 > 0$. Substituting into (32) and collecting terms gives:

$$\frac{d^2 \ln h}{d\mu^2} = \frac{2\sigma(\sigma-1)^2(3\sigma-2)}{\eta^2} - \frac{2\sigma^2(3\sigma-2)^2\mu^2}{\eta^2} - \frac{2\sigma^2}{\alpha\beta} - \frac{4\sigma}{\sigma-1} \frac{1}{(1-\mu^2)^2} \quad (35)$$

This is zero at $\mu=0$ and approaches minus infinity as μ approaches the black-hole threshold ($\beta=0$); its second term is zero at $\mu=0$ and negative otherwise; and each of the other three terms is decreasing in μ .²² Hence (35) is negative throughout the admissible region. To sum up, h equals unity and is stationary in μ at $\mu=0$, and it is concave in μ throughout. It follows that $h(\mu) < 1$ for $\mu > 0$. Hence, at the break threshold of transport costs, there is no incentive to deviate from the agglomerated equilibrium, so $T^B < T^S$, as required.

To show that the cost linkage is crucial for Proposition 3, consider two new thresholds, $T^{B'}$ and $T^{S'}$, defined respectively as the break and sustain levels of trade costs which equalise

²² This follows since η is increasing in μ and $\alpha\beta = (\sigma-1)^2 - \sigma^2\mu^2$ is decreasing in μ .

nominal rather than real wages between countries. Then:

Proposition 4: $T^{B'}$ equals T^S , and both are less than T^B and T^S , provided $\mu > 0$.

Proof of Proposition 4: From the discussion following (14), $T^{B'}$ is defined by the condition $Z = \mu$. Substituting for Z from (10) gives: $(T^{B'})^{\sigma-1} = (1+\mu)/(1-\mu)$. As for T^S , it is evident from (28) that it is defined by the condition $H=1$, or $(1+\mu)T^{1-\sigma} + (1-\mu)T^{\sigma-1} = 2$. This is a quadratic in $T^{\sigma-1}$. Solving, and ignoring the trivial root $T=1$, gives $(T^S)^{\sigma-1} = (1+\mu)/(1-\mu)$. This proves that $T^{B'} = T^S$. But we have already seen that (provided μ is strictly positive) the level of trade costs just sufficient to yield equal nominal wages *under-estimates* the incentives to deviate from the symmetric equilibrium, so $T^{B'} < T^B$; while it *over-estimates* the incentives to deviate from the agglomerated equilibrium, so $T^S < T^S$.

References

- Abdel-Rahman, Hesham M. 1988. "Product Differentiation, Monopolistic Competition, and City Size," *Regional Science and Urban Econ.* 18:1, pp. 69-86.
- Abdel-Rahman, Hesham M. and Ping Wang. 1988. "Toward a General Equilibrium Theory of a Core-Periphery System of Cities," *Regional Science and Urban Econ.* 25:4, pp. 529-46.
- Arthur, W. Brian. 1986. "Industry Location Patterns and the Importance of History," reprinted in *Increasing Returns and Path Dependence in the Economy*. W. Brian Arthur, Michigan: Michigan U. Press, pp. 49-68.
- Audretsch, David B. and Maryann P. Feldman. 1996. "R&D Spillovers and the Geography of Innovation and Production," *Amer. Econ. Rev.* 86:3, pp. 630-40.
- Baldwin, Richard E. 1999a. "Agglomeration and Endogenous Capital," *Eur. Econ. Rev.* 43:2, pp. 253-80.
- Baldwin, Richard E. 1999b. "The Core-Periphery Model with Forward-Looking Expectations," CEPR Discussion Paper No. 2085, February.
- Brakman, Steven, Harry Garretsen and Charles van Marrewijk. 2001. *An Introduction to Geographical Economics*, Cambridge: Cambridge University Press (forthcoming).
- Brander, James A. 1981. "Intra-Industry Trade in Identical Commodities," *J. Int. Econ.* 11:1, pp. 1-14.
- Brander, James A. 1995. "Strategic Trade Policy," in *Handbook of International Economics, Vol. 3*. Gene M. Grossman and Kenneth Rogoff, eds. Amsterdam: North-Holland, pp. 1395-455.
- Braunerhjelm, Pontus, Riccardo Faini, Victor Norman, Frances Ruane, and Paul Seabright. 2000. *Integration and the Regions of Europe: How the Right Policies can Prevent Polarization*, London: CEPR.
- Christaller, Walter. 1933. *Central Places of Southern Germany*, Jena, Germany: Fischer. (English translation, London: Prentice Hall, 1966.)
- Combes, Pierre-Philippe. 1997. "Industrial Agglomeration under Cournot Competition," *Annales d'Economie et de Statistique.* 45, pp. 161-82.
- Combes, Pierre-Philippe and Miren Lafourcade. 2000. "Transportation Costs and Regional Employment Inequalities: Evidence from France, 1993," mimeo., CERAS.
- Davis, Donald R. 1998. "The Home Market, Trade and Industrial Structure," *Amer. Econ. Rev.* 88:5, pp. 1264-76.

- Davis, Donald R. and David E. Weinstein. 1996. "Does Economic Geography matter for International Specialization?" NBER Working Paper No. 5706, August.
- Davis, Donald R. and David E. Weinstein. 1998. "Market Access, Economic Geography and Comparative Advantage: An Empirical Assessment," NBER Working Paper No. 6787, November.
- Davis, Donald and David Weinstein. 1999. "Economic Geography and Regional Production Structure: An Empirical Investigation," *Eur. Econ. Rev.* 43:2, pp. 379-407.
- Deardorff, Alan V. 1998. "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?", in *The Regionalization of the World Economy*. Jeffrey A. Frankel, ed. Chicago: Chicago U. Press, pp. 7-22.
- Dixit, Avinash K. and Victor Norman. 1980. *Theory of International Trade: A Dual, General Equilibrium Approach*, London: Cambridge U. Press.
- Dixit, Avinash K. and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity," *Amer. Econ. Rev.* 67:3, pp. 297-308.
- Ethier, Wilfred J. 1982. "National and International Returns to Scale in the Theory of International Trade," *Amer. Econ. Rev.* 72:3, pp. 389-405.
- Falvey, Rodney E. 1976. "Transport Costs in the Pure Theory of International Trade," *Econ. J.* 86:343, pp. 536-50.
- Fujita, Masahisa, Paul Krugman and Tomoya Mori. 1999. "On the Evolution of Hierarchical Urban Systems," *Eur. Econ. Rev.* 43:2, pp. 209-51.
- Fujita, Masahisa and Jacques-François Thisse. 1996. "Economics of Agglomeration," *J. Japanese and Int. Econs.* 10:4, pp. 339-78.
- Fujita, Masahisa and Jacques-François Thisse. 2000. "Agglomeration and Market Interaction," presented to the World Congress of the Econometric Society, Seattle.
- Gabaix, Xavier. 1999. "Zipf's Law for Cities: An Explanation," *Quart. J. Econ.* 114, pp. 739-767.
- Gallup, John Luke and Jeffrey D. Sachs with Andrew D. Mellinger. 1998. "Geography and Economic Development," in *Annual World Bank Conference on Development Economics*. Boris Pleskovic and Joseph E. Stiglitz, eds. Washington D.C.: The World Bank, pp. 127-78.
- Hanson, Gordon H. 1998. "Market Potential, Increasing Returns, and Geographic Concentration," NBER Working Paper No. 6429, February.
- Harris, Richard. 1995. "Trade and Communication Costs," *Can. J. Econ.* 28, pp. 46-75.

- Head, K., T. Mayer and J. Ries (1999): "On the pervasiveness of home market effects," mimeo., University of British Columbia.
- Helpman, Elhanan. 1997. "The Size of Regions," in *Topics in Public Economics: Theoretical and Applied Analysis*. David Pines, Efraim Sadka and Itzhak Zilcha, eds. Cambridge: Cambridge U. Press.
- Helpman, Elhanan and Paul R. Krugman. 1985. *Market Structure and Foreign Trade*, Cambridge, Mass.: MIT Press.
- Jones, Ronald W. 1965. "The Structure of Simple General Equilibrium Models," *J. Polit. Econ.* 73, pp. 557-72.
- Kim, Sukkoo 1995. "Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in U.S. Regional Manufacturing Structure, 1860-1987," *Quart. J. Econ.* 110:4, pp. 881-908.
- Krugman, Paul R. 1979. "Increasing Returns, Monopolistic Competition and International Trade," *J. Int. Econ.* 9:4, pp. 469-79.
- Krugman, Paul R. 1980. "Scale Economies, Product Differentiation and the Pattern of Trade," *Amer. Econ. Rev.* 70, pp. 950-9.
- Krugman, Paul R. 1991. "Increasing Returns and Economic Geography," *J. Polit. Econ.* 99:3, pp. 483-99.
- Krugman, Paul R. 1995. *Development, Geography and Economic Theory*, Cambridge, Mass.: MIT Press.
- Krugman, Paul R. 1999. "Was it All in Ohlin?", at <http://web.mit.edu/krugman/www/ohlin.html>.
- Krugman, Paul R. and Anthony J. Venables. 1990. "Integration and the Competitiveness of Peripheral Industry," in *Unity with Diversity in the European Community*. Christopher Bliss and Jorge Braga de Macedo, eds. Cambridge: Cambridge U. Press, pp. 56-75.
- Krugman, Paul R. and Anthony J. Venables. 1995. "Globalization and the Inequality of Nations," *Quart. J. Econ.* 110:4, pp. 857-80.
- Leamer, Edward E. and James Levinsohn. 1995. "International Trade Theory: The Evidence," in *Handbook of International Economics, Vol. 3*. Gene M. Grossman and Kenneth Rogoff, eds. Amsterdam: North-Holland, pp. 1339-94.
- Leamer, Edward E. 1997. "Access to Western Markets and Eastern Effort," in *Lessons from the Economic Transition: Central and Eastern Europe in the 1990s*. Salvatore Zecchini, ed. Dordrecht: Kluwer Academic Publishers, pp. 503-26.

- Lösch, August. 1940. *The Economics of Location*, Jena, Germany: Fischer. (English translation, New Haven: Yale U. Press, 1954.)
- Lucas, Robert E., Jr. 1988. "On the Mechanics of Economic Development," *J. Mon. Econ.* 22:1, pp. 3-42.
- Martin, Philippe P. and Gianmarco I.P. Ottaviano. 1999. "Growing Locations: Industry Location in a Model of Endogenous Growth," *Eur. Econ. Rev.* 43:2, pp. 281-302.
- Martin, Ron. 1999. "The New 'Geographical Turn' in Economics: Some Critical Reflections," *Camb. J. Econ.* 23:1, pp. 65-91.
- Midelfart-Knarvik, Karen Helene, Henry G. Overman and Anthony J. Venables. 2000. "Comparative Advantage and Economic Geography: Estimating the Location of Production in the EU," mimeo., LSE.
- Mussa, Michael. 1978. "Dynamic Adjustment in the Heckscher-Ohlin-Samuelson Model," *J. Polit. Econ.* 86, pp. 775-92.
- Neary, J. Peter. 1978. "Dynamic Stability and the Theory of Factor-Market Distortions," *Amer. Econ. Rev.* 68:4, pp. 671-82.
- Neary, J. Peter. 1999. "Discussion" of "Economic Policy and the Manufacturing Base: Hysteresis in Location," [by Anthony J. Venables], in *Dynamic Issues in Commercial Policy Analysis*. Richard E. Baldwin and Joseph F. Francois, eds. Cambridge: Cambridge U. Press, pp. 196-201.
- Neary, J. Peter and Dermot Leahy. 2000. "Strategic Trade and Industrial Policy Towards Dynamic Oligopolies," *Econ. J.* 110:463, pp. 484-508.
- Ohlin, Bertil. 1933. *Interregional and International Trade*, Cambridge, Mass.: Harvard U. Press.
- Ottaviano, Gianmarco and Diego Puga. 1998. "Agglomeration in the Global Economy: A Survey of the 'New Economic Geography'," *The World Econ.* 21, pp. 707-31.
- Ottaviano, Gianmarco, Takatoshi Tabuchi and Jacques-François Thisse. 2000. "Agglomeration and Trade Revisited," mimeo., CORE, Université Catholique de Louvain, Belgium.
- Puga, Diego. 1999. "The Rise and Fall of Regional Inequalities," *Eur. Econ. Rev.* 43:2, pp. 303-34.
- Puga, Diego and Anthony J. Venables. 1996. "The Spread of Industry: Spatial Agglomeration and Economic Development," *J. Japanese and Int. Econs.* 10:4, pp. 440-64.
- Quah, Danny T. 1997. "Increasingly Weightless Economies," *Bank of England Quart. Bull.* 37, pp. 49-56.

- Redding, Stephen and Anthony J. Venables. 2000. "Economic Geography and International Inequality," mimeo., LSE.
- Smith, Alasdair and Anthony J. Venables. 1988. "Completing the Internal Market in the European Union: Some Industry Simulations," *Eur. Econ. Rev.* 32:7, pp. 1501-25.
- Solow, Robert M. 1962. "Comment" on "On a Two-Sector Model of Economic Growth" (by Hirofumi Uzawa), *Rev. Econ. Studs.* 29, pp. 255-7
- Venables, Anthony J. 1996. "Equilibrium Locations of Vertically Linked Industries," *Int. Econ. Rev.* 37:2, pp. 341-59.
- Venables, Anthony J. 1999. "Economic Policy and the Manufacturing Base: Hysteresis in Location," in *Dynamic Issues in Commercial Policy Analysis*. Richard E. Baldwin and Joseph F. Francois, eds. Cambridge: Cambridge U. Press, pp. 177-95.
- Ventura, Jaume. 1997. "Growth and Interdependence," *Quart. J. Econ.* 112:1, pp. 57-84.

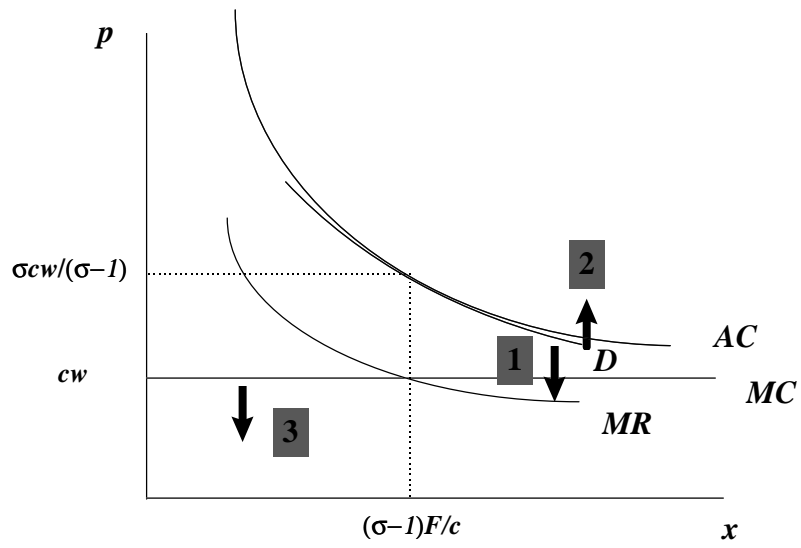


Figure 1: Chamberlin-Dixit-Stiglitz Equilibrium and Effects of Entry by a New Firm

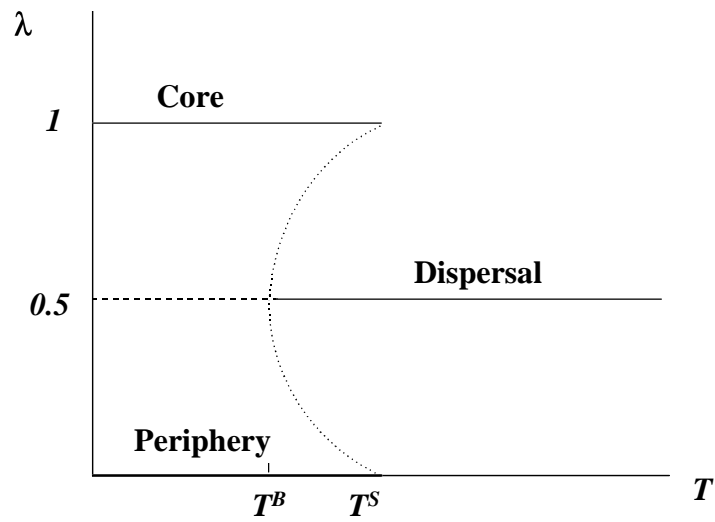


Figure 2: Agglomerated and Dispersed Equilibria as a Function of Trade Costs

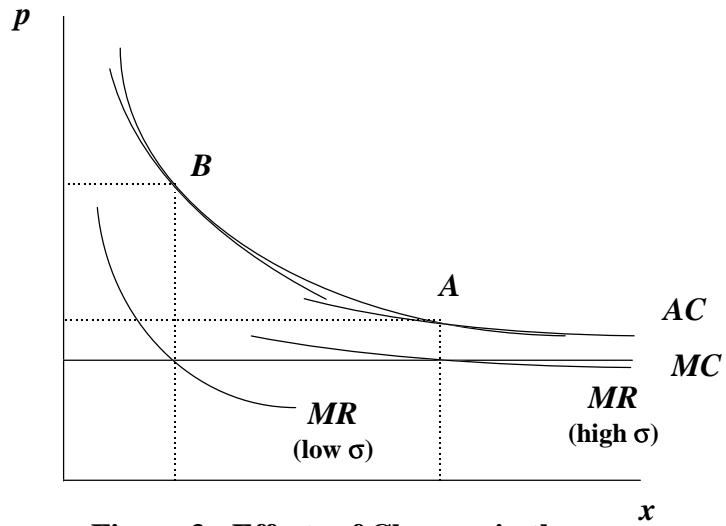


Figure 3: Effects of Changes in the Elasticity of Substitution