Chapter 2
The Spatial Perspective

Abstract In this chapter, we start showing that development processes are spatially uneven. In advanced regions there are factors that lead to increasing returns: economies of specialization, economies of scale, and external economies. Opening markets might thus bring about regional divergence as a result of a cumulative inflow of mobile resources to the more advanced regions. We then observe that in lagging regions there may be, however, untapped immobile resources, and their valorization justifies place-based policies. Current profitability may be in favor of a certain spatial distribution of activities, but potential profitability may be in favor of a different distribution. The possible movements that may arise will depend on the formation of ex ante expectations. It is therefore reasonable to think that intentional actions such as place-base policies—supporting the best exploitation of untapped, immobile resources where they exist—are justified and may produce significant results. We will lastly examine whether and how realistic it is to assume that these resources are untapped, taking into account a strong objection: if resources are available, they will be spontaneously exploited in a market capitalism system.

2.1 Regional Imbalances

In Chap. 1 we saw that place-based policies are appropriate if they are designed to effectively exploit existing, untapped local resources. Otherwise they are less efficient than real people-based ones. Place-based policies, therefore, require special organizational tools and detailed impact evaluation to learn from experience. We then saw that since the 1990s European cohesion policy has aimed to be a place-based policy, especially with the reform approved in December 2013. We also saw that this reform did not entirely succeed in strengthening its organizational and evaluative system. These considerations could lead to a certain pessimism regarding the chances of success of a place-based policy. We will now consider what objectives and challenges place-based interventions require. We will start

G. Seravalli, *An Introduction to Place-Based Development Economics and Policy*,
DOI 10.1007/978-3-319-15377-3_2

this survey using the NEG models that inspired the 2009 World Bank report. NEG models can lead also to place-based policies. In fact, the mechanism shaping the spatial location of production activities also depends on local, immobile resources. In order better to understand the significance of immobile resources, however, we must first examine the issue of mobile resources, and why they do not go in the "right" direction.

Over the last 40 years, capital has become more mobile worldwide. Such mobility has been even more marked in the Euro zone. Kaminsky and Schmukler (2002) constructed an index of capital flows liberalization ranging from one (no liberalization) to three (maximum liberalization), and found that from 1973 to 2001 in G-7 countries the index has grown continuously from 2 to 3 while in other European countries it has increased from 1.5 to 3 and in South-East Asian countries it has grown from 1 to 2.7. Even in Latin America the index has grown almost to the same extent, recovering the serious fall that came with the 1983–1988 financial crisis. Considering Quinn et al. (2011), these trends continued between 2001 and 2007, while other indicators showed different results in terms of liberalization levels, but similar trends.

These openings in capital markets should have reduced regional and country disparities in Europe and around the world. That is, income gaps lead to much greater gaps in the rate of returns on capital, dictating—in free capital markets—that large flows of capital should be transferred from more advanced to backward areas, thus reducing these disparities. In Europe, however, while disparities between countries first increased and then (in the last 15 years) decreased, disparities between regions, which are much larger, showed no reduction at all. (In 2010, the GDP per capita of the richest country, Luxembourg, was 7 fold higher than the poorest, Bulgaria; but the richest region, Inner London, had a GDP per capita 12.4 times higher than the poorest region, Severozapaden—Bulgaria). We can add that, between 1995 and 2009, in terms of indicators of well-being such as life expectancy and mortality (normally more converging than the per capita income), a recent research did not find any sigma convergence between European regions as a whole (Maynou et al. 2014).

In the rest of world, according to 2012 World Bank data, the per capita income ratio in purchasing power parities between the 20 % of richest countries and the 20 % of the poorest, in a sample of 184 countries, is 30. The largest country among the 20 % of the richest is U.S., and the largest among the 20 % of the poorest is Bangladesh. The ratio between U.S. and Bangladesh per capita income is 20. However, considering the leading area in U.S. (New Jersey) and the most lagging area in Bangladesh (Rajshahi), the ratio of New Jersey's per capita income and Raishahi's is 117.[1] The proportional income gap between Mayer-Foulkes's (2002)

[1] Calculation is based on data from page 338 and 339 of the World Bank's 2009 report. Leading area is defined as the area with the highest measure of welfare per capita (income or consumption or GDP) as percentage of country's average welfare measure. Lagging area is defined as the area with le lowest measure of welfare as percentage of country's average welfare measure.

richest and poorest groups of countries (a total of five groups) grew by a factor of 2.6 between 1960 and 1995. International inequality continued to rise after 1995 among world countries (excluding China) until about 2000. It then began to decrease, this time including China. However, «income disparities among global citizens seem to have remained unchanged, despite a reduction in international (between countries) inequality. Most important, this widening gap between international and global inequality appears to have resulted from increased income disparities *within* countries—notably in large emerging Asian economies such as China, India and Indonesia, as well as many OECD countries» (Olinto and Saavedra 2012). «Disparities in economic performance across OECD countries are often smaller than those prevailing among regions of the same country. Further, these regional disparities have persisted over time, even when economic disparities among countries were falling» (OECD 2010). In sum, world globalization and the integration process in Europe seem to be accompanied by a significant economic divergence among regions. This is difficult to understand if you consider that convergence was favored by an increase in capital mobility. Mobility of capital, of course, is greater between regions than between countries.

This demonstrates that there is something else to deal with, beyond the availability of capital and free markets. Recent growth economics has taken human capital into account, seen as a complement of material capital. In principle, this could explain the persistence of the gap. Material capital does not move in the right direction owing to a lack of human capital in lagging economies, which makes material capital less productive and less profitable than it should be if we consider its scarcity. However, it has been shown that only half of the gap between rich and poor countries depends on both material and human capital (Caselli 2005). A calculation breaks down the value of 24 (existing ratio between the 20 % richest and 20 % poorest countries' GDP per capita based on 2005 Penn World Tables, on a sample of 84 countries) into the *product* of two components: 4.8 points relate to the endowment of labor, material and human capital, while 5 points are a *residual*, called total factor productivity (*tfp*).[2] At macro and sectorial level, countless empirical works, for countries and regions, essentially confirm the same results, i.e. that a significant part of per capita income differences cannot be explained by different endowments of labor, material and human capital.

Many studies have tried to identify factors affecting the residual *tfp*.[3] They have shown that: (i) these factors are significantly different and country-, region-, and sector-specific; (ii) some regularities emerge suggesting five major influences: initial conditions, cumulative processes due to complementarities (giving increasing returns) between material, human capital and economic organization, technological change, specific resources, factor and product mobility; (iii) there are

[2] Klenow (2006) divides the same ratio (24) into four multiplicative components: the employment rate of the population (1 point), physical capital per output unit (2 points), human capital (from 2 to 4), and residual *tfp* (from 3 to 6).

[3] For valuable reviews see: APO (2004), Isaksson (2007), Danquah et al. (2011).

interrelationships among these five factors, and it is very difficult to identify which of them prevails.

As it is difficult to distinguish empirically what the prevailing factors are, two different positions could be justified. The first claims that resources and mobility of factors and products are what count most. It also claims that initial conditions, complementarities and cumulative mechanisms (history and geography) may affect development processes, although to a lesser extent in the long-term (when eventually diminishing returns overbear) or are offset in the whole economic system. The second position claims that "long term" and the "whole economic system" do not provide a useful perspective. «The purpose of the exercise in comparative dynamics is not to show how some hypothetical economy can expand indefinitely over time given certain underlying conditions. It is rather to explain why, as the historical record bears such strong witness, the expansion path of a free enterprise economy is likely to be so erratic» (Eichner and Kregel 1975, 1296).

From a territorial point of view, the most evident phenomenon linked to history and geography is the persistent spatial concentration of economic activities in some places and not in others. Spatial economics, interested in studying these agglomerations, believes it is useful to explain the different («erratic») growth of different places, considering the growth of the country as a whole as a result of these different growth patterns. In fact, spatial economics and economic geography have managed for a long time to explain the permanent gaps between regions that depend on the geographical advantages of some places (natural resources, historical concentrations of populations, favorable geographic positions), and on concentration forces such as *economies of specialization* (Smith 1776; Babbage 1832), *economies of scale* (Smith 1776; Young 1928) and *external economies* (Marshall 1890) even if they are contrasted by a number of divergent forces leading to dispersion, such as transportation and congestion costs. Economies of specialization, economies of scale and, in part, even externalities are linked to *indivisibilities* that are at the root of increasing returns and significant parts of *tfp*, which in turn imply imperfectly competitive markets[4] and cumulative processes of development by means of agglomeration.

[4] Increasing returns are conceived at the firm level or at the level of linked firms complex, as pointed out by Leijonhufvud (1986). It follows that the perfect competition is excluded, since firms or systems of firms with increasing returns undergo strong incentives to increase their size and consequently their market power. This, on the other hand, is what you need to have a definite geographical distribution of productive activities, which, according to the Starret's theorem (Starrett 1978), is not possible under perfect competition if there are indivisibilities and non-zero transportation costs.

2.1.1 Economies of Specialization

As Edwards and Starr (1987, 192) pointed out, Adam Smith's famous phrase «division of labor is limited by the extent of the market» would be false if labor were not "indivisible", because of "nonconvexity".[5] This term means that «the elasticity of output with respect to inputs is greater than 1 and the [production] function is not concave» (Romer 1990b, 4). This happens when increasing production also increases the division of labor and consequently requires skilled workers, who—through learning—execute a task better and in less time than unskilled workers. By doubling, for example, the work input, production is more than doubled. With division of labor, the addition of one unit of work gives rise, for example, to three product units, while without division of labor, each work unit corresponds to one product unit. This is the meaning of indivisibility, which is related to nonconvexity and economies of specialization. By employing specialized labor, you cannot gain an additional unit of production by employing one additional unit of labor; you gain necessarily a minimum of three product units. By contrast, unskilled labor is divisible, giving one product unit for each unit of work.

As Mill (1848) claimed, there are several causes of economies of specialization and «by Adam Smith they are reduced to three. First, the increase of dexterity in every particular workman; secondly, the saving of the time which is commonly lost in passing from one species of work to another; and lastly, the invention of a great number of machines which facilitate and abridge labor, and enable one man to do the work of many» (Mill 1848, I.8.15). Leaving aside for now mechanization, the two factors (dexterity and time loss reduction) do not include what Mill considered the most important source of economies of specialization, i.e. the assignment of different tasks to workers based on their different skills, as emphasized by Babbage (1832). «The greatest advantage (next to the dexterity of the workmen) derived from the minute division of labor which takes place in modern manufacturing industry, is one not mentioned by Adam Smith, but to which attention has been drawn by Mr. Babbage; the more economical distribution of labor, by classing the work-people according to their capacity» (Mill 1848, I.8.21). Babbage and Mill, therefore, assume that workers acquire different skills, and this has central importance in the process of division of labor and increasing returns. As we will show, this consideration implies the possibility of involuntary structural unemployment. It also allows interesting developments in our analysis, providing insight into the underlying causes of these different acquisitions of skill.

[5] «Indivisibilities typically imply local nonconvexities» (Romer 1990b, 3).

2.1.2 Economies of Scale

The division of labor, however, requires a costly, purpose-built organization. Thus, the division of labor is useful only if the size of the market is such that the entire production is sold to sustain organizational costs. If the market is not able to absorb the whole production, operators should avoid division of labor. Productivity will be lower, but they will not have to support organizational costs. This principle was pointed out by Charles Babbage, taken up by John Stuart Mill[6] and then by Nicholas Kaldor.

> It is to be remembered that it was under the assumption of "perfect divisibility" where all economies of scale are absent that the conception of equilibrium of the Lausanne School was elaborated. […] in a world where the scale of operations offers no technical advantages, economies could be gained by reducing that scale further and further until the need for co-ordination (i.e. the need for a specialized function of control, of decision-making) was completely eliminated. […]. In such a world, therefore, there would be no organization of production into firms, or anything comparable to it. (Kaldor 1934, 72-73, note 2).

We can now see that these concepts are applied in a numerical example in the spirit of Alex Leijonhufvud (1986) and using Miura's (2005) contribution on Babbage. There are drastic simplifications that nonetheless do not betray its logic and results. Let us assume, as in Smith's example, the production of pins requiring several activities, which Babbage schematized in seven processing stages[7]: (1) drawing wire, (2) straightening wire, (3) pointing, (4) twisting and cutting

[6] «[…] "it will have been found necessary to establish an accountant's department, with clerks to pay the workmen, and to see that they arrive at their stated times; and this department must be in communication with the agents who purchase the raw produce, and with those who sell the manufactured article" (Babbage). It will cost these clerks and accountants little more time and trouble to pay a large number of workmen than a small number; to check the accounts of large transactions, than of small. If the business doubled itself, it would probably be necessary to increase, but certainly not to double, the number either of accountants, or of buying and selling agents. […]. If an increased quantity of the particular article is not required, and part of the labourers in consequence lose their employment, the capital which maintained and employed them is also set at liberty» (Mill 1848, I.9.5, I.9.7).

[7] The stages were many more in Smith's example, so the seven indicated by Babbage were probably a result of its clustering. «To take an example, […] one in which the division of labor has been very often taken notice of, the trade of the pin-maker; a workman not educated to this business […] make one pin in a day, and certainly could not make twenty. But in the way in which this business is now carried on […] it is divided into a number of branches, of which the greater part are likewise peculiar trades. One man draws out the wire, another straightens it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations; to put it on, is a peculiar business, to whiten the pins is another; it is even a trade by itself to put them into the paper; and the important business of making a pin is, in this manner, divided into about eighteen distinct operations, which, in some manufactories, are all performed by distinct hands. The business of making a pin is divided into about eighteen distinct operations» (Smith 1776, 3).

Table 2.1 Working hours of workers (A, B, C) in three production phases (I, II, III) for making one lb. of pins

Workers	Phases		
	I (wire processing)	II (heading)	III (finishing)
A	2	3	3
B	3	1	4
C	3	3	2

Table 2.2 Production in parallel

Products (lb. of pins)	Working hours							
	First	Second	Third	Fourth	Fifth	Sixth	Seven	Eighth
	Workers and phases							
1	AI		AII			AIII		
2	BI			BII	BIII			
3	CI			CII			CIII	

heads, (5) heading, (6) tinning, or whitening, (7) papering (Babbage 1832, 184). To simplify, let us assume that the production process can be further clustered into three phases: wire processing, heading, and finishing. In fact, as pointed out by Miura (2005, 10), heading and finishing should be considered relevant stages (bottlenecks). Thus the division into three stages of production represents a strong but rational simplification.

Let us assume that three workers A, B, C are equally capable of producing one lb. of pins in each hour of eight hours of work,[8] but differently capable in the individual stages of production, as shown, for example, in the Table 2.1.[9]

Consequently, in isolation and working in *parallel*, each of them produces their lb. of pins every eight hours of work, implementing the three stages of processing and sells them on the market at the same price. The production arrangement is as follows (Table 2.2).

Worker A attends phase I in the first two hours, goes onto phase II in the following 3 h, and finally moves to phase III during the last 3 h. Worker B attends phase I in the first 2 h, then goes onto phase II in the 4th h, and so on.

These workers, however, understand that there is a way to reorganize their work among themselves to obtain greater efficiency. This is to assign to each

[8] In Babbage's Table, the time for making one lb. of pins was 7.6892 h. One lb. was the weight of 5,546 pins.

[9] In Babbage's Table, the time for making the individual stages of production for one lb. of pins was 1.3436 h for drawing wire, straightening wire, pointing, twisting and cutting heads; 4 h for heading, 0.1071 h for tinning or whitening, 2.3456 h for papering. In order to make simpler the representation, we assume a narrower definition of heading so that the three production stages can appear more similar in average time of execution.

Table 2.3 Series (lame) production

Products (lb. of pins)	Working hours							
	First	Second	Third	Fourth	Fifth	Sixth	Seven	Eighth
	Workers and phases							
1;2	AI		BII		CIII		AI	
2;3	BII		CIII		AI		BII	
3;4	CIII		AI		BII		CIII	

Note We assume that this is the generic working pattern. The production has already begun before the 8 h work represented here. At the beginning there are ready: one lb. Of pins that lacks the Phase II and III, another lb. that lacks only the phase III. The same two semi-finished lb. of pins can also be found at the end of the 8 h, in addition to the four finished. *Dark* areas indicate hours of inactivity/unemployment

Table 2.4 Series production

Products (lb. of pins)	Working hours							
	First	Second	Third	Fourth	Fifth	Sixth	Seven	Eighth
	Workers and phases							
2;4	2(AI)		BII	BII	2(CIII)		2(AI)	
4;6	BII	BII	2(CIII)		2(AI)		BII	BII
6;8	2(CIII)		2(AI)		BII	BII	2(CIII)	

worker the production phase in which he/she is more efficient. The new arrangement may look like a *lame series* production. Each worker is now dedicated to the phase in which he/she is more efficient. The A worker implements only phase I, the B worker phase II, the C worker phase III, respectively. As shown in the next table, this production system involves a waste of time which would become unemployment in a larger scale of production. Three men are now making 4 lbs. of pins working a total of 20 h with 4 h of inactivity. This means that 12 lbs. of pins require 60 h of work with 12 h of inactivity, i.e. one worker/day of unemployment.[10] This is the reason why we used the term "lame". The same waste of time/labor, can be seen as a saving of time/labor if the production system is organized by an employer (Table 2.3). And it is from this viewpoint that Babbage defines the principle of division of labor, according to the different skills he introduced. «That the master manufacturer, by dividing the work to be executed into different processes, each requiring different degrees of skill or force, can purchase exactly that precise quantity of both which is necessary for each process».[11]

While Babbage's principle of the division of labor has generally been considered only from this labor-saving point of view, most likely because this is the way

[10] The working time was 12 h a day, as reported by Babbage.

[11] Babbage (1832, 175), as quoted by Miura (2005, 3).

Marshall used it, Babbage also formulated a second distinct principle. «Babbage himself implicitly indicated another principle of the division of labor that the number of employees must be based on the synchronization among processes» (Miura 2005, 1). In fact, as we can see in the next table, the example we have given shows that a third system of working is possible, i.e. the most efficient in terms of productivity. It requires, however, two more workers (one type A, and one type C) and a broader market of the final product (8 lbs. of pins sold every 8 h instead of 4). We will see later that this condition assumes considerable importance for the possible development of our analysis.

Let us assume for now that these workers are available, so that the productive framework may be a *series* production (Table 2.4). The waste of time and unemployment disappear by adding new skilled labor resources, which are *complementary* to the existing ones, the very lack of which gave rise to lame series production and unemployment. In this way the hourly productivity of the overall workforce increases.

We now have 8 lbs. of pins produced by 5 workers. Productivity has increased. In the parallel system, without division of labor, the productivity was 3 lbs. of products divided by 3 workers, equal to 1. In the lame series production, the productivity was 1.3. In complete series production productivity was 1.6. We thus have *increasing returns due to economies of scale* at a system level.

Furthermore, the number of exchanges also increased. Not considering the purchase of raw materials, in the parallel system there were 3 products which reached the final market. In the lame series system, we find 4 products in the final market, and 6 from one worker to another with a total of 10 exchanges. In the series system, the exchanges are 8 in the final market and 12 from one worker to another, i.e. equal to 20. The exchanges entail organization: to establish and continuously monitor the quantity and quality of deliveries and to ensure appropriate timing (if a phase slows, the whole process of production is delayed). Moreover, in the parallel system each worker had his own contact with the end customer. In the series production system this contact is held only by one worker (C), who bears greater responsibility because he is holding the economic fortunes of all the others in the balance. These series system complications imply substantial costs to be borne by each worker since they are not specialized in these functions. It is less costly to introduce division of labor into productive and organizational tasks. This means arranging production within a *firm*, where a specialized coordination function that is able to deal with all the exchanges can be made available, as Kaldor (1934) stressed.

Let us imagine, for simplicity's sake, that without a firm, workers would be forced to spend all the productivity gains made through the division of labor on organization costs. In this case, parallel and series systems would be economically the same. It is the firm that can implement the division of labor, since it is able to obtain a consequent *surplus* if the final market is large enough. Introducing a firm also brings about positive feed-back between division of labor and the size of the market. «The division of labor depends upon the extent of the market, but the extent of the market also depends upon the division of labor» (Young 1928, 539). The reason for this reverse causation is that the firm, with the surplus that it gets, has the means to increase its production scale (with additional 5, 10, 15 workers and so on) in order to introduce powerful *production tools*. «It would be wasteful to make a hammer to drive a single nail; it would be better to use whatever awkward implement lies conveniently at hand […]; the principal advantage of large-scale operation […] is that it […] make methods [the hammer] economical which would be uneconomical if their benefits could not be diffused over a large final product» (Young 1928, 530).

The fact that these methods are «uneconomical if their benefits could not be diffused over a large final product» indicates the existence of *economies of scale at firm level*, alongside economies of scale at production system level. The resulting productivity gains can be sufficient to support an adequate expansion of demand (market size) in an open partially export-driven economy, as in the Beckerman (1962) model. As to domestic demand, we can add that the same increasing returns logic may be applied to both firms producing tools and firms using pins to make clothes. This logic entails economies of scale also at the *economy level*, i.e. a process of general income growth. The pin internal market will also grow.

2.1.3 External Economies

So far, nothing has been said about the spatial dimension. We will now see that *external economies*, which are also linked to the division of labor, constitute important factors of agglomeration because they give rise to increasing returns of productive activities located *close* to one another.

Dedicated tools and machines, brought into use in production processes with division of labor and large economies of scale, need to be invented, produced, and continually updated. They require specific activities, which cannot usually be conducted by the same firm that uses the equipment, whose production requires different knowledge and organization. In fact, the experience of countless industrial districts would indicate that industries of final goods grow side by side with industries that produce machinery and equipment for the same final goods industries. The reason is that there are reciprocal know-how *spillovers* coming from a close interplay between dedicated equipment producers and users, which have a significant role in supporting the continuous improvement in production processes. We may say that this kind of productivity gains are due to *external economies*: one firm's benefits created externally by the very existence of other firms.

However, it is important to specify that the supply and purchase of dedicated, customized equipment require mutual exchange of relevant information concerning customer's special needs as well as supplier's technical and economic capabilities and constraints. This exchange of information is normally included in the contract. The buyer and the supplier mutually benefit from information exchange, but this is fully accounted for in the price-fixing. Information exchange does not constitute an *externality*, to the extent that buyer and seller could find themselves far apart and exchange information through means of distance communication. Knowledge spillovers as positive externalities refer to another mechanism. If buyer and supplier may closely interact, they can directly learn what the other's technical and cultural operating environment, where its *tacit* knowledge is embodied, is (Rosemberg 1982; Nelson and Winter 1982; Von Hippel 1994).

This discovery of the other party's *tacit* knowledge may suggest possible improvements in their reciprocal "languages", which can, in its turn, lead to possible improvements in reciprocal products and processes. Still, there is no contract for this learning. The reciprocal benefits are often involuntary, always uncertain, and never quantifiable. In fact, it is possible that this learning precedes and makes feasible a *subsequent* contract between firms that develop explicit, durable, and collaborative relationships. Thus, knowledge spillovers of this kind can be considered externalities. This is the third type of positive (local) externalities indicated by the «Marshall's trinity: specialized providers of industry inputs, thick markets for specialized labor skills, and information spill overs» (Krugman 2010, 1). Knowledge spillovers between producers and users of dedicated equipment thus join information spillovers between nearby firms producing the same or similar goods (Feldman 2000), which can adopt the technical and manufacturing solutions introduced by others.

Holding together these two forms of spillover seems to be the best interpretation of the famous passage from Marshall, usually interpreted as including only the second form.

> When an industry has thus chosen a locality for itself, it is likely to stay there long: so great are the advantages, which people following the same skilled trade get from near neighborhood to one another. The mysteries of the trade become no mysteries; but are as it were in the air, and children learn many of them unconsciously. Good work is rightly appreciated, inventions and improvements in machinery, in processes and the general organization of the business have their merits promptly discussed: if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further new ideas. And presently subsidiary trades grow up in the neighborhood, supplying it with implements and materials, organizing its traffic, and in many ways conducing to the economy of its material. Again, the economic use of expensive machinery can sometimes be attained in a very high degree […]. For subsidiary industries devoting themselves each to one small branch of the process of production, and working it for a great many of their neighbours, are able to keep in constant use machinery of the most highly specialized character, and to make it pay its expenses, though its original cost may have been high, and its rate of depreciation very rapid. (Marshall 1890, IV.X.7 and IV.X.8).

In order to see how the first two types of externalities (inputs' specialized providers, thick labor market for specialized skills) have their root in division of labor and, consequently, in increasing returns of scale, we have to consider that there are limits to the expansion of the firm size. The fact that organizational advantages turn into disadvantages over a certain threshold (a principle that is usually referred to Williamson 1975) was already clear in a passage from Stuart Mill. «Whether or not the advantages obtained by operating on a large scale preponderate in any particular case over the more watchful attention, and greater regard to minor gains and losses, usually found in small establishments, can be ascertained, in a state of free competition, by an unfailing test» (Mill 1848, I.9.7).

The market cannot be dominated by one large firm even if there are economies of scale at firm level. The result of the feedback between division of labor and the extent of the market will lead to the birth of *many* new firms. More firms entail more differently skilled workers. Going back to the previous numerical example, the possibility of finding workers "of type A" especially able in phase I (wire processing), and of "type C" especially able in phase III (finishing), depends on the existence of a broad pool of differently specialized workers. The existence of many firms nearby is *one of the two* conditions that makes this differentiation possible.

If we consider the series production as a firm, we may say that this firm's "core" activity is phase II (heading) in which it is particularly efficient thanks to the specialized worker B. Other firms may have different peculiar capacities. This can happen in part, as the example shows, in the production of pins. We have seen that the most critical production phases are heading (phase II) and finishing (phase III). It is therefore reasonable to imagine that besides the firm specialized in phase 2, there will be another firm, or other firms, specialized in phase III. These firms will therefore provide skilled workers of type B and type C to the local labor market. In order to obtain the availability of workers of type A, however, another condition must be true.

As mentioned before, the specialization in stages two and three is because these are crucial phases in pin making, owing to the technology of the production process. Thus, the meaning of the term 'crucial' is 'particularly difficult from a technical point of view'. For example, producing the pin head requires extreme precision in size and movement. Finishing requires sharpening and polishing with exact pressure and timing in the application of friction materials, otherwise the pin breaks. Consequently, only "experts" can perform these activities without too many mistakes in an appropriately short time, namely those, which have undergone specific learning curve. Phase one (wire processing) is less technically difficult and therefore there is no need to develop specific skills.

In order to have skilled workers in phase I, the firm has to invest in its learning process purposely. It will therefore be probable that firms forced to invest in learning processes in phase I, will ensure such investment against economic fluctuations. They implement this strategy by diversifying the final production, starting, for example, with needles, wherein phase I (wire processing) is crucial from a technical point of view, while raw materials are the same as for pins, and the end markets are similar or related. Needles do not have heads but eyes. If heading (with finishing) is central in the pin production cycle, in needle production wire processing is crucial because the eyelets can break the stem of the needle if the wire processing is not accurate.

We can therefore assume that, in general, a variety of firms implies a variety of forms of division of labor, as well as a diversification of final products. We can imagine that each firm uses and reproduces specifically differentiated specialized workers, through learning. The result is a diversified labor market of specialized skills, as workers move from one company to another in search of better earnings and working conditions.

It is essential to find out whether *all* specializations are actually available. In our example, the point is to know whether skilled workers of type A and of type C are actually available on the labor market as a result of existing firms with a "core" activity in phase I (wire processing) and others with a "core" activity in phase III (finishing). This availability can be said to exist if the firms are maximizing actors and respond to *all the opportunities* available. As we will see later, a more realistic assumption sees firms as actors with a narrower set of available options. In this way, it is realistic to consider that the specializations available could be incomplete. In this case, the feasible production system will not be optimal. It will become an intermediate form between a lame series and an in series. There will also be involuntary structural unemployment due to the lack of *complementary skills* required for full employment.

Let us assume for now the simplest hypothesis that all specializations are actually available. In *every* firm it is possible to complete the in series production systems by hiring the needed complementary skilled workers, hiring them from the local thick labor market. By doing this, we may also have a second form of system in series completion. The firm, instead of hiring one more worker of type A and (one more) of type C, may purchase phase products (of phase I and of phase II) from other specialized firms to complete their production cycle. This is the second type of external economy, i.e. «specialized providers of inputs».

In conclusion, the existence of indivisibility—and hence economies of specialization—is linked to economies of scale and external economies which, taken together, can be considered strong factors of agglomeration. Therefore, persistence and expansion of economic gaps between regions despite the mobility of factors and products ultimately depends on the indivisibility.

Of course, all this takes place in industrial production that are able to agglomerate (because they are not tied to the land) and have productive processes divisible in elementary tasks in which learning and specialization can be carried out. John Stuart Mill observed this.

> The division of labor is also limited, in many cases, by the nature of the employment. Agriculture, for example, is not susceptible of so great a division of occupations as many branches of manufactures. […]. To execute a great agricultural improvement, it is often necessary that many laborers should work together; but in general, except the few whose business is superintendence, they all work in the same manner. A canal or a railway embankment cannot be made without a combination of many laborers; but they are all excavators, except the engineers and a few clerks. (Mill 1848, I.8.24).

2.2 New Economic Geography

2.2.1 Purpose

Increasing returns and external economies (that give rise to spatial industry agglomeration) were discovered, as we have seen, as far back as the end of the eighteenth century. Subsequently, they were not entirely neglected, although the marginalism that was becoming dominant built its models without taking them into account. Immediately after World War II, the agglomeration processes returned to the fore, especially with François Perroux, Gunnar Myrdal, and a considerable literature of spatial economics inspired by their contributions. Perroux was especially known for his contribution on asymmetric influences between firms[12] depending on their dimension, industrial sector, and local context. He claimed that some of them, in some places, may be "propulsive", that is, able to drive a "pole" of growth. Myrdal, on the other hand, was a sociologist and economist engaged in high-level direct political responsibilities (Twice Member of U.S. Chamber of Representatives from 1933 to 1947, he served as Trade Minister in Tage Erlander's government, and was Executive Secretary of the United Nations Economic Commission for Europe in 1947). He received the Nobel Prize in Economic Sciences with Friedrich Von Hayek in 1974 for «Research on the interrelations between economic, social, and political processes». In his famous essay "An American Dilemma: The Negro Problem and Modern Democracy", published in 1944, the "cumulative causation" mechanism was applied. He had already encountered this kind of circular process in Knut Wicksell's theories of endogenous money, and then he proposed other applications, also to explain spatial agglomeration processes. With "Economic Theory and Under-Developed Regions" (1957), "Asian Drama" (1968) and "The Challenge of World Poverty" (1970), «Myrdal showed how cumulative causation was fundamental in explaining international and interregional disparities of income. As such disparities are reflections of agglomeration of production and wealth in geographical space, […] it can be fairly said that he developed a theory of spatial agglomeration even if he did not use the term» (Meardon 2001, 45).

Poles of growth and cumulative causation concepts were conceived from the outset as part of an analysis taking history, society, culture, and institutions explicitly into account. It was widely used in a vast literature of spatial economics, which for 30 years stood in opposition to orthodox economics where space had no scope. The approach also gained considerable practical importance having inspired a season of "policies for the development poles" throughout the world (Meardon 2001, 38 and 40). The fact remains that several scholars justified Mark Blaug's severe judgment on Perroux's methodological weakness: «Unfortunately, the theory is unsatisfactory […] being in principle non-falsifiable: it is simply a slogan masquerading as a theory» (Blaug 1964, 563, cit. in Meardon 2001, 42).

In the mid and late 1970s, these theories became much less popular. At the same time, there was widespread disappointment for the failures of development that poles policies had brought about. According to Parr (1999a, b), these failures

[12] Perroux clearly makes use of the concept of external economies: «[…] la situation où le profit d'une firme est fonction de son débit, de ses achats de services, du débit d'une autre firme, des achats de services d'une autre firme. Dans cette situation, les deux firmes ne sont plus reliées entre elles par le seul prix; elles le sont aussi par le débit et par les achats de services, c'est-à-dire, puisque ces éléments dépendent de la technique et de ses changements, par la technique pratiquée par les firmes et par leurs changements» (Perroux 1955, 309).

were the result of «[…] insufficient attention […] paid to the question of whether planned poles were equivalent to poles that had evolved independent of planning, or for that matter whether poles could be planned successfully in the first place» Parr (1999a, b, 1198, cit. in Meardon 2001, 40). Planning the poles meant providing an answer to the question "what if?"[13] Is it true, and depending on what further conditions, that a process of localized development can be self-sustained by cumulative causations? The answer could only be given with a model in which the main currently *operable* variables and mechanisms are well identified. Furthermore, in order to be truly operable, these variables must normally be quite distant from the effects which they give rise to. It is not enough to say that the firm must be of a "propulsive"-type and that, after its propulsion, growth will continue supported by a circular causation process. This does not guarantee that the requirements have a framework that help answer the question "what if?"

In this regard, we can recall Krugman's list of requirements for a model suitable for this purpose: (1) Microfoundation (we must understand how everything in set in motion by agents, their actions and reactions); (2) No dormitive properties («the reference is to Moliere's doctor, who triumphantly explains that opium puts people to sleep because of its dormitive properties»). (3) Distance between assumptions and conclusions («you're learning more from a model if the rabbit isn't stuffed too visibly into the hat just before the theorist pulls it out»). (4) More than one possible outcome. (5) The model must be tractable («something you could analyze with pencil and paper and understand what was going on») (Krugman 2010, 9).

However, you may argue (against all this) that it might be a too high price to pay for a rigorous and formalized model fulfilling these five requirements. In fact, without and before formal models fulfilling Krugman's prescriptions, «there is a long geographical tradition of using externalities, increasing returns and cumulative causation in urban and regional analysis», able to explore «industrial agglomeration, […] uneven regional development, […] "industrial districts", […] inertia effects in the rise and decline of urban and regional economies, […] and the importance of labor and technology in regional development» (Martin 1999, 70–71). To be designed according to Krugman's list, the models would require highly unrealistic exclusions or oversimplifications concerning geography, culture, society, and institutions, all of which have a considerable impact in the real world.

Finally, a model that answers the question "What if?", but with the wrong answer, is not much use. We could say, then, that a tension between two opposing

[13] «When the Depression struck, there was a desperate need for answers—and the answers wanted were to the question, "What do we do?" not "How did we get here?". Faced with that question, the institutional economists couldn't deliver; all they could offer was, well, persuasive discourse on the complex historical roots of the problem. […] what mainstream economists want is the ability to answer "what if" questions: if something were different, how would that change the economic outcomes? That's a kind of question that's almost by definition impossible to answer if your approach emphasizes the uniqueness of each individual case and the specifics of history» (Krugman 2010, 5–6).

and irreconcilable methodological poles takes place. The first accuses the second of proposing interesting historical reconstructions that are useless from a policy point of view, while the second accuses the first of proposing models so abstract as to be misleading. The result is that mainstream economics completely ignored spatial economics and economic geography for a long time.

> What you have to understand is that in the late 1980s mainstream economists were almost literally oblivious to the fact that economies aren't dimensionless points in space – and to what the spatial dimension of the economy had to say about the nature of economic forces. (Krugman 2010, 1).

> This neglect is surprising. The facts of economic geography are surely among the most striking features of real-world economies, at least to laymen. For example, one of the most remarkable things about the United States is that in a generally sparsely populated country, much of whose land is fertile, the bulk of the population resides in a few clusters of metropolitan areas; a quarter of the inhabitants are crowded into a not especially inviting section of the East Coast. (Krugman 1991, 483).

New Economic Geography (NEG) aimed to convince mainstream economists to consider space as a decisive dimension of economic analysis. «I have no problem with people investigating local specificity and engaging in discursive persuasion. But the new economic geography was designed to attract the attention of mainstream economists. And mainstream economics decided long ago that devising abstract models is an essential part of being a useful profession» (Krugman 2010, 5).

NEG assumed that spatial distance is always surmountable, but that you have to pay to do so. Selling products in a different place from where they were produced leads to packaging, preserving and transport costs, as well as to documentation and administration fees. These fees are especially high when the buyer has no direct control over the quality and quantity of the product, and the seller does not receive the money immediately. This situation is highly probable when trade occurs between distant places. As a whole, "transport costs" imply that, from a firm's point of view, different locations represent different markets, leading the firm to provide the same goods at different prices. From a consumer's point of view, therefore, buying a product far from the production site would be more expensive. Increasing returns and non-competitive markets give rise to different spatial patterns of agglomeration depending on the results of contrasting forces: concentration due to increasing returns and dispersion due to transport costs.

2.2.2 First Empirical Evidence

It is thus expected that when transport costs are reduced in relation to production costs, a spatial concentration of producers, compared to an existing dispersion of consumers, may occur. We can observe that industrial spatial concentration, for 11 EU Members by sector (17 manufacturing sectors), tended to increase in the 1980s while transport costs certainly decreased thanks to the process of integration. In Fig. 2.1, in fact, the majority of the points are located above the bisector.

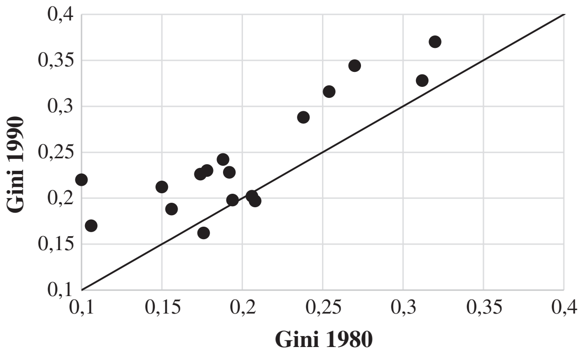


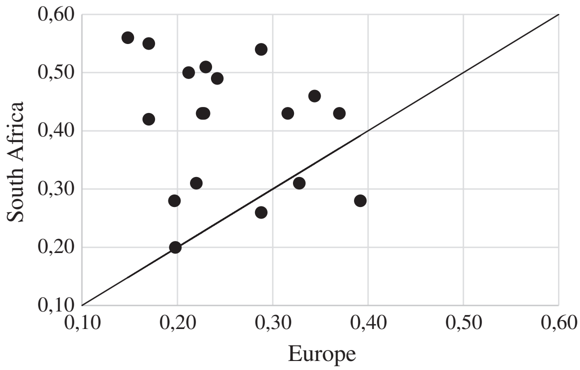
Fig. 2.1 Spatial concentration of 17 manufacturing sectors, Gini index 1990 and 1980, EU-11. *Source* Figure obtained by processing data from Brülhart (1998)

The European integration process started in March 1979 when European Monetary System (EMS) was created to stabilize exchange rates between members of the European Community. At the same time a common currency (not coined), the ECU European Currency Unit, was introduced. This was a basket of currencies that prevented movements above 2.25 % around parity in bilateral exchange rates, by reducing the exchange rate risk, which is at the basis of significant transaction costs. The common market was then enlarged. On January 1981, Greece joined the common market. In January 1986, Spain and Portugal joined (EU-12). In February 1986, the SEA (Single European Act) introduced the first major revision of the Treaty of Rome to harmonize laws and resolve discrepancies, allowing the free movement of goods, labor, services, and capital. In December 1991 the Maastricht Treaty was signed, which had three main objectives: to develop a single market with economic and monetary union through the creation of a single currency by 1999, define and consolidate cooperation in foreign policy and security, find modalities of collaboration between police and the judiciary power.[14] This example, and others related to other countries (Desmet and Fafchamps 2005), could help us define a rule whereby the concentration of productive activities is greater where transport costs are lower.

However, if we compare 11 Southern African with 11 European countries, this rule does not apply. It can be safely assumed that in industrialized countries, such as in Europe, transport costs are significantly lower than in Africa because infrastructures are more developed. We could therefore expect industry in less developed countries, such as in Southern Africa, to be more dispersed than in Europe. However, this is not the case as Fig. 2.2 shows.

[14] In December 1995 the third phase was launched with the Euro definition, which would enter in circulation on January 1, 2002. On January 1997, Austria, Finland and Sweden joined (EU-15). In May 1998, the European Central Bank was established. In March 2002, national currencies were no longer legal tender.

Fig. 2.2 Comparison between Gini indices (Southern Africa and Europe) for territorial concentration of manufacturing industry in the early 1990s. *Source* Figure obtained by processing data from Petersson (2000) and from Brülhart (1998)



The graph compares the Gini indices for the concentration of manufacturing industries[15] and shows that nearly all sectors are more territorially concentrated in Southern Africa than in Europe: 14 points out of 17 representing pairs of Gini values are located above the bisector.

This puzzle may be solved by observing that the number of cities is far lower in Southern Africa than in Europe, as is population density. Since urban areas are also industrial areas, there are very few cities in Africa that show greater industrial concentration than in Europe, where the number of urban areas and industrial centers is higher. This observation shows that many variables must be taken into consideration when analyzing the location of production activities. NEG models, in which history and geography are taken into account in a very limited, stylized way, often neglects to take these variables into account.

Nevertheless, NEG models have something to contribute. While «empirical tests have […] tended to be of an indirect kind» (Martin 1999, 70), NEG models have been confirmed empirically. «These studies may provide some support for the role of increasing returns and externalities in spatial agglomeration» (Ibidem), even if they «neglect […] important forces that also influence the geographical distribution of industry and economic activity (such as the role of local infrastructure, local institutions, state spending and intervention, regulatory arrangements, foreign investment and disinvestment, and global competition)» (Ibidem).

[15] The eleven Southern African countries are: Tanzania, Namibia, Zambia, Malawi, Zimbabwe, Mozambique, Botswana, Lesotho, Mauritius, Swaziland and South-Africa (Petersson 2000). The eleven European are: Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Netherlands, Portugal, Spain and United Kingdom. In this case, the Gini index is based on the Lorenz curve, which plots the proportion of the total manufacturing workers (y axis) that is cumulatively placed in the x % of the regions. The line placed at 45° thus represents perfect equality distribution of workers among regions. The Gini index can then be thought of as the ratio of the area that lies between the line of equality and the Lorenz curve A over the total area under the line of equality (A and B); i.e., G = A/(A + B), and G = 2A being A + B = 0,5. The Gini index can theoretically range from 0, when all the regions have the same proportion of manufacturing workers, to 1 when all the workers are in only one region.

NEG models prove, in particular, that even if firms and consumers are rational maximizing agents, a Pareto efficient state[16] of distribution of economic activities and, consequently, of income (Ottaviano and Thisse 2002), does not necessarily follow, as would be expected in neoclassical models—at least in the long run. The value of this demonstration cannot be underestimated. It legitimates intentional interventions designed to affect the distribution of income and opportunities among different ways of thinking and ideological positions.

A second contribution of the NEG models concerns the concrete possibility of success of these policies. As we will show, they illustrate an important potential role of local *immobile untapped* resources, coupled with different expectations. This implies that policies are not only justified but should also be considered able to give valuable results at reasonable costs. They can in fact produce remarkable outcomes with a limited effort *affecting expectations* when local immobile resources are present. It remains true, however, that NEG models similarly justify spatially-blind policies, as we saw in the first chapter. These policies assume that, if there are local resources, they will have already been used, or will be used spontaneously. This is why, after having highlighted the contribution NEG models can give to our subject, we are forced to consider other more useful approaches to explain the existence and persistence of local *untapped* resources.

In order to illustrate these issues, and to move on to other approaches we will now examine the simplest NEG set-up.

2.2.3 Basic Model: The Distance and Location Effects

The objectives of NEG's basic analysis can be summarized as follows: (1) to identify the effects of distance between production sites and buyers and examine the consequences of location choices; (2) to identify the forces that lead to agglomeration of firms in certain areas rather than others, as well as the forces that work against this concentration; (3) to identify the factors underlying these forces.

In order to examine the effects of distance between a production site and its markets, let us consider two economies (A and B) that produce two products: one "traditional" x, and one "modern" y.

We will begin by examining a case in which *distance is irrelevant* and NEG models *do not apply*. As we have already seen, the distance between production and markets is not relevant with zero transport costs and perfect competition. In this case, the following assumptions hold.

Both traditional and modern goods are produced with different technologies but with *constant* returns of scale. Simplifying, we may imagine a production function

[16] A Pareto non efficient distribution means that a strategy exist by which it is possible to improve one party's situation without making another party' situation worse. We will examine again this issue later.

with labor as its only input. With only one input, the hypothesis of constant returns of scale implies that the ratio between output and input is constant. In this case, the production functions of traditional (x) and modern (y) goods, with l_x e l_y representing the labor used to produce each good, are as follows:

$$x = \alpha l_x \quad \alpha > 0 \tag{2.1}$$

$$y = \beta l_y \quad \beta > 0 \tag{2.2}$$

Each of the two economies produces each of the two goods and firms make their choices based on a given market price. In this case, production takes place in perfect competition.

The optimal choice for each consumer is based on personal utility.

$$U = U(x, y) = x^s y^{1-s} \quad 0 < s < 1 \tag{2.3}$$

This formula describes the degree of utility associated with each basket of goods x and y consumed. Let us also assume that all consumers are identical and that consumers perceive as identical the units produced by different companies. In this scenario, the quantity of industrial good y produced is simply the sum of the quantities produced by each company, so that, if the total number of companies is N:

$$y = \sum_{i=1}^{N} y_i \tag{2.4}$$

$y_i =$ the quantity produced by the ith company.

The goods produced in economy A can be consumed by economy B, and vice versa, without incurring costs other than production costs (zero transport costs).

The consequence of all this is that if a given quantity of goods is produced by economy A rather than by economy B has no effect on production costs, on the firm's profits, and on sales prices. Economies A and B represent a single market in which the distance between producers and buyers, as well as the location of production companies, are irrelevant. It is as if all companies and consumers were concentrated in a single point.

The analysis conducted by NEG erases this picture.

First eliminated assumption: the hypothesis that goods produced in one economy can be consumed in the same economy or in a different one with the same costs. NEG stresses that goods produced in one place can be consumed in another place only if they are transported and this entails costs, such as packaging, preservation, and transport services (fuel, vehicle amortization, labor costs for transporters, information and contracts, enforcement of contracts). In this view, every area represents, a different market with different sales costs including transport costs. In our two economies model, this implies that for each firm there are two markets: an internal market, where sales prices are the same as production costs (including profit), and an external market, where sales prices are equal to the sum of production and transport costs. This approach makes it possible to take into account a change in transportation costs over time. Economic historians have always underlined the importance of technology progress in this field. Conservation has improved drastically, vehicles have grown in size allowing for scale economies and a more efficient transfer of goods; communications have improved to an unexpected extent. All these facts have greatly reduced the cost per unit of goods transported.

Second eliminated assumption: the hypothesis that both x and y are produced with constant returns of scale on the basis of production functions (2.1) and (2.2). In NEG models, this assumption is maintained as far as traditional goods are concerned, but it is assumed that the technology used to produce modern goods is:

$$y = \beta l_y - \gamma \quad \gamma > 0 \tag{2.5}$$

According to this hypothesis, the relationship between output and input is:

$$\frac{y}{l_y} = \beta - \frac{\gamma}{l_y} \tag{2.6}$$

Modern production thus has increasing returns of scale. This difference between traditional and modern production functions establishes the simplest way to study improvement in technologies as a change in the sectorial composition of the regional production system. Increasing the weight of the modern sector entails increasing productivity, given the number of production factors employed, by means of an organization that can spread the division of labor and consequently the use of more productive technologies.

Third eliminated assumption: the hypothesis that the location of firms is irrelevant and that consumers perceive goods produced by different companies as identical. NEG analysis assumes that consumers perceive goods produced by different companies as different varieties of modern good y (though the traditional good is still assumed to be homogeneous). It also takes into account the possibility that, if total quantities are equal, consumers prefer to consume goods of different varieties. This hypothesis can be summarized by assuming that:

$$y = \sum_{j=1}^{M} \left(y_j^r\right)^{\frac{1}{r}} \quad 0 < r < 1 \tag{2.7}$$

M = the number of varieties[17];
y_j = the quantity produced by the jth variety

The parameter r measures the intensity of preferences for variety. This intensity would be zero if r were equal to 1 [in which case (2.7) would be the same

[17] With increasing returns of scale, the optimal choice for a company would be specializing in the production of a single variety of manufactured good. The number of varieties of goods produced thus coincides with the number of companies N.

as (2.3)], and infinite if r was equal to 0, these two cases being excluded, while it is assumed that preference for variety would decrease if r increases. Where there is preference for variety, the market for modern good y is segmented into many different markets for each variety. For each of these segments, the producer's market power is limited by the fact that consumers may choose another variety if the one they have previously chosen becomes too expensive. It is assumed, therefore, that each company chooses the unit price of the variety of goods it produces. This hypothesis, along with the previous ones, means that the modern products market is no longer perfectly competitive; it is rather characterized by 'monopolistic competition'. It is possible to demonstrate (Dixit and Stiglitz 1977) that, in this case, the optimal behavior for a company is to fix sale prices by imposing a mark-up on production costs. This assumption on the preference for variety allows demand drive to be introduced into this simple model. The increase of the weight of the modern sector feeds itself because it increases productivity, owing to the hypothesis on the economies of scale, while increasing demand owing to the hypothesis on the preference for the variety.

Fourth eliminated assumption: the hypothesis that both goods are produced using homogeneous labor in both sectors. NEG assumes that workers employed to produce modern goods have different skills compared to those employed for traditional goods and that workers in the modern sector are completely mobile, in the sense that they are perfectly willing to move from one economy to another if they are paid more. Workers in the traditional sector, on the other hand, are assumed to be immobile.[18] It also supposes, for simplicity's sake, that each worker has a given set of skills, belonging therefore either to the modern or to the traditional sector.

There is a fifth assumption to eliminate, that we will examine shortly.

Let us now consider the results of this fourfold elimination of assumptions, which provides the necessary conditions stated by Starret's theorem. That is, to have an equilibrium in the spatial distribution of productive activities with indivisibilities, i.e. imperfect competition in the final market. In this scenario, profit levels in each of the two regional economies A and B depend on the number of firms that have located production there. As a result, a firm that shifts from one location to another modifies the location's profit condition.

This modification takes place with *three* different ways.

First, a new firm's arrival in a location changes the competitive context of the firms working there, which have to face more competition. Greater competition alters the market condition, lowering the price of goods produced and reducing profit levels. This is called "price effect" (*PE*) and tends to make the market *less* profitable for other firms seeking to enter.

Second, if a firm decides to move (for example from region A to region B), there will also be a shift of workers between the two regions with an increase in

[18] The assumption introduces typical features of the production of goods that depends on immobile factors in a simplified context as exemplified by the agricultural production, which uses immobile land as one of its productive factors.

employment and income in region B. Since part of this income is necessarily spent inside that economy, this will increase demand and generate greater profits in B. This phenomenon is called "market effect" (*ME*), and tends to make the market *more* profitable for other firms seeking to enter.

Taking into account these two effects (we will soon see a third effect), the decision of a firm to locate its production in one area rather than another generates an overall effect on profits:

$$\Delta\pi = PE + ME \quad PE < 0 \;\; ME > 0 \tag{2.8}$$

The overall effect $\Delta\pi$ is positive if market effect $ME > 0$ prevails over price effect $PE < 0$ and negative if the opposite occurs. If price effect prevails over market effect, a new entry will discourage other firms to do the same. If, however, market effect prevails over price effect, a new entry will increase the profitability of the area, persuading other firms to make the same move. When either price or market effects prevail, therefore, two different models of development tend to evolve. In the first one ($PE > ME$), the modern sector expands in parallel in both regions. In the second, the modern sector agglomerates in one of the two regions. This agglomeration becomes the "center" where goods are produced for its own market and for the other region, which becomes "peripheral".

A *third effect,* which depends on the elimination of a fifth simplification of the model of perfect competition and space irrelevant, is a useful addition to the *tfp-underlying* mechanisms.

Fifth eliminated assumption. So far it has been assumed that labor is the only input used in production. In order to make the hypothesis more realistic, two inputs should be taken into consideration: labor and goods produced by other firms. As we have seen considering the organizational consequences of indivisibilities, this hypothesis—that there are firms producing phase goods for other firms—follows from the same division of labor. For simplicity's sake, let us assume that goods used as input to produce other goods, and those bought by consumers, coincide. In this case, the demand for goods of each firm would derive in part from consumers, and in part from other firms. Firms are thus mutually connected through input-output links. Consequently, a new entry in an economy generates a third effect since a decrease of sales price (price effect) reduces the expenses of firms that use the goods as input. The effect on production costs caused by the arrival of a new firm in a marketplace is called "cost effect" (*CE*) (Venables 1996).

The overall effect of a new entry, thus, becomes:

$$\Delta\pi = PE + ME + CE \quad PE < 0; \quad ME > 0; \quad CE > 0 \tag{2.9}$$

The overall effect $\Delta\pi$ is still unclear. A new entry increases the profitability only if market and cost effects are greater than the price effect.

Note that this result, already complex enough, derives from a model that has been drastically simplified. On the grounds of geography, only two regions are assumed (as we will see, the results thus obtained cannot be applied directly to a case of multiple regions). Moreover, they have no size (they are points in space).

Furthermore, they are structurally homogeneous, except for the proportion of modern activities, which may be different.

2.2.4 Basic Model: Dynamics

A general analysis is complex even though it uses the basic model we have outlined (two punctiform homogeneous economies, two sectors, elementary production functions and preferences). It implies examining the consequences on market, price and cost effects of the whole range of possibilities opened up by the different parameter values. These are: α (productivity of traditional labor), β (productivity of modern labor), s (a parameter that regulates the utility of consumers according to the combination of modern and of traditional goods), γ (a parameter that regulates scale economies in modern sector), r (consumer preference for a variety), as well as other parameters indicating transport costs and the importance of input-output links among firms.

It is therefore preferable to perform a qualitative analysis of some results and a few main trends. The dynamics we examine are not the only economic development paths outlined by NEG; they represent some possible trends of particular interest. It can be shown that: (1) greater preference for variety in modern products tends to reinforce the market effect and cost effect rather than the price effect and tends to *favor* agglomeration; (2) given the total workforce, a higher proportion of modern workers tends to reinforce both the market and cost effect rather than the price effect, and tends to *favor* agglomeration; (3) stronger scale economies reinforce the market effect but weaken the cost effect against the price effect, with nonlinear consequences, *normally in favor* of agglomeration; (4) lower transport costs tend to reinforce both the market and cost effect rather than the price effect and tend to *favor* agglomeration. NEG is thus in a position to explain industrial agglomeration as the result of strong attracting forces (*high preference for variety, greater proportion of workers in modern sectors, scale economies*), and weak resisting forces (*lower transport costs*), both of which are significant in developed countries.

Before considering in more detail the dynamics of possible agglomeration, it may be useful to report some preliminary data confirming this conjecture.

Table 2.5 shows that the Gini index has increased in relation to the concentration of locations in manufacturing industry in Europe at country level. The same is true for unemployment at regional level in the 1980s and in the early 1990s (rows 1 and 2). In the same period, reduction in transport costs was significant, explaining a regular 1 % increase in the international trade share. This occurred in the 1970s, as well as in the 1980s and early 1990s (row 4). On the other hand, reduced tariffs and quotas explain 2 % of international trade increase in the 1970s, 4 % in the 1980s and 0 % in the early 1990s (row 5). The residual (4 % in the 1970s, 9 % in the 1980s and 3 % in the early 1990s) is therefore considerable. It relates to the effects of a country's economic growth, and to the vertical

Table 2.5 Spatial concentration of industry and of unemployment rates in Europe and causes

		1970s–1960s	1980s–1970s	1995–1980s
1	Spatial concentration of manufacturing industry by country (EU-11), Gini indices	(a)	0.156	0.188
2	Regional concentration of unemployment rates (168 European regions) Gini indices	(a)	0.284	0.302
3	International exchange of goods in % of industrial value added	+7	+14	+4
4	As in row 3 dependent only on reduced transport costs	+1	+1	+1
5	As in row 3 dependent on reduced tariffs and quotas	+2	+4	0
6	As in row 3, residual (dependent on growth and the dis-integration of productive processes)	+4	+9	+3

Note (a) Unavailable. *Sources* Table obtained by processing data from Feenstra (1998), Baier and Bergstrand (2001), Brülhart (1998), Eurostat data-base

disintegration of productive processes that could be associated with the increasing proportion of the modern sector and an increasing preference for variety. It is possible to state, as NEG indicated, that in Europe the forces of attraction towards central areas were particularly strong and accompanied by increasingly weak resistance factors.

Let us now complete our analysis on agglomeration dynamics. The link between transport costs and the location of firms is shown in the following figure referring to one region (for example A), assuming that forces of agglomeration prevail, *net of transport costs*.

Let us define: $T =$ transport costs, in a broad sense, as a percentage of the value of traded goods ($0 \leq T \leq 1$); $\psi =$ degree of integration (linked to transport costs in the opposite direction) ($0 \leq \psi \leq 1$; $\psi = 1$ if $T = 0$); $\sigma =$ share of the A region's modern employment ($0 \leq \sigma \leq 1$).

In Fig. 2.3, a point along one of the continuous or dotted bold lines represents a possible equilibrium of a stationary state in which the system can come to a halt. A point along the continuous lines represents a *stable* stationary state, towards which the system tends to move in a disequilibrium situation. A point along the curved dotted line, on the other hand, represents a state of *instable* equilibrium, from which the system tends to move away in case of disequilibrium. Let us start with a situation in which half the modern workers are active in economy A and half in economy B ($\sigma = 0.5$) and transport costs are high (low integration). If integration rises to ψ_1 (transport costs reduce to T_1) the system turns into a state of possible instability. However, a block of firms would need to re-locate in order to create effects of increased profitability sufficient to compensate high transport costs. In fact, at that level (T_1 and ψ_1), re-location would be sustainable only if *all* the firms in one region moved. If integration continues to increase, the dimension

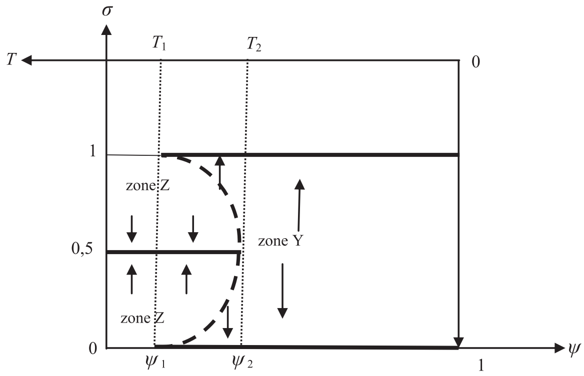


Fig. 2.3 Relation between integration (transport costs) and location, region A

of the necessary block reduces along the dotted line, which separates zone Y from zone Z. At degree of integration ψ_2 and transport costs T_2, only one firm would need to move, and all the others would follow, moving production and modern workers to either region A or B (the top and bottom continuous horizontal lines). This minimal shift is enough to determine an increase in demand and a significant reduction in production costs. If we assume that the probability of movement is inversely proportional to the dimension of a requisite block of firms, concentration will proceed quickly or slowly depending on the integration (and transport cost) level. Moreover, if at the start there are *more* firms in region A than in region B, a reduction in transport costs towards T_2 would cause a gradual, parallel shift of firms towards region A, until region B would become completely emptied. The opposite shift would take place if initially there had been more firms in region B than in region A.

In summary, if transport costs are high, the forces that prevent firms from re-locating tend to prevail, and production is distributed over different regions. If transport costs are reduced, the forces that persuade firms to re-locate prevail, and production tends to agglomerate. In the latter case, the fact that production will concentrate in region A or B depends, in the simplified model structure, on one element: the initial distribution of firms. On the other hand, the speed of the agglomeration process will depend on the level of integration (transport costs).

The long lasting agglomeration process of Italian industry after Unification in 1860 is a suggestive example.

The initial concentration process of Italian industry lasted 60 years, from 1861 to 1921 (Fig. 2.4); its slowness can be explained with the slow pace of the increase in rail and motorway national networks. There is, indeed, a close correlation between the share of industrial employment in the Center-North and the number of kilometers of the country's railways and highways (Fig. 2.5).

Given the extent of this 60 year process, of course, a different explanation could be offered. The increase in the share of the North's industrial employment is normally considered the effect of the birth of modern industry in those regions, while in the South the traditional craftsmanship that was especially important in large cities, such as Naples and Palermo, declined. According to this

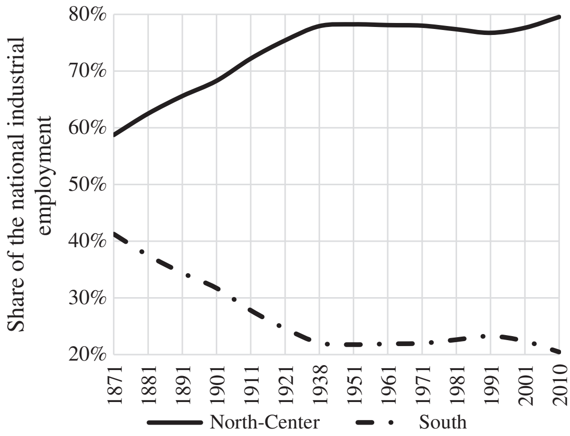


Fig. 2.4 Regional distribution of industrial employment in Italy 1871–2010. *Source* Figure obtained by processing data from: Felice (2009, Table 8, 14) for total workforce as a share of the population 1871–1971; Istat (2012) for the population 1871–1971, for share of workforce in the industrial sector 1871–1971, and for industrial share of employment 1981–2010

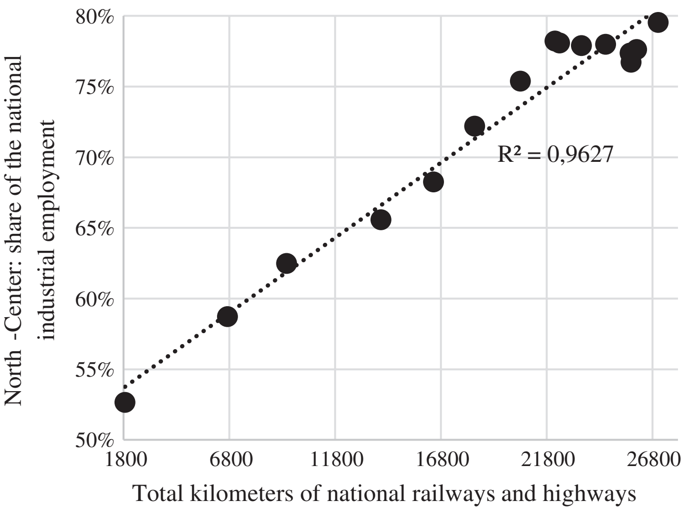


Fig. 2.5 Correlation between industrial employment spatial concentration and railway-highway network. *Source* Figure obtained by processing data from Istat (2012)

interpretation, the transfer of mobile resources (labor and capital) from the South to the North was not significant before the end of World War 1, and infrastructures for transportation grew as an effect (not as a cause) of this uneven development process. Until the eve of World War 1, the development process was uneven, but the growth of the industrial North did not damage the South. It was claimed that no significant transfers of capital and people occurred and the State did not interfere. Between the two wars a further process of industrial concentration took place, now supported by the State, which favored the North in the distribution of resources to the detriment of the South (Iuzzolino et al. 2011).

Historians and politicians from southern Italy complained at the time about the role of the State, even in the first industrialization stage.

> The Italian unification has been and will be - I have indomitable faith - our moral redemption. But it was, unfortunately, our economic ruin. We were, in 1860, in a flourishing condition for an healthy and profitable economic revival. The unity has condemned us. And as if that were not enough, there is evidence, contrary to the opinion of all, that the Italian Government lavishes its financial benefits in the northern provinces to a greater extent than in the southern. (Fortunato 1899, 65).[19]

Given the lack of data for the period before World War 1 it is impossible to exclude one or the other of these interpretations. In any case, a sixty-year time lapse is so long that circular effects may very well have occurred, implying that the causes became effects and vice versa.

2.2.5 The Bell-Shaped Model

We have seen the implications of the NEG's model dynamics assuming only two regional economies and simple price, market, and cost effects (a *core-periphery* model). «In these core-periphery models, agglomeration of economic activity is generally catastrophic [cumulative causation], complete and permanent. […]. Models with bell-shaped agglomeration patterns overcome the extreme and arguably unrealistic implication of catastrophic, complete, and permanent agglomeration of the core-periphery models. In a nutshell, these models add additional dispersion forces to a core-periphery model that break the dominance of the linkage effects at lower trade costs or higher levels of agglomeration» (Bickenbach and Bode 2013, 126).

Adding agglomeration costs such as, for instance, increasing housing costs (congestion) we easily obtain the *bell shaped* model shown in Fig. 2.6, where (in region A) we have again:

$t =$ time,
$\psi =$ degree of integration (linked to transport costs in the opposite direction),
$\sigma =$ share of the region A modern employment ($0 \leq \sigma \leq 1$).

[19] L'unità d'Italia è stata e sarà—ne ho fede invitta—la nostra redenzione morale. Ma è stata, purtroppo, la nostra rovina economica. Noi eravamo, nel 1860, in floridissime condizioni per un risveglio economico sano e profittevole. L'unità ci ha perduti. E come se questo non bastasse, è provato, contrariamente all'opinione di tutti, che lo Stato italiano profonde i suoi benefici finanziari nelle province settentrionali in misura ben maggiore che nelle meridionali (Fortunato 1899, 65).

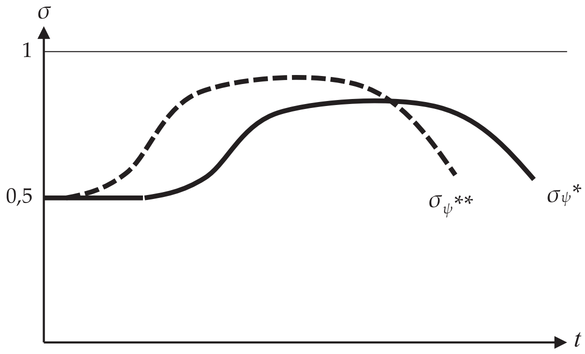


Fig. 2.6 The bell-shaped model

Starting with an equal share of modern workers in regions A and B ($\sigma = 0.5$), the concentration process in A may be of type σ_ψ** with high degree of integration (low transport costs) or of type σ_ψ* with a low degree of integration (higher transport costs). In this model, however, there is no longer a single situation towards which the agglomeration process tends ($\sigma = 1$). The integration pace affects the speed of a process that reaches a maximum concentration level and then enters a new dispersion phase. This phase of dispersion is caused by the congestion that takes place in the area where the concentration of productive activities takes place. This is essentially a new (*fourth*) effect that reduces the profitability of the area in which firms and workers concentrate. The most frequently analyzed factor is the increase in the price of housing, which leads to higher wage demands and therefore the gradual convenience of re-locating production activities to peripheral areas where wages are lower.

An example of the bell-shaped model application refers to German reunification after 1990.

> [NEG] offers a plausible story for why the blooming landscapes in East Germany have not appeared, and may not appear in the near future. At the same time, it offers a plausible story for why the catastrophic scenario of East Germany being depopulated has not, and likely will not, become reality either. In our view, the most plausible NEG view of German integration suggests that Germany may currently be somewhere close to the peak of the bell curve that describes the equilibrium relationship between integration and agglomeration in NEG. (Bickenbach and Bode 2013, 150).

If we take into consideration Fig. 2.7, we will realize that Germany has already gone beyond the peak of the bell curve. In the former East Germany the share of employment in industry and advanced services[20] decreased very soon after reunification and recovery soon followed, supported by the State. When this support ended, the share of employment went down again, reaching its minimum around 2005. In recent years, however, with very little public support, there have been

[20] Industry, not including construction, plus business, finance, insurance, communication and information services.

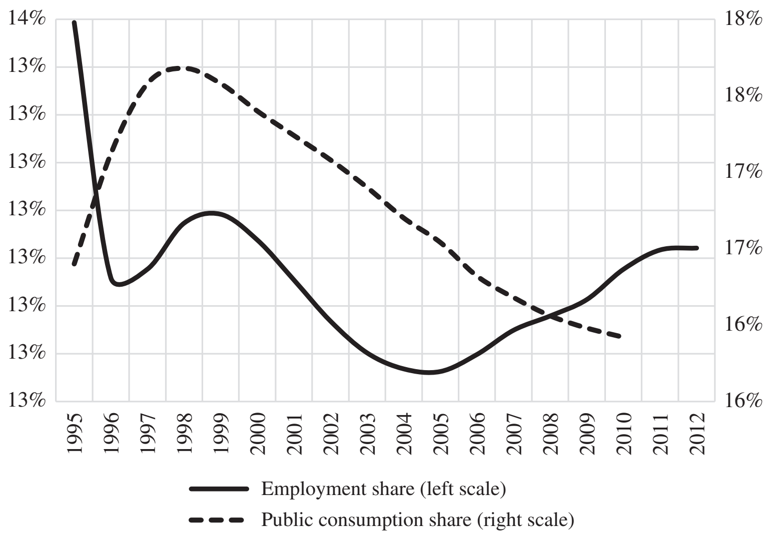


Fig. 2.7 Former DDR Germany: industry and advanced services employment share and public consumption share on the country's total, moving average five-membered. *Source* Figure obtained by processing data from Statistische Ämter des Bundes und der Länder: R1B1. http://www.vgrdl.de/VGRdL/tbls/R0B0.asp?rev=RV2005&tbl=R1B1

significant signs of recovery. As pointed out by Bickenbach and Bode (2013, 150), this may be considered evidence in favor of the bell-shaped model, even though other facts—both political and cultural—have intervened, and the long lasting consequences of the active support to former DDR regions, even when reduced, have left a positive mark.

We will now go on to analyze some framework improvements introduced to overcome the simplistic assumption that only two structurally homogeneous regions, with no dimensions were to be taken into consideration. The structural homogeneity assumption may be eliminated in different ways, but one of the most significant is related to the different endowment of immobile resources, as we will see in the next section. We will then consider some of the consequences of the two-or-more-region assumption. Finally, we will examine, especially in the case of metropolitan regions and cities, the significance of the size of the economies considered.

2.2.6 Immobile Resources

One particularly interesting form of heterogeneity is related to the number of immobile workers, which also represents one share of demand for modern goods. Since the other share of demand is mobile with mobile workers, the fact that one region has more immobile workers than another means that this region has a greater potential demand for modern goods. In the basic model, immobile workers are employed in the traditional sector. However, the same holds for any other resource that is useful for production and can be considered immobile. For example, if we use a more complex production function that includes skills linked to a territory (such as those resulting from a long history of collective, localized, specific know-how, or existing in local universities or research centers) natural resources (such as climate, landscape or cultural heritage), or socially and institutionally molded organizational frames, these resources work in the same way as for traditional workers. When exploited, they give an income that can be spent locally and sustains the demand for modern goods.

Let us imagine that transport costs are low enough to guarantee agglomeration. Modern firms tend to concentrate in one of two areas, and if the two economies are homogeneous, firms will always locate where modern firms are already settled. But where heterogeneity is related to immobile workers' stock (or immobile resources), the results are quite different. Let us assume that there is a higher number of immobile workers in region B than in region A. The simplest situation is where the initial number of modern firms is only slightly higher in region A, so that the total number of workers, mobile and immobile, is greater in B. In this case, the incentive for a modern firm to re-locate from B to A will be less pressing than the incentive to stay put. Incentive is significant, indeed, for firms in region A to move to region B where there is a greater number of immobile workers. In region B, in fact, both present and potential demands for modern goods are probably higher. We can thus say that agglomeration will normally take place in regions with a higher number of workers (including immobile workers) even if this region initially presented fewer modern firms.

The result would be different if there were a greater difference in the initial number of firms within the two economies. If, for historical reasons, there were far more modern firms operating in region A, it would be possible that the *present* demand for their products could be greater in region A than in region B, even though there were fewer immobile workers in region A. On the other hand, the *potential* demand for modern goods would be greater in region B where most of the immobile workers live. Thus, if the effect of higher *present* demand prevails in determining location, then agglomeration will tend towards region A. If, on the contrary, *potential* demand prevails, then agglomeration will tend towards region B. The result depends both on the conjectures made by each firm and on the behavior of others. If firms were able to coordinate their decisions and come to an agreement on where to locate production, they would clearly decide to concentrate in economy B. If all firms moved there, economy B would effectively become the region with a higher number of workers, given that more traditional workers are available there. Nevertheless, this kind of collaboration is not spontaneously possible and each firm would only decide to locate production in economy B if it thought that most firms in economy A were going to move toward B. The same, however, would be true in relation to the move toward A. Briefly, where there is homogeneous conjecture, firms tend to concentrate either in A or

B and there is equilibrium in their distribution. Where, on the other hand, there is heterogeneous conjecture, firms tend to localize in many different ways and there is no equilibrium in their distribution. Therefore, there is a way to shape expectations intentionally, i.e. increasing the potential demand for modern goods. If we have immobile unemployed or employed resources at low productivity, a deliberate place-based policy may play an incisive role.

2.2.7 More Regions and Convergence Clubs

We can eliminate the two economies hypothesis where firms can locate production, as they do in some NEG models, by considering *more than two* regions. In this case, however, the analysis becomes highly complex. Nevertheless, some of its conclusions are easy to summarize. One of the main conclusions states that, when transport costs are sufficiently low, agglomeration usually occurs in a *subset of regions* rather than in a single region alone.

> [In a multiregional asymmetric model] we find, on the basis of […] extensive simulations […], that […] at a certain level of integration, agglomeration starts, with a number of core regions attracting activity from nearby regions […] until […] industrial activity only takes place in the centrally located core regions. A further falling of trade costs eventually reverses this process, with industrial activity gradually spreading from the core, at first to nearby regions (not to the peripheral ones!) […]. (Bosker et al. 2010, 811).

Let us take the Italian case again, distinguishing North-West, North-East-Center and Southern areas, and considering the share of workforce. See Fig. 2.8.

The two phases, before and after the 1911 break, are clearly different. In 1871 the South had more industry and more workforce than the North-West. In the following decades, industrial concentration should have occurred in Southern regions while it actually took place in the North-West. Either NEG models do not apply to nascent industries, or Giustino Fortunato might have been right to stress the role of central government that supported the North against the South in order to subvert forces that should have favored a spontaneous industrialization in the South.

In 1911 the situation changed (see Fig. 2.9). The share of workforce was the same in the three areas and the industrial concentration process took place in the North-West where the share of industrial employment was, at that point, the highest. In this case, the model's predictions are fulfilled. Moreover, the industry agglomeration process extended also to the North-East-Center, as the more-then-two-regions model predicts. World War II represented a new turning point. The concentration process in the North-West reached its peak and dispersion forces started to work in favor of the North-East-Center area, rather than in favor of the South. Intuitively, by using the bell-shaped model, we may conjecture that this took place because congestion costs in the North-West generated location opportunities in the North-East-Center where a higher share of the workforce was located, while the share of industry was lower in the North-East-Center than in the North-West, but not as low as in the South.

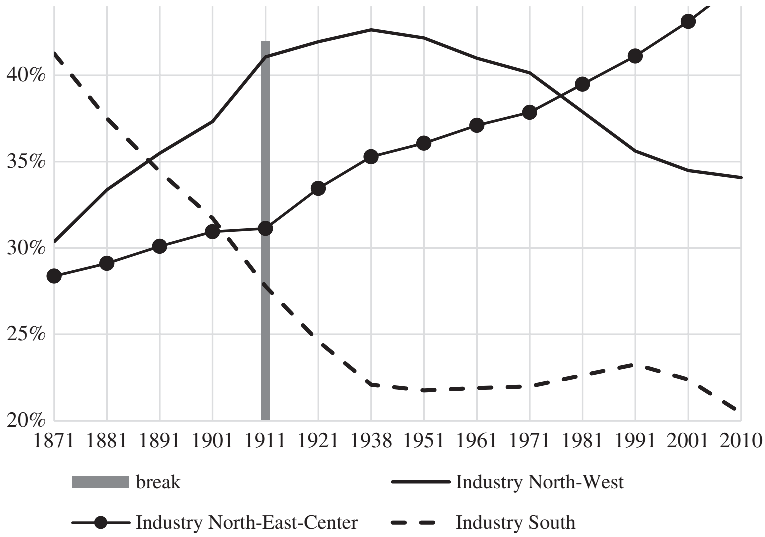


Fig. 2.8 Regional distribution of industrial employment in Italy (1871–2010) in three areas. *Source* Figure obtained by processing data from: Felice (2009, Table 8, 14) for total workforce as a share of the population 1871–1971; Istat (2012) for the population 1871–1971, for share of workforce in the industrial sector 1871–1971, and for industrial share of employment 1981–2010

We can also consider, as another example, the regional scenario of European construction. In recent years, the effects of integration in Europe have been extensively examined. We have already said that research has unequivocally shown convergence among countries and divergence among regions within countries (Geppert and Stephan 2008, 208; Doran and Jordan 2013, 29–30) (Fig. 2.10).

The reason underlying this complex dynamic of growth processes in progressively integrated areas lies in the formation of "convergence clubs". European Nuts 2 regions form six different convergence clubs[21] based on the 1990–2002 data analyzed by Bartkowska and Riedl (2012) by means of the Phillips and Sul's (2007) clustering algorithm. A country effect is only apparent in Switzerland, Austria and Finland: in these countries some regions within the country tend to cluster together. Convergence clubs generally spread among different countries, explaining the overall convergence between countries. Difference between clubs are very high (in club 6, GVA per worker is four times that of club 1). This explains the within-country divergence of regions. As the theory predicts, initial conditions determine the probability that will converge into clubs. Regions with

[21] The GVA per worker in 2002 was: club 1 €65,450; club 2 €48,180; club 3 €39,030; club 4 €31,520; club 5 €25,620; club 6 €15,590 (Bartkowska and Riedl 2012, 24).

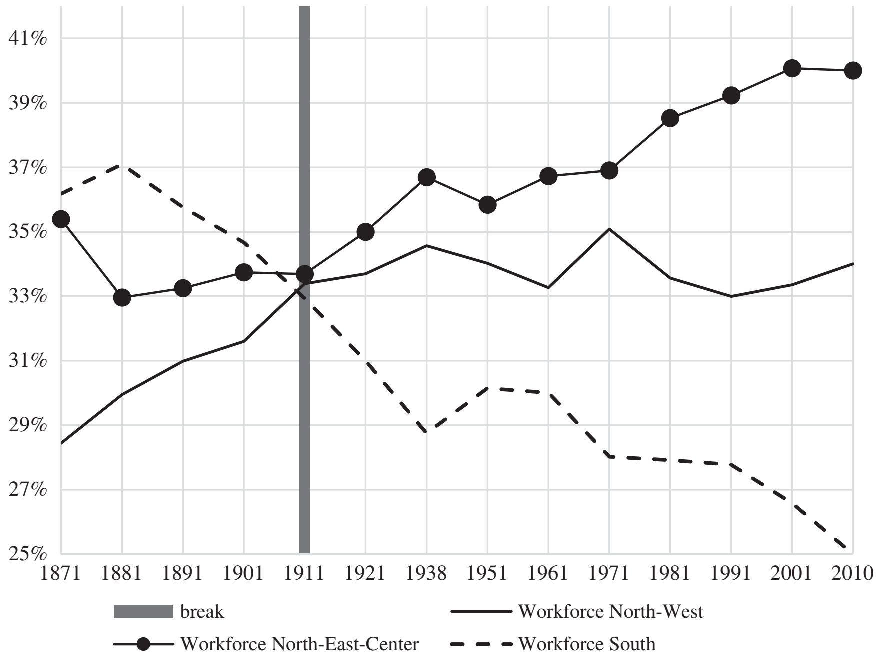


Fig. 2.9 Regional distribution of total workforce in Italy (1871–2010) in three areas. *Source* Figure obtained by processing data from: Felice (2009, Table 8, 14) for total workforce as a share of the population 1871–1971; Istat (2012) for the population 1871–1971, for share of workforce in the industrial sector 1871–1971, and for industrial share of employment 1981–2010

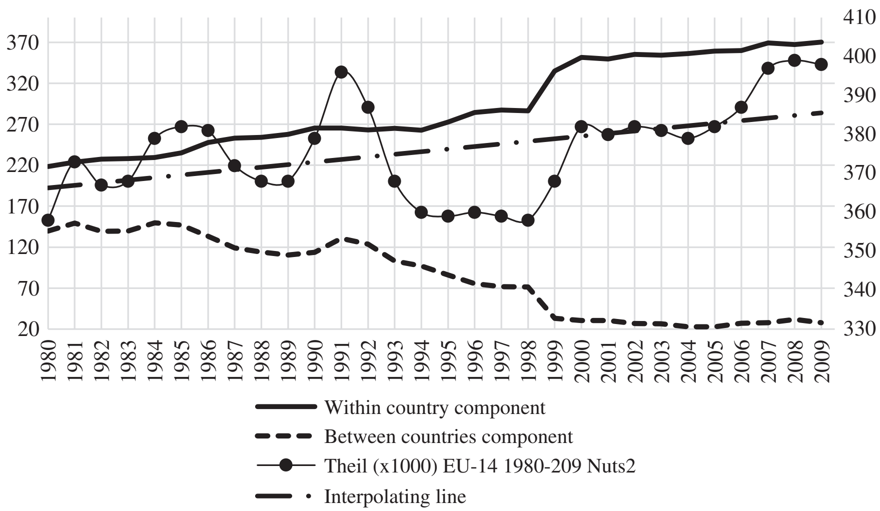


Fig. 2.10 Inequality in GVA per capita constant prices and PPP Nuts 2 European (E-14) regions, components within and between countries 1980–2009. *Source* Figure obtained by processing data from Doran and Jordan (2013, 34)

an initial high endowment of human capital, experience a higher probability of belonging to a high-income club compared to regions with a low initial endowment of human capital.

2.2.8 Size and Density

One remarkable phenomenon is that regions including the country's capital city, appear to belong to a higher club than the neighboring regions. Examples of this are Île de France (including Paris), Inner London, Lisbon, Madrid and Vienna. The special prominence of metropolitan areas in the polarization processes has often been stressed. Geppert and Stephan (2008) showed initial and Markov-chain limiting distributions on the 1980–2000 regional income per capita in EU-15. They classified regional per-capita income (EU = 1) into 8 classes, calculated transition probability matrices and the limiting distribution, distinguishing non metropolitan and metropolitan regions.[22] Non metropolitan regions show a convergence process across Europe. Metropolitan areas are, by contrast, clearly subject to a polarization process. They tend to concentrate (according to the limiting distribution taking 1980, 1990, 2000 data into account) in the highest per capita income category, while in 1980 two poles existed, concerning high and lower income, respectively (Figs. 2.11 and 2.12).

Simple NEG models are unsuited to explain this growing gap between metropolitan areas and other European regions. The high congestion costs in metropolitan areas (high housing costs and consequent high wages and labor costs) should work as a factor of dispersion, reducing their growth rates. If, however, the opposite is taking place, it means that in metropolitan areas there are powerful factors of agglomeration that may overcome congestion costs. Indeed, a burgeoning literature stresses that these factors are linked to a new role of cities as a favorable context for innovation due to knowledge spillovers and other external economies encouraged by dimension and density. Feldman and Audretsch (1999), for example, pointed out that the majority of innovative products were located in cities: innovation is, therefore, an urban activity. More precisely, they showed that the number of innovations was correlated to the dimension of the cities. This is also the result of Duranton's (2006) model built on Romer's (1990a) endogenous growth model, as pointed out by Duranton and Puga (2013, 51). The number of innovations is linked to the research activity favored by a high variety of local products that, in turn, is proportionate to the population. This observation,

[22] «We apply a four-level typology. The top group (type 1) is formed by 'large agglomerations', i.e., regions with an urban core of more than half a million inhabitants (in the year 2000). At the second level (type 2), we have 'urbanised regions' with a population in the core between 300,000 and 500,000. […]. We use the term 'metropolitan regions' to denote large agglomerations (type 1) and urbanized regions (type 2) as a group, and the term 'non-metropolitan regions' to denote intermediate regions (type 3) and rural regions (type 4) as a group» (Geppert and Stephan 2008, 199–200).

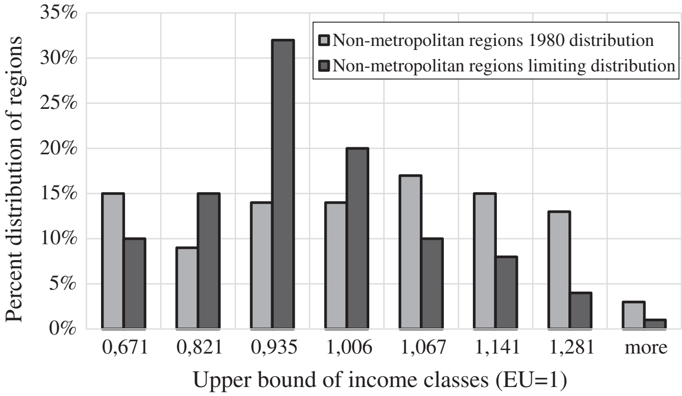


Fig. 2.11 Distributions of non-metropolitan regions EU-15. *Source* Adapted from Geppert and Stephan (2008, 205)

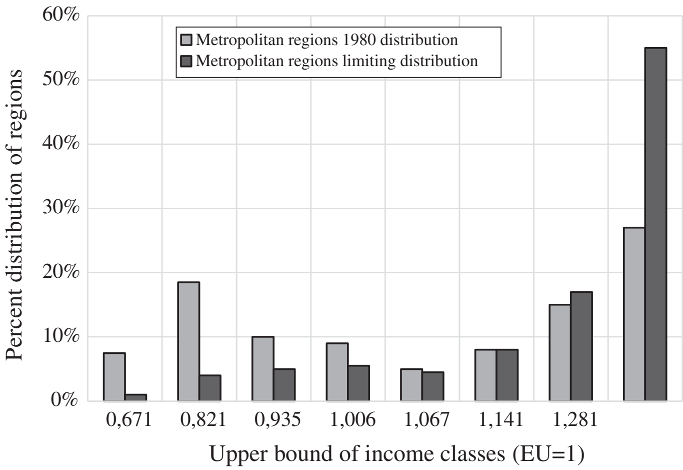


Fig. 2.12 Distributions of metropolitan regions EU-15. *Source* Adapted from Geppert and Stephan (2008, 205)

however, is a first approximation. As the appendix on the relationship between growth and size of cities in Europe highlights, this relationship appears nonlinear and requires a more complex model.

2.3 Summary

NEG's most significant results, shedding light on the forces behind the location of firms, the possible dynamics and the elements that influence firms' choices, suggest that forces fostering agglomeration, and leading to dispersion, work together. Nevertheless, all other things being equal, it is probable that improvements on infrastructure and transport services (both important for general development) widen the gap between regions, because these improvements support the spatial concentration of modern productive activities by reducing transport costs. On the other hand, concentration in one area rather than another, may entirely depend on the initial spatial distribution of modern firms, thus indicating that local development could indeed be exclusively history-dependent and proceed in a cumulative way. However, this is not the end of the story. A «catastrophic, complete and permanent» agglomeration will probably not take place and we may expect that congestion costs entail strong dispersion forces with more complex results, when considering groups of regions and convergence clubs. The initial distribution of firms is not the only significant factor. The endowment of immobile resources is also relevant. A local economy with few modern activities can attract other modern firms if there are valuable immobile resources. In fact, even if the actual demand for modern goods is scanty, the potential demand is much larger. This suggests that policies may have an important role to play.

In this regard, what has emerged from our analysis may sound puzzling. The bell-shaped model and its applications could deny the importance of public intervention. In backward regions, it seems, you should just sit and wait for the beneficial effects of dispersion due to the cost of congestion in advanced regions to rain down on you. The German and Italian cases would lead us to believe that these effects tend to be significant. Yet, apart from the social costs of this wait—and nobody knows how long for—we have also found completely different results. The processes of agglomeration and dispersion have timing, direction and intensity that may also be attributed to initial movements, given the role of expectations in conditions of unstable equilibria. Current profitability may be in favor of a certain spatial distribution of activities, but potential profitability may be in favor of a different distribution. The possible movements that may arise will depend on the formation of *ex ante* expectations. It is therefore reasonable to think that in situations like this, intentional actions such as place-base policies are justified and may lead to significant effects by supporting the best exploitation of untapped, immobile resources where they exist.

It is therefore necessary to examine whether and how realistic it is to assume that there are untapped immobile resources, taking into account the strong objection that comes from most of the literature on economic growth which assumes full employment. If, they say, resources are available, sooner or later they will be spontaneously exploited in a market capitalism system.

This subject will be discussed in Chap. 3, which starts by considering an interesting alternative to policies supporting the employment of untapped resources. This consists in policies for turning resources which are normally *mobile* into *immobile* ones.

Appendix: Size and Growth of European Cities

Introduction

> The number of cities in a given territory is always in an inverse geometrical progression of their size (for instance, there are about 23,000 urban agglomerations larger than 10,000 inhabitants in the world, 2,000 are larger than 100,000 inhabitants and about 200 above 1 million, after Moriconi-Ebrard 1993). This persistent scaling behaviour has been questioned for more than one century, giving rise to a large variety of interpretations. (Pumain 2004, 1).

The essay by the German physicist, Felix Auerbach (*Das Gesetz der Bevölkerungskonzentration*), published a century ago, is always cited. He observed that the probability of finding a city larger than x was inversely proportional to x. This kind of distribution is found in many natural and social processes, but Auerbach was the first to apply it to the urban concentration of population. The American linguist George Kingsley Zipf, who discovered the same distribution in word frequency, evidenced this: «The first person to my knowledge to note the rectilinear distribution of communities in a country was Felix Auerbach in 1913» (Zipf 1949–2012, 374). The term rectilinear refers to the fact that, taking cities in order of decreasing size, the second city tends to have half of the population compared to the first, the third one-third, the fourth one-fourth and so on. In this way, the logarithms of rank as a function of the logarithms of size, follow a straight line with an angular coefficient equal to minus one.

The studies on this subject have never ceased, and thanks to the wider availability of new statistics, they have become numerous in the late 1990s. There are many tests today that show the applicability of this distribution, or its variants. There are also different interpretative proposals concerning the mechanisms that may lie behind it, as shown in the Berry and Okulicz-Kozaryn (2012) review.

As to Zipf's distribution applicability to different contexts and different samples of cities, a specific problem has always hindered the easiest exercise, which could be the regression of the rank logarithm to the size logarithm, in order to verify whether the estimated coefficient is significantly different from -1. The reason is fairly obvious, underlined a long time ago (Quandt 1964; Rapoport 1978) and mentioned on several occasions (Gabaix and Ioannides 2004; Gan et al. 2006). The variable rank is not the result of independent observations but it is obtained by ordering variable size. It is endogenous. Estimates are therefore biased and inconsistent. This was tested by using Monte Carlo simulations. With generated random numbers and applied econometric rank-size estimates, there are significant results in favor of Zipf-type distribution, while this should not happen for construction (Ganet al. 2006, 259, 262). The problem is serious and the different solutions offered (including the remarkable one Gabaix and Ibragimov (2011) proposed), confirmed Urzua's judgment (2000, p. 260): «strictly speaking, Zipf's law cannot hold except for a certain sample size». Generally, after a great deal of theoretical-methodological and empirical work, a substantial consensus emerges (Malevergne et al. 2009) on expected distributions. For large samples of cities, we can expect

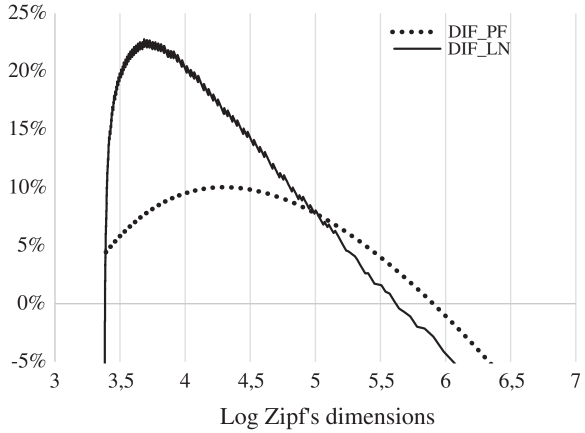


Fig. 2.13 Percentage deviations, respect Zipf's log of cities' size, of that corresponding to lognormal (DIF_LN) and parabolic fractal (DIP_PF) distributions. Artificial simulation 2,000 cases

(even in different contexts like the U.S., Europe, and emerging countries) a lognormal distribution in the range of medium-small cities. A Zipf distribution should prevail among large cities. A parabolic fractal[23] distribution is expected in the middle (see also Giesen and Suedekum 2012; Rybski et al. 2013). To illustrate the distinctive features of these distributions, Figure 2.13 (based on an artificial sample of 2,000 cases) shows the percentage deviations, with respect to the linear Zipf, of the logarithms of the lognormal and parabolic fractal distribution size.

The lognormal distribution presents deviations from Zipf's, which take the shape of a heightened bell displaced to the left, while the deviations of the parabolic fractal are lower.

A second problematic aspect concerns the definition of a city. It seems now undisputed (Rozenfeld et al. 2011; Veneri 2013) that the results of the analysis concerning the shapes of the distributions also depend on the criteria adopted to identify what a city is. It often emerges that the indicated rank-size rule is best applied to urban realities defined by substantive criteria rather than administrative boundaries. It is therefore appropriate to consider temporal data over not too long periods relying on statistical data based on administrative boundaries. In this way, it is at least possible to conjecture a constant error.

[23] This term indicates the discrete probability distribution in which the logarithm of the frequency of a dimension of an entity is a quadratic polynomial of the logarithm of the rank. It is said fractal with reference to the fact that the distribution of "fractal objects" (Benoit Mandelbrot) takes this form. Fractal objects have forms that repeat similar at different scales (self-similarity). Note that the Zipf distribution is also indicated as fractal linear and that the lognormal distribution is approximated by a polynomial of third degree in the logarithms of the size and rank. You could then say that the sequence for the city of decreasing size is of fractal distributions of first then of second then of third degree.

Finally, concerning the mechanisms that can lead to these distributions, Gabaix (1999) and Malevergne et al. (2009) showed that they may be the result of the long term growth rates, independent of size (Gibrat law).

In the following section, we will introduce an exercise on data concerning 352 European cities. We will see that the conjecture concerning distribution forms is verified, indicating that in the past city growth rates were independent of size. We will then see that in recent times the growth rates, by contrast, appear to increase—in a nonlinear manner—in relation to size. This requires a specific interpretation which we will attempt to provide.

Data

The explanations for city growth usually evoke mechanisms that affect productivity (and hence the degree of utilization of resources, in an open economy): economies and diseconomies of scale, human capital, innovation, public goods. It seems therefore appropriate to consider employment, rather than population. Employment size and growth are more immediately linked to productivity and its effects on the use of resources. Eurostat (with additions[24]) provides, for a total of 352 European cities,[25] the average employment levels for the periods 1999–2002, 2003–2006, 2010–2012. The city sample (determined by the availability of data but fairly representative also of medium-small size cities) appears as shown in Table 2.6. From the same source, we have data on area size and population, and we can calculate the population density relative to 2012. Finally, with Google Maps, we have, obtained distances (in minutes of travel by car) between the towns of various countries and their capitals. Size, growth, density, distances are, in fact, considered essential in the recent debate on the analysis of urban area evolution.

[24] Some data not available from Eurostat were obtained from the National Institutes of Statistics as in the case of Bulgaria and Latvia (1999–2002), Spain, Latvia, the United Kingdom (2010–2012). All data concerning the Italian cities are derived from Istat. Those for the 1999–2002 period are taken from the 2001 Census of Population. As to the following periods, they are calculated using 2001 data, rates of change in employment in the local labor systems (SSL), considering cities in the center of SSL with more than 100 thousand inhabitants in 2001, a sample of 74 Italian cities comparable to those of other countries.

[25] The cities in the Urban Audit database are more than 900, but for many the data for this exercise are missing.

Table 2.6 European cities in the sample

Country	Number of cities in the sample	Percentage of employment in the sample on national employment 1999_2002	Percentage of employment in the sample on national employment 2010_2012	Smallest city (thousand employed 2012)	Larges city (thousand employed 2012)
Belgium	7	30.5	34.2	63.4 (Namur)	675.3 (Brussels)
Bulgaria	8	43.0	53.0	14.4 (Vidin)	832.0 (Sofia)
Germany	88	41.0	41.3	30.7 (Weimar)	1,636.1 (Berlin)
Spain	26	29.7	29.0	51.4 (Badajoz)	1,956.2 (Madrid)
France	31	19.6	21.1	28.2 (Creil)	1,797.7 (Paris)
Latvia	2	52.7	52.0	30.9 (Liepaja)	392.9 (Riga)
Italy	74	26.8	27.1	14.0 (Nocera Inf.)	1,196.2 (Roma)
Lithuania	3	31.5	36.7	39.9 (Panevezys)	256.3 (Vilnius)
Hungary	7	34.2	26.0	32.7 (Nyíregyháza)	719.0 (Budapest)
Netherlands	15	29.0	31.7	57.8 (Heerlen)	542.0 (Amsterdam)
Poland	28	39.0	38.0	11.4 (Zory)	820.0 (Warszawa)
Slovenia	2	29.9	32.9	57.5 (Maribor)	204.9 (Ljubljana)
Slovakia	8	35.2	36.0	39.0 (Trencin)	358.1 (Bratislava)
Sweden	9	49.5	52.2	59.8 (Umeå)	912.1 (Stockholm)
United Kingdom	28	25.4	27.6	26.7 (Gravesham)	3,432.1 (London)
Norge	6	33.9	34.1	39.7 (Tromsø)	435.9 (Oslo)
Switzerland	10	28.2	27.2	33.8 (Biel)	359.9 (Zurich)
Totale	352	35.5	36.4	11.4 (Zory)	3,432.1 (London)

Source Table obtained by processing data as indicated in note 50

The Exercise

Let us consider first (Fig. 2.14) the differences between the logarithms of the actual city size and those that would occur applying a Zipf distribution (taking Gabaix-Ibragimov's correction).

As expected, the distribution is lognormal with transition to the parabolic fractal starting at about 100,000 employees, while the tail of large cities drops rapidly to zero indicating that, in this range, size distribution becomes Zipf-type. From 1999–2002 to 2010–2012, the shift of distribution to the right and down is evident. This displacement will be in part a statistical result, owing to the fact that the sample of 352 cities is constant. New cities do not enter in this sample in the period considered, producing an increase in average size. We cannot exclude, however, that this may also depend on a change of the relative weight structure, regardless of this statistic effect. In fact, there is evidence for this when we consider frequency by size groups. The percentage frequency of the cities with less than 40,000 employed people decreases (from 18 to 15 %). However, the percentage frequency of the cities between 40,000 and 250,000 employees does not increase and stays at 69 %, while the share of the largest cities increases from 13 to 16 %. There is confirmation and explanation for this change considering ratios between 2010–2012 and 1999–2002 cumulative employment (Fig. 2.15).

Figure 2.15 shows that cities in the size class 40–250,000 employed (logarithms from 4.6 to 5.4) had the lowest gains comparing the actual data with a linear interpolation. Smaller cities had higher increases and large cities even more so.

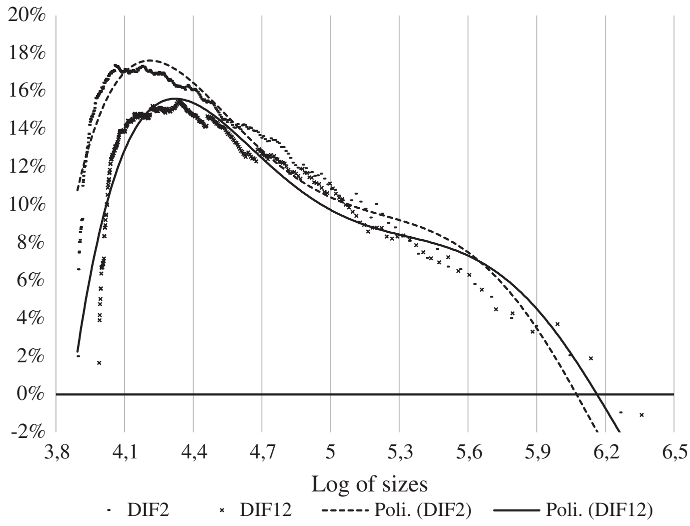


Fig. 2.14 Percentage deviations, respect Zipf's log of cities' size, of actual log sizes: DIF2 = 1999_2002, DIF12 = 2010_2012. *Source* Figure obtained by processing data as indicated in note 50

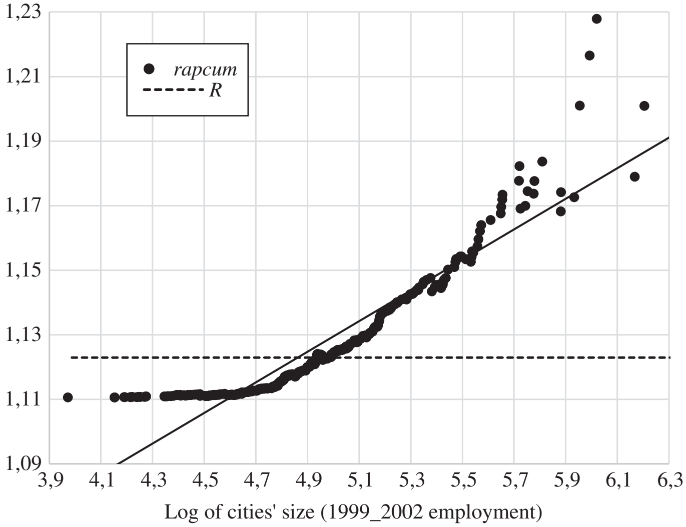


Fig. 2.15 Ratios between cumulative employment 2010_2012 on 1999_2002: *rapcum* actual figures; *R* simulation assuming proportional growth (Gibrat's law). *Source* Figure obtained by processing data as indicated in note 50

The interpolating line indicates employment growth of individual cities regularly increasing with the size-related logarithm. If we make a comparison considering constant growth rates independent of size (R line), however, the transition between the upper and lower growth is around the city size of 100,000 employees.

Summing up: (1) The relationship between growth and size in the ten-year period considered deviates from Gibrat's law, which probably operated in the long run in the past. (2) One part of the cities grows less and one part more than a constant growth rate; the dividing threshold is around the size of 100,000 employed. (3) A recent positive correlation between growth and size is not linear and the cities that are growing relatively less are those between 40 and 250,000 employed.

These results, however, may arise from a spurious correlation between size and growth. Employment could have increased more, independently of city size, in countries where the sample share of small and large cities is greater. While we cannot completely exclude this possible distortion using the available data, we can limit it considering the growth rates of cities net of countries' growth rate and other possible effects. An equation is therefore estimated and the results are shown in Table 2.7.

Some country dummies were significant, indicating city employment growth rates larger than the overall nationwide growth (in Belgium, France and Italy), and smaller instead in Hungary and Spain. Distance from the capital and population density (in logarithms) are significantly negative, while positive distance divided by the maximum distance is not very significant. The negative effect of population density probably depends on congestion costs, intended in the broadest sense, which detract from the benefits of agglomeration. The negative effects of distance from the capital, while positive in terms of relative distance, are more difficult to

Table 2.7 Regression results

Dependent variable: yearly cities' rate of growth of employment 1999_2002 to 2010_2012 net of corresponding national growth			
Independent variables	Estimated coefficient	*t*	Standard error
Constant	3.3	3.9	0.8
30–40 thousand employed	0.7	2.5	0.3
40–250 thousand employed	0.4	1.8	0.2
250–400 thousand employed	0.9	2.8	0.3
More than 400 thousand employed	1.0	2.7	0.3
Belgium	1.2	2.6	0.5
Spain	−1.3	−5.4	0.2
France	0.8	3.5	0.2
Hungary	−3.9	−8.5	0.5
Italy	0.4	2.0	0.2
Log(DIST)	−0.6	−3.8	0.1
Log(DEN)	−0.2	−3.0	0.1
DR	0.6	1.9	0.3
Adj. R2	0.33		
SE regression	1.1		
SD dependent	1.4		

DEN Population density 2012, inhabitants for square Km; *DIST* Distance from the city capital, minutes by car; *DR* DIST divided the maximum of it

interpret. Distance probably captures an adverse effect, related to the physical size of the country, as the cities grow more in small countries. Relative distance, however, captures a positive effect as the net result of conflicting mechanisms, with the prevalence of dispersion over concentration. The non-linear trend of the relative rates of growth of employment in relation to the size of the city seems to be confirmed. The estimated value of dimensional dummy coefficients, in the range of 40–250,000 employed, is the smallest. The coefficient is greater in the range of 30–40,000, and over 250,000. The standard estimate errors, however, are high enough to make these differences not very significant. Nevertheless, the significance of the non-linear trend is clear considering the residuals. We computed, for size ranges, averages of residue growth rates after removing effects of variables not related to the size: country dummies, population density, and distance. We obtained the following residues which measure cities' annual relative growth rates. These residues depend on the size of the cities, on other unknown variables, and on a random component. They are: 1.6 % (30–40,000 employed), 1.5 % (from 40 to 250,000), 2.0 % (from 250 to 400,000), 2.3 % (over 400,000). Finally, we compared actual data and simulated data by means of these residues as ratios of cumulative employment in 2010–2012 and 1999–2002. The non-linearity in the relation between growth and size is now less strong but not canceled. The growth of large cities is also reduced. An increasing relationship between growth and size of cities remains however (Fig. 2.16).

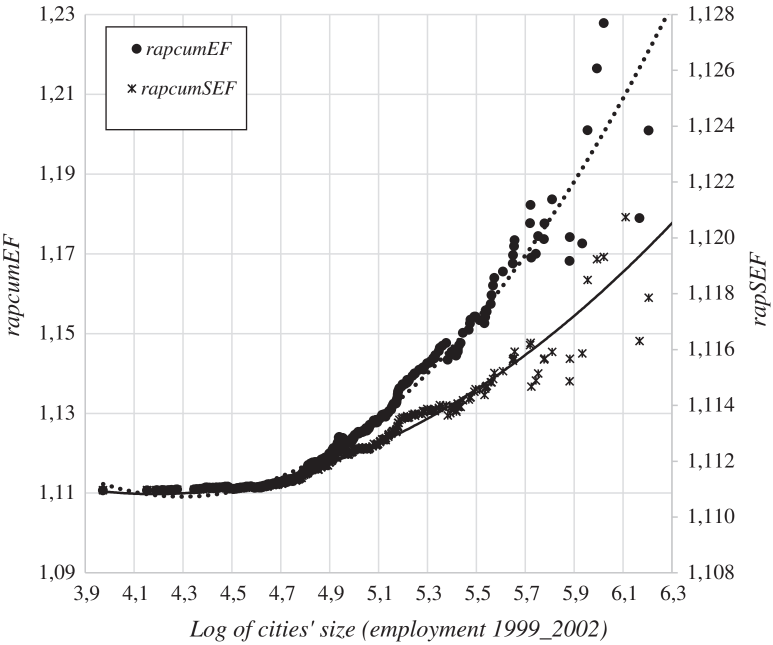


Fig. 2.16 Ratio 2010_2012 to 1999_2002 cumulative employment: *rapcumEF* actual data; *rapcumSEF* simulated data depending on cities' sizes. *Source* Figure obtained by processing data as indicated in note 50

Suggested Interpretation

The exercise shows that, during the last 10 years in Europe, medium-sized cities appear to have grown less than in the past, while larger and, to a lesser extent, smaller cities have grown relatively more. It is not easy to give an interpretation for this intervened non-linear relation between size and growth.

The Acemoglu and Zilibotti's growth model (1997) suggests considering it a possible result of an increase in uncertainty, which in fact appears to have taken place, fueled by intensified international competition. In Fig. 2.17, we represent unemployment rates as moving averages of three four-month terms from 1970 to 2013. Since the beginning of the 1990s unemployment has not only remained high (having progressively increased since the 1970s), but it has also become more unstable.

The Acemoglu and Zilibotti model seems suitable to explain how this increased instability and uncertainty might have penalized growth in medium-sized cities. The microfounded model contains several assumptions, the most important of which are the following three: (1) Savings are used by agents to make investments, which are classified into two types: risky investments with greater expected earnings, and safe investments with lower returns. (2) Different projects madepossible by risky investments are imperfectly correlated so that agents have the possibility of reducing the

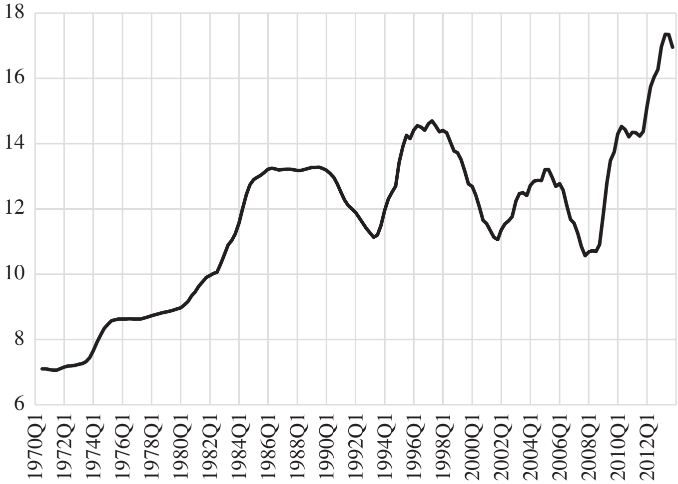


Fig. 2.17 Moving three terms averages of unemployment percentages on work force (Europe 12). *Source* Figure obtained by processing data from Drèze e Bean (1990) and from Eurostat data base

risk through portfolio diversification. (3) The allocation problem is not trivial (it would be so if all agents could invest in all projects by diversifying most of the risks) as there is a minimum investment threshold for each project which gives rise to a trade-off between expected returns and risk insurance.

The authors suggest applying the model to national economies. The reason is that the considered economic system must have boundaries. However, the model can also be applied to a city if it is admitted that the city constitutes a system with limits. This does not exclude the city having large external relations. It implies that the *considered* possibilities for agents to act for potential projects are internal. An external wider system of relations may exist, but we assume that the system is constant while internal activities change.

Figure 2.18 shows the dynamics of accumulation of productive capital in the city under uncertainty conditions (standardized variables).

G(Kt + 1) and B(K + 1) indicate the stock of capital at time t + 1 in the case of "good news" and "bad news", respectively. At initial accumulation level (small town) the ease of finding small-scale, not too risky projects allows growth in every case, with both good and bad luck. The two curves are both, in fact, to the left of the bisector. However, as the process goes on, in the case of bad news, growth stops (in the figure to the level Kt = Kt + 1 = 0.11). It remains positive in the case of good news. Therefore, the system may remain, for the initial phase of its growth at least, subject to a condition oscillating between stop and go. Such a condition would be overcome and growth could restart without interruption over a certain threshold (in the figure above Kt = Kt +1 = 0.23), when the curve indicates that the effect of bad luck goes to zero and when the G curve is always to the left of the bisector.

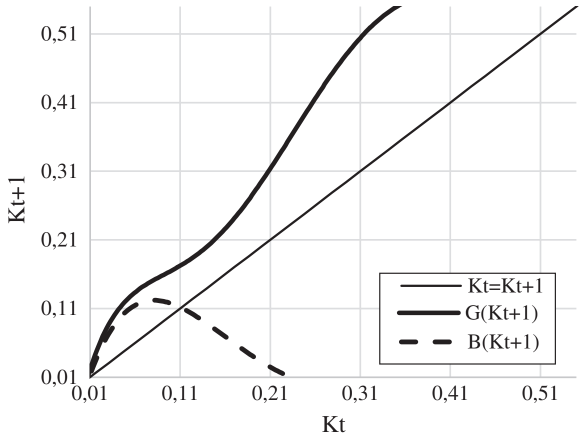


Fig. 2.18 Economic growth under condition of uncertainty

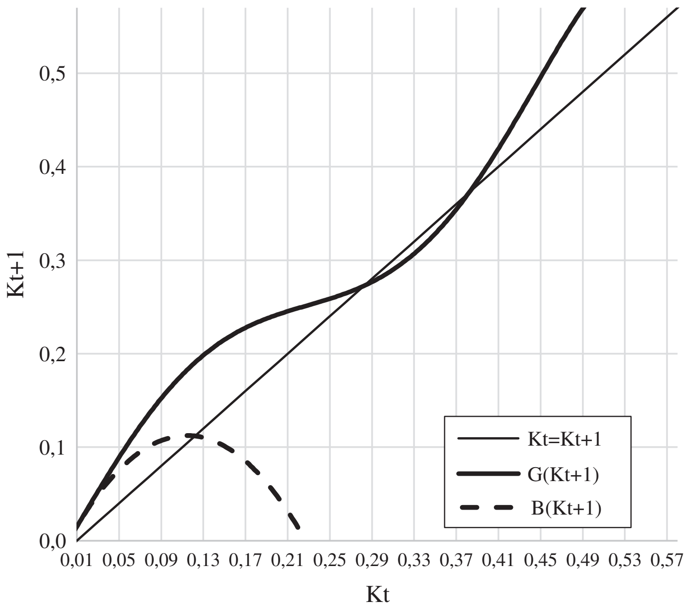


Fig. 2.19 Economic growth with uncertainty and reduced difference between risky and safe projects' returns

This result, for a given risk aversion [which determines the saddle shape of the curve G(Kt +1)] depends on the *difference* between the returns on risky projects and safe projects. Reducing this parameter by half, we obtain Fig. 2.19, where circumstances are very different.

Now a condition exists (Kt between 0.29 and 0.37) in which there is no growth, as the curve G(Kt + 1) is not on the left of the bisector. Here the city is fairly big and take-off difficulties have already been surmounted. These difficulties of medium-size cities are born because of a low yield spread between safe and risky assets owing to an increase of all risks resulting from the increased instability and uncertainty. In a medium-size city, this lower return differential does not find a balance in risk distribution, while in larger cities it is guaranteed by a broad diversification of assets. In a medium-size city, furthermore, lower return differential is not offset by the lower cost of the projects as in the case of small cities.

References

Acemoglu D, Zilibotti F (1997) Was prometheus unbound by chance? Risk, diversification and growth. J Polit Econ 105(4):709–751

APO (2004) Total factor productivity growth: survey report. Asian Productivity Organization, Tokyo

Babbage C (1832) On the economy of machinery and manufactures. Charles Knight, Pall Mall East London

Baier SL, Bergstrand JH (2001) The growth of world trade: tariffs, transport costs, and income similarity. J Int Econ 53:1–27

Bartkowska M, Riedl A (2012) Regional convergence clubs in Europe: identification and conditioning factors. Econ Model 29(1):22–31

Beckerman W (1962) Projecting Europe's growth. Econ J 72(288):912–925

Berry BJL, Okulicz-Kozaryn A (2012) The city size distribution debate: resolution for US urban regions and megalopolitan areas. Cities 29(1):S17–S23

Bickenbach F, Bode E (2013) New economic geography and reunified Germany at twenty: a fruitful match? Spat Econ Anal 8(2):120–153

Blaug M (1964) A case of Emperor's clothes: Perroux' theories of economic domination. Kyklos 17(4):551–564

Bosker M, Brakman S, Garretsen H, Schramm M (2010) Adding geography to the new economic geography: bridging the gap between theory and empirics. J Econ Geogr 10(6):793–823

Brülhart M (1998) Economic geography, industry location and trade: the evidence. University of Manchester, mimeo for The World Economy Special Issue

Caselli F (2005) Accounting for cross-country income differences. In: Aghion P, Durlauf S (eds) Handbook of economic growth. North Holland—Elsevier, Amsterdam, London, pp 679–741

Danquah M, Moral-Benito E, Ouattara B (2011) TFP growth and its determinants: nonparametrics and model averaging. Banco de Espagna Working Paper 1104

Desmet K, Fafchamps M (2005) Changes in the spatial concentration of employment across US counties: a sectoral analysis 1972–2000. J Econ Geogr 5(3):261–284

Dixit AK, Stiglitz JE (1977) Monopolistic competition and optimum product diversity. Am Econ Rev 67:297–308

Doran J, Jordan D (2013) Decomposing European NUTS2 regional inequality from 1980 to 2009 national and European policy implications. J Econ Stud 40(1):22–38

Drèze J, Bean C (1990) Europe's employment problem: Introduction and synthesis. CORE Discussion papers 1990041, Université Catholique de Louvain, Center for Operations Research and Econometrics

Duranton G (2006) Some foundations for Zipf's law: product proliferation and local spillovers. Reg Sci Urban Econ 36(4):542–563

Duranton G, Puga D (2013) The growth of cities. CEPR Discussion Papers 9590

Edwards BK, Starr RM (1987) A note on indivisibilities, specialization, and economies of scale. Am Econ Rev 77(1):192–194

Eichner AS, Kregel JA (1975) An essay on Post-Keynesian Theory: a new paradigm in economics. J Econ Lit 13(4):1293–1314

Feenstra R (1998) Integration trade and disintegration of production in the global economy. J Econ Perspect 12(4):31–50

Feldman MP (2000) Location and innovation: the new economic geography of innovation, spillovers and agglomeration. In: Clark GL, Gertler MS, Feldman MP (eds) The oxford handbook of economic geography. Oxford University Press, New York

Feldman MP, Audretsch D (1999) Innovation in cities: science-based diversity, specialization and localized competition. Eur Econ Rev 43:409–429

Felice E (2009) Estimating regional GDP in Italy (1871–2001): sources, methodology, and results. Universidat Carlos III Madrid, Working papers in Economic History

Fortunato G (1899) Lettera a Pasquale Villari 2 settembre 1899. In: Giustino Fortunato e Emilio Gentile, Carteggio 1865–1911. Laterza, Bari 1978, p 65

Gabaix X (1999) Zipf's law for cities: an explanation. Q J Econ 114(3):739–767

Gabaix X, Ibragimov R (2011) Rank−1/2: a simple way to improve the OLS estimation of tail exponents. J Bus Econ Stat 29(1):24–39

Gabaix X, Ioannides YM (2004) The evolution of city size distributions. In: Henderson JV, Thisse JF (eds) Handbook of regional and urban economics, vol 4., Cities and geographyNorth-Holland, Amsterdam, pp 2341–2378

Gan L, Li D, Song S (2006) Is the Zipf law spurious in explaining city-size distributions? Econ Lett 92:256–262

Geppert K, Stephan A (2008) Regional disparities in the European union: convergence and agglomeration. Papers Reg Sci 87(2):193–217

Giesen K, Suedekum J (2012) The size distribution across all cities: a unifying approach. CESifo Working paper Series 3730

Isaksson A (2007) Determinants of total factor productivity: a literature review. Unido, Vienna

Iuzzolino G, Pellegrini G, Viesti G (2011) Convergence among Italian regions 1861–2011. Banca d'Italia Economic History Working Papers 22

Kaldor N (1934) The equilibrium of the firm. Econ J 44(173):60–76

Kaminsky G, Schmukler S (2002) Short-run pain, long-run gain: the effects of financial liberalization. World Bank Working paper 2912

Klenow P (2006) Income differences across countries. Society of economic dynamics conference, 6 July 2006, Vancouver, Canada

Krugman P (1991) Geography and trade. MIT Press, Cambridge

Krugman P (2010) The new economic geography, now middle-aged. Paper prepared for presentation to the Association of American Geographers, April 16, pp 1–19

Istat (2012) L'Italia in 150 anni. Sommario di statistiche storiche 1861–2010. Roma

Leijonhufvud A (1986) Capitalism and the factory system. In: Langlois RN (ed) Economic as a process: essays in the new institutional economics. Cambridge University Press, New York, pp 203–223

Malevergne Y, Pisarenko V, Sornette D (2009) Gibrat's law for cities: uniformly most powerful unbiased test of the Pareto against the lognormal. Swiss Finance Institute Research Paper Series 09/2009

Marshall A (1890) Principles of economics. An introductory volume. MacMillan, London

Martin R (1999) The new 'geographical turn' in economics: Some critical reflections. Camb J Econ 23(1):65–91

Mayer-Foulkes D (2002) Global divergence. Center for Research and Teaching of Economics (CIDE)—Division of Economics, Mexico

Maynou L, Saez M, Bacaria J, Lopez-Casasnovas G (2014) Health inequalities in the European union: an empirical analysis of the dynamics of regional differences. Eur J Health Econ Line

Meardon SJ (2001) Modelling agglomeration and dispersion in city and country—Gunnar Myrdal, Francois Perroux and the new economic geography. Am J Econ Sociol 60(1):25–57

Mill JS (1848) Principles of political economy with some of their applications to social philosophy, 7th edn. (1909). Longmans, Green and Co, London

Miura T (2005) What has the division of labour brought? A lesson from Charles Babbage. Fukuoka Univ Rev Commercial Sci 50(2–3):119–141

Moriconi-Ebrard F (1993) L'urbanisation du Monde depuis 1950. Editions Anthropos, Paris

Nelson R, Winter SG (1982) An evolutionary theory of economic change. Harvard University Press, Cambridge

OECD (2010) Factbook 2010: economic, Environmental and Social Statistics. OECD, Paris

Olinto P, Saavedra J (2012) An overview of global income inequality trends. World Bank Inequality Focus 1(1):1–11

Ottaviano GIP, Thisse J-F (2002) Integration, agglomeration and the political economics of factor mobility. J Public Econ 83:429–456

Parr JB (1999a) Growth-pole strategies in regional economic planning: a retrospective view part 1. Origins and advocacy. Urban Stud 36(7):1195–1215

Parr JB (1999b) Growth-pole strategies in regional economic planning: a retrospective view part 2. Implementation and outcome. Urban Stud 36(8):1247–1268

Perroux F (1955) Note sur la notion de pôle de croissance. Economique Appliquée 7:20–307

Petersson L (2000) The theory of new economic geography and industrial location in SADC. Paper presented at the XXIX annual conference of the Australian Society of Economists

Phillips PC, Sul D (2007) Transition modelling and econometric convergence tests. Econometrica 75(6):1771–1855

Pumain D (2004) Scaling laws and urban systems. Sfi Working paper 2004-02-002, Santa Fe Institute

Quandt RE (1964) Statistical discrimination among alternative hypotheses and some economic regularities. J Reg Sci 5:1–23

Quinn D, Schindler M, Toyoda AM (2011) Assessing measures of financial openness and integration. IMF Econ Rev 59(3):488–522

Rapoport A (1978) Rank-size relations. In: Kruskal H, Tanur JM (eds) International encyclopedia of statistics, vol 2. Free Press, New York, pp 847–854

Romer PM (1990a) Endogenous technological-change. J Polit Econ 98(5):S71–S102

Romer PM (1990b) Are nonconvexities important for understanding growth? Working paper 3271, NBER working paper series

Rosenberg N (1982) Inside the black box: technology and economics. Cambridge University Press, New York

Rozenfeld HD, Rybski D, Gabaix X, Makse HA (2011) The area and population of cities: new insights from a different perspective on cities. Am Econ Rev 101(5):2205–2225

Rybski D, Garcìa Cantù Ros A, Kropp JP (2013) Distance weighted city growth. Phys Rev. doi:10.1103/PhysRevE.87.042114

Smith A (1776) An inquiry into the nature and causes of the wealth of nations. Strahan W, Cadell T, London

Starrett D (1978) Market allocations of location choice in a model with free mobility. J Econ Theory 17:21–37

Urzua CM (2000) A simple and efficient test for Zipf's law. Econ Lett 66:257–260

Venables AJ (1996) Equilibrium locations of vertically linked industries. Int Econ Rev 37:341–359

Veneri P (2013) On city size distribution: evidence from OECD functional urban areas. OECD regional development working papers 2013/27, OECD Publishing

Von Hippel E (1994) "Sticky information" and the locus of problem solving: implications for innovation. Manage Sci 40(4):429–439

Williamson OE (1975) Markets and hierarchies: analysis and antitrust implications. The Free Press, New York

Young AA (1928) Increasing returns and economic progress. Econ J 38(152):527–542

Zipf GK (1949–2012) Human behavior and the principle of least effort: an introduction to human ecology. Reprint of 1949 edition, Martino Publishing, Manfield CT

http://www.springer.com/978-3-319-15376-6

An Introduction to Place-Based Development
Economics and Policy
Seravalli, G.
2015, X, 144 p. 46 illus., Hardcover
ISBN: 978-3-319-15376-6