The "new" economic theories

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Additional Information:

- This is Working Paper; no. 104. Working papers (Trabalhos em curso) from Faculdade de economia, Universidade de Porto: http://www.fep.up.pt

Metadata Record: [https://dspace.lboro.ac.uk/2134/182](https://dspace.lboro.ac.uk/2134/182)

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/
THE “NEW” ECONOMIC THEORIES

Helena Marques
THE “NEW” ECONOMIC THEORIES*

HELENA MARQUES
Department of Economics, Claremont Tower
University of Newcastle upon Tyne
Newcastle upon Tyne NE1 7RU
United Kingdom
email: H.I.Marques@ncl.ac.uk

ABSTRACT
This paper has two main goals. The first is to study the links between the “new” economic theories, this is, the “new” trade theory, the “new” growth theory and the “new” economic geography. These are three apparently distinct strands of economics, yet they have a common motivation: the role of increasing returns and the consequent market structure (imperfect/monopolistic competition). The second goal is to present the “new” economic theories as case studies in what concerns the debate over modelling and its role in the progress of economics. Since these theories contribute fundamentally by applying new modelling techniques to old real world problems, they add something to economic knowledge to the extent that we accept formalisation as a source of progress in economics.

RESUMO
Este artigo tem dois objectivos principais. O primeiro consiste no estudo da relação entre as “novas” teorias económicas, isto é, a “nova” teoria do comércio, a “nova” teoria do crescimento e a “nova” geografia económica. Estes são três ramos da teoria económica aparentemente distintos que apresentam, contudo, elementos comuns: o papel dos rendimentos crescentes e a estrutura de mercado utilizada (concorrência imperfeita/monopolística). O segundo objectivo é a apresentação das “novas” teorias económicas como exemplos do debate sobre modelização e seu papel no progresso da economia. Uma vez que a contribuição destas teorias consiste fundamentalmente na aplicação de novas técnicas de modelização a problemas reais já antigos, elas incrementam a compreensão dos fenómenos económicos na medida em que aceitarmos a formalização como fonte de progresso científico.

Keywords: Krugman, new growth theory, new trade theory, new economic geography, history of economic thought.

* Research paper presented at the University of Porto for the History of Economic Thought Doctoral Module under the supervision of Prof. Roger Backhouse (University of Birmingham, UK).
1. INTRODUCTION

This essay has two main goals. The first is to study the links between the “new” economic theories, this is, the “new” trade theory, the “new” growth theory and the “new” economic geography. These are three apparently distinct strands of economics, yet they have a common motivation: the role of increasing returns and the consequent market structure (imperfect/monopolistic competition). The second goal is to present the “new” economic theories as case studies in what concerns the debate over modelling and its role in the progress of economics. Since these theories contribute fundamentally by applying new modelling techniques to old real world problems, they add something to economic knowledge to the extent that we accept formalisation as a source of progress in economics.

Ron Martin (1999) considers essentially four steps in the evolution of what Stephen Meardon (1999) calls “geographical economics”: German location theory (Johann von Thünen 1826, Alfred Weber 1929, Walter Christaller 1933, August Lösch 1939), regional science (Walter Isard 1956, 1960), economic geography (he points to the work of several geographers since the 1970s) and the “new” economic geography (Paul Krugman 1991a, 1991b, 1993a, 1993b, 1994a, 1995a, 1996a, 1996b)². A detailed description of the German location theory is provided in Meardon (1999).³ In what concerns regional science, Martin (1999) describes it as “a highly mathematical and esoteric theory of abstract, equilibrium economic landscapes, in effect the formalised successor to the German “location economics” tradition”. This is clearly in opposition to economic geography, “a more eclectic and empirically-oriented subject, in which formal neoclassically-oriented location theory had been largely displaced by concepts imported from other branches of economics”.

Finally, the new economic geography is “a theory of economic localisation based on increasing returns ... long on mathematical modelling but exceedingly short on empirical application”. This last observation constitutes a controversial point⁴ and will

---

¹ Although the terms do not have exactly the same meaning in the original Robinson/Chamberlin versions, the main “new” theories authors and their critics use them indifferently.
² Other important contributions come from Michael Porter, Brian Arthur, Robert Barro, Xavier Sala-i-Martin, Barry Eichengreen, Olivier Blanchard, Lawrence Katz, Anthony Venables, Danny Quah and others, the first one being an exception for his descriptive approach.
³ See also Blaug (1996).
⁴ Ron Martin (a geographer) has been very critical of Paul Krugman’s (an economist) work.
be dealt with later in the paper. Martin (1999) further points to two main directions of research in the new economic geography: the dynamics of regional growth and convergence, and the spatial agglomeration of economic activity. While the former focuses on long-run regional growth and convergence and is linked to the “new growth theory”, the latter focuses on industrial and urban location and is linked to the “new trade theory”. The story is basically that, when increasing returns stem from externalities, economies of scale and imperfect competition arise through a process of regional or local economic agglomeration. “Thus to understand trade it is necessary to understand increasing returns and to understand increasing returns it is necessary to study regional economic concentration and specialisation”, writes Martin (1999).

Indeed, a major contribution of the “new” theories is the clarification of the role played by increasing returns. Until the 1980s, economics was heavily dominated by what Krugman (1995b) calls “the Ricardian Simplification”, this is, the assumption of constant returns and perfect competition. Admitting increasing returns bears two consequences: the existence of plausible and relevant multiple equilibria and explaining how the economy picks one of these, which involves dynamic analysis. We then go from “static models in which equilibrium is uniquely determined by tastes, technology and factor endowments” to “dynamic models in which the choice of equilibrium also reflects history”.

Naturally Krugman was neither the first nor the only economist to defend increasing returns: Nicholas Kaldor attacked constant returns in the 1960s, Thomas Schelling talked about dynamics and multiple equilibria in the 1970s and Paul Romer applied increasing returns to economic growth in the 1980s. Growth, trade and location issues had faded or stagnated mostly due to the absence of a formalised theoretical framework that was able to treat them in the presence of market structures characterised by increasing returns and monopolistic competition (Krugman 1995a). It was the introduction of the missing analytical structure that brought such theories back into the research agenda. Furthermore, the study of location and spatial concentration of economic activity was fostered by the progress of regional economic integration in recent decades, with special attention to the European case. In fact, economic integration is intertwined with the new economic geography literature in two different ways: integration of goods markets diminishes transport costs *latu sensu* while integration of factor markets increases factor mobility.
The remaining of the essay briefly presents the main characteristics and developments of the “new” theories and discusses the arguments for and against their approach, with special focus on the “new” economic geography. In a way, these arguments particularise the ones that are usually advanced for and against formalism in economics.

2. The New Growth Theory

Krugman (1994b) defines “high development theory” as “the view that development is a virtuous circle driven by external economies – that is, that modernization breeds modernization”. According to this view, there is a multiplicity of equilibria, namely a high and a low development equilibrium level. If a country fails to reach the virtuous circle critical level, it remains underdeveloped, stuck in a low-level trap. In general, both growth and stagnation have a cumulative and self-reinforcing nature, thus showing that increasing returns are central to development theory. In fact, the virtuous circle stems from an interaction between economies of scale at the firm level and market size. This interaction is accompanied by economic dualism – the economy has two sectors, traditional and modern, the latter paying higher wages. In the seminal paper by Rosenstein-Rodan (1943) we find both the assumption of economies of scale and the assumption of dualism.

Similarly to trade and geography, the multiple equilibria feature leaves a scope for government intervention, which can lead the economy to that particular equilibrium in the possible set that is considered the most desirable – in this case the high level one. In the literature there were essentially two opposite views. On the one hand, Rosenstein-Rodan and others defended a co-ordinated and broadly based investment program – the Big Push. On the other hand, Hirschman argued that the correct policy would be that of “balanced growth”: promoting first those key sectors with stronger linkages, then correcting the disequilibria generated in the other sectors by these investments. Again increasing returns are fundamental to the definition of forward and backward linkages: these concepts involve an interaction between scale and market size. Since economies of scale were crucial to high development theory, yet very difficult to introduce into the increasingly formal models of mainstream economic theory, development theory faded.

---

5 Krugman (1995a) recognises that “there are obvious affinities between the concepts that arise naturally in geographic models and the language of ... the “high development theory””.

6 A backward linkage implies that an upstream industry is able to produce at least at the minimum economic scale. Forward linkages involve the reduction of downstream industries costs.
Krugman claims that “the glory days of “high development theory” spanned about 15 years, from the seminal paper of Rosenstein-Rodan (1943) to the publication of Hirschman’s Strategy (1958)”. In short, it was highly influential in the 1940s and 50s, but faded until the 1980s. The crisis of development theory in the 1950s was above all methodological, this is, failure in expressing its central concept of increasing returns through the formal models that were becoming increasingly popular. The essential problem was that of market structure, of the perfectly competitive model embodied in the Ricardian Simplification being incompatible with scale economies. Consequently development theory faded out and some authors like Myrdal and Hirschman even rejected formalism.

Moreover there was an increasing dissociation between growth theory and development economics. The former became more and more abstract and formal, while the latter focused mainly on empirical studies, not always backed by theory. In addition to the methodological problem, there was a basic incompatibility of application of growth theory to the underdeveloped countries. In the 1950s it was widely believed that people simply behaved differently in developed and underdeveloped countries, so that the sophisticated growth models used for the former could not be applied to the latter. This conception was not changed till the 1970s and it was also responsible for driving apart growth and development theories.

In the 1970s economists came up with what Krugman calls a “bag of tricks” that allowed the modelling of market structures other than perfect competition in a tractable manner. This breakthrough lead to a revival of international trade, economic growth and development, and finally economic geography. The new “bag of tricks” was in fact used in all the “new” economic theories of growth, trade and geography. The benchmark models of the new growth theory, incorporating the concepts of increasing returns, imperfect competition and/or externalities are Romer (1986, 1987, 1990) and Lucas (1988). It was also possible to give a formal treatment to ideas that had been previously exposed informally (the importance of increasing returns for growth had already been recognised by Smith, Marshall, Young, Kaldor). An interesting example is the model presented by Murphy, Shleifer and Vishny (1989) on Rosenstein-Rodan’s Big Push.

---

7 The main growth models are surveyed in Barro and Sala-i-Martin (1995).
8 See Section 5 and the case of the new economic geography in theory versus reality.
9 I thank Prof. Backhouse for pointing this out.
10 In fact, they went further than Rosenstein-Rodan, showing under which conditions the Big Push occurs.
3. THE NEW TRADE THEORY

The traditional theory of international trade had several implications (Dixit 1993). First, trade would occur mainly between countries with different factor endowments, according to what Krugman (1993b) called *first nature* advantages (technology, production factors availability). Hence the traditional trade theory is unable to explain the existence of different production structures in similar regions. Second, trade should lead to conflict between factors of production (workers in capital-exporting countries face greater market competition and thereby loose income). Finally, countries having complementary factor endowments were the best candidates to the formation of trading blocs, so that they will specialise in different commodities. However, these implications did not accord with post-war facts: trade among similarly endowed countries, intra-industry trade, and the formation of EEC. In addition, the traditional trade theory performs poorly when there is high mobility of production factors. Krugman then raises the possibility of predominance of *second nature* advantages, such as historical concentration of population in a given region, over *first nature* ones.

Krugman (1996c) provides a description of the thinking of trade theorists before the rise of the new trade theory: “The observation that increasing returns could be a reason for trade between seemingly similar countries was by no means a well-understood proposition ... The idea that trade might reflect an overlay of increasing returns specialization on comparative advantage was not there at all: instead, the ruling idea was that increasing returns would simply alter the pattern of comparative advantage. Indeed, as late as 1984 many trade theorists still regarded the main possible contribution of scale economies ... as being a tendency for large countries to export scale-sensitive goods. The essential arbitrariness of scale-economy specialization, its dependence on history and accident, was hardly ever mentioned.”

However, Krugman recognises that this was not Ohlin’s view in 1933 and suggests that Ohlin already acknowledged the important role of increasing returns\(^\text{11}\) and talked about a “unified field theory” of factor-based and scale-based trade. This unified theory is a clear antecedent of the “integrated economy” approach that ended up playing a central role in post-1980 trade theory – a unification of trade theory and location theory. As Ohlin puts it in Chapter III of *Interregional and International Trade*:

\(^{11}\) Unfortunately it disappeared with Samuelson through his modelling efforts.
“[T]he advantages of producing a large quantity of a single commodity instead of a little of all commodities must lead to interregional trade ... To demonstrate the importance of this, assume that a number of regions are isolated from each other, and that their factor endowments and their demand are so balanced that the relative prices of factors and commodities are everywhere the same. Under the [constant returns] assumptions of Chapter I, no trade is then possible. As a matter of fact, insofar as the market for some articles within each region is not large enough to permit the most efficient scale of production, division of trade and labour will be profitable. Each region will specialize on some of these articles and exchange them for the rest ... The tendency toward specialization because of differences in factor endowments is reinforced by the advantages of large-scale production. The location of an industry in one region and not in another might simply be due to chance ... The conclusion that interregional trade reduces the disadvantages of indivisibility corresponds to the previous conclusion that trade mitigates the disadvantages of an unequal geographical distribution of productive agents ... Thus, all interregional trade, whether due to the one cause or the other, might be regarded as a substitute for geographical mobility of productive factors.”

Krugman (1999) comments: “that view [on the role of increasing returns] remained hidden in plain sight for nearly 50 years: in the late 1970s ... few trade theorists thought of increasing returns as a potential independent source of trade”. There are several reasons for this temporary neglect of increasing returns. Firstly, in the Samuelson’s trade models that replaced the original source as widespread reading, the endowment-based HOS model appeared to follow naturally from the technology-based Ricardo, while increasing returns, “in which differences [in resources] are the result rather than the cause of trade”, were a completely different story. Secondly, Ohlin considers increasing returns to be a subsidiary cause for trade, “carrying the division of labour and trade a little further than it would otherwise go, but not changing their characteristics” (Ohlin 1933). Even though this was true in Ohlin’s days, when intra-industry trade among advanced countries was still insignificant, this attitude towards increasing returns led subsequent modellers to set them aside. Thirdly, Krugman mentions the generic market structure problem that also led to a “temporary forgetting of insights in location theory and development economics”. Ohlin failed in providing the clear distinction between internal and external economies that formal modelling
required, so his intuition was left aside until “with the development of tractable models of monopolistic competition ... it became easy to think of increasing-returns trade as a beneficial overlay on comparative advantage”. Finally, the difficulty of inserting increasing returns into the two-good framework of Heckscher-Ohlin delayed its introduction until one reasoned in terms of, say, three goods, like in Helpman and Krugman (1985).

Since the tools that were required to study the real world had not yet been developed, the traditional theory in its modelled version failed to explain fully the causes of trade. After a peak in the 1960s, it reached a dead end. It was only after Spence (1976) and Dixit and Stiglitz (1977) introduced manageable models of monopolistic competition that in the eighties Krugman (1979, 1980, 1981)12, Ethier (1982), Helpman and Krugman (1985)13 proceeded to build up a “new” theory of international trade. They noted that increasing returns, transport costs and the asymmetric distribution of resources prevented backyard capitalism and were the main determinants of the concentration of economic activity14: “In the new trade theory, the basic point was that increasing returns are a motive for specialization and trade over and above conventional comparative advantage, and can indeed cause trade even where comparative advantage is of negligible importance ... among industrial countries with similar resources and technology” (Krugman 1995b).

Accordingly, the main goal of the founders of the new trade theory was to explain trade patterns in the presence of increasing returns and imperfect competition, thereby finding a theoretical justification for the increasingly observed intra-industry trade. Krugman (1980) argued that firms tend to agglomerate in order to benefit from scale economies and simultaneously would locate close to the market so that transport costs were minimised. According to this home market effect, exporting countries would be the ones possessing large home markets. In fact, a synthesis of the old and the new views of

---

12 These papers laid the basis of the new trade theory. They explained, through the lens of monopolistic competition theory, the expansion of post-war trade between countries with similar factor endowments in the absence of distributive struggles between “wining” and “loosing” factors of production.

13 This book became the “magnum opus” synthesising what had become known as the “new trade theory”, the merger of industrial organisation and trade theory.

14 Scotchmer and Thisse (1992) call it “the folk theorem of spatial economics”.
trade was achieved\textsuperscript{15}. Consider two sectors, a Chamberlinian one that expands through increase in the number of firms (greater product variety) and the size of each firm (greater scale economies), and another operating under constant returns to scale and perfect competition. There is both inter-industry trade (homogeneous good traded against the differentiated good) still governed by the factor endowment differences and intra-industry trade (different countries produce different varieties and trade them).

Another important innovation was the introduction of transport costs in theoretical models. The traditional theory considered them to be either zero or prohibitive. On the contrary, the “new” theory considers these costs explicitly, under the form suggested by Samuelson (1954): iceberg costs (a part of the product “melts” during transportation). This way of modelling is extremely useful since it avoids the incorporation of an additional transport sector into the model. In addition, it integrated perfectly within the models of monopolistic competition.

The study of other market structures, such as oligopoly associated with Cournot competition (Brander 1980, Krugman and Brander 1983), followed naturally that of monopolistic competition. Here trade still occurs despite the absence of comparative advantage. In addition, there is intra-industry trade with gains for both countries. The results were later re-examined in other settings (Venables 1990, Ben-Zvi and Helpman 1992). Another extension concerned trade policy under imperfectly competitive markets (Flam and Helpman 1987, Venables 1990, Brander and Spencer (1985) over export subsidies; see survey in Dixit 1987). Krugman (1984) developed this ideas showing that, with oligopoly, import protection may act as export promotion.

Ottaviano and Puga (1998) remarked that the new trade theory approach is rather incomplete because it considers exogenous geographical advantages. Countries differ in their market dimension, but it lacks an explanation of how such differences arise. The path followed by economic geography was then to go beyond the pure trade theory and indigenise geographical advantages. Krugman declared that “geography matters” in determining trade patterns, even if increasing returns are absent. Therefore, it makes no

\textsuperscript{15} With respect to these, Krugman (1999) considers that there are five big ideas in international trade theory: comparative advantage, determination of the terms of trade by reciprocal demand, the interaction between factor abundance and factor intensity, the interaction between domestic distortions and trade policy and arbitrary specialisation driven by increasing returns. In Development, Geography and Economic Theory, Krugman identifies the “five lost traditions” of economic geography: Germanic location theory, social physics (gravity and potential models), cumulative causation, land use and land
sense to view nations as dimensionless points and trade patterns as spaceless flows. In fact, in his 1993a article he argued that international trade is simply an aspect of the more general theory of location of production: “I view this paper as a part of a larger project, which is to recast at least some of the theory of international trade as simply an aspect of the field of economic geography”. And so he led his research from increasing returns and market structure – international trade theory – straight into the role of external economies and the accidents of history in explaining the locational dimension of economic activity. It was the rise of the “new economic geography”.

4. THE NEW ECONOMIC GEOGRAPHY

As Helpman and Thisse (1999) put it, economic geography tries to answer questions like: “Can we expect convergence between different economic areas or, instead, a more agglomerated pattern of economic activity at the regional and global levels? What will be the impact on international trade of the fading of national borders? What is the role of cities in the growth and trade processes? How will increased openness to trade and factor mobility affect policy making?”

A decade spanned “between showing how the interaction of transport costs and increasing returns at the level of the plant could lead to the "home market effect" (Krugman 1980) and realizing that the techniques developed there led naturally to simple models of regional divergence (Krugman 1991b)” (Krugman 1999). In the tradition of trade theory since Ohlin (1933), trade in goods substitutes for factor mobility, while in economic geography trade in goods complements and interacts with factor mobility. According to Krugman (1999):

“Unlike the integration of increasing returns with comparative advantage, which in effect reinforced his [Ohlin] basic vision, the interaction of increasing returns with factor mobility actually tended to run counter to that vision [that] the movement of goods and factors is the way the world economy tries to overcome the limitations placed upon it by the fragmentation of its resources, and the effect of that movement is one of convergence in prices. The logic of modern economic geography models, in which trade and factor mobility are often complements, and factor movements often lead to rent models, and local external economies. The new versions would then add to, rather than eliminate, the old ones.
divergence in factor and even goods prices, is something that Ohlin realized could happen, but it ran counter to his main theme.”

The new economic geography has in common with the new trade theory the features of increasing returns and multiple equilibria, as well as the technical tricks needed to make the models tractable. However, there is a difference in emphasis – “the trade models were largely focused on internal economies of scale, while geography is largely about external economies” (Krugman 1995b).

The major breakthrough of the new economic geography relative to the traditional location theory consisted in including in the concept of geographical advantages the economic advantages that result from the interaction of economic agents, along with the traditional physical ones. This interaction translates itself into both centripetal forces that promote concentration (like increasing returns), and centrifuge forces that cause dispersion of economic activity (like competition). With respect to the former forces, Krugman (1991a, 1998a) clearly borrows Marshall’s (1920) concept of localised external economies, as well as the previously mentioned Hirschman’s (1958) backward and forward linkages. In what concerns the latter forces, Krugman (1998a) follows the tradition of urban economics in considering immobile factors (such as land, natural resources and labour at the international level), rents and external diseconomies (such as congestion costs).

In his models, Krugman weights market size effects against transport costs, having into account the love of variety that characterises models of the Dixit-Stiglitz type and giving a central role to iceberg costs. The general result is that clusters of activity are first established through a pattern based on the principles of chaos – of non-linear dynamics. As Krugman (1998a) recognizes, economic geography models deal with very complex realities, leading to complex solutions. What comes out are multiple equilibria and non-linearities, which require a numerical approach and simulations carried out by sophisticated software. However, the arbitrariness of industrial location found in the new trade theory is maintained in the new context (Krugman 1991a, 1991b, 1991c). In addition, this literature, together with Krugman (1991d), also emphasises the

---

16 The original Marshallian classification refers to the ability of producers to share specialised providers of inputs, labour market pooling and localised spillovers of knowledge, especially through personal interaction.

17 The existence of non-linearities makes the new economic geography “analytically intractable” and consequently it “must be explored via the computer”. See Fujita, Krugman and Venables (1999).
importance of history\textsuperscript{18} and expectations in explaining observed patterns of industry location and growth, supporting the work of David (1985) and others. Afterwards, we see a cumulative and self-reinforcing concentration when there already is one.

Fujita and Thisse (1996) consider three types of economic geography models according to the motivation for agglomeration: externalities, increasing returns or spatial competition. As an example of the first type of models we can point out Henderson (1974), based on the definition of technological externalities (Scitovsky, 1954). However, this approach tells us nothing about the way agglomeration forces relate to microeconomic conditions. The second group forms what could be called economic geography models \emph{strictu sensu} and is further divided into urban models (Fujita 1988), and models that demonstrate the possibility of regional divergence (Krugman 1991b, Fujita and Krugman 1995, Venables 1996). In his 1998a paper, Krugman provides a somewhat different classification of increasing returns models: bridge-building between “new” economic geography and traditional location theory (Fujita and Krugman 1995), or using new economic geography to give international trade a spatial framework (Venables 1996). Krugman (1998a) only takes into account post-Krugman models,\textsuperscript{19} while Fujita and Thisse (1996) present both the pre-Krugman and post-Krugman view. Finally, spatial competition models try to overcome the absence of strategic interaction in the Dixit-Stiglitz-Spence increasing returns approach and stem from the already mentioned research in strategic trade.

Ottaviano and Puga (1998) distinguish four mechanisms leading to circular and cumulative spatial concentration of economic activities: labour migration, input-output linkages due to intermediate goods, factor accumulation and inter-temporal linkages, history and expectations.

Firstly, let us consider demand linkages induced by labour migration. Krugman and Venables (1990) attempted to study the impact of the European integration process in the context of a two-region centre/periphery model. Each region initially had an industrial and an agricultural sector respectively producing a tradable differentiated good subject to increasing returns and a non-tradable homogeneous good subject to constant returns. The main conclusion of this work is the non-monotonicity of the relationship between agglomeration and integration under the form of a U-shaped curve:

\textsuperscript{18} History matters too!
\textsuperscript{19} Not surprisingly!
the industrial sector concentrates in the centre only for intermediate transport costs, viewed as trade barriers. In fact, at very high transport costs, there cannot be agglomeration: the world consists of self-sufficient peasants. At very low transport and communication costs, there is little incentive for agglomeration: necessary inputs can be delivered to wherever the factor costs are lowest.

Krugman (1991b) presents a periphery/periphery model based on scale economies, as concentration forces, and on transport costs, as determinants of location near the larger markets. There is no inter-sector mobility of labour (the only production factor in the model). While farmers are totally immobile, industrial labour shows inter-regional mobility. The reason why any firm eventually decides to move to another region is totally unexplained. However, once that happens, workers migrate along, increasing demand in the recipient region. Other firms and workers follow. As Krugman (1999) wrote: “Suppose that there are strong advantages to concentration of factors – where these advantages may take the form of true external economies, but may also be due to "linkage" effects arising from the effect of concentration on the size of markets and the availability of inputs. And suppose also that some factors are more mobile than others. Then factor mobility will tend to increase differences among regions rather than reducing them, and instead of substituting for regional specialization will promote it.” The demand linkages presented here are similar to the home market effect in Krugman’s 1980 paper on new trade theory, except for the self-reinforcing element now introduced.

Secondly, it follows the study of demand and cost linkages induced by consumption and supply of intermediate goods. The economic geography model in Krugman (1991b) was developed in the light of the United States experience. However, Europe’s economic reality is rather different from that of the United States in what concerns integration and labour mobility. As a result, in the former, economic activity is much less concentrated and income disparities are much wider than in the latter. The US cumulative process is well explained by Krugman’s model, but in Europe the location problem is more adequately tackled by the models in Krugman and Venables (1995) and Venables (1996). It is argued that an industry’s change of location would be followed, not by workers, since even industrial labour is immobile, but by intermediate goods industries (Krugman and Venables 1995) or upstream and downstream industries (Venables 1996). These models, besides a higher adequacy to international rather than regional contexts, introduce the important linkages established among several different
industries. In other words, it is recognised the heterogeneity of the industrial sector. Since industry is not homogeneous, it is even possible that different industries locate in different regions according to comparative advantages, this is labour intensive industries locate in labour abundant regions and similarly for capital.

Thirdly, there is an agglomeration mechanism that works through endogenous growth and inter-temporal linkages. This agglomeration mechanism arises from the merger of the new economic geography (Krugman 1991b, Venables 1996) with endogenous growth (Romer 1990, Grossman and Helpman 1991). Through growth linkages, growth and agglomeration are mutually reinforcing. In Martin and Ottaviano (1996) an R&D sector is introduced which uses the composite differentiated good as an input to produce new varieties of itself. This sector is then the engine of growth with industry locating near most R&D activities. Martin and Ottaviano (1999) introduce a second production factor, capital, which was not explicitly present in Krugman’s former models.20 There are learning effects that locate the capital-producing sector near the rest of the industry, contributing to concentration. It is interesting to mention that Baldwin (1999), despite using a neo-classical instead of an endogenous growth model, concludes that integration leads to long run income divergence.

Finally, since economic geography models predict multiple equilibria, both history and expectations determine which among them will be the effective outcome. History (for example, the amount of industrial employment) may cause small asymmetries in initially identical regions and give rise to an agglomeration process. Krugman (1992) talks about the concept of catastrophe: small changes in the key parameters of the model (elasticity of substitution among varieties, share of industry in the economy, transport costs) may lead to jumps, or discontinuous changes, in location. However, it is possible that self-fulfilling expectations outweigh history and regions neglected by history manage to attract economic activity. This will happen if the lock-in effects (initial advantage), transport and/or migration costs are not too large. Integration increases the power of expectations through the reduction of such costs. Naturally, in other cases, expectations reinforce history. Although this is an essentially dynamic issue, Krugman (1991d), Matsuyama (1991) and Ottaviano (1996, 1999) have conducted static analyses in which they concluded that the relative importance of history and expectations

20 Baldwin and Forslid (1998) added capital to Krugman’s (1991b) center/periphery model with inter-regional labour mobility.
depends on the *underlying structure* of the economy, in particular the adjustment costs. History will be the more important the more the future is discounted, the smaller are external economies and the slower is the adjustment process. However, Krugman himself recognises in his 1991d paper that “we should try to focus on the kinds of external economies that can be modelled other than by assumption” and “knowledge flows ... are invisible; they leave no paper trail by which they may be measured and tracked”.

In general, the existence of multiple equilibria is a weakness of economic geography models and confers them little predictive power. The question is not only to know whether agglomeration will occur but also, should it occur, to be able to know where. In addition, there are several factors that may induce a reversal in the location pattern.

The first factor to be considered is labour immobility. It delays agglomeration during an integration process and cushions its effects when it finally happens. If as a consequence real wages are lower in the peripheral regions, there may be a motive for firms to move back to periphery, provided transport costs are low enough. In fact, in Europe labour is much less mobile than in the US and industry is less concentrated, regional wage differentials being substantially higher.

Secondly, we should consider the existence of non-tradable goods. Agglomeration increases the prices of non-tradable goods, due to higher demand in more densely populated regions, so that the desire to escape such higher prices may create a flow of return to less crowded areas, where housing and certain services, for instance, are cheaper. Non-tradable goods prices will be weighted against the greater number of varieties of tradable goods in the central regions. In addition, if the homogeneous good is also subject to transport costs, the tendency to concentrate decreases and in the limit (when transport costs are too high) the homogenous good becomes non-tradable.

Thirdly, the technological spillovers may be global instead of local. When the R&D spillovers are global, all regions benefit from the invention of a new variety. Poorer regions have an opportunity to industrialise using knowledge created in richer regions. On the other hand, there may be scale economies at the firm level, but not at the plant level. Hence, multinational companies may operate different plants in different countries. In fact, intra-industry trade decreases with trade costs and national incomes,

---

21 The effect is parallel to that of *congestion costs* in urban economics.
but also with scale economies at the firm level relative to individual plants, since firms are induced to produce at different locations instead of exporting their products.

Finally, similarly to the “high development theory”, the multiple equilibria feature allows the government to influence the outcome, leading market forces to the preferred equilibrium in terms of policy or welfare. In fact, inefficient equilibria may arise due to co-ordination failures between firms and workers/consumers. We can consider three broad forms of government intervention that favour dispersion by manipulating market size and transport costs within a country: public expenditure, redistributive policies and public infrastructures. In what concerns the relations with other countries, an important instrument is trade policy (Helpman and Krugman, 1989). A country forgotten by history can use it to overcome its gap. Further, transport costs are increased by protectionism, and high transport costs induce the location of industry near smaller markets.

In general, trade is both influenced by and in turn influences the process of geographical industrial specialisation within nations. Further links between trade and the new economic geography are provided also in the urban systems version of the new economic geography. Suppose two cities equidistant from international markets, one of which concentrates all production due to agglomeration benefits. Krugman (1996b) argued that the opening to international trade might change urban concentration by creating two cities of equal size focused on international exports.

In the previous section it was discussed the (relative) importance Ohlin (1933) attached to increasing returns, these being paramount in the “new” economic theories. Having reviewed the new economic geography, it is time to ask what then was not in Ohlin, this is, what was innovative about the new trade theory in the 1980s and the new economic geography in the 1990s. Krugman (1999) answers: “The first [aspect] is the appreciation of the importance and distinctiveness of imperfect competition ... A second ... is the distinction between equilibria and optima ... Finally, ... the idea of qualitative, discontinuous change ... that small changes in underlying parameters - say, in transportation costs - sometimes bring about large changes in behaviour”.

---

22 See the Canada example in Krugman (1991a).
23 Import substitution and unilateral trade liberalisation in economic geography models are analysed by Puga and Venables (1999) and illustrated by the asian NICs.
24 We refer to Meardon (1999) for a discussion of urban systems and the Krugman/Henderson debate and simply illustrate such links through the Krugman (1996b) model.
5. THE “NEW” THEORIES’ MODELLING APPROACH

In the preceding sections we attempted to highlight the main features of the “new” economic theories: the new growth theory, the new trade theory and the new economic geography. Through the analysis of those features it becomes clear that there are indeed links between them. They share the study of increasing returns and imperfect competition, as well as an outcome of multiple equilibria that allows for government intervention. It is also argued that they all arose after a period of neglect or stagnation in their respective older versions. This was due to the failure to incorporate such more realistic concepts into the formal analysis that became the trademark of the mainstream.

In the Ohlin lectures which formed the book “Development, Geography and Economic Theory” Krugman (1995a) defended that “a growing emphasis on formal modelling led economists to “forget” insights about the role of increasing returns in industrialization and economic location, only to rediscover those insights when modelling techniques became sufficiently advanced”. In fact, the “new” theories largely owe their existence to the use of formal mathematical models often with Dixit-Stiglitz as a specific functional form. In this section we intend to use that feature as a case study of a broader debate – that of formalism in economics.

The following discussion concerns the “new” economic geography, but what will be said about it can be applied both to the “new” growth and to the “new” trade theories. The former was chosen as a case study for three main reasons. First, to shorten the paper. Second, because it has been the most attacked of the three, thereby giving rise to a debate which opposed mainly economists and geographers. Finally, the same person was to a great extent responsible for the expansion of new trade and new economic geography (exactly: Krugman), so that to talk about his methodology is to a great extent to talk about both trade and geography, and even growth.

Martin (1999) argues that the new economic geography presents two main drawbacks: it is not new and it is not geography. Krugman (1995a) claims that the developments in mathematical economics allowed economists to “integrate spatial issues into economics through clever models ... that make sense of the insights of the geographers in a way

25 See the reference to Paul Krugman (Cambridge, MA) and Ron Martin (Cambridge, UK) in The Economist (March 13th 1999, page 104). Should we call it “The Two Cambridges Are Back”? It is ironic, however, that Isard’s formalised regional science was successful among geographers in the 1950s, when today those same geographers disclaim Krugman’s new economic geography as containing too little region and too much mathematics. They have already been there and back, they argue.
that meets the standards of the economists”. Martin replies that geographers deliberately abandoned mathematical models. Further, the new economic geography simply dressed old ideas in a formalised suit. In fact, the argument that the “new” economic geography is not new simply because its ideas are not new is also valid for growth and trade, with a qualification: new growth models were generally accepted as new, while geographers contended that Krugman’s new economic geography models brought no novelty. Being based on the role of increasing returns, these “new” theories emphasise and formalise an old idea. As it was shown, the concept was present in Ohlin, but also in Smith, Marshall, Young, Kaldor.26

Martin and Sunley (1996) point out the differences between Krugman’s geographical economics and economic geography: (1) The latter is carried out by geographers; (2) Krugman uses formal models, while geographers have abandoned models and are more concerned with “reality”, this is, the political, economic, institutional and social bases of regional development and industrial agglomeration; (3) The former emphasises continuity in the forces responsible for agglomeration, while the latter focus on historical patterns of restructuring. The new economic geography uses the same model “to explain spatial agglomeration and specialisation at vastly different scales, from the international level, to broad core-periphery patterns within nations, to local urban industrial concentrations and even intra-urban neighbourhoods. Processes are thus assumed to be largely scale-independent. For economic geographers, however, the issue of spatial scale is central ... The spatial agglomeration models may well predict that, under specific assumptions, industrial localisation and specialisation will occur, but they are unable to tell us where it actually occurs, or why in particular places and not in others” (Martin 1999).

From the evolution of regional science Isserman (1996) draws three main lessons. First, “economic geography cannot and will not be reformed on the basis of principles from economics”. Therefore, the term “economic geography” is misleading and Krugman ought to replace it with “geographical economics”. Second, the new economic geography should combine the study of the real world with the study of mathematical economics. Finally, theoretical issues in regional economics are predominantly empirical questions. Therefore, the new economic geography needs to share regional

26 I thank Prof. Backhouse for calling my attention to this point.
science’s emphasis on empirical testing, a “niche between theoretical excess ... and descriptive excess”.

Krugman’s models are usually simple: he uses the “minimum necessary model” approach. He further argues that models are necessarily based on assumptions that only sketch reality, yet may prove useful: “in economics we are always making silly assumptions ... some of them have been made so often that they come to seem natural. And so one should not reject a model as silly until one sees where its assumptions lead”\(^\text{27}\) (Krugman 1993c). Sometimes the results may even surprise us; “it is exactly when the conclusions produced by our economic models surprise us that economic theory is most useful”, he wrote to James K. Galbraith in November 5th 1996. Despite its usefulness (precision, transparency, conclusive demonstration), Krugman recognises the relativity and trade-offs implicit in model building: “Modelling is all about loosing some information so that new insights can be gained. In fact, we are all builders and purveyors of unrealistic simplifications. Some of us are self-aware: we use our models as metaphors. Others ... are sleepwalkers: they unconsciously use metaphors as models” (Krugman 1994b); “always remember that you may have gotten the metaphor wrong, and that someone else with a different metaphor may be seeing something that you are missing” (Krugman 1993c).

It should be noted that the very same simplicity that made Krugman’s models famous is simultaneously one of their most criticised features: “while the claim that “history matters” is certainly correct, the treatment of history in the new economic geography is more metaphorical than real and, despite the importance assigned to path-dependence, this notion remains a conceptual and explanatory black box” (Martin 1999). In fact, Krugman does not deny that models are indeed metaphors: “I am a strong believer in the importance of models ... they greatly extend the power and range of our insight ... I have no sympathy for those people who criticize the unrealistic simplifications of model-builders, and imagine that they achieve greater sophistication by avoiding stating their assumptions clearly. The point is to realize that economic models are metaphors” (Krugman 1993c). Still, he claims that they are useful metaphors, since they provide new insights. And in economics we are bounded to trade-offs. After all,

\(^{27}\) Weintraub (1998) says that formalism requires consistency of meaning (the conclusions are true if the assumptions are true).
economics is a science of choices. It is at least partly about quantities and their relationships.

Two important points should be made. First of all, Krugman identifies formalism with mathematical modelling. Backhouse (1998) defines formalism as “breaking down arguments into a series of steps, each of which can be analysed on its own”. As he notes, there are three different interpretations of formalism: axiomatisation (“reducing a body of knowledge to a set of independent axioms, with all propositions being derived from those axioms using well-defined logical rules”); mathematisation (“the use of mathematical techniques ... in economic arguments”); and methodological formalisation (“the use of an agreed set of methods for the solution of certain types of problem”).

Formalism clearly means simplification, but simple models may leave behind a complex reality. If we view the economy as an isolated decomposable system, formalism is enough. On the other hand, if we see the economy as a complex intricate system, then there are arguments which cannot be proved, yet constitute “reasonable knowledge” that may be useful. Accepting formalism as the basis for economics implies setting aside such arguments and rending the analysis incomplete. What Backhouse (1998) wrote about mathematics can be extended to economics: “The point here is not to argue that the attempt to be precise and consistent (two features of formalism in mathematics) does not lead to progress in mathematics. That would clearly be nonsense. The point is rather that there is more to mathematics than deriving the properties of formal systems ... Formalism may have played a major role in the development of mathematics, but it leaves out a lot – possibly even the most interesting aspects of mathematics.”

Second, another major critique to Krugman’s work (and the new economic geography in general) is the lack of empirical studies that can build a bridge between theory and reality. Backhouse (1998) defends that “theory [should] be kept close to its empirical roots”, which requires: “(1) that economists put sufficient effort into empirical work; 28 (2) that they take empirical evidence seriously; (3) that they be aware of the conceptual gap between theories and reality; (4) that they be aware that mathematical imperatives may result in changes in the questions that are being addressed”. In fact, Krugman (1999) argues that “getting each of [trade theory’s] Big Ideas into mainstream thinking required a major intellectual struggle, in general involving new techniques (offer
curves, Edgeworth boxes, Dixit-Stiglitz) and a painful process of changing not only the way one answered questions but the questions themselves”. The danger to change questions because of models, instead of using them to answer new questions, is to be avoided.

6. CONCLUDING REMARKS

From the three “new” economic theories, which constituted the subject of this paper, the most controversial one is certainly the “new economic geography”. Its first controversial element is the name itself, as was discussed above. Nevertheless, Krugman (1995b) states that his “intention is to establish economic geography as a branch of economics that is taken as seriously as international trade” and he is confident that he will succeed: “it’s a reasonable prediction that ten years from now the new economic geography will be as firmly established as the new trade theory”.

Meanwhile, the gap between descriptive and more formal languages should be narrowed, not only in the new economic geography, but also in what concerns the whole body of economic theory. One appropriate method is Marshall’s, as described by Krugman (1998b):

“(1) Use of mathematics as a shorthand language, rather than an engine of inquiry. (2) Keep to them till you have done. (3) Translate into English. (4) Then illustrate by examples that are important in real life. (5) Burn the mathematics. (6) If you can’t succeed in 4, burn 3.”

In short, check intuition mathematically, but spread it in words. However, Krugman diverges from Marshall by distinguishing between outsiders – English should be used when conveying them economic concepts – and insiders to the economics profession, whether scholars or students, to who should be taught “methods, not answers”.

Krugman then proposes a revised version of Marshall’s rules:

“(1) Figure out what you think about an issue, working back and forth among verbal intuition, evidence and as much as you need. (2) Stay with it till you are done. (3) Publish the intuition, the math and the evidence – all three – in an economics journal. (4) But also try to find a way of expressing the idea without the formal apparatus. (5) If you can, publish that where it can do the world some good.”

28 Krugman (1993d) concedes that he has “never engaged in really serious empirical work”.

21
If we follow Krugman’s prescription of “public differentiation”\textsuperscript{29}, this is, to distinguish economics scholars and students from the general audience, we should use the nasty bits with the former, but save a listener or reader-friendly approach for the latter. In fact, “once you have stripped an idea down to its essence, it is often surprisingly possible to express that essence without any visible display of technique” (Krugman 1995b). Then it is not a matter of difficulty or practical impossibility, but of radicalism in attitudes, much due to inertia in behaviour. As Krugman (1993c) acknowledges, “the clarity and power of economic analysis can spoil you: once you have a taste of what it means to have a really insightful model, you tend to be inhibited about looser speculations”, what he calls “speaking the wrong language”. Such fundamentalism leads formalists and non-formalists astray, impeding the progress of knowledge. In short, formalisation has advantages, yet it “must be tempered by an understanding of how theoretical concepts might, or might not, relate to the real world” (Backhouse 1998).

\textsuperscript{29} In his book “The Age of Diminished Expectations” (1990), Krugman states there are three kinds of economics writing: Greek letter, up-and-down and airport. The first is the way professional economists communicate with each other, the second is the stuff of the nightly business report and the third are paperbacks that fill airport bookstores.
REFERENCES


Krugman, P. (1999), “Was it All in Ohlin?” (forthcoming),


VON THÜNEN (1826), *The Isolated State*


