DISCUSSION PAPER SERIES

No. 7063

NEW ECONOMIC GEOGRAPHY: AN APPRAISAL ON THE OCCASION OF PAUL KRUGMAN'S 2008 NOBEL PRIZE IN ECONOMICS

Masahisa Fujita and
Jacques-François Thisse

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7063.asp**

ISSN 0265-8003

NEW ECONOMIC GEOGRAPHY: AN APPRAISAL ON THE OCCASION OF PAUL KRUGMAN'S 2008 NOBEL PRIZE IN ECONOMICS

Masahisa Fujita, Konan University and RIETI
Jacques-François Thisse, Université Catholique de Louvain, Zhejiang University and CEPR

Discussion Paper No. 7063
December 2008

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INTERNATIONAL TRADE AND REGIONAL ECONOMICS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

CEPR Discussion Paper No. 7063

December 2008

ABSTRACT

New Economic Geography: an appraisal on the occasion of Paul Krugman's 2008 Nobel Prize in Economics*

Paul Krugman has clarified the microeconomic underpinnings of both spatial economic agglomerations and regional imbalances at national and international levels. He has achieved this with a series of remarkably original papers and books that succeed in combining imperfect competition, increasing returns, and transportation costs in new and powerful ways.Yet, not everything was brand new in New Economic Geography. To be precise, several disparate pieces of high-quality work were available in urban economics and location theory. Our purpose in this paper is to shed new light on economic geography through the lenses of these two fields of economics and regional science.

JEL Classification: F12, L13 and R12
Keywords: economic geography, location theory, trade and urban economics

Masahisa Fujita
President and Chief Research Officer
Research Institute of Economy, Trade and Industry (RIETI)
1-3-1 Kasumigaseki Chiyoda-ku
Tokyo 100-8901
JAPAN
Email: fujita-masahisa@rieti.go.jp

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=118328

Jacques-François Thisse
Université Catholique de Louvain
34 Voie du Roman Pays
B-1348 Louvain-la-Neuve
BELGIUM

Email: jacques.thisse@uclouvain.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=104425

Submitted 27 November 2008

* We thank T. Smith for helpful comments and suggestions. This paper was written when the second author was P.K. Professor of Economics at Zhejiang University.

1. Introduction

Economic geography seeks to explain the riddle of unequal spatial development. The most salient feature of the spatial economy is in effect the presence of a large variety of economic agglomerations. Hundreds of books and papers have been devoted to this topic but it was not until Paul Krugman's (1991a) seminal paper that a full-fledged general equilibrium model became available to explain why, how and when the economic activity may be agglomerated in a few places.

Although using "economic agglomeration" as a generic term is convenient at a certain level of abstraction, it must be kept in mind that this concept refers to very distinct real world situations. At one extreme of the spectrum lies the North-South divide. At the other extreme, agglomeration arises when restaurants, movie theaters, or shops selling similar products are clustered within the same neighborhood, not to say on the same street.

What distinguishes those various types of agglomeration is the *spatial scale*, or the spatial unit of reference, chosen in conducting one's research, very much as there are different levels of aggregation of economic agents. Whatever the scale of analysis retained, the emergence of economic agglomeration is naturally associated with the emergence of inequalities across locations, regions or nations. Such inequalities are often at the origin of strong tensions between different political bodies or jurisdictions, or even social, religious or ethnic groups when they are geographically concentrated. Understanding how spatial inequalities in living standards arise is thus a fundamental challenge for economists and regional scientists.

Paul Krugman has clarified the microeconomic underpinnings of both spatial economic agglomerations and regional imbalances at national and international levels. He has achieved this with a series of remarkably original papers and books that succeed in combining imperfect competition, increasing returns, and transportation costs in new and powerful ways.

Yet, not everything was brand new in Krugman's New Economic Geography. To be precise, several disparate pieces of high-quality work were available in urban economics and location theory. Our purpose in this paper is to shed new light on economic geography through the lenses of these two fields of economics and regional science.

2. On the relationships between economic geography, urban economics, and location theory

How is economic geography related to other fields such as **urban economics** and **location theory?** The former aims to explain the internal structure of cities, that is, (i) how land is distributed among plants, offices, dwellings, and infrastructure, and (ii) why cities have one or several central business districts. The basic concept of urban economics is the *land market,* which serves to allocate both economic agents and activities across space. Alonso (1964) and Mills (1967) may be considered as the founders of this field. Fujita (1989) has provided what remains the best overview of the state of the art up to the 1990s. More recently, urban economics has addressed a large number of new issues, ranging from the microeconomic foundations of urban agglomeration economies to the impact of neighborhood effects and other spatial externalities on the social stratification of cities (Duranton and Puga, 2004; Zenou, 2009).

The latter is concerned with the location of economic activities, with a special attention paid to the geographical distribution of firms and the geographical variations in prices and costs. This field has its roots in Hotelling's (1929) path-breaking paper "Stability in Competition". The value and importance of this contribution was brought to light in the 1980s thanks to the emergence of non-cooperative game theory. Its use exceeds the original geographical interpretation to accommodate almost any dimension that differentiates firms and consumers in a given market. To be precise, the spatial framework proposed by Hotelling may serve as a powerful metaphor for dealing with issues involving heterogeneity and diversity across agents in a host of economic, political and social domains (Rosen, 2002). Examples include the specification of products by firms competing for customers in industrial organization and the choice of a political platform by parties competing for votes in political science.

Whatever their respective merits, it is fair to say that neither urban economics nor location theory has reached the level of visibility and interest achieved by economic geography. Why is this so? Though the scope of urban economics is perhaps too narrow for general economists, the wider field of location theory is not perceived as a well-defined domain of research, probably because of its dissemination across many economic fields. In contrast, economic geography or, to use the most common terminology, **New Economic Geography** (henceforth, NEG) has a well-defined and yet broad objective: *it is the first body of economics that seeks to provide a detailed description of spatial inequalities that emerge as the outcome of a full-fledged general equilibrium model*. Such inequalities are at the heart of many contemporary debates, both between developed and developing countries and within developed countries.

To be sure, the existence of interregional inequalities has long attracted the attention of economists, especially in the area known as **regional economics**. For a long time,

however, regional concepts, models and techniques were often a mere extension of those used at the national level, with an additional index identifying the different regions [think of interregional input-output matrices]. Despite valuable contributions made by Myrdal and Perroux in the 1950s, no one before Krugman had been able to show how regional imbalances could arise within the realm of economic theory. But why did it take so long? As shown in the next section, the answer lies in the inability of the dominant paradigm of economic theory to handle the "regional question".

3. The Spatial Impossibility Theorem

To start with, it is useful to summarize briefly what the neoclassical model has to offer regarding the spatial distribution of activities. In a frictionless world, production factors will act to remove any inequalities in earnings by moving from regions where their remunerations are lower to those where they are higher. If, as assumed in the neoclassical model, there is perfect competition and constant returns to scale, then the marginal productivity of each factor must rise in the regions of origin and fall in the regions of destination. At spatial equilibrium, the remuneration of each production factor must then be the same across regions. As a result, in a world characterized by the "death of distance", it would suffice to wait for this prediction to occur. However, this argument has a second and more far-reaching implication: each region becomes *autarchic* as it produces only for its local market. Such a prediction would be very hard to sell.

The difficulty encountered by using the competitive paradigm to study the location of firms and households was definitely made clear by Starrett (1978), in a paper that deserves far more recognition than it has achieved. To motivate his result, note first that firms tend to produce in only a few places, and likewise, that households typically have a single residence. Hence it is reasonable to assume that each firm (each household) chooses a single location and engages in production (consumption) activities where it is established. However, firms and households are free to choose any location they wish in a given set of places. This set is said to be *homogeneous* when (i) the preferences of each household are independent of its location, and similarly that (ii) the set of production technologies available to a firm is independent of its location. In other words, consumers' and producers' fundamentals are not influenced by their locational choice. Within this setting, Starrett proved the following unsuspected result, which has been called the *Spatial Impossibility Theorem*:

Consider an Arrow-Debreu economy with a finite number of agents and locations. If space is homogeneous and transport is costly, then there is no competitive equilibrium involving transportation.

What does this mean? If economic activities are perfectly divisible, then a competitive equilibrium exists in which each location operates as a separate autarchy.[1] For example, when households and firms are identical, all locations have the same relative prices and the same production structure. In particular, if the assumptions of the neoclassical model were true, factor mobility and interregional trade would be *perfect substitutes*, as shown by Mundell (1957). In retrospect, this hardly seems surprising in a world where activity can operate at arbitrarily small levels. At equilibrium, firms and households naturally disperse their activity all over the places, thus reducing transport costs to zero.

However, when economic activities are *not* perfectly divisible, the transport of some goods between some places becomes unavoidable. In this case, the Spatial Impossibility Theorem tells us that *no competitive equilibrium exists*. This is a very strong result in that it shows the impossibility of having trade once agents are free to choose their location within a featureless space. When regions are not autarchies, one should keep in mind that the price system must serve two purposes simultaneously: (i) to support trade between regions (while clearing the markets in each region), and (ii) to prevent firms and households from relocating. The Spatial Impossibility Theorem says that, in the case of a homogeneous space, it is impossible to hit two birds with one stone: *the price gradients supporting trade bear wrong signals from the viewpoint of locational stability*. Formally, the reason for the Spatial Impossibility Theorem lies in the non-convexity of the set of feasible allocations. This non-convexity is caused by the combination of two elements: the existences of positive transport costs and the fact that agents are not ubiquitous, but rather have distinct and specific locations in space (Fujita and Thisse, 2002). Therefore, ignoring transport costs in economic theory is far being an innocuous assumption despite "the sanguine hope that if included they would not materially affect the results" (Deardorff, 1984, p.470),

4. What are the alternative modeling strategies?

Thus, if we want to understand something about the spatial distribution of economic activity and, in particular, the formation of major economic agglomerations as well as regional specialization and spatial inequalities, then the Spatial Impossibility Theorem tells us that we must make at least one of the following three assumptions: (i) space is heterogeneous, (ii) externalities in production and consumption exist, or (iii) markets are imperfectly competitive. Of course, in reality, the space-economy is the outcome of different combinations of all three ingredients. However, it is convenient to consider each separately in order to disentangle the various effects at work. Note also that some of these ingredients may be relevant at one spatial scale but less at another one.

[1] This does not mean, however, that no transactions take place within each location and that locations are identical closed economies.

(i) ***Comparative advantages***. Models of comparative advantage focus on the economic consequences of spatial heterogeneities among regions such as the uneven distribution of immobile resources (natural harbors) and amenities (climate). There are also important spatial heterogeneities with respect to the locations of transport nodes (transshipment points) and trading places (central business districts). Under the assumptions of constant returns and perfect competition, these heterogeneities generate comparative advantages across space, which in turn give rise to specialization and trade [as the central business district in the monocentric-city model of urban economics or the different endowments of production factors in the Heckscher-Ohlin theory of international trade]. Even though such spatial heterogeneities are important, they cannot by themselves explain the formation of large agglomerations and spatial inequalities. Furthermore, some of these heterogeneities [think of trading places or the supply of various amenities] are not given by nature and should be treated as being endogenous.

(ii) ***Agglomeration externalities***. Unlike comparative advantage models, the basic forces for spatial agglomeration and trade in models of agglomeration externalities arise from non-market interactions that yield increasing returns external to firms. They include knowledge spillovers, business communications, face-to-face communication, and other spatial externalities. Again, this approach allows one to appeal to the "constant return/ perfect competition" paradigm, which probably explains why it appeared long before the development of NEG. The pioneering contributions here are due to Ogawa and Fujita (1980) as well as Fujita and Ogawa (1982). Traditionally, spatial externalities have been treated in a "black box" manner that tends to hide the actual micro-interactions giving rise to such externalities. The study of micro interaction behavior is now an active area of research. But in spite of these efforts, we still know very little about how such behavior leads to those aggregate external effects that are central to this family of urban models. Furthermore, such externalities seem to be critical in the small but less essential in the large. This opinion is backed by the many empirical analyses undertaken during the last decade and all suggesting that spillover effects are much localized (Audretsch and Feldman, 2004; Rosenthal and Strange, 2004).

(iii) ***Imperfect competition***. Models of imperfect competition no longer treat firms as price-takers, but rather assume that their pricing policies depend on the spatial distribution of both consumers and other firms. This in turn leads to significant interdependencies between the location choices made by firms and households. For under imperfect competition, the location of firms can be explained primarily in terms of the search for privileged access to customers and the desire to relax competitive pressures imposed by other firms. In this context, it is useful to distinguish two modeling approaches to market competition.

A. **Oligopolistic competition**. Models of oligopolistic competition assume the existence of few "large" agents (firms, local governments, land developers) who interact strategically by taking into account their respective shares of market power. They belong to the realm of game theory. The prototype is given the spatial competition model in which a small number of firms compete for dispersed consumers.

B. **Monopolistic competition**. Models of monopolistic competition allow certain interdependencies beyond the competitive model in that firms are allowed to set their own prices and to produce differentiated goods under conditions of increasing returns to scale. But even though each firm is a price-maker, there are no real strategic interactions because firms are many. Thus many issues, such as the existence of equilibrium, are less problematic in this setting than under oligopolistic competition.

In supplying (or sourcing from) a market a firm can choose between shipments from (or to) some pre-existent distant plant or sales from (or to) a new local plant. The choice is not trivial when transportation is costly and there are increasing returns to scale at the plant level. Increasing returns lead the firm to concentrate its production in a few plants, whereas transport costs raise the issue of where to locate those plants. Increasing returns and transport costs are, therefore, the two fundamental ingredients to take on board in any model that aims to explain the space-economy. This was already known to Koopmans's (1957, p. 154): "without recognizing indivisibilities - in human person, in residences, plants, equipment, and in transportation - urban location problems, down to those of the smallest village, cannot be understood."

Adding increasing returns to scale internal to firms has some fundamental implications for the nature of competition in space. This point warrants further discussion. In his review of Chamberlin's book, Kaldor (1935) claims that a firm affects the sales of its neighboring firms, but not distant ones. This suggests that competition across locations is inherently oligopolistic (Eaton and Lipsey, 1977; Gabszewicz and Thisse, 1986). Unfortunately, such interdependencies can lead to the nonexistence of pure strategy equilibria. Thus models allowing such interdependencies must either (i) assume mixed-strategy behavior by firms or (ii) assume that such interdependencies are "sufficiently weak" as in monopolistic competition. For the sake of simplicity, most of the economics profession has followed the second option.[2]

In the Dixit and Stiglitz (1977) setting, *monopolistic competition* emerges as a market structure determined both by consumers' preferences for variety and firms' fixed requirements for limited productive resources. On the demand side, consumers' exhibit

[2] This is a reasonable approach as long as we address spatial issues at the macro-spatial level. However, at the micro-spatial level, Kaldor's criticism is still valid and shows that spatial competition cannot be dismissed.

a preference for variety, i.e., their utility increases not only with the amount of each (horizontally differentiated) good, but also with the total number of such goods available.[3] On the supply side, production exhibits internal economies of scale for each good, but no economies of scope across goods. Hence it is assumed that each firm supplies one and only one good ("monopolistic"), and thus that there is a one-to-one correspondence between firms and goods. However, there are no entry or exit barriers, so that prices are just enough to cover average costs ("competition"). Finally, firms are assumed to be so numerous that interactions between any small numbers of firms have no effect on overall supply levels. More formally, there is assumed to be a *continuum* of firms that interact only indirectly in terms of their overall response to aggregate demand levels. In particular, the price-setting behavior of each firm need only take into account the aggregate behavior of its competitors.[4]

The continuum approach not only provides a natural way to capture Chamberlin's intuition regarding the behavior of "large group" industries, it also serves to simplify the analysis of individual firms. Since each firm is negligible to the market as a whole, it responds only to aggregate residual demand and hence is indifferent between price and quantity strategies. This continuum assumption also allows one to maintain the indivisibility of each agent's location ("address"), while at the same time avoiding the strong non-convexities resulting from interactions among influential (positive mass) agents. As a final simplification, it allows descriptions of the regional shares of economic and demographic magnitudes in terms of continuous variables.

5. The beginnings of NEG

There seems to be a large consensus among economists and geographers that the space-economy can be viewed as the outcome of a process involving two types of opposing forces: *agglomeration* (or centripetal) *forces* and *dispersion* (or centrifugal) *forces*. The resulting spatial distribution of economic activities is thus a complex balance between these forces that act on both consumers and firms. By identifying the main forces at work in NEG, two groups of papers have played a major role in the late 1970s and early 1980s.[5] However, the approaches followed in these contributions deal only with the location of firms, and assume that consumers' locations are fixed.

[3] This is achieved by assuming that the sub-utility involving the differentiated product is of the CES type. It is worth noting here that the CES is a discrete choice model which is a close relative of the logit model much use in regional science (Anderson, de Palma and Thisse, 1988).

[4] The first applications of modern monopolistic competition to urban economics are Abdel-Rahman (1988) and Fujita (1988).

[5] The corresponding two families of models have led to several extensions and applications. They are studied by Fujita and Thisse (2002) in the former case and by Baldwin et al. (2003) in the latter.

5.1 The Principle of (Minimum) Differentiation

As mentioned above, there is an old tradition in location theory, going back to Hotelling (1929), which suggests that spatial competition leads to the agglomeration of firms. In the typical example of two vendors selling a homogeneous product, each firm gains by locating close to its competitor on the more populated side of the market. Hence, two firms competing for clients choose to minimize their spatial differentiation, a result that became known as the Principle of Minimum Differentiation. The proponents of this approach, however, overlooked the fact that firms selling a homogeneous product always want to locate far apart in order to avoid the devastating effects of price war. Indeed, when two firms are located back to back, they get trapped into a Bertrand situation where they find themselves with zero operating profits.

When consumers visit firms to buy products, and when the marginal disutility generated by a trip rises with distance, the equilibrium prices decrease with the inter-firm distance, thus reflecting the fact that spatial proximity reduces firms' market power. Hence, as shown by d'Aspremont, Gabszewicz and Thisse (1979), firms always choose locations that are far apart. Such a spatial dispersion is the result of a trade-off in which price competition pushes firms away from each other, whereas competition for market area tends to pull them together. Thus, in contrast to Hotelling's claim, firms selling a homogeneous good tend to separate spatially at equilibrium in order to reduce price competition. This shows that *price competition is a strong dispersion force*, a result that has been called the Principle of Differentiation (Tirole, 1988).

However, it is fair to say that the Principle of Differentiation assumes extreme price-sensitivity of consumers: if two firms are located side by side with identical prices, a small price reduction of one firm will attract all consumers. Such a sharp behavior seems unwarranted. When products are differentiated and when consumers have different tastes or like product variety, the aggregate response to a price cut will not be so abrupt because prices are no longer the only relevant attribute of goods. Stated differently, product differentiation alleviates price competition. This modification of the spatial competition model has one major implication for economic geography: when consumers' preference for variety is sufficiently strong or, alternately, when transport costs are sufficiently low, de Palma et al. (1985) have shown that the Nash equilibrium is characterized by firms agglomerating at that location providing the best access to the entire market. Consequently, product differentiation appears to be an agglomeration force, whereas transport costs can act either as an agglomeration force (when they are low) or as a dispersion force (when they are high). Interestingly enough, transport costs play a similar role in NEG.

When transport costs drastically fall, the benefits of geographical separation are reduced and prices are lower. Firms might then choose to reconstruct their profit margins by differentiating products in terms of non-geographical characteristics,

either tangible or intangible. Stated differently, *product differentiation is substituted for geographical dispersion* (Irmen and Thisse, 1998). In this case, firms are less concerned with the effects of price competition (weakened by product differentiation) and more concerned with locating close to those consumers whose preferences best match their products. Thus when consumers are spatially dispersed, the best strategy for firms is to agglomerate at the market center, thereby minimizing their geographical differentiation. This result also agrees with market potential theory, as developed by Harris (1954) in classical economic geography, according to which firms tend to locate where they have the best access to those markets in which they can sell their product.

5.2 The Home Market Effect

In one of the first contributions belonging to the new strand of literature, called *New Trade Theories* [initiated by Krugman (1979) and developed further by Helpman and Krugman (1985) and others], Krugman (1980) investigates the impact of market size on the location of firms in a two-region, two-sector, two-factor economy. The industrial sector produces a differentiated good under increasing returns, whereas the agricultural sector produces a homogeneous good under constant returns. What makes this problem new and interesting is that it combines the mobility of both commodities and capital. In contrast to classical trade theory, firms are now free to choose their locations. But consumers/workers continue to be immobile. Furthermore, the mobility of industrial goods is imperfect because their shipments incur positive transport costs. It is therefore tempting to conclude that the region with the larger market will always attract firms, because this location minimizes total transport costs to both markets. However, this argument ignores the fact that when more firms locate within the same region, local competition is intensified and profits are lower. For even though competition is reduced by product differentiation, it does not disappear altogether.

In order to study how firms distribute themselves between the two regions, Krugman models the industrial sector by combining the Dixit-Stiglitz model of monopolistic competition with Samuelson's (1954) iceberg model of transport costs in which only a fraction of the goods shipped reach their destination. In contrast, the agricultural sector is treated as purely Walrasian. The resulting equilibrium distribution of industrial firms across regions again arises from a balancing of agglomeration and dispersion forces. Here the agglomeration force is generated by each firm's desire for *market access* (corresponding to the market-area effect in spatial competition). The dispersion force is in turn generated by each firm's desire to avoid *market crowding* (corresponding to the price effect in spatial competition). This modeling strategy is ideally suited for the spatial scale at which NEG is designed to operate.

When one region is larger in terms of population and purchasing power, Krugman shows that this region attracts a more than proportional share of firms – a result known as the *Home Market Effect*. Because of its comparative advantage in terms of size, it is of

course natural that this region should attract more firms. What is less expected is the fact that the share of firms exceeds the relative size of this region. Moreover, this effect is amplified by decreases in transport costs, which serve only to improve access to the smaller region. This suggests that deeper economic integration within an industry might actually lead to more regional imbalance in the spatial distribution of that industry.

The home market effect implies some degree of regional specialization, since the larger region will be a net exporter of the manufactured good and the smaller region a net exporter of the agricultural good. Another important consequence of this result is that size difference gives rise to spatial inequalities among consumers. For even though all consumers will benefit from deeper industrial integration, those in the larger region will achieve the greatest benefits from their direct access to a wider array of products. Hence *economic integration among regions of different sizes is likely to foster spatial inequalities between them*.

To sum-up, we know that competition is a strong dispersion force. However, competition is relaxed by product differentiation, thus allowing firms *to seek the location with the highest market potential where demand is high and transport costs low*. We also know that *large markets are relatively more attractive to firms*. More precisely, any initial advantage in market size is further amplified by the home market effect. But while this may help to explain why many spatial inequalities emerge, it implicitly assumes the existence of an exogenous comparative advantage, namely market size. These results also assume that consumers are immobile. So what happens if consumers are also free to choose their location? Krugman has addressed this question in his core-periphery model.

6. The core-periphery structure with labor mobility

While the movement of capital to a region brings with it the benefits of added production capability, it is not true that the returns from this capital need be spent in the same region. In contrast, when workers move to a new region, they bring with them *both their production and consumption capabilities*. As a result, their movements affect the size of labor and product markets in both the origin and destination regions. This is one of the main differences between capital and labor mobility. Another is that while the mobility of capital is driven by differences in *nominal* returns, workers move when there is a positive difference in *real* wages. In other words, differences in living costs matter to workers but not to owners of capital.

This is the starting point of Krugman's celebrated 1991 paper, in which he observes that the spatial economy is replete with *pecuniary externalities*. In particular, when some workers choose to migrate, their move affects the welfare of those who do not move.

Indeed, their migration may actually change the relative attractiveness of both origin and destination regions. These effects have the nature of pecuniary externalities because migrating workers do not take them into account in their decisions. Moreover, such externalities are of particular importance in imperfectly competitive markets, where prices fail to reflect the true social value of individual decisions. Hence the full effects of migration are best studied within a general equilibrium framework, where one can capture not only the interactions between product and labor markets, but also the dual role of individuals as both workers and consumers.

At first sight, the task seems to be out of reach. Yet, as shown by Krugman, most of the effects above can be combined and studied within a simple and elegant general equilibrium model that has come to be known as the *core-periphery model*. This is essentially a monopolistic competition model of the Dixit-Stiglitz-iceberg genre that focuses on an economy with two regions, two sectors, and two factors of production. One factor (farmers) is immobile and is used as an input in the agricultural sector. The second factor (workers) is spatially mobile and is used as an input in the industrial sector. Whereas trade theory typically assumes that factors are spatially immobile but mobile between sectors, Krugman takes the opposite view by assuming that workers do not move between sectors but are able to move across regions.

We are now equipped to see how a "snowball effect" can emerge that leads to industrial agglomeration. Two interacting effects are at work: one involves firms and the other workers. Assume that one region becomes slightly bigger than the other. Then on the one hand, this increase in market size leads to a higher demand for the industrial good, and thus generates the establishment of more firms. But by the home market effect, this increase in firms should be more than proportional, thus pushing nominal wages up. On the other hand, the presence of more firms means a greater variety of local products and, therefore, a lower local price index (since transport costs are lower). Accordingly, real wages should rise, and this region should attract a new flow of workers. The combination of these two effects should thus trigger a feedback process leading to the eventual agglomeration of all firms in a single region - the *core* of the economy. Stated differently [as noted by Krugman himself (1991a, p.486)], there is *circular causation* (à la Myrdal) in which these two effects reinforce each other: "manufactures production will tend to concentrate where there is a large market, but the market will be large where manufactures production is concentrated."

Even though this process seems to generate a cumulative dynamical process, it is not obvious that it will always develop according to that prediction. Indeed, the foregoing argument has ignored several key impacts of migration on the labor market. On the one hand, the increased supply of labor in the region of destination will tend to push wages down. On the other hand, since new workers are also consumers, there may be an increase local demand for the industrial good that leads to a higher demand for labor. But the story does not end here. As more firms enter the local market, there is

increased competition to attract workers. So the final impact on nominal wages is hard to predict. Likewise, there is increased competition in the product market, which makes the region less attractive to firms (the market crowding effect). The combination of all these effects may in fact lead to a "snowball meltdown", which results in a spatial dispersion of firms and workers.

The great accomplishment of Paul Krugman was to integrate all these effects within a single framework and to determine precisely the conditions under which the Myrdalian prediction holds or not. Stated differently, starting from an arbitrarily small difference between regions, perhaps due to a small accident, Krugman has been able to single out those cases under which there is agglomeration of the manufacturing sector. Furthermore, *his approach relies only on market interactions.* This is to be contrasted to most of the existing literature dealing with the causes of agglomeration, which typically appeals to externalities of various types (see section 4). Note finally that, as with the New Trade Theories, this modeling strategy is well suited to the spatial scale considered by NEG. By contrast, the impact of spatial externalities being often localized, the externality approach underscores the need to study agglomeration at the micro-spatial level.

Turning next to the specific conditions for agglomeration identified by Krugman (1991a), he has shown that the value of transport costs is the key-determining factor. On the one hand, if transport costs are sufficiently low, then all manufacturers will concentrate in a single *core* region, while the other *peripheral* region supplies only the agricultural good. In this way, firms are able to exploit increasing returns by selling more goods in the larger market without losing much business in the smaller market. Thus, *the home market effect is magnified through the mobility of workers/consumers*. It is worth stressing here that this core-periphery structure emerges as the equilibrium balance of a system of opposing forces. Hence any spatial inequalities that arise are the involuntary consequence of decisions made by a large number of economic agents pursuing their own interests.

On the other hand, if transport costs are sufficiently high, then interregional shipments of goods are discouraged. Hence the economy displays a symmetric regional pattern of production that focuses on local markets. This core-periphery model thus allows for *the possibility of convergence or divergence between regions*, whereas the neoclassical model, based on constant returns and perfect competition in the two sectors, would predict convergence only. Consequently, Krugman presents a synthesis of the polarization and neo-classical theories. In addition, he shows that multiple equilibria may arise, thus implying a need for further refinement. As usual, unstable equilibria are dismissed. But for intermediate values of transport costs, several stable equilibria can still exist, so the market outcome is likely to depend on initial conditions. Less formally, this means that history matters – even a small accident may suffice to generate a lock-

in effect in the process of spatial development. Hence *the space-economy seems to be of a putty-clay nature*.

The process of circular causation above can lead industrial firms to be tied to the same region for long time periods. It also provides a concise formalization of an idea put forward by Myrdal (1957, pp. 26-27) long ago: "Within broad limits, the power of attraction of a centre has its origin mainly in the historical accident that something was once started there, and not in a number of other places where it could equally well have been started, and that start met success."

By the same token, Krugman has provided a unifying framework that allows one to study "the location of all economic activities, with attention paid to the geographical distribution of inputs and outputs and the geographical variations in prices and costs" (Isard, 1949, p.505), an objective that Ohlin (1933, 1968, p.97) had stressed long before: "international trade theory cannot be understood except in relation to and as part of the general location theory, to which the lack of mobility of goods and factors has equal relevance." Yet, Krugman was not aware of Isard's ideas and read Ohlin much later on (Krugman, 1999). It is, therefore, tempting to adapt a remark of Picasso: "it is so hard when you are all on your own. What a terrible idea it was to bring another economist into the studio."

Krugman's core-periphery model also provides a powerful analytical framework for studying the historical development of many spatial economies. For example, it helps to illuminate the spatial consequences of the Industrial Revolution as described, for example, by Pollard (1981, p.11): "the industrial regions took from them [their agricultural neighbors] some of their most active and adaptable labour and they encouraged them to specialize in the supply of agricultural produce, sometimes at the expense of some pre-existing industry running the risk thereby that this specialization would permanently divert the colonized areas from becoming industrial themselves."

7. Input-output linkages and the bell-shaped curve of spatial development

So far, agglomeration has been considered as the outcome of a circular causation process fed by the mobility of workers. However, agglomeration of economic activities also arises in contexts in which labor mobility is very low, as in most European countries. This underscores the need for alternative explanations of industrial agglomeration.

One strong contender is *the presence of input-output linkages between firms*: an output of one firm can be an input for another, and vice versa. In such a case, the entry of a new firm in a region not only increases the intensity of competition between similar

firms; it also increases the market of upstream firm-suppliers and decreases the costs of downstream firm-customers. This is the starting point of the well-known paper by Krugman and Venables (1995).

Their idea is beautifully simple and suggestive: *the agglomeration of the final sector in a particular region occurs because of the concentration of the intermediate industry in the same region, and conversely*. Indeed, when firms belonging to the final sector are concentrated within a single region, the local demand for intermediate inputs is very high, thus making this region very attractive to firms producing intermediate goods. Conversely, because intermediate goods are made available at lower prices in the core region, firms producing final goods find that region very attractive. Thus, a cumulative Myrdalian process may still develop that leads to industrial agglomeration within the core region.[6]

In this alternative setting, new forces arise. Indeed, if firms agglomerate in a region where the supply of labor is inelastic, then wages must surely rise. This in turn has two opposite effects. On the one hand, *consumers' demand for the final product increases because they have a higher income*. This is again a market expansion force, triggered now by higher incomes rather than larger populations. On the other hand, such wage increases also generate a dispersion force. When the wage gap between the core and the periphery becomes sufficiently large, some firms will find it profitable to relocate in the periphery, even though the local demand for their output is lower than in the core. This is especially true when transport costs are low, because asymmetries in demand then have a weaker impact on profits.

The set of equilibrium patterns obtained in the present setting is much richer than in the initial core-periphery model. In particular, if a deepening of economic integration triggers the concentration of industrial activities in one region, then beyond a certain threshold, an even deeper integration may lead to a reversal of this tendency. Some firms now relocate from the core to the periphery. In other words, the periphery experiences a process of *reindustrialization*. Simultaneously, the core might start losing firms, thus becoming *deindustrialized*. As Fujita, Krugman and Venables (1999, p. 260) put it, "declining trade costs first produce, then dissolve, the global inequality of nations". Therefore, economic integration would yield a bell-shaped curve of spatial development à la Kuznets. And indeed, such a curve may be obtained in several extensions of the core-periphery model (Combes, Mayer and Thisse, 2008). In particular, the agglomeration of activities often materializes under the form of a city

[6] Toulemonde (2006) has identified another relevant mechanism of agglomeration, which bears some strong resemblance with the vertical linkage setting. When workers are a priori unskilled immobile, some of them choose to become skilled in order to be able to work in the manufacturing sector. As a result, they earn a higher income and, therefore, have a higher demand for manufacturing goods, making their region a larger and more attractive market to firms. At the same time, the installation of new firms within this region gives a stronger incentive to workers to improve their skill. As above, we obtain a mechanism of cumulative causation in which spatial mobility is replaced by sector-based mobility.

where competition for land acts as a strong dispersion force. By providing some form of reconciliation between the market outcome and the concern for more spatial equity, such a bell-shaped curve of spatial development lends strong support to a deeper integration of modern economies.

8. City formation and urban system

It should also be emphasized that much interest generated by Krugman's landmark paper in 1991 is due to the fact that it offers an analytical approach to a wide range of related issues, including city formation, urban system and regional growth.

Cities, and more generally metropolitan areas, constitute the most extreme form of economic agglomerations within nations -- blending a very high share of their economic activities within a very small share of their land. Both the metropolitan areas of Tokyo and of New York have a GDP that exceeds the GDP of India, while the metropolitan area of Paris accounts for 29% of the French GDP but only 2.2% of its territory. An added dimension here is the wide variety of agglomerations that exist within urban hierarchies. On the one hand, some cities may be specialized in a very small number of industries, such as in many medium-size American cities. On the other hand, large metropolises like New York or Tokyo are often highly diversified, in that they contain many industries not related through direct linkages. How and why does this arise?

Fujita and Krugman (1995) appear to have been the first to offer a model in which both urban and agricultural land-use patterns are determined endogenously. In many ways, this model completes the classical work of von Thünen (Samuelson, 1983) who postulated the existence of a predetermined monocentric city in his setting. This work not only fills in a major gap in the literature but, more importantly, it suggests a microeconomic approach that allows one to study the emergence of an urban system in the spatial economy. Indeed, the identification of conditions under which a monocentric economy is sustainable as an equilibrium outcome is a formal way of addressing the fundamental question: *where and when do new cities emerge?*

The framework used by Fujita and Krugman is closely related to that of the core-periphery model described above. However, here all workers are assumed to be identical and perfectly mobile between locations and sectors (typically, agriculture and industry). The location space is now continuous and one-dimensional. In this setting, the centrifugal force acting on firms now arises from the existence of a land market in the agricultural sector. Since production of the agricultural good requires both land and labor, and since farmers are also consumers of manufactured goods, this leads to a spatial dispersion of demand for manufactured goods. In the presence of no other forces, this would lead firms to disperse throughout the agricultural landscape. However there are countervailing centripetal forces driven by the presence of scale

economies of industrial production. This together with the demand for product variety by workers/consumers creates forward and backward linkages that lead to urban concentrations of both firms and workers. Such a setting allows for a synthesis of the classic von Thünen model and of the Dixit-Stiglitz-iceberg model used in NEG. Fujita and Krugman show that *a monocentric economy is a spatial equilibrium provided that the population size does not exceed some threshold value* depending on the structural parameters of the economy. In particular, as the population continues to grow, more cities will emerge to form an urban system.

Despite its importance for economic geography, this first contribution fails to recognize the fact that cities typically have different sizes. What is probably the most striking feature of the space-economy is that cities form a hierarchical system exhibiting some regularity in terms of both their size and the array of goods they supply. According to Christaller (1933), Lösch (1940) and their successors, different markets are arranged in a way such that a city in which a good is supplied also provides the goods made available in a larger number of urban centers. The same authors also argue that, in a featureless space, the urban system would be formed by a family of nested and regular lattices, with one lattice per good, each vertex of a lattice accommodating a city that supplies consumers situated in its vicinity. These considerations are only interesting if they are based on microeconomic foundations. If there are no economic forces that lead firms of different types to cluster, it is hard to see why such a nested system of cities would be more likely to emerge than any other configuration.

Henderson (1974, 1988) has developed a compelling and original approach that allows him to describe how a hierarchy of cities emerges. In each city, there is again a tension between two forces. On the one hand, there are external economies associated with the agglomeration of firms within a city. On the other hand, there are diseconomies generated by the need to commute in a more or less large city. Hence, in equilibrium, each city has a well-defined size that depends on the type of firms it accommodates. As cities vary in their industrial mix, they have different sizes because industries differ in the external economies they are able to create. However, in this approach, cities are like floating islands because nothing is said about city locations.

Fujita, Krugman and Mori (1999) take a first step in this direction by introducing several industrial and differentiated goods in a NEG-like model. As the population increases, they show that a more or less regular hierarchical central place system emerges within the economy. In this urban system, higher-order cities provide a larger number of groups of goods. In addition, there is two-way trade between cities because these cities supply differentiated goods. This leads to a more intricate pattern of trade in which horizontal relations are superimposed onto the pyramidal structure of central place theory. As expected, higher-order cities export a larger variety of goods than lower-order cities. However, horizontal relations between cities of the same order

may be more important than trade with lower-order cities. Put simply, the equilibrium pattern is consistent with an urban system involving large and diversified cities along with small and specialized cities, a well-documented empirical fact (Duranton and Puga, 2000).

Those urban hierarchies are more complex than in the simple Christaller model of central places in that they combine both the hierarchy of various centers with the existence of networks of cities exchanging specialized goods and services. They seem to fit well the description of modern space-economies such as those provided by geographers (Pred, 1966) and historians (Hohenberg and Lees, 1985).

9. Regional growth

As mentioned above, the core-periphery model of Krugman (1991a) can also be used to address questions of regional growth. Of particular interest is the question of whether regional discrepancies increase or decrease over time. Moreover, if these discrepancies do change, then what are the main reasons for such divergence or convergence? To answer this question, we must consider the complex issue of possible interactions between spatial and temporal decisions. Since both agglomeration and growth are complex phenomena by themselves, one must overcome many conceptual and analytical hurdles to relate the two. Not surprisingly, the field is thus still in its infancy, and relevant contributions are few. Yet the task is not out of reach. Because both the "new" theories of growth and NEG share the same basic framework of monopolistic competition, there exists a solid foundation for cross-fertilization between the two fields. Exploiting this formal analogy, R. Baldwin, Ph. Martin, G. Ottaviano and others have explored several interactions between growth and agglomeration (Baldwin and Martin, 2004). However, this work has focused mainly on the mobility of physical capital, and ignores the fact that human capital has become increasingly more mobile that leads to a growing concentration of skilled workers within a few areas.

Fujita and Thisse (2002) have attempted to fill this gap by developing a dynamic generalization of Krugman's core-periphery model. They propose a new and simple model of endogenous growth for a two-region economy, which represents a natural combination of a Krugman-type core-periphery model and a Grossman-Helpman-Romer-type model of endogenous growth with horizontally differentiated products. To do so, they add to the core-periphery model a research-and-development sector that utilizes skilled labor to create new varieties of goods for the industrial sector, so that the number of firms producing in this sector becomes variable. More precisely, they combine (i) the *demand effect* generated by the migration of skilled workers [as in core-periphery models] and (ii) the *productivity effect* generated by the existence of spillovers [such as those studied in endogenous growth theories]. The interaction of these two effects leads to the creation of new varieties of goods, which (together with consumer preferences for

variety) gives rise to a secondary demand effect. One of the most interesting results obtained by Fujita and Thisse is that the growth rate, measured by the variation in the number of varieties, changes with the spatial distribution of skilled workers. In other words, they show that *the growth of the global economy depends on the spatial distribution of the innovation sector across regions*. This analysis thus lends support to the existence of a trade-off between growth and spatial equity, as suggested by recent economic policy debates in a number of industrialized countries.

But even when growth leads to increased spatial inequalities, the above analysis suggests that the overall outcome may often be *Pareto superior*. Specifically, when the economy moves from dispersion to agglomeration, the rate of innovation tends to increase. Consequently, *if the growth effect triggered by the agglomeration is strong enough, then even those who remain in the periphery will be better off.* Hence it can be argued by Rawls' principle that there is no real conflict between growth and equity here, in the sense that *all* workers are made better off. It should also be stressed that this Pareto-optimality property does not require any transfer whatsoever: it is a pure outcome of market interaction.

But to be sure, unskilled workers who live in the core of the economy will enjoy a higher level of welfare than those in the periphery, and indeed, the gap between those who live in the core and the periphery will increase. In other words, *agglomeration gives rise to regressive effects in terms of income distribution*. Such increasing welfare gaps may call for corrective policies, even though such policies might actually impede growth, and thus reduce individual welfare. For example, in the case of the European Union, Article 130a of the Amsterdam Treaty states that the "Community shall aim at reducing disparities between the levels of development of the various regions and the backwardness of the least favoured regions or islands, including rural areas". This suggests that social cohesion and economic growth might well be conflicting objectives.

Note finally that these regional income discrepancies reflect, at least to a large extent, *the spatial distribution of skills*. The welfare gap between the core and periphery arises because of the additional gains from growth generated by the agglomeration of skilled workers in the core region. This in turn increases the welfare of unskilled workers residing in the core region, even though their productivity is the same as those unskilled workers in the periphery. Thus, these redistribution effects are significantly more complex than they might first appear. Such findings help to illustrate how the core-periphery model can help shed light on issues that are fundamental for society.

10. Where do we stand?

The foregoing results are very important for the way the space-economy is organized. This is why it is critical to know how they depend on the specificities of the framework

used. In his contribution to the Symposium "Urban Agglomeration" published in the Journal of Economic Perspectives, Krugman (1998) set up indirectly the agenda for future research by summarizing NEG with the slogan "Dixit-Stiglitz, iceberg, evolution, and the computer." Let us discuss each part of the slogan.[7]

(i) **The computer**. Introducing transport costs to trade models seems both natural and simple. Natural because shipping commodities across space requires resources; simple because transport costs could just be one more type of cost to take into account. As experienced by Samuelson (1954) himself, this appears to be a very hard task, the reason being that transport costs are associated with general equilibrium effects across spatially separated markets. To be precise, the price of a good that prevails at one place depends on the price of the same good in another location through the cost of shipping the good between these two places. As shown by Samuelson (1952), this is something that can be done when firms and consumers are immobile and when transport services are provided by a perfectly competitive sector.

Things become much more complex once agents are mobile. This is because supply and demand schedules are shifted downward and upward by their relocation across places. It is no surprise, therefore, that it is not possible to come up with the analytical solution of the core-periphery model. This is what led Krugman to resort to numerical analysis to uncover the impact of decreasing transport costs on the location of economic activity. Yet, we need not the explicit solution of a model to determine its main features. Different mathematical tools may be used to study how the equilibrium outcome varies with the structural parameters of a model. And indeed, subsequent developments confirm Krugman's results but it has taken quite a while to prove them all. The formal stability analysis was developed in Fujita, Krugman and Venables (1999) while it was not until Robert-Nicoud (2005) that a detailed study of the correspondence of spatial equilibria was provided.[8] Put differently, we don't need the computer anymore to study the core-periphery model.

(ii) **Evolution**. The core-periphery model and its many relatives display several spatial equilibria. Stability is then used as a refinement to dismiss unstable equilibria as these ones are very unlikely to emerge. Studying stability requires specifying a dynamic adjustment process. Krugman and others use myopic evolutionary dynamics, meaning that workers care only about current utility levels. This is fairly restrictive to the extent that migration decisions are typically made on the grounds of discounted future utility flows and costs (e.g. search, mismatch and homesickness). In addition, this approach is not consistent with fully rational forward-looking behavior. It is, therefore, important to determine if and how workers' expectations about the evolution of the economy may influence the process of agglomeration. The key-issue here is the

[7] There are countless extensions of the core-periphery model. To the best of our knowledge, the most general and rich framework remains the work of Puga (1999).

[8] Note also that the existence and uniqueness of a short-run equilibrium (i.e. all markets clear when workers' and firms' locations are fixed) were shown by Mossay (2006).

specification of a particular dynamic process that would fit the observed migration behavior of workers.

Among the few existing results, one is worth mentioning. When the two regions host different numbers of workers, one may wonder whether or not their common belief about the future of the space-economy leads them to agglomerate in the currently smaller region, thus reversing the historically inherited advantage of the larger region (Ottaviano, 2001)? Interestingly enough, the answer is positive provided that future matters to workers, i.e. they have a low discount rate. Two cases are then worth mentioning. In the former, the population of farmers is larger in one region than in the other, thus endowing one region with a given size advantage. Even though, for whatever reason, all workers would be located in the initially disadvantaged region, there exist self-fulfilling expectations that lead the economy toward full agglomeration in the other region, but the reverse does not hold. Likewise, when one region benefits from higher trade barriers, there are self-fulfilling expectations that lead workers to agglomerate in this region, but again not vice versa (Oyama, 2009).

(iii) **The iceberg transport cost**. Samuelson's trick is extremely ingenious but it treats freight rates as exogenously given parameters. This is more of an issue in NEG than in trade as the main message delivered by NEG is that changes in the costs of shipping goods have a major impact on industry location. So one would like to know how these costs are formed. Ideally, the two-way interactions between transport costs and the spatial distribution of economic activity must be analyzed within a general equilibrium framework where locations, prices and freight rates are endogenously determined by market mechanisms. The iceberg transport cost also implies that any variation in mill prices is accompanied with a proportional variation in transport cost, an assumption that seems hard to justify.

Using a different framework to model monopolistically competitive markets, Ottaviano, Tabuchi and Thisse (2002) assumes that trading the manufactured good requires a given amount of the numéraire per unit shipped between regions. This agrees with common firms' practices, which absorb a fraction of their freight costs in order to supply distant markets (Phlips, 1983). We will return to this setting below but it is worth stressing here that this alternative setting leads to conclusions that are similar to Krugman's core-periphery model. Hence, the main properties of NEG do not seem to depend on the iceberg specification used to model transport costs. This is reassuring.

(iv) **Dixit-Stiglitz**. As discussed in section 4, working with imperfectly competitive markets appears to be a very sensible approach to tackle the issue of spatial inequalities at the interregional level as we expect the working of product and labor markets to be critical in the emergence of regional disparities. Using the Dixit-Stiglitz model of monopolistic competition is the natural research strategy since this model integrates both imperfect competition and increasing returns at the firm level.

However, by combining the CES and the iceberg transport cost, the elasticity of a firm's aggregated demand is constant and equal to the elasticity of substitution across varieties. This elasticity is, therefore, independent of the spatial distribution workers and firms between regions. Although analytically convenient, this fails to capture the idea that lower transport costs unleashes competition between spatially separated markets, while large markets are likely to be more competitive than small markets, thus squeezing mark-ups. This modeling strategy reduces the intensity of both agglomeration and dispersion forces. Indeed, the dispersion force is weaker because firms' markups are the same on the periphery and in the core. Likewise, the agglomeration force is also weakened, as workers' real income is not positively affected by the price drop that should occur in the core. It is, therefore, quite possible that the predictions of this model depend on this absence of pro-competitive effects. This is why it is worth checking whether the main conclusions obtained so far hold.

This is what Ottaviano, Tabuchi and Thisse (2002) have undertaken by using an alternative model of monopolistic competition in which firms' demand are linear in their own price as well as in the price index. Such a framework allows one to capture the two foregoing effects. It also permits to work with the standard specification of transport costs used in location theory and yields analytical solutions. Its main shortcoming lies in the absence of income effect in goods' demands due to the quasi-linear upper-tie utility used.[9] The good news is that, within this fairly different framework, all the qualitative results obtained by Krugman (1991a) hold true. In other words, we may conjecture that most of the main results obtained so far in NEG remain valid for a whole class of models.[10]

That said, given the success of the Dixit-Stiglitz model of monopolistic competition in various economic fields (Matsuyama, 1995; Brakman and Heijdra, 2004), the economics profession should strive to develop more general models of monopolistic competition. The idea is that a firm has no impact on the market outcome – the continuum approach - but chooses its strategy on the basis of a few statistics – such as the price index - obtained from the distribution of prices set by all firms. Ultimately, this should lead to a price theory that would be general enough to compete on equal footing with the general competitive setting (Arrow and Hahn, 1972).

This is not the end of the story, however: there remain other unsatisfactory features to be discussed. The following list is not exhaustive but it covers a fair deal of issues. (i) NEG models overlook many costs whose origin lies in the space-economy (for example, the congestion costs generated by the emergence of an agglomeration) and, conversely, overlooks other agglomeration benefits, such as a better matching on labor markets, the proximity of intermediate inputs and knowledge spillovers. (ii) These

[9] Note, however, that wages enter the utility differential that drives the mobility of workers.

[10] Behrens and Murata (2007) propose a new framework of monopolistic competition with both price competition and income effects. Its future application to NEG may constitute another step in the direction of exploring the robustness of this theory.

models only accounts for two sectors and two regions. (iii) The agricultural sector is given a very restricted role, its job being to guarantee the equilibrium of the trade balance. Along the same line, it is hard to see why trading the agricultural good is costless in a model seeking to ascertain the overall impact of trade costs.[11]

All these features have already attracted a lot of attention. There is no doubt, however, that the "dimensionality problem" is likely to be the most challenging one. This is not a brand new issue for trade people. For example, Deardorff (1984, p. 468) wrote accurately that "the Heckscher-Ohlin theorem is derived from a model of only two of each of goods, countries, and factors of production. It is unclear what the theorem says should be true in the real world where there are many of all three."

The same holds true for NEG (Behrens and Thisse, 2007). But why should one bother about the existence of many regions instead of two? In addition to the need for a better theoretical understanding of spatial interdependencies to guide the empirical analysis, as emphasized in the quotation above, the new fundamental ingredient that a multi-regional setting brings about is that the accessibility to markets varies across regions – the "three-ness effect" of Krugman (1993). In other words, spatial frictions between any two regions are likely to be different, which means that the relative position of the region within the whole network of interactions matters. In this respect, it is worth recalling that even the simplest firm location model accounts for the fact that the access to several markets is the key-issue faced by a firm making its location choice (Thomas, 2002). Although location theory accurately stresses this fact, most trade theorists are still reluctant to the idea of assuming that different countries have a different access to each other. Yet, armchair empirical evidence shows that a good access to markets is a major determinant for the location of economic activity.

Another key insight one can derive only in a multi-region economy is that any change in the underlying parameters has in general complex impacts which vary in non-trivial ways with the properties of the graph representing the spatial economy. When there are only two regions, any change in structural parameters necessarily affects directly either one of the two regions, or both. On the contrary, when there are more than two regions, any change in parameters that directly involves only two regions now generates spatial spillover effects that are unlikely to leave the remaining regions unaffected. This in turn further affects the other regions and so on.

By way of conclusion, we would like to add a few remarks. The first one concerns the deep impact of NEG on the economics profession, as reflected by two small books

[11] For Neary (2001), another shortcoming of NEG lies in the absence of strategic considerations among firms. This should not jeopardize the main conclusions, however. Indeed, the linear model allows for pro-competitive effects that are absent from the standard setting. Furthermore, comparing spatial competition and economic geography models, Combes, Mayer and Thisse (2008) observe that strategic behaviours do not seem to vastly affect the conclusions obtained in settings based on monopolistic competition, even though the underlying market mechanisms differ.

published in the 1990s. First on the list is Krugman's (1991b) *Geography and Trade*. What makes this small book so appealing is its informal and intuitive presentation of most of the main issues in economic geography and related fields. Though it was not intended to present original research, it has nonetheless contributed enormously to the visibility of field, both in economics and elsewhere. In *Development, Geography, and Economic Theory*, Krugman (1995) goes one step further and explores the (common) methodological challenges faced by geography and development. He argues in a very lucid and clear way why the market structure problem discussed in section 4 has confined for too long these two fields at the periphery of mainstream economics.

Second, one feature of NEG that makes it so exciting is its strong connection to other recent developments in modern economics. First of all, while NEG is closely related to trade theory, it is also very much connected to modern industrial organization. In particular, the economic geography of a territory is recognized now to be related to the way firms organize their production activities, management, and outsourcing, within and across firms. There are also strong links to recent intellectual developments such as the new growth theories. Another distinctive feature of NEG is that it has generated an impressive flow of empirical studies of high quality, which are covered in various chapters of the fourth volume of the *Handbook of Regional and Urban Economics*.

Last, spatial inequalities are a fundamental issue that has been too long neglected by economists. Paul Krugman has succeeded in clarifying the nature of these inequalities in simple and yet powerful ways. His work has truly inspired all of us working in this field. Yet, it is not totally surprising that the surge of NEG took place a few years after the revival of monopolistic competition and industrial organization, from which it borrows many ideas and concepts. Ever since Alfred Marshall, everyone knows that "The mysteries of the trade become no mysteries; but are as it were in the air." While many researchers have contributed to economic geography, there is no question that the most prominent figure in its development has been Paul Krugman.

References

Alonso, W. 1964. Location and Land Use. Cambridge (Mass.), Harvard University Press.

Abdel-Rahman, H.M. 1988. Product differentiation, monopolistic competition, and city size. Regional Science and Urban Economics 18, 69-86.

Anderson, S.P., A. de Palma and J.-F. Thisse. 1988. The CES and the Logit: two related models of heterogeneity. Regional Science and Urban Economics 18, 155-164.

Arrow, K.J. and F. Hahn. 1972. General Competitive Analysis.

Audretsch, D. and M. Feldman. 2004. Knowledge spillovers and the geography of innovation. In: J.V. Henderson and J.F. Thisse, editors, Handbook of Regional ad Urban Economics. Volume IV. Amsterdam: North-Holland, 2714-2739.

Baldwin, R.E., R. Forslid, Ph. Martin, G.I.P. Ottaviano and F. Robert-Nicoud. 2003. Economic geography and public policy. Princeton, Princeton University Press.

Baldwin, R.E. and Ph. Martin. 2004. Agglomeration and regional growth. In: J.V. Henderson and J.F. Thisse, editors, Handbook of Regional ad Urban Economics. Volume IV. Amsterdam: North-Holland, 2671-2711.

Behrens, K. and Y. Murata. 2007. General equilibrium models of monopolistic competition: a new approach. Journal of Economic Theory 136, 776–787.

Behrens, K. and J.-F. Thisse. 2007. Regional economics: a new economic geography perspective. Regional Science and Urban Economics 37, 457-465.

Brakman, S. and B.J. Heijdra. 2004. The Monopolistic Competition Revolution in Retrospect. Cambridge, Cambridge University Press.

Christaller, W. 1933. Die Zentralen Orte in Süddeutschland. Jena, Gustav Fischer Verlag.

Combes, P.-P., T. Mayer and J.-F. Thisse. 2008. Economic Geography. The Integration of Regions and Nations. Princeton, Princeton University Press.

d'Aspremont, C., J.J. Gabszewicz and J.-F. Thisse. 1979. On Hotelling's "Stability in Competition". Econometrica 47, 1045-1050.

Deardorff, A.V. 1984. Testing trade theories and predicting trade flows. In: R.W. Jones and P.B. Kenen (eds.) Handbook of International Economics, volume 1. Amsterdam, North Holland, 467-517.

de Palma, A., V. Ginsburgh, Y.Y. Papageorgiou and J.-F. Thisse. 1985. The principle of minimum differentiation holds under sufficient heterogeneity. Econometrica 53, 767-781.

Dixit, A.K. and J.E. Stiglitz. 1977. Monopolistic competition and optimum product diversity. American Economic Review 67, 297-308.

Duranton, G. and D. Puga. 2000. Diversity and specialization in cities: why, where and when does it matter? Urban Studies 37, 533-555.

Duranton, G. and D. Puga. 2004. Micro-foundations of urban increasing returns: theory. In J.V. Henderson and J.-F. Thisse (eds.) Handbook of Regional and Urban Economics, volume 4. Amsterdam, North Holland, 2063-2117.

Eaton, B.C. and R.G. Lipsey. 1977. The introduction of space into the neoclassical model of value theory. In M. Artis and A. Nobay (eds.) Studies in Modern Economics. Oxford, Basil Blackwell, 59-96.

Fujita, M. 1988. A monopolistic competition model of spatial agglomeration: a differentiated product approach. Regional Science and Urban Economics 18, 87-124.

Fujita, M. 1989. Urban Economic Theory. Land Use and City Size. Cambridge, Cambridge University Press.

Fujita, M. and P. Krugman. 1995. When is the economy monocentric? von Thünen and Chamberlin unified. Regional Science and Urban Economics 25, 505-528.

Fujita, M., P. Krugman and T. Mori. 1999. On the evolution of hierarchical urban systems. European Economic Review 43, 209-251.

Fujita, M., P. Krugman and A.J. Venables. 1999. The Spatial Economy. Cities, Regions and International Trade. Cambridge (Mass.), The MIT Press.

Fujita, M. and H. Ogawa. 1982. Multiple equilibria and structural transition of non-monocentric urban configurations. Regional Science and Urban Economics 12, 161-196.

Fujita, M. and J.-F. Thisse. 2002. Economics of Agglomeration. Cities, Industrial Location and Regional Growth. Cambridge, Cambridge University Press

Gabszewicz, J. J. and J.-F. Thisse. 1986. Spatial competition and the location of firms. In: J.J. Gabszewicz, J.-F. Thisse, M. Fujita and U. Schweizer. Location Theory. Chur, Harwood Academic Publishers, 1-71.

Harris, C. 1954. The market as a factor on the localization of industry in the United States. Annals of the Association of American Geographers 64, 315-348.

Helpman, E. and P.R. Krugman. 1985. Market Structure and Foreign Trade. Cambridge (Mass.), The MIT Press.

Henderson, J.V. 1974. The sizes and types of cities. American Economic Review 64, 640-656.

Henderson, J.V. 1988. Urban Development. Theory, Fact and Illusion. Oxford, Oxford University Press.

Hohenberg, P. and L.H. Lees. 1985. The Making of Urban Europe (1000-1950). Cambridge (Mass), Harvard University Press.

Hotelling, H. (1929) Stability in competition. Economic Journal 39, 41-57.

Irmen, A. and J.-F. Thisse. 1998. Competition in multi-characteristics spaces: Hotelling was almost right. Journal of Economic Theory 78, 76-102.

Isard, W. 1949. The general theory of location and space-economy. Quarterly Journal of Economics 63, 476-506.

Kaldor, N. 1935. Market imperfection and excess capacity. Economica 2, 35-50.

Krugman, P. 1979. Increasing returns, monopolistic competition, and international trade. Journal of International Economics 9, 469-479.

Krugman, P.R. 1980. Scale economies, product differentiation, and the pattern of trade. American Economic Review 70, 950-959.

Krugman, P.R. 1991a. Increasing returns and economic geography. Journal of Political Economy 99, 483-499.

Krugman, P.R. 1991b. Geography and Trade. Cambridge (Mass.), The MIT Press.

Krugman, P. 1993. The hub effect: or, threeness in international trade. In: W.J. Ethier, E. Helpman and J.P. Neary (eds.) Theory, Policy and Dynamics in International Trade: Essays in Honor of Ronald Jones. Cambridge (Mass.), Cambridge University Press, 29-37.

Krugman, P. 1995. Development, Geography, and Economic Theory. Cambridge (Mass.), The MIT Press.

Krugman, P. 1998. Space: the final frontier. Journal of Economic Perspectives 12(2), 161-174.

Krugman, P. 1999. Was it all in Ohlin? Unpublished notes.

Krugman, P. and A. Venables. 1995. Globalization and the inequality of nations. Quarterly Journal of Economics 110, 857-880.

Lösch, A. 1940. Die Räumliche Ordnung der Wirtschaft. Jena, Gustav Fischer.

Matsuyama, K. 1995. Complementarities and cumulative process in models of monopolistic competition. Journal of Economic Literature 33, 701-729.

Mills, E.S. 1967. An aggregative model of resource allocation in a metropolitan area. American Economic Review 57, 197-210.

Mossay, P. 2006. The core-periphery model: a note on the existence and uniqueness of short-run equilibrium. Journal of Urban Economics 59, 389-393.

Mueller, D. 1989. Public Choice II. A Revised Version of Public Choice. Cambridge, Cambridge University Press.

Mundell, R. 1957. International trade and factor mobility. American Economic Review 47, 321-335.

Myrdal, G. 1957. Economic Theory and Underdeveloped Regions. London, Duckworth.

Neary, J.P. 2001. Of hype and hyperbolas: introducing the new economic geography. Journal of Economic Literature 39, 536-561.

Ogawa, H. and M. Fujita. 1980. Equilibrium land use patterns in a non-monocentric city. Journal of Regional Science 20, 455-475.

Ohlin, B. 1933. Interregional and International Trade. Cambridge (Mass.), Harvard University Press. Revised version published in 1968.

Ottaviano, G.I.P. 2001. Monopolistic competition, trade, and endogenous spatial fluctuations. Regional Science and Urban Economics 31, 51-77.

Ottaviano, G.I.P., T. Tabuchi and J.-F. Thisse. 2002. Agglomeration and trade revisited. International Economic Review 43, 409-436.

Oyama, D. 2009. History versus expectations in economic geography reconsidered. Journal of Economic Dynamics & Control, forthcoming.

Phlips, L. 1983. The Economics of Price Discrimination. Cambridge, Cambridge University Press.

Pollard, S. 1982. Peaceful Conquest. The Industrialization of Europe 1760-1970. Oxford, Oxford University Press.

Pred, A. 1966. The Spatial Dynamics of US Urban-industrial Growth. Cambridge (Mass.), The MIT Press.

Puga, D. 1999. The rise and fall of regional inequalities. European Economic Review 43, 303-334.

Robert-Nicoud, F. 2005. The structure of simple `New Economic Geography' models. Journal of Economic Geography 5, 201-234.

Rosen, S. 2002. Markets and diversity. American Economic Review 92, 1-15.

Rosenthal, S. and W. Strange. 2004. Evidence on the nature and sources of agglomeration economies. In: J.V. Henderson and J.F. Thisse, editors, Handbook of Regional ad Urban Economics. Volume IV. Amsterdam: North-Holland, 2120-2171.

Samuelson, P.A. 1952. Spatial price equilibrium and linear programming. American Economic Review 42, 283-303.

Samuelson, P.A. 1954. The transfer problem and transport cost, II: analysis of effects of trade impediments. Economic Journal 64, 264-289.

Samuelson, P.A. 1983. Thünen at two hundred. Journal of Economic Literature 21, 1468-1488.

Starrett, D. 1978. Market allocations of location choice in a model with free mobility. Journal of Economic Theory 17, 21-37.

Thomas, I. 2002. Transportation Networks and the Optimal Location of Human Activities: A Numerical Geography Approach. Cheltenham (U.K.), Edward Elgar.

Tirole, J. 1988. The Theory of Industrial Organization. Cambridge (Mass.), The MIT Press.

Toulemonde, E. 2006. Acquisition of skill, labor subsidies, and agglomeration of firms. Journal of Urban Economics 59, 420-439.

Zenou, Y. 2009. Urban Labor Economics. Cambridge, Cambridge University Press, forthcoming.