THE SPATIAL ECONOMY: HIGH-TECH GLOSSARY OR NEW REGIONAL ECONOMICS?

Ricardo Machado Ruiz


Palavras-chave: economia regional; geografia econômica; aglomerações industriais; concorrência monopolística.

Abstract: The Spatial Economy – Cities, Regions and International Trade, by Masahisa Fujita, Paul Krugman and Anthony J. Venables, has been seen as the best synthesis of the so-called new economic geography. The purpose of this paper is twofold: first, to review the book; and second, to place it in the current debate on regional economics. The first part of the paper is a brief introduction that situates the book in the history of regional economics. In the second part, a description and analysis of the basic models, their variations and applications are presented. The debate on regional economics concerning the theory proposed in the book is in the third part. Two aspects are stressed: the criticisms on the core-periphery model and how it incorporates regional diversity. The fourth topic is a critical assessment of the achievements and originality of the new economic geography analysis...

1 Assistant Professor at Centro de Desenvolvimento e Planejamento Regional – CEDEPLAR (Center for Regional Development and Planning) of the Universidade Federal de Minas Gerais – UFMG (Federal University of Minas Gerais), Brazil. Granted a CAPES (a federal fellowship agency) for his Ph.D. program at the Graduate Faculty of Political and Social Science – New School University, New York. E-mail: rmruiz@cedeplar.ufmg.br
Key words: regional economics; economic geography; industrial agglomerations; monopolistic competition.

Classificação JEL/JEL Classification: D43, D51, R12

1 A MISSING GENERAL THEORY?

Paul Krugman, Masahisa Fujita and Anthony Venables are the most important representatives of the so-called new economic geography. Since the beginning of the 90’s, they have discussed the most important and traditional topics of regional economics, such as the von Thünen’s monocentric city model, Lösch and Christaller’s theories of urban systems and hierarchies, the size distribution of cities, the determinants of industrial clustering and interregional analysis by input-output matrices.

As far as the recent debate on regional economics is concerned, Spatial Economics can be seen as the most accurate and complete synthesis of the arguments proposed by the new economic geography. The authors claim that all themes and disparate models of urban and regional economics, as the ones above, can be viewed as variations of a broad theory that explains why economic activities are unevenly distributed across space. Therefore, the main task of the book is to enunciate a general theory of the spatial organization of the economy.

Fujita (1999) says that this book can be qualified as a general theory of location and space-economy that fulfills Walter Isard’s theoretical work, which is depicted in his Location and Space-Economy (1956). In Isard’s models of firm location and regional development, transport costs, geographic cost pattern and input substitution are important theoretical tools and play similar rules as transport costs, peripheral markets and increasing returns in the new economic geography. In both approaches, there are spatial preferences and asymmetric distribution of economic agents in the landscape. In short, Spatial Economics can be seen as a continuation, and possibly a validation, of Isard’s project.

Isard (1999) evaluates the new economic geography as a gigantic improvement in developing general equilibrium models applied to regional economics. He also agrees that the approach is part of his own research agenda:
“Their results are well beyond what I imagined could be achieved when I wrote Location and Space Economy with monopolistic competition and a dynamic hierarchical urban system embodied.” (Isard, 1999, p. 383).

However, he also believes that there are several other aspects that are not considered in Fujita, Krugman and Venables' work, such as the regional economic diversity that is presented in any economy. Scott (2000, p. 487) also emphasizes the continuum between Isard and Krugman:

“(…) despite its originality [Krugman's work], it can perhaps be best seen as a continuation of the tradition of spatial analysis and regional science. Better yet, we might call it a 'new' regional science.”

2 THE SPATIAL ECONOMICS

2.1 The fourth wave and the redemption

The first part of the book is a very brief description of influential models in urban and regional economics. In urban economics, the traditional Von Thunen (1966), Alonso (1964), and Henderson (1974, 1988) models are considered to be theoretical references. In regional economics, the most important approaches are the central-place theory (Christaller, 1933, and Lösch, 1954), the base-multiplier analysis (Pred, 1966), and the market potential analysis (Harris, 1954).

There is no doubt that these models are at the core of regional economics. Nonetheless, the authors argue that all these theories, and most of their restatements, have several theoretical drawbacks. The urban theories do not have a plausible story on the forces that drive agglomeration. For example, the von Thünen model just assumes the concentration of manufacturing in a single central city, but does not explain the relations between the city (its size and economic structure) and the hinterland. The central-place theory is not a causal model, but a mere description of a pattern of organization that still needs to be explained. Regarding the basic-multiplier and market potential models, these do not have a consistent model of how competition among different agents in many regions can produce the predicted outcomes: the spatial agglomerations.
Their conclusion is categorical: in general, all these models have no clear explanation of how the agents spread in regions in order to create spatial distributions of production and markets. They miss the rules of “microrganization”. Furthermore, none of these models is a general theory, and the recent restatements have too many ad hoc arguments. Because of all these theoretical limitations, until recently, regional science had been prevented from becoming an integrated part of mainstream economics.

The rediscovery of links between geography and economy in the 90’s is attributed to the most recent wave of innovation produced by the theory of increasing returns. The first wave was in industrial organization; the second and third in international trade and economic growth, respectively. Hence, the new economic geography would be the fourth wave of increasing returns in economics. The new theories on increasing return, in particular the Dixit and Stiglitz’s (1977) model on imperfect competition, were the redemption of the field. The core-periphery model, which is fully described in the second part of the book, is a very good example of how the fourth wave could rescue regional economics from the backstage of mainstream economics.

2.2 The basic machinery: the core-periphery model

The second part of the book is its theoretical heart. Chapters 4 – 7 describe the workhorse of the new economic geography: the core-periphery model. The first version of this model can be found in Krugman (1991). The 1999 model is basically a restatement and refinement of that approach.

The core-periphery model is an adaptation of Dixit and Stiglitz’s (1977) work: a modern formalization of Chamberlin’s concept of monopolistic competition. The model has two interlinked sets of rules: the first set defines how the consumers spend their income (the demand side), and the second one specifies how firms choose the level of production and prices (the supply side). Firms and consumers are spread in a landscape with costly transport of goods, and they optimize their localization and expenditure taking into account all other agents’ positions and respective economic conditions.

The consumer’s problem is to maximize the utility subject to a budget constraint. Given the utility function, delivered prices and

---

2 The model can also be seen as a clear development of models of imperfect competition with increasing returns and product differentiation applied to international trade. See Krugman (1979, 1980).

---
income, households minimize the cost of purchasing a composite set of industrial and agricultural goods. Consumers have a preference for variety (the desire to utilize differentiated goods), thus, as the number of varieties increases, demand decreases, and thus the expenditure on one single good. A positive preference for variety guarantees that consumers do not spend all income consuming one single cheap good, which means that there is no monopolization on the supply side. Consequently, if the supply is not concentrated in one location, the interregional demand for each good can be determined by consumer behavior.

The supply side of the economy has two sectors: agriculture and manufacturing. In the simplest version of the model, agriculture represents fixed markets and a competitive sector (constant returns of scale and homogeneous production). Manufacturing produces a large collection of differentiated goods with increasing returns of scale. Each firm has fixed and constant marginal costs, thus, as production increases, unitary costs decrease since the fixed costs per unit produced falls. There is no input substitution, and the physical input-output coefficients of any particular good are constant.

In spite of increasing returns, the firm does not monopolize the market. Because of the preference for variety, the demand for a specific good decreases as the local markets increase: local consumers increase the demand for non-local goods, and regional imports. The declining demand lowers production and price for each good, which limits the size of the firm and blocks monopolization. A surprising outcome of the model is that, as the regional market grows, the number of varieties increases and the size of firms remains unchanged.

In the core-periphery model demand and supply are defined simultaneously, such as in the conventional general equilibrium framework. Following traditional microeconomics models, a particular firm producing a specific good maximizes profits by taking into account a perceived elasticity of demand. Firms know the demand and can define the production that maximizes profits. As a result, there is an instantaneous achievement of monopolistic competition equilibrium at all points in time (demand is equal to supply), expectations are always met, and all factors of production are allocated since all markets clear. Furthermore, firms are fully mobile, technology is given, and there are no external economies of scale.

Equilibrium on the supply side does not imply that the system is in a steady state. Full equilibrium is achieved only when consumers (or labor supply) are also in equilibrium. The workers (or consumers) are fully mobile and search for higher real wages, so, only if real wages are equalized is the system in equilibrium. In the Fujita et al.'s (1999) model,
since firms are always in equilibrium, the labor market adjustment is what drives spatial organization.

Some of the typical results of the model are:

1) in equilibrium, all firms are the same size;
2) all workers have the same real wages;
3) any increase in the size of regional markets changes only the number of varieties produced in that market and does not affect the size of the firm; and
4) there is no bankruptcy, lost of capital, inventory oscillation or unemployment.

Regardless of all the restrictive hypotheses, the simplest version of the model still generates asymmetric agglomerations, even in a "clean landscape" (i.e., no differences in endowments, technologies, or policies regimes).

The tale which was told us by the model can be understood in its simplest version. Imagine an economy with two sectors: a competitive and fixed agriculture sector spread in a landscape, and a monopolistically competitive manufacturing sector. Both industries have firms fully integrated (no intersectoral or intrasectoral trades) and the only factor of production used by industry is mobile (workers are mobile and farmers are fixed). Finally, all regions are a priori identical.

The regional scenario is in equilibrium until some workers decide to move to another place, which creates a small difference among regions that sparks a spatial reorganization of firms. The extra supply of labor increases regional expenditure and allows the entry of new firms or additional local production of varieties. As consumers value variety, the location with a larger manufacturing sector also has a lower price index, simply because a smaller proportion of the consumers' income bears transport costs (price index effect). Given nominal wages, the lower price index raises real wages and fosters migration of workers. The location with the larger labor supply is also the one with a larger manufacturing sector and regional market (home market effect). In addition, the nominal wages are positively correlated to the demand, so locations with a higher demand for manufactures may pay a higher nominal wage, at least during the period of adjustment.

The home market and price index effects are strong centripetal forces that drive up real wages in industrialized regions. Without any centrifugal force, a complete concentration of the industrial production in a single place would be a natural outcome of the model (the "black-hole
location"). Nevertheless, there is an important countervailing force: fixed agricultural populations, which push the manufacturing sector through the landscape.

The agricultural populations distributed over the landscape are called peripheral markets and play a crucial role in the model. Because the production of agricultural goods demands a non-tradeable input – land – the farmers cannot be concentrated in one single location. They have to search for locations where land is available. From these places (agricultural regions), they ship the excess production to cities (industrial regions) and import manufactures. The interregional trade links the peripheral regions to the industrialized core.

Nonetheless, if the peripheral markets are large enough and transport costs for industrialized goods is expensive, it could be profitable for firms in the core to move into these regions and replace the interregional trade by the local supply. In the case of urban consumers, large cities have to import food from remote regions. If the transport of agricultural goods were costly, this would reduce real wages. Therefore, large and distant peripheral markets are upper bounds to the spatial concentration of the industry.

Based on the relations above, it is possible to summarize the basic story of the core-periphery model. Suppose that an economy has two similar industrial regions with identical agriculture and very high transport costs. In this case, no complete agglomeration is possible because each region is, in fact, an autarky: there is no demand for any good produced in other regions, and firms supply only their local markets. Further reductions in transport cost can change the landscape if firms start to supply other markets (exports) and consumers begin to spend part of their income in other regions (imports). So, for intermediate levels of transport cost, agglomeration takes place and small regional differences can bring on regional and irreversible divergences: some regions become industrialized and others are deindustrialized (only agriculture remains). If transport costs are lower, the opposite happens. Consumers have cheap access to all varieties, and firms can sell to all regional

---

3 Consequently, it is not a coincidence that in the simulations the share of income spent in manufactures is not above 60%. Such an upper bound means that the peripheral markets (40% of the expenditure!) are large enough to block the “black-hole location”. For empirical researchers, this limit leads to some tricky questions: How would researchers manage applied studies based on a model that requires such parameterization? In other words, in which country (and when) the agriculture goods or fixed sectors have had such importance in the expenditure or aggregate income? Is it the case of industrialized countries?
markets. Once more, agglomerations are not likely. The overall effect of centripetal and centrifugal forces (i.e., a “black hole” or total concentration) depends on transport cost, a parameter exogenously given, that can produce the “bifurcated tale” described above. Nevertheless, without peripheral markets, full concentration would be impossible to avoid.

The other two parameters of the model (expenditure share and preference for variety for industrial and agricultural goods) also influence spatial organization in the same fashion. Given transport costs, agglomeration takes place earlier the stronger the preference for variety and larger the share of manufactured goods in expenditure. Large preference for variety means that price effect is strong. Thus places with many varieties have higher real wages and attract more workers. If consumers expend a great share of their income on manufactured goods, their reaction to the regional supply of industrial goods is greater. Hence, agglomerations occur earlier in economies with consumers valuing variety and spending a large proportion of their income industrial goods.

2.3 Urban hierarchies and the Zipfs’ Law challenge

In the third part, the core-periphery model is redefined in order to generate the spatial structures described by von Thünen (1966), Christaller (1933) and Lösch (1954): the monocentric city and the urban hierarchies characterized by many cities with different sizes and production structures. The basic tool for this task is the market potential function.

A traditional market potential function measures the purchasing power of a region, which depends on the weighted sum of purchasing power of all other sites surrounding that region. The increasing distance between industrial regions and the surrounding markets (agriculture sector) reduces the export of local varieties and restricts the size of the industry, since transport cost increases delivered prices. However, if a city has an industry with internal economies of scale, it can overcome the “tyranny of space” and monopolize the industrial production in a region. On the contrary, if the peripheral markets are large enough and far enough away, firms can move out of the industrial core and a new city emerges. The tensions among transport costs, increasing returns of scale and size of peripheral markets are the central point of the urban system.

Neary (2001, p. 548–549) has a more technical approach to the core-periphery model. His excellent summary shows, among other facts, that the normalizations (e.g., links between tastes and technologies) are not harmless to the model. They can indeed lead to questionable interpretations.
In the core-periphery model, any regional manufacturing is feasible if it can pay a real wage that guarantees a minimum supply of labor. Thus, in Chapter 9, the market potential function is redefined as a potential real wage curve. The basic idea is as follows: as the distance between the regions increases, transport costs rise and the demand for imported varieties decreases. After some critical distance, the potential market in peripheral areas is so large that local production breaks even. This means that firms can pay a higher nominal wage that guarantees a local real wage equal to the one in the industrialized core, so that industrial production can take place.

This model has a similar structure to the previous one. There are two sectors, industry and agriculture. Industry has economies of scale, produces differentiated goods, and demands only one input: labor. Agricultural production has fixed technical coefficients and requires two inputs: labor and land. Land rents are also added to the model, but landlords are only defined as receiving income that is consumed where it is produced. Landlords and/or developers do not play any other role in the model.

Nonetheless, an important change is introduced in the labor market. All workers are now homogeneous (fully mobile labor market), and land is the only fixed factor of production (or the only centrifugal force). Given this set-up, the main question of Chapter 9 is as follows: Is a von Thünen type geography (the monocentric city) sustainable?

The von Thünen city is stable only if the population is small, transport costs are lower, manufacturing share in expenditure is larger, manufacture of goods is differentiated, and firms have significant economies of scale. When the monocentric city is stable, the small population does not demand a large supply of agricultural goods, so peripheral markets are relatively close to industry, and increasing returns overcome transport costs. However, for large populations (and a large demand for agricultural goods), the hinterland increases and, after a certain point, producers have an incentive to relocate into the periphery. Then, the monocentric city structure becomes unsustainable, and a new city emerges. For that reason, a monocentric city is sustainable only if the population is less than some critical value. In other words, the peripheral markets cannot exceed an upper limit.

In Chapter 10, these insights of Chapter 9 serve as the basis for the multiple-city system or the Christaller-Lösch urban hierarchy. There is only one change: the population growth by a given rate. The von Thünen monocentric city is seen as an extreme and simple case in which, in a lesser-populated region, the increasing returns prevail completely over transport costs. The urban hierarchy is a more complex and realistic
case: as population grows, many cities pop up. Each city rises having a
different industrial structure and its own hinterland where no industrial
activity is profitable; only agriculture can be present. The city sizes and
their hinterlands are endogenous, created by a complex interplay of
different transport costs, preference for variety, and internal economies
of scale.

The explicit solutions of the model, either analytically or
numerically, are shown to be quite difficult. For this reason, the authors
use a simulation to discuss urban structuring. The results are very
interesting:

- first, the emergence of a new city changes the size of
  established cities;
- second, cities do not have similar sizes or economic struc-
  tures (different number of varieties and hinterlands);
- third, cities have different rates of growth;
- fourth, the urban frontier is very unstable, and urban cores
  tends to stabilize; and
- fifth, there is no clear urban geometry, such as that pre-
  dicted by traditional urban models.

In short, there is an emergence of a complex and organized
hierarchy of central places.

In Chapter 11, the model is improved by the inclusion of
many industries with different transport costs, substitution parameters,
and economies of scale. The simulations show the von Thünen and
Christaller-Lösch’s models as particular cases of a spatial general theory.
The inclusion of many industries allows a more differentiated regional
structure, and the hierarchical urban system shows first, second, and
third order cities. The highest order cities have a complete economic
structure, and the third order cities have just light industries (large
substitution parameters and high transport costs). There is no doubt that
this is the high point of the second part of the book, and maybe of the
whole book. The authors considered this general equilibrium model as
their most important achievement5.

---

5 In their interesting paper on why cities diversify, Duranton and Puga (1999,
p. 20) also state that this model represents “a landmark in the modeling of
urban systems, since this is the first paper able to deal with central-place theory
in a general equilibrium framework with well-specified micro-economic foun-
dations”.

18 Nova Economia | Belo Horizonte | v. 11 | n. 1 | jul. 2001
In spite of this apparent success, in Chapter 12 the authors warn that the model has a major flaw: it cannot generate a robust Pareto distribution (Zipf's Law), which is a remarkable empirical regularity in any modern urban hierarchy. The rank-size rule states that, for cities within a given country, the outcome of a city's rank and its population is approximately a log-linear regression near linearity with an approximately 45-degree slope. The urban model described above does not automatically produce anything similar to the rank-size rule. In fact, Fujita et al.'s (1999) model is now part of a broad set of models that cannot easily generate such a distribution from the interaction of economic variables. Zipf's Law still remains a challenge for urban economists.

2.4 International trade and economic development

In the fourth part of the book, the focus changes from regions to countries. Once again, the core-periphery model is expanded to incorporate an international scope: regions become countries; exchange rates, tariffs, import quotas and all other obstacles of foreign trade are boiled down in transport costs, and fully mobile workers are replaced by a domestic supply of labor which is mobile exclusively inside each country. There are two relevant innovations: input-output matrices add inter-industrial relations, and a persistent increase in the industrial expenditure share replaces the initial Cobb-Douglas consumer preferences with fixed parameters of expenditure. The more complex industrial structure that emerges is discussed in four chapters.

In Chapter 14, one of the most polemical issues on international economics is debated: the spread of industries from developed to developing countries. The basic theoretical tool resembles the market potential function: given transport costs, the industrialization in developing countries breaks even when rising wages in the core push industries to move to peripheral markets where labor is cheap (nontradable factor of production). However, cheap labor in peripheral areas also means small markets. This works as a disincentive to industrial relocation, since the local production would have limited economies of scale. So, a cheap labor force is not a sufficient condition for industrialization and can indeed reinforce agglomerations. As industrialization in the core proceeds, nominal wages increase and the gap between rich and poor countries increases: there is a divergence in income per capita.

6 Brakman et al.'s (1999) paper is a good example on the difficulties met by urban economists when trying to simulate Zipf's Law using the new economic geography tools. For some researchers, Zipf's law constitutes a minimum criterion of admissibility for any model of cities. (Garbáč, 1999).
The core-periphery structure is stable up to a certain point, when centrifugal forces (expensive labor force in the core and cheap wages in the peripheral areas) overcome the centripetal forces (the economies of scale and the size of central markets relative to those in peripheral countries). The rupture of equilibrium is abrupt and industries start to move to developing countries or industrialization becomes possible in the agricultural regions.

Industrialization in peripheral areas follows the same steps as industrialization in regions of a country:

- high transport costs means autarkies and no agglomeration;
- intermediate transport costs create global agglomeration and international markets;
- low transport costs produce the opposite: agglomeration does not happen; and
- real wages are equalized.

Other changes in the parameters of the model have the same effects as those described for the core-periphery model.

In Chapter 15, a similar model on the industrialization of peripheral markets is presented. In this version, there are many industries with different backward and forward linkages, and the driving force is the expansion of demand from agriculture to manufacturing. It is supposed that technical progress steadily augments all primary factors, which is labor supply. The most interesting aspects of the model are the waves of industrialization and the spread of different industries throughout peripheral countries. The first industries to relocate are those with heavy labor costs or limited industrial backward and forward linkages (light industries). The first wave of industries increases wages as well as the size of peripheral markets. It also creates conditions for the production of linked industries, which can lead to waves of growth until heavy industrialization (industries with high forward and backward linkages) takes place.

Two interesting questions concerning these two models can be asked. First, in Chapter 14, which are the determinants of decreasing transport cost? The time-path of this variable is crucial for the model, but there is no observation at all with regard to how and why it is changed. In Chapter 15, a second and also obvious question refers to how the increasing expenditure in manufactured goods is related to an exogenous technical progress that affects only the supply of labor. How might one explain this? A simple and easy answer would be to consider investment...
in human capital, since firms are quite passive agents and cannot be responsible for technological progress. If this is the case, the core-periphery model depends upon the validation of another theory.

The next three chapters of part three are a generalization of the model developed in Chapters 14 and 15. Chapters 16 and 17 discuss the clustering of industries with different input-output coefficients and transport costs among countries. The conclusions recall those in the simplest version of the core-periphery model:

- at high trade costs, industries operate in all countries;
- for intermediate values, clustering is possible and asymmetric economies emerge;
- and for low trade costs, clustering is not expected.

Chapter 18 is peculiar, since it is the most questionable application of the core-periphery model and the only chapter in which a clear policy is proposed: trade openness can bring spatial decentralization of industry and spatial clustering of particular industries. The first drawback in the model is the assumption that trade openness is correlated with economic growth (in this case, growth means an increase in the real wages). There is no discussion at all as to whether increasing imports affect the industrial sector, if there are external markets for domestic firms or even if imports take the place of the domestic production. The second shortcoming is the removal of peripheral markets and their replacement by congestion cost functions. As a result, any negative or positive impact of trade openness in the peripheral markets is canceled out. Third, the outcome depends on very specific and unrealistic transport costs among domestic regions and foreign markets.

The peculiar transport cost structure in Chapter 18 suggests that transport costs among all cities and the outside world are equal. However, this argument runs against the main conclusion of Chapter 13, in which the largest city is also that with the lowest transport cost. In such a case, trade liberalization could reinforce some advantages of the main city. For example, trade liberalization reduces the cost of agricultural goods, which means decreasing income in the hinterland and declining living costs in cities. Thus, migration to the main cities is stimulated, and the outcome is contrary to that suggested. The opposite

7 This argument is analogous to that defended by David Ricardo in the “Corn Question”: the import of food allows reductions in the cost of living and increases industrial growth. Here, the import of agricultural goods reduces the agricultural price index in the major cities and reinforces urbanization.
can also happen when trade openness increases income in the hinterland. The growth of peripheral markets attracts new industries, and the increasing price of agricultural goods in main cities stimulates migration to the countryside, all of which can happen without any type of ad hoc congestion cost function. In brief, the model is too simple, it leads to misleading policies, and its conclusions cannot be sustained even in the core-periphery model style.

3 CRITICAL REFLECTIONS ON THE NEW ECONOMIC GEOGRAPHY APPROACH

The review above indicates that, despite its flaws, the new economic geography has a very consistent approach to regional and urban economics. However, it does not guarantee any consensus among economists or geographers. In fact, it is possible to identify a marked division that creates two “theoretical agglomerations.” The first agglomeration results from the project proposed by Isard during the 50’s, which was the creation of a regional science. The main objective was to include space in the neoclassical models; consequently, the new economic geography is a renovation of the research agenda proposed by the old regional economists. The leading names in the new economic geography are Brian Arthur, Masahisa Fujita, Paul Krugman, Gianmarco Ottaviano, Diego Puga, Phillippe Martin, and Anthony Venables. All of these authors share the same methodological approach: rational-choice / equilibrium model, such as the core-periphery model described above.

Geographical economics is the second agglomeration which is represented by researchers using a different methodology. One of these researchers is Martin (1999), who argues that positivist models, such as those presented by Isard (1956) and Fujita et al. (1999), have long disappeared from the research frontier because they could not explain the multiplicity of cases in which political, economic, and institutional aspects were interconnected in very specific ways. Martin states that

8 In the Brazilian case, Diniz (1999) shows that trade openness has a dubious impact on regional structures, and it appears to reduce the centrifugal forces. This results in an outcome opposite to that predicted by the new economic geography. Haddad and Azoni (1999) also say that trade liberalization is likely to harm peripheral regions.

9 Barro and Sala-i-Martin (1995) are important for the debate on the convergence of regions, but their works are much less based on traditional location theory and regional science. Michael Porter could also be part of the new economic geography group, but his models have a different methodological perspective.
“Geographers became more interested in real economic landscapes, with all their complex histories and local contexts and particularities, and less entranced by abstract models of hypothetical space economies.”

The main difference between the new and old economic geography and geographical economics lies

“on philosophical and epistemological grounds, as part of the large-scale movement away from logical positivism that occurred in geography at that time (late 70s)”. (1999, p. 81).  

In the next two topic sections, the recent debate between these two groups of regional economists is recreated. Some commentary on the relevance of the critiques is also pointed out, since it seems that there are some misunderstandings and other overlooked topics.

3.1 Analyzing the Basic-Machinery

Critics of the new geographical economics have claimed that models - such as the core-periphery model have several failures. Dymski (1996) criticizes Krugman’s models asserting that they have three main problems. First, the incorporation of the interdependence between firms’ decisions on how much to produce and where to produce is limited. Second, the model is vulnerable to the capital critique. Third, the footloose firms do not consider sunk costs and costly transport of capital. So, several “frictions” (relatively nontradable factors or sluggish adjustments) that could justify shifts in technology, restructuring of industrial districts, different pricing, and existence of fixed markets and suppliers are not considered.

David (1999) and Neary (2001) also claim that the supply side is not stressed in Krugman’s models and that there are several effects at the micro level that are quite important in the agglomeration/decentralization processes, such as localized knowledge spillovers, intangible input markets (service sector), and public infrastructure. Isard (1999) is confident that the introduction of more imperfect competition would produce significant changes. The use of game theory, coalition formation and

10 Other authors who have similar opinions are M. Castells, D. Harvey, A.M. Isserman, A. Markusen, M. Piore, C.F. Sabel, E.S. Sheppard, M. Storper and P.J. Sunley. See Scott (2000) for an intellectual history of economic geography.
disruption theory could lead to some interesting outcomes different from those of the standard model. Fujita and Thissse (1996) and Ottaviano and Puga (1997) also agree with this criticism and say that the main limitation of the monopolistic competition models lies in the absence of strategic behavior\footnote{Anas (2001, p. 611) is more emphatic, he says: “In effect, the FKV [Fujita, Krugman and Venables] apparatus is a monopolistic competition model, with strategic behavior squeezed out of it and thus made to work much like a model of perfect competition would!”}.

Fujita et al. (1999, p. 52) are aware of some of these criticisms. On the oligopolistic interaction, they believe that any inclusion of inter-depenendency would only change the price rule without any important change in the model. In their opinion, some interaction could be a second force operating in the same direction.

The inexistence of sunk costs brings some problems. The simulations show that their absence produces undesirable results, such as the constant creation and destruction of new cities in the frontier. An ad hoc solution proposed by the authors to eliminate this “odd phenomenon” is the introduction of urban infrastructure (Fujita et al., 1999, p. 213, footnote 11). However, since there is no government, the footloose firms should pay for this extra-cost, which would in turn make them relatively fixed. They do not discuss the consequences of such a change in the model, but it seems that the existence of urban infrastructures would delay the spread of firms to the hinterland and would change the pattern of regional growth\footnote{See also Duranton and Puga (1999, p. 19-21) for different considerations of the problems of this urban model.}.

Isserman (1996) is much more emphatic in his concern with regard to the limits of the models. Debating the impact of trade openness in the internal geography, Isserman argues that it is not possible to accept the main conclusions defended by Krugman and Livas (1996), which are also presented in Chapter 18 of The Spatial Economics. He stressed two points:

1) the transportation infrastructures are a “mystery”; and

2) the asymmetric economic structures that characterize different regions are not discussed at all.

Again, the homogenization of space, firms, and consumers is the center of criticism.
David (1999) and Isard (1999) also propose changes in transportation costs. They say that a first advance would involve dropping the assumption of iceberg costs and including an endogenous transport cost index. There are also major indivisibilities, cumulative effects, and increasing returns in transportation, which would affect the whole model. In their opinion, without a clear understanding of the transport structure, any model could be named “unrealistic” or lead to erroneous conclusions.

Martin and Sunley (1996) argue that there is no doubt that differentiated regional structures are not captured by homogenous agents in a clean landscape, which is the basis for the new economic geography. Additionally, they also affirm that in the new economic geography the reallocation mechanisms are overstressed, and Schumpeterian waves of innovation, which is the core of economic diversity, are not considered.

Of all these criticisms, however, there are two that must be emphasized. First, the core-periphery model is static; its dynamic is completely concentrated in a sluggish adjustment of the labor market. Firms set prices, buy inputs, pay wages (income), and produce, knowing how much of each good the consumers will buy when spending their income (wages). Employment, income, prices and output are defined instantaneously. Say’s Law is also part of the model: firms and consumers spend all income, and there are no savings, injection of demand or unemployment. Thus, a potentially richer and dynamic economic process is put aside in order to get a tractable model.

Second, firms are passive agents. There is complete capital mobility, which means that there are no sunk costs, and an implicit perfect capital market solves any financial problem. In such an environment, firms focus solely on demand and invest (if applicable?) without any constraint. There is no specific asset (nontradable input) that can create economic asymmetries among producers. Neither learning processes nor imitation are present, technology is given, and product differentiation is part of the “consumer’s world.” In short, firms have “no-memory” or active behavior. This is why varieties are not expelled, firms do not fail, and there is no loss of capital (in urban economic jargon: “dead cities”). An astonishing result of the model is that there are economies of scale and spatial concentration, but no economic concentration. In addition, the preference for variety, which is the bedrock of the demand side, is a very questionable assumption.

In the new economic geography literature, there is a limited debate on technological change. In general, relocation is the most stressed aspect.
In fact, there is no individual firm in the model; regions or cities are the “microagents”. Each region or city has an aggregate production function characterized by increasing returns with the number of varieties given by the amount of workers. This methodological option restricts the introduction of “micro-asymmetries” and blocks the emergence of a truly self-organizing system based on the agents (firms and consumers). Quoting Fujita et al. (1999, p. 27):

“I in economic modeling we try to show how a phenomenon emerges (there’s one of those words again) from the interaction of decisions by individual families and firms; the most satisfying models are those in which the emergent behavior is most surprising given the players’ micromotives.”

So, what is disappointing about their model is that it gives a limited account along these lines. Furthermore, what is astonishing is that this is precisely their dissatisfaction with Lösch and Christaller’s central-place theory.

The third criticism is obvious: in the basic model land is absent. Since von Thünen, any model of regional economics has to take into account, even briefly, the existence of land rents as well as their impacts on the regional distribution of incomes and activities. Hence, it is quite shocking to realize that there is no discussion of the determinants of land cost or the behavior of landlords and developers. There are just brief notes, such as where the landlords spend their rents. Nothing is said about the behavior of these agents: how they invest, how land is “created” and appropriated, the cost of the land in urban and rural areas, how landlords and developers extract income from firms and households etc. Fujita et al (1999) are conscious of the importance of these players, but they just avoid the analysis saying that it is an abstraction used to simplify the model.

The fourth criticism is not usually stressed, but is decisive for the theory: the way that the authors deal with peripheral regions (fixed production and demand). The peripheral market structure and transport costs are analogous to Isard’s geographic cost pattern, which encompasses a fixed and relatively stable spatial supply of factors of production and their costs (Isard, 1956, p. 138). Isard considers these factors as given and Fujita et al. (1999) see the peripheral markets as predetermined and/or determined by exogenous factors (e.g. population growth and land).
In the core-periphery model, peripheral markets are the most relevant (maybe the only) centrifugal force. Consequently, sustainability of industrial cores depends on peripheral markets. However, there is no systematic discussion concerning the determinant of their growth rate, productivity, or intersectoral relations. There are just a few hypotheses about their structures. For example, the size of the agricultural sector depends on population size, and the demand for agricultural goods decreases as income level increases. Beside these notes, nothing else is stated about the nature of the peripheral markets.

Actually, the polemic on the peripheral markets starts when it is just conceptualized:

“Of course, the label ‘agriculture’ need not always to be interpreted literally; the sector’s defining characteristic is that it is the ‘residual’, perfect competitive sector that is the counterpart to the action taking place in the increasing-returns, imperfectly competitive manufacturing sector.” (Fujita et al., 1999, p. 45).

However, throughout the whole book, the peripheral markets emerge as a more specific competitive sector than the one described above. They are a competitive sector that requires a nontradable factor of production: land. The peripheral markets are, in fact, the agriculture sector.

The relevance of the specific competitive sector in the theory is clear in Chapter 18, where the authors change the model to discuss the relation between external trade and internal geography. The basic modification is that there is no agricultural sector in the economy. The authors say:

“As we have set it up so far, this model contains no form of diminishing returns. Because there is only one factor of production, and it is mobile between the two domestic locations, there is no apparent reason why all labor should not concentrate in one location or the other. To produce an interesting tension between centripetal and centrifugal force, then, we must introduce some counter-vailing force. Once such force would be the existence of immobile factors, such as land; indeed, that is how we have created a tension in other models.” (Fujita et al., 1999, p. 331–332).
Maybe they are not conscious of the importance of the sentence, because with this statement they are indeed back to the Isard's geographical cost pattern described above. If this is true, the increasing return models applied to regional economics depend on new ad hoc assumptions:

1) the model needs large peripheral markets that are immune to those effects that come from industries with increasing returns, and/or

2) it needs fixed factors of production that bring to regions specific industries that create local markets.

Neither assumptions are fully explained by the theory proposed in the book.

The importance of the theoretical status of the peripheral markets is also clear in Pines (2001), who defines them strictly as an agricultural sector:

“In this sense, the present study [The Spatial Economy] provides an excellent explanation for the emergence of urban structure... during the surge of urbanization.... Notwithstanding the above, the explanation provided by the present study for the emergence of urban structure is less relevant to what happened in advanced economies in the 20th century and, especially after the Second World War” (Pines, 2001, p. 144).

This doubt about the relevance of the theory is correct if fixed markets are just the agricultural sector. If so, Fujita et al.’s (1999) theory is not a general theory, as they propose. It is just a theory that explains urbanization during a specific historical period: that which took place in developed countries during the 19th century!

3.2 Diversity in a “Black Box?”

A focal criticism on the new economic geographic approach is the limited incorporation of “real places.” Martin (1999) and Martin & Sunley (1996) say that in these models regions and locations are just points without any particular characteristic; the economic space is homogeneous. Hence, these models cannot take into account patterns of uneven regional development and structural changes. In their opinion,
these models fail when incorporating the diversity of regional structures that mainly characterize developing countries, but also those developed ones.

Isard (1999) seems to agree with this critique and suggests that different kinds of aggregation of economic activities have to be theorized. The traditional classification of scale, localization and urbanization could be replaced by different spatial connections of activities as meaningful aggregations, which he called “subaggregates” or “mini-complexes”. This improvement would be an exploration of “self-organization and co-evolution of dual spaces consisting of the traditional economic space and the knowledge mind-culture space”14.

The diversity demanded by these authors is evident in the discussion of industrial districts and growth in metropolitan regions. Markusen (1995, 1996) argues that there are at least four types of industrial districts or agglomerations:

- the traditional Marshallian one (spatial concentration of small and similar firms);
- the hub and spoke agglomeration (in which large firms play an important role);
- industrial enclaves (in which subdivisions of corporations search for low costs); and
- state-centered industrial districts (capitals of states, military installations, and public firms work as anchors of regional development).

These districts differ in their origin, production structure, growth rate, institutional organization and social relations. These different types of agglomerations have no clear place in the new economic geography.

Pollard and Storper (1996) study the employment change in twelve major American metropolitan areas and suggest that there is a multiplicity of urban trajectories. For many reasons, some cities were able to diversify to the most dynamic activities, others remained dependent on stagnant industries, and a third group connected its original structures to specific branches of the new industries. Technological changes seem to be the most important determinant of the asymmetries, and cities engage in the new technological path with varying degrees of success.

14 Fujita (1999) agrees that microorganizations would be an improvement in their general model, but he does not explicitly show how to include them.
which leads to differences in growth rates as well as in regional and urban structures.

In short, the differences between Krugman's clean space and the space claimed by the critics seem to go deeper. The critics stress that relevant spatial differences do not result only from natural or given aspects. The most important differences are the ones related to the dynamic of the "social geography": the differences among regions and cities that are local, nontradable and specific to the organization of agents (institutions, unions, networks of firms and so on)\textsuperscript{15}.

Many of these criticisms might be correct, but there is some confusion. For example, Martin and Sunley (1996), Martin (1999) and many other nonmainstream regional economists criticized the new economic geography because it does not take into account spatial asymmetries (e.g. regions are points organized in a circle or line in a clean landscape). This is an incorrect critique and a theoretical misunderstanding. The racetrack economy (regions organized in a circle) does not represent Euclidian space (physical distance among regions). The distance between two regions means all difficulties that firms have in delivering their products outside their local markets or all costs paid by consumers when importing products from nonlocal firms. These costs are not a function of any Euclidian distance. They are a function of all frictions that makes trade among regions costly. Hence, there are asymmetric spaces in the new economic geography and they play a crucial role in the models.

The agents are also asymmetric. In the new economic geography, the asymmetries among agents are mainly related to:

1) different substitution parameters;
2) transport costs;
3) input-output coefficients;
4) increasing returns; and
5) degree of factor mobility.

These four factors together are responsible for diversity in the landscape. Besides these factors, all locations are assumed to be identical "as a way of isolating the pure forces of geographic self-organizing which results in growth rates as well as in regional and urban structures.

\textsuperscript{15} Glaeser \textit{et al.} (1992) also show that there were differentiated growth rates among the most important US cities after the 50's. They conclude that the cities with high growth were those with more diversified industrial structure. Thus, patterns of local interindustrial relations play a crucial role in regional and urban growth.
zation rather than adulterating our analysis with inherent locational distinctions" (Fujita et al., 1999, p. 322). This is a quite acceptable way of stressing some determinants of spatial organization. In addition, Krugman et al. (1999) could reply to these critiques through their concept of transport costs:

“(…) the transport costs that appears in our models are only metaphorical. We are really interested in all of the costs of doing business over geographical space. In other words, we want for the theory a measure of the full cost, including all the costs of doing business at a distance – lack of face to face contact, more complex and expensive communications and information gathering, and possibly also different languages, legal systems, product standards and cultures” (Fujita et al., 1999, p. 98).

Nevertheless, there are two problems with this reply. First, transport costs in the core-periphery model are not discussed at all and there is no theoretical justification of its main determinants. It is an exogenous variable! Moreover, any spatial organization reflects the level of this holistic transport cost. In fact, in many simulations, a small change in transport costs produces a totally different spatial organization. Thus, when critics say “regions are just a circle in a plain landscape”, they are misunderstanding the relevance and the concept of transport costs. The correct criticism should ask why circles and lines are not just regions separated by Euclidian distances in plain landscapes. There is no doubt that, in the new economic geography, transport cost is a theoretical “black box” that has to be opened.

Second, it is quite difficult to reproduce all regional diversity in the core-periphery model without questioning the relevance of some of its theoretical aspects. It is not only a technical problem related to mathematical tools, but also a very deep theoretical impasse. In the new economic geography, “social space” and its intrinsic diversity would only be a set of secondary variables that do not need to be included in the model. “Clean space” is a simplification and also a selection of forces that drive the spatial organization. Perhaps Martin (1999) is correct when he says that the new economic geography and geographical economics have indeed irreconcilable discourses, since they have different views on the forces that drive spatial organization.

16 See Martin (1999, p. 80-84).
4 NEW, OLD ... WHAT IS IT?

There are some questions that still need to be answered. After all, is the new economic geography a new regional science or just a high-tech glossary of the old regional science conceived by Walter Isard in the 50’s? If it is a new regional science, what is new? The best answers for these questions are: the new economic geography is not that new, but there are some novelties in its approach.

The positive aspects are clear. First, the model merges into a single framework the most traditional and relevant models of regional economics: Lösch, Christaller and von Thünen’s approach are now part of a well-defined theoretical system. Second, it allows the emergence of stable and asymmetric regional structures in an environment with increasing returns. Traditional regional science stressed differences in endowments, technologies and policy regimes; the new economic geography adds to the theoretical menu the monopolistic competition with increasing returns and reemphasizes that space is not homogenous, even in a “clean landscape”. Consequently, it is not a general theory; it is just a continuation of a research agenda. Third, many variations of the basic model incorporate valuable insights from theories that, until recently, were not seen as relevant for mainstream economics. The outcomes are interesting descriptions on how industries spread and shape the economic space in an irreversible and discontinuous way. Fourth, because the core-periphery model is not as rigid as a general equilibrium model, it allows many adaptations. Given regional diversity, any plasticity is a welcome aspect. Fifth, there is no general conclusion as to which is the best regional structure or the optimum spatial organization. Each spatial organization has initial conditions and historical accidents that create specific path-dependences and landscapes. There is a multiplicity of equilibria. In the new economic geography, history can matter!

In spite of the noted flexibility, for some purposes the rational choice / equilibrium models are still very rigid. First, they are static models and, as for the core-periphery model, its dynamic is fully concentrated in a sluggish process of adjustment in the labor market. The firm – a key player in the regional and international game – is a passive actor: there are no disequilibria whatsoever on the supply side, no growth or endogenous technological changes. The firm is a very simple agent, maybe too simple for the rich regional structure that characterizes any country. In short, the model still has a lack of “micromotives” and “microdiversity”. As a result, it allows only limited theoretical studies with asymmetric agents, local learning, competitive selection and many other cumulative processes that are peculiar to regions and agents in a
dynamic environment. Following Schumpeter's (1934) view on economic development, the new economic geography overstresses the relocation of factors and underestimates structural changes.

What is not new in the new economic geography are those aspects related to its main concepts and theories. The models of Lösch, Christaller, Pred, and von Thünen are very traditional models, and their conclusions are simply restated. Because of this, the majority of empirical studies based on the so-called new theories do not add any new variables, nor do they establish different relations or reach original interpretations. In fact, many of the empirical studies are quite conventional, some of them are questionable regarding their proof of theoretical conclusions, and the majority of case studies are very "contaminated" by particularities.17

The waves of industrialization and the patterns of agglomeration in Part III are not new at all, in fact what is presented in this section is a quite standard view of the industrialization in developing countries, at least in some nonmainstream economics. For example, Gerschenkron (1962), Hirschman (1958), Prebisch (1950), Chenery (1975), and Myrdal (1957) saw development as a sequence of abrupt ruptures of equilibria. Almost all these authors stressed the importance of economies of scale, backward and forward linkages, and the size of domestic markets in industrialization. Concerning these aspects, the new economic geography is just a high-tech glossary of some very famous and insightful theories of urban and regional development.

5 REFERENCES


17 Ottaviano and Puga (1997) do a survey of the new economic geography and conclude that after almost a decade of research, there are only a small number of empirical tests that strictly corroborate the proposed theories. In fact, they mention less than ten works. Neary (2001) reaches a similar conclusion.


LÖSCH, A. The economics of location. New Haven, Conn.: Yale University Press, 1954.


36 Nova Economia | Belo Horizonte | v. 11 | n. 1 | jul. 2001