Neither easy nor impossible: Local development economics and policy

Seravalli, Gilberto

University of Parma (Italy)

10 August 2011

NEITHER EASY NOR IMPOSSIBLE:

LOCAL DEVELOPMENT ECONOMICS AND POLICY

By

Gilberto Seravalli

Translation from the Italian to English by

Clarissa Botsford

Introduction

This assai leads the reader on an analytical path to development, with institutions as its central focus. While its main emphasis is on the problematic nature of development in lagging or backward regions, contributions from the literature on developing countries have also been taken into account, and European countries, Italy and the South of Italy (the Mezzogiorno) also feature in the debate. There is no doubt that lagging regions in advanced countries are very different from backward regions in poor countries. I think that there are two basic assumptions that hold for both. First, the most important constraints on development are internal; second, in order to combat underdevelopment strong actions that break with tradition are needed. These two assumptions are often ignored by development literature. These studies may be well written, fun to read and able to communicate faithfully their authors views on social and economic life, but their perspective is necessarily different from that of people who are actively involved in social and economic development efforts or tackling underdevelopment. The assai offers some interpretative starting points. The path distances itself from five concepts that are widely held but superficial.

The first is that economic development in a backward or lagging area depends on growth in advanced areas, given that capital will flow where there is an abundant supply of labour and because technical progress is easier for latecomers than for innovators. The common view is that there is spontaneous economic convergence simply because capital has low returns in advanced areas and high returns in backward areas. In advanced areas, moreover, already on the threshold of improved technology, technical progress is more of a struggle, whereas in backward areas all that is needed is to imitate others. Since this idea has no real empirical basis, those who hold this view claim that the differences between regions persist because markets work badly, segmented as they are by rules and behaviour that allow unjustified returns. Without these artificial distortions, they argue, resources would be spontaneously used where they were most productive. In this view, the most relevant development policy (and for many, the only possible policy) is to reduce curbs to maximum competition in all sectors and in all walks of life.

There are many reasons for thinking that this is irrelevant. Many crucial resources are in fact attracted to advanced areas, not because there is over-regulation. If

anything, the opposite. Opening up markets that work well would therefore seem to be a result of development rather than a way to achieve it. In order to achieve development what is needed is more intentional public intervention: more State, not less.

The path undertaken to reach this conclusion is arduous, however. Not many people today share the view that the State can play a relevant role. The path requires a critical reappraisal of four other common views. The second widespread view that we need to demolish, then, is the idea that -. even if we accept that economic development in a backward area is impeded by the fact that resources are attracted to advanced areas - there will always be immobile resources that can spontaneously trigger development and sustain it. This view is supported by the limits on mobility of resources. Taking into account that important resources such as the environmental, historical, cultural heritage of a region are substantially immobile, those who hold this view believe that these resources cannot be lost as a result of the capacity for attraction of an advanced area. They believe that these immobile resources sooner or later will stimulate those who know how to exploit them effectively and profitably. The reason why this view is unhelpful is that existing immobile resources can and often do remain inaccessible. Potentially profitable resources never actually become useful for development because the conditions for exploiting them have not been achieved, or nobody has ever thought of using them. A case study illustrating how this can happen is presented in the Appendix of Chapter Three.

Proceeding along our path, a third commonly held view must be set aside. Even if they accept that the process of valorising immobile resources is not always so easy, economists often say that this is not a problem if in the backward area the cost of labour is very low. The objection is that the difficulty of valorising immobile resources is made up for by a particularly advantageous condition of low labour costs. If there were no institutions keeping wages artificially high, unemployment or under-employment would be solved by salaries so low that the labour cost per production unit would be advantageous even in situations of low productivity. This, people claim, would help overcome other curbs on development. This approach is on the surface quite correct, but it neglects to take into account one feature. Low labour costs could well get an area out of underdevelopment, but there are no guarantees that if will keep that area out.

It is thus necessary to deal with the fourth widely help view. Admitting, at this point, that underdevelopment in a backward area cannot be overcome spontaneously, even where there are low labour costs, and accepting that development needs an action that breaks with tradition, then, it is claimed, bottom-up collective action on the part of civil society can provide precisely this break. Those who hold this view believe that in order to trigger and sustain a development process collective goods such as, for example, material and immaterial infrastructures, insurance, training, and promotional services, are essential. They do not feel, however, that corresponding public action is needed. They are convinced that, collectively, civil society can achieve the goods and services needed. This idea ignores the serious difficulties that bottom-up collective action meets with if it is not supported by public action, as Chapter Three illustrates.

This brings us to the fifth and final point of view that we need to deal with. Admitting that bottom-up collective action is unsuccessful, a break with tradition can be guaranteed simply by decentralising institutions, providing local authorities with the sufficient skills and resources to take care of their own needs. We argue that decentralizing on its own is not a solution. Development policies require State intervention to be organized in a complex institutional framework, as Chapter Five shows.

The assai's five chapters are each devoted to one of these themes, while the sixth provides a conclusion and completes the arguments set out in the Introduction. The themes are: territorial imbalances, underdevelopment traps and social traps, collective goods and services, institutional action and institutional frameworks. Every theme is supported with an analytical argument and development policy proposals that emerge as a result.

In Chapter One, territorial imbalances are explained. Space is an obstacle, but not an absolute barrier to the free movement of goods, people and firms. All that is needed, then, is for a process of economic growth to be triggered, after which it becomes cumulative and attracts activities and resources from other places. Backward areas not only do not receive new resources, they lose the few resources they have, remaining underdeveloped even when at a national level there is growth. New Economic Geography (NEG) illustrates these findings well by means of models of general economic equilibrium. The approach linked to traditional regional economics

based on partial economic equilibrium did not do enough to explain these cumulative processes. Individual agglomeration factors were identified without placing them in context. On the other hand, the findings of NEG imply that general equilibrium is inefficient when territorial development gaps are not temporary. If it is not efficient, how are these gaps to be closed? NEG makes one suggestion, based on its own models. In backward areas there are usually unexploited immobile resources that could be put to use in order to trigger development. The literature on industrial districts gives an interesting indication of the nature of these immobile resources and what mechanisms translate into factors of development. These include professional and organisational resources, which in other areas are mobile but which in these districts take on a special cohesion (or functional coherence), to the extent that they adopt the same features as immobile resources in the sense that if they were used elsewhere they would not be as productive.

The first part of Chapter Two considers underdevelopment traps. While reallocating resources from less productive to more productive uses may trigger a process of growth, the related costs may be so high that growth is unlikely to take place – either spontaneously or gradually – because a block of a significant size would have to be reallocated at the same time. This argument can be countered by the observation that low labour costs compensate for these costs. If labour costs are low owing to unemployment, then total reallocation costs are not very high. Reallocation takes place gradually, without any thresholds and without any underdevelopment traps. But this means ignoring the problem of social traps, which are examined in the second half of Chapter Two. Low labour costs may facilitate the exit of resources from less productive uses but in the end they will stop them from staying in more productive uses. And yet, low labour costs imply an absence of barriers to entry for more productive uses, with the result of rapidly exhausting existing potential by impeding sufficient accumulation. Escaping from combined underdevelopment and social traps requires actions that markets are unable to support, even taking into account opportunities for exploiting existing resources and the low cost of labour. What is needed is to allow labour costs to rise and thus overcome reallocation costs that have thus become so significant. What is needed are collective goods, which provide an injection of benefits such that people and organisations are able to take on the reallocation costs of moving to modern and more productive market sectors, and deal

with the risks and difficulties inherent in transformation so that they can distance themselves from the static but guaranteed structure of backward societies.

Chapter Three is devoted to the theme of collective goods in underdeveloped contexts. The first part of the chapter examines the possibility that collective goods are realised by those who need them, and concludes that this only happens under very strict conditions, which do not apply to backward areas. The main obstacle is the lack of information, which is a problem even in advanced economies but it is exacerbated in situations of underdevelopment. The inevitable result is that collective goods need to be public goods. Institutions that contribute to development by establishing public goods, not just by enforcing regulations, thus become central to any analysis of how underdeveloped areas can trigger the growth process. The chapter argues, however, that these active institutions cannot be considered a solution, even assuming that they are virtuous and not predatory. If imperfect information and uncertainty call for institutional intervention, this intervention will in its turn be a victim. In particular, as the chapter goes on to show, a central State that adopts vertical and sectorial interventions is especially inefficient.

In Chapter Four the question of decentralisation is analysed, taking as given that institutions should be more accountable and closer to the citizenry. Applying the criterion of subsidiarity, it would be easy to believe that the more resources local institutions are endowed with, the more effective their actions will be for development, and the better they will be able to sustain and support the role of civil society. The widely held belief, in this case, is that by pursuing pluralism and democracy, development will automatically follow. Those who reject the orthodox economic view that all that is needed for development to take off is a functioning market often adopt this paradigm. And yet, as Chapter Four demonstrates, in conditions of uncertainty, the proximity of local institutions to specific interest groups generates a problem of consensus and increases the complexity of the institutional action needed. The result is that institutional action, which is anyway conditioned by the history, skills and experience of its managers, is undetermined and leaves plenty of space for 'institutional entrepreneurship', which may or may not be successful.

Chapter Five is based on the idea that institutional entrepreneurship can be encouraged, in a decentralised context, by a multi-level organisation of the State, where every level (multinational, national, regional, and local) interacts with every other. Local actions that break with tradition, as well as innovation, require skilled

personnel and functions that allow local systems to communicate constantly with the outside world in order to obtain additional resources and grasp new opportunities. This requires a complex institutional framework, which on one hand is essential for gaining resources, but on the other generates learning costs that can easily lead to actions that are ineffective for development. The institutional framework also provides a solution, however, if certain conditions are fulfilled. These are: carefully designed rules, a realistic expectation of results, continuous monitoring and evaluation. This path is very different from pure and simple decentralisation because it implies a strong input from the centre, and an emphasis on inter-institutional coordination rather than bottom-up policies per se.

The assai has various readers in mind. It is aimed at students in further education, and in particular in graduate programmes. But it is also intended for people actively involved in local development.

Chapter One - Territorial disparities[1]

Traditional growth economics is based on the principle that differences in per capita income among countries or among regions of the same country depend mainly on differences in capital endowment per worker. And yet, in a context of open economies, with mobile production factors, and in particular mobile capital, neoclassical growth economics has not been able to explain why these per capita differences live on, or even degenerate. In the neoclassical approach, income gaps lead to even gaps in the rate of returns on capital, dictating that large flows of capital should be sent from more advanced areas to more backward areas in order rapidly to reduce these disparities. With a production function in which labour and capital are substitutes, even when substitution elasticity is equal to one[2], and returns of scale are constant, the rate of return on capital of a backward area whose per capita income is a tenth[3] of an advanced area would be 215 times greater[4]. The rate of return on capital income would be double in an area whose per capita income is only slightly lower (80%) than the more advanced area.

Evidently, the principles of neoclassical economics, especially constant returns of scale and decreasing returns of capital, do not reflect real life. In more recent growth economics, these two factors have been removed and new production factors have been introduced, in particular human capital. Thus up to 60% of per capita income differences can now be explained, whereas the traditional model only justified 25% (Lasagni, 2006; Psacharopoulos-Patrinos, 2002; Caselli, 2004). Regional economics[5] is better suited to our purposes because it focuses on the relationships between open economies at a territorial level. Before Krugman's contribution (1991), regional economics did not have a convincing answer to the question of why production tends to agglomerate in certain areas, leaving other areas behind and preserving disparities

---

[1] Sections 1,2 and part of 5 were written with Mario Menegatti
[2] With elasticity less than one, which is more realistic to assume, the results would be even more surprising.
[3] The poorest country in a sample of 78 nations had an average in the period between 1990-2000 of 3% of the per capita income of the richest country with purchasing power parity (Lasagni 2006). From 1995-2000 in Italy, Calabria – the poorest region – had an average of 44% of the per capita income of the richest region, Emilia-Romagna. Bank of Italy data.
[4] With a share of capital income of 30% and thus a labour income of 70%.
[5] The appendix on European territorial imbalances contains further considerations on economic theory concerning growth in open economic systems.

in regional development.[6] Partial equilibrium analyses identified a number of agglomeration factors. There were not only the natural advantages of certain regions to take into account - greater natural resources, historical concentrations of populations, favourable geographic positions – but many other factors. These included Marshall's externalities,[7] Hirschman's linkages in economic activities,[8] cumulative processes of increasing returns of scale, a series of divergent forces leading to dispersion, such as the various categories of congestion costs (Christaller, 1933; Henderson, 1974), as well as transport costs. On the other hand, general equilibrium analyses were anchored in mainstream international trade theory, which attempted to explain sectorial specialisation rather than development.[9] General equilibrium models yielded indeterminate results for transport costs, non-competitive markets and increasing returns. In these models, restrictive (or unrealistic) hypotheses were often adopted: returns of scale were considered constant, space was either considered insurmountable (for productive factors) or, by contrast, irrelevant (for products).[10] Thus, while regions in partial equilibrium analyses were allowed to be historically, geographically, socially and economically specific - and thus have a variegated profile - these specificities in general equilibrium analyses were reduced essentially to the different relative endowments of production factors. In the standard model, production factors were assumed to be immobile. This made for a rather peculiar vision of history. Before there was any external trade, autarchic regions were explained by their history; their specific history was both the cause and the

---

[6] The statement that "productive activities tend to concentrate in certain areas... thus preserving differences in regional development" should be discussed. In highly industrialised countries, and in the United States in particular, the process of economic development took place at the same time as a marked reduction in regional differences. It is important then to consider what conditions contribute to regional convergence and then see how permanent the disparities are. The Appendix contains a brief contribution with specific reference to Europe.

[7] Marshall, 1920

[8] Hirschman, 1963

[9] The importance assigned to concepts of open economies and international specialisation in international trade theory justified these applications for the simple reason that regions within a country are, by administrative definition and for objective de facto reasons, open economies. The title of one of the most important contributions to international trade theory (Ohlin, 1933: *Interregional and International Trade*) indicated that the application of his theoretical framework was possible for regions as well as internationally.

[10] On a similar basis, general equilibrium analysis showed that international and inter-regional trade could be considered a "perfect substitute" for the mobility of productive factors. The consequence was that powerful economic forces leading to convergence in *levels* of remuneration of productive factors (and per capita income), and therefore to development, were born precisely from the consideration of diverse goods produced, as well as diverse factors and diverse open economies of exchange. This was especially true in regions.

justification for their economic and social differences. Once there was external trade, however, history was no longer relevant. A new route was taking shape where deeply lying forces made those differences disappear.

New Economic Geography (NEG), from Krugman's contribution in 1991 onwards, reconciled general equilibrium theory with the well-known results of partial analyses.[11] In NEG models, space no longer had the simplified dichotic relevance whereby in some markets mobility was perfect and in others it did not exist. NEG assumed, more realistically, that spatial distance always costs something to overcome. The result was a richer and less simplistic definition of regional differences that envisaged imperfect mobility both of products and factors. Agglomeration processes of different production factors and products, and therefore externalities, increasing returns of scale and finally non-competitive markets, could thus take place. At the same time, solutions depended on the equilibrium of economic forces both in terms of concentration and in terms of dispersion.[12]

To sum up, mainstream international trade theory, from Riccardo to Ohlin, tended to interpret exchanges between economies as the result of their production specialisations. These specialisations were explained either on the basis of different technologies, or on the basis of different endowments of factors - such as, for example, natural resources and populations – or again, on the basis of relative advantages in producing different goods. Goods markets were generally characterised by full mobility, while production factor (or resource) markets had zero mobility. It

---

[11] It is easy to see why NEG has at times been considered "old wine in new bottles" (Smutzler, 1999, p.358); but it is also clear why it is important.

[12] "Reduced to its essence, the new economic geography is a theory of the emergence of large agglomerations which relies on increasing returns to scale and transport costs, and emphasizes linkages between firms and suppliers as well as between firms and consumer. The basic story underlying this type of analysis can be sketched as follows. Increasing returns to scale tend to foster geographical concentration of production of each good. When transportation costs play a role, attractive locations for production are those which are close to markets and suppliers, other things being equal. Finally, concentration of production in some location tends to attract the mobile factors of production. Workers have better job and consumption opportunities where production is concentrated. The resulting concentration of the labour force leads to more demand for consumption goods in that location, which makes this region more attractive for producers. Once a region has a high share of production, this pattern is likely to reinforce itself: a so-called second-nature advantage for the dominant region develops, that is, the region becomes attractive for firms because so many other firms already produce there (rather than because of superior resource endowment). In other words, success bread success. Working against these centripetal forces which strengthen agglomerations are centrifugal forces. For instance, concentration of productive activities in one region may drive land rents and housing prices up, and may lead to environmental problems. Also, in immobile factors of productions remain in peripheral areas, firms from the centre may want to move there to serve these areas. The population and production patterns result from a balancing of these centrifugal ad the centripetal forces. (Schmutzler, p. 356).

was still not clear what role transport costs and company location played in markets, although these factors were recognised as being important.[13] Traditionally, international trade theory focused less on agglomerations of production and the effects of reducing transport costs, and more on the international division of labour in terms of specialisation in order to stress its greater economic efficiency compared to autarchic structures.

Starting with Krugman's contribution in the early 1990s, NEG attempted to find reasons for geographic location and for the tendency of industrial production to concentrate in some areas rather than others. NEG's main departure from the literature regarding international (and inter-regional) trade was to attribute importance to the distance between production location and markets in which output is sold and inputs are bought. Firstly, NEG took account of the fact that selling goods in a different location from where they are produced incurs costs related to packaging, preserving, and transporting goods, as well as to the documentation and administration of the contracts required when the buyer has no direct control over the quality, quantity or nature of the product, and the seller does not receive the revenues directly. These "transport costs" mean that from a company's point of view, different locations represent different markets, in the sense that the company provides the same goods at differing costs. Secondly, NEG recognised the fact that there are scale economies and that the structure of markets is not perfectly competitive, meaning that the level of demand for goods produced by one company is influenced by the location of other producers. Thirdly, these assumptions, together with the existence of input-output links between companies, imply that the cost of buying intermediate goods used in production depends on the location of suppliers.

These intuitions allowed NEG to identify the forces at work behind the agglomeration of production in certain areas, as well as to recognise the divergent forces that contribute to dispersion. When forces that induce firms to concentrate in certain areas prevail, a self-reinforcing mechanism - whereby links between production and demand for goods, and between production of final goods and production of intermediate goods, are often forged - leads to further agglomeration.

---

[13] "International trade cannot be understood except in relation to and as part of the general location theory, to which the lack of mobility of goods and factors has equal relevance." Ohlin, 1933, pp.141-2

NEG was able to recognise when this mechanism had been set in motion and at the same time to identify the factors that initiated or hindered the process.

The agglomeration mechanism described by NEG played a significant role in the way local development was analysed for two reasons. First, NEG evidenced that mobile resources tended to migrate towards more developed areas; in under developed regions, NEG concluded, it was preferable to concentrate interventions on immobile resources. Second, it stressed that, once the mechanism was set in motion, the process could become self-reinforcing by attracting new companies. The contribution of NEG to the debate on developing backward regions was by no means definitive, but it did forge new analytical paths. On the one hand, there was the matter of determining how to set the mechanism in motion; on the other, the question of identifying what conditions favour or hinder the mechanism once it has been triggered. The innovative factor in the debate, in contrast to mainstream ideas of economic development, was the idea that economic development in a region is a process that is difficult to set in motion, but that once it has been triggered, it can become self-reinforcing and even accelerate, unless obstacles come in its way, or attrition slowly blocks the mechanism. This idea contrasted with prevalent thinking in mainstream economic development approaches, which considered the country as a whole. The accepted view was that historical circumstances, more than anything else, determined whether local development could be set in motion very easily or with great difficulty. According to this view, once it had started, development proceeded in a regular fashion, without any particular acceleration or any relevant slow downs. The justification for this view was that a local system was more open and inter-dependent. Every local system was susceptible to being hindered or encouraged by other local systems. It could be hindered if it were weak,14 because it would lose resources to others; it could be encouraged as long as it were strong, because the resources of other systems would naturally flow towards it.15 When a 'national' economic system was considered as a whole, the positive or negative effects of inter-dependence in its regions were not taken into account.

---

[14] A weakness at the beginning, as we shall see, can also become a strength if there exists a potential to exploit and if it is effectively exploited. A historical strenght can also become a weakness if, for example, congestion costs become prevalent.

[15] The NEG approach is anyway preferable for national economies in a world that is increasingly inter-dependent.

Sections one and two present a simplified version of the issues examined by NEG, with reference respectively to the structure of models and the consequences of firms' location choices. Section three examines some important extensions. Section four gives two concrete examples, while section five summarises the conclusions.

The basic model: the effects of distance and location

The objectives of NEG's analysis of development can be summarised as follows:

- to identify the effects of distance between production and its buyers and suppliers and examine the consequences of location choices
- to identify the forces that lead to agglomeration of companies in certain areas rather than others, as well as the forces that work against this concentration
- to indicate the factors behind these forces.

In order to examine the effects of distance between production and its markets, it is important to identify the elements that make location economically significant. With this aim in mind, let us simplify the matter and consider two economies (A and B) that produce two goods: one agricultural or 'traditional' product (such as wheat) x, and one industrial or 'modern' product (such as cars), y. We will begin by examining a case in which distance is irrelevant. It can be demonstrated that the distance between production and markets is not relevant if the following assumptions hold.

First assumption. Both goods, wheat and cars, are produced with different technologies but with constant returns of scale. Put simply, a production function can be imagined in which the only input is labour. When there is only one input, the hypothesis of constant returns of scale means that the relationship between output and input is constant. The production functions of wheat (x) and cars (y) in cases like this are:

$$x = \alpha l_x \qquad \text{with} \quad \alpha > 0 \qquad\qquad [1]$$

$$y = \beta l_y \qquad \text{with} \quad \beta > 0 \qquad\qquad [2]$$

where l_x e l_y represent the labour used to produce each good.

Second assumption. Each of the two goods is produced by each of the two economies and they make their choices on the basis of a given market price. In this case production takes place in perfect competition.

Third assumption. The optimal choice for each consumer is made on the basis of personal choices. Put simply, these can be summarised as follows:

$$U = U(x, y) = x^s y^{1-s} \quad \text{with} \quad 0 < s < 1 \qquad [3]$$

This describes the degree of utility associated with each basket of goods *x* and *y* consumed. We also assume that all consumers are identical and that consumers perceive the units produced by different companies, for each of the two goods, as identical.

Fourth assumption. The goods produced in economy A can be consumed by economy B, and vice versa, without incurring costs other than production costs.

If these four assumptions hold, the fact that 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, or on sales prices. In the situation we have just described, economies A and B represent a single market inside which the distance between producers and buyers, and the location of production companies, is irrelevant. It is as if all companies and all consumers were concentrated in a single area.

The analysis conducted by NEG eliminates these assumptions.

First eliminated assumption: the hypothesis that goods produced in one economy can be consumed in the same economy or in another one incurring the same costs. NEG stresses that goods produced in one place can be consumed in another only if they are transported. Some goods are impossible to transport, just as some goods cannot be consumed in a different place from where they are produced (take, for example, services provided by a hospital). If goods cannot be transported, they cannot be ceded to other economies. However, if both goods (agricultural and industrial) were impossible to transport, economies A and B would become closed markets with no trade between them. When goods can be transported there are certain costs, such as packaging, preservation and transport services (fuel, vehicle amortisation, labour costs for transporters, information and contracts). For a firm, every area represents a different market with different sales costs that include transport costs. In our model, with only two economies, this implies that for each company there are two markets: an internal market, where sales costs are the same as production costs, and an external market, where sales costs are equal to the sum of

production and transport costs.[16] It should be stressed that this approach makes it possible to take account of changes in transportation costs over time. Economic historians have always underlined the importance of progress in technology in this field. Conservation has improved drastically; goods are transported faster and with greater security. Vehicles have grown in size, allowing for scale economies in transferring goods, and communications have improved in untold ways. All these factors have greatly reduced the cost per unit of goods transported. As we shall see, this dynamic has played an important role in explaining industrial agglomerations.

Second eliminated assumption: the hypothesis that both goods are produced with constant returns of scale on the basis of production functions [1] and [2]. In NEG models this assumption is maintained as far as agricultural goods are concerned, but it is assumed that the technology used to produce industrial goods is as follows:

$$y = \beta l_y - \gamma \qquad \text{with} \quad \gamma > 0 \qquad\qquad [4]$$

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

$$\frac{y}{l_y} = \beta - \frac{\gamma}{l_y} \qquad\qquad [5]$$

This relationship increases as input increases and as the quantity produced increases. Production thus has increasing returns of scale.

Third eliminated assumption: the hypothesis that the location of firms is irrelevant and that consumers perceive goods produced by different companies as identical. 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:

$$y = \sum_{i=1}^{N} y_i \qquad\qquad [6]$$

In [6], y_i is the quantity produced by the *i'*th company.

NEG analysis assumes that consumers perceive goods produced by different companies as *different varieties* of industrial good *y* (though the agricultural good, wheat, is still assumed to be homogeneous). It also takes into account the possibility

---

[16] Transport costs are highly simplified in NEG analysis. It is assumed that, given a quantity of a good q, if the good is transported to another economy, the quantity that reaches its destination is equal to Tq where 0<T<1 . These are called "iceberg costs" because the good produced is like an iceberg that partially melts during transportation from one area to another. Moreover, in many NEG models it is assumed, for simplicity's sake, that agricultural products can be transported at no cost.

that, if total quantities are equal, consumers prefer to consume goods of different varieties. This hypothesis can be summarised analytically by assuming that:

$$y = \sum_{i=1}^{M} \left(y_i{}^r \right)^{\frac{1}{r}} \quad \text{with} \quad 0 < r < 1 \qquad [7]$$

In [7], M is the number of varieties and y_i is the quantity produced by the *i*'th variety. In this structure the parameter *r* measures the intensity of preferences for variety. This intensity would be zero if *r* were equal to 1 (in which case [7] would be the same as [4]), and infinite if *r* were equal to 0, these two cases being excluded, while it is assumed that it would decrease if *r* increases between 0 and 1, both values being excluded. Where there is preference, the market for each industrial good (cars) is segmented into many different markets for each variety. In each of these segments, the producer has "market power" limited by the fact that consumers choose another variety if the one they had previously chosen becomes too expensive. It is assumed, then, that each company chooses the unit price of the variety of goods it produces,[17] keeping the general price index unchanged. This hypothesis, together with the previous ones, means the market in industrial products is no longer perfectly competitive; it is characterised rather by a form of competition that is called *monopolistic*. It is possible to demonstrate that optimal behaviour for a company is to fix the sales price by imposing a mark up on production costs.[18]

Fourth eliminated simplification: the hypothesis that both goods are produced using only labour, and that workers have the same characteristics in both sectors. NEG assumes that workers employed to produce industrial goods have different skills than those employed for agricultural goods, and that workers in the modern sector in each of the two economies are completely mobile, in the sense that they are perfectly willing to move from one economy to another if, by doing so, they are paid more. (Agricultural workers, on the other hand, are assumed to be completely immobile.[19]) It also assumes, for simplicity's sake, that each worker has a given set of skills and cannot acquire other competencies, meaning that every worker belongs either to the

---

[17] With increasing scales of return, the optimal choice for a company would be to specialise in the production of a single variety of manufactured good. The number of varieties produced thus coincides with the number of companies N.

[18] The demonstration of this result is omitted because it goes beyond the aim of this book. It can be found in Dixit and Stiglitz (1977).

[19] This is easy to understand in the light of the characteristics of agricultural production, which uses land, which is immobile, as one of its productive factors. The assumption proposed is thus a means of introducing typical features of the production of goods that depend crucially on immobile factors in a simplified context with a single input.

modern or to the traditional sector, and cannot move from one to the other. Finally, it is assumed that labour markets for both type of worker are in equilibrium (i.e. that there is no involuntary unemployment).

Having presented the assumptions proposed by NEG, we can now begin to examine the effects of production location and of distance between firms. In this new scenario, firm location is an economically relevant factor because profit levels in each of the two economies depend on the number of firms that have located production there, and because a firm that shifts from one location to another modifies the profit conditions of the new location. This modification takes place in two different ways. First, a new firm's arrival in a location changes the competitive context of the firms working there. Greater competition, in its turn, alters the characteristics of market equilibrium, lowering the price of goods produced and reducing profit levels. This so-called "price effect" (PE) tends to make the market less profitable for other firms seeking to enter. Second, a new entry modifies the demand side because it increases the size of the market in which the firm has decided to produce. If a firm decides to come into a market, for example by shifting production from region A to region B, there will be a subsequent shift of workers between the two regions with an increase in employment and income in region B. Since part of this income is necessarily spent inside that economy, this will increase demand in B and generate greater profits there. This second so-called "market effect" (ME), which is negligible in perfect competition but significant in monopolistic competition, tends to make the market more profitable for other firms seeking to enter.

Thus, the decision of a firm to locate its production in one area rather than another generates an overall effect on profits. The following can represent this effect:

$$\Delta\pi = PE + ME \qquad\qquad [8]$$
$$PE < 0 \text{ e } ME > 0.$$

The overall effect is unclear: it is positive if market effect prevails over price effect, but negative if price effect prevails over market effect. Whether market or price effect prevails is an important signal to determine the future development of an economy. If price effect prevails over market effect, a new entry will tend to discourage other firms from coming in, impelling them to find a different location. In the case of the

two economies, if a firm chooses to locate its production in region B, shifting its production from region A, it will reduce the profitability of market B, making it less advantageous for other firms. In this case, both economy A and economy B tend to develop industrial goods production. 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. In the case of the two economies, then, if a firm chooses to locate its production in region B it will make it advantageous for other firms to make the same choice, causing a gradual migration of firms from region A to region B. In this case, economy B tends to develop a significant industrial sector that produces the industrial goods necessary for both regions, while economy A becomes solely agricultural.

When either price or market effects prevail, two different models of development tend to evolve in the two regions. In one, the modern sectors expand in parallel in both regions. In the other, the modern sector agglomerates in one of the two regions. This agglomeration becomes the "centre" from where goods are produced for its own market and for the other region, which becomes "peripheral".[20]

A third effect linked to location identified by NEG needs to be mentioned before analysing in more detail the features of these two models, and before examining which elements contribute to one prevailing over another. So far it has been assumed that the only input used in production was labour. In order to make the hypothesis more realistic, a *fifth simplification* could be eliminated and two inputs could be assumed: labour and goods produced by other firms. A further assumption, again for simplicity's sake, could be that goods produced by a firm and used as input to produce other goods, and industrial goods bought by consumers, are one and the same thing. 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 connected to one another by input-output links, meaning that the revenues of a firm selling the goods it has produced are equal to the costs incurred by other firms to produce the same goods. As a consequence, a new entry in an economy generates a third effect.

As we have seen, the arrival of a new firm boosts internal demand for goods and increments profits (market effect). It also stimulates market competition that cuts

---

[20] N.B. The expressions "centre" and "periphery" are not used geographically; they refer to the different economic roles of the two economies.

sales prices and reduces profit margins (price effect). The third effect, caused by input-output links, is that a cut in sales price reduces the expenses of firms that use the goods they produce as input. For each variety of manufactured good, the negative effect on profit margins of a drop in revenues is partially compensated by the positive effect of the lower expenses incurred by the other firm. The effect on production costs brought about by the arrival of a new firm in a marketplace is known as the "cost effect" (CE).[21] The overall effect of a new entry, then, is as follows:

$$\Delta \pi = PE + ME + CE \qquad\qquad [9]$$

$$PE < 0, ME > 0 \text{ and } CE > 0.$$

The overall effect is still unclear. A new entry increases the profitability of the market if market and cost effects are greater than the price effect. If the price effect prevails, however, profitability decreases. The possible effects of the location of firms are thus similar to those described in equation [8].

The basic model: dynamics

We have seen how the choice of a firm to locate production in one region rather than another can have different effects on the profitability of the region. Market effect and cost effect tend to increase profitability, and are thus factors of attraction for other firms, leading to agglomeration of industrial production. Price effect, on the other hand, tends to reduce profitability, and is thus a factor of repulsion for other firms, leading to dispersion in industrial production. Having described the origin of these effects, it is now possible to determine which elements make certain effects prevail over others.

A general analysis, even using the simplified model outlined above (two economies, two sectors, production functions and elementary preferences) is complex. It would involve examining the consequences on market, price and cost effects of the whole range of possibilities opened up by the different values of the parameters: $\alpha > 0$ (productivity of agricultural labour), $\beta > 0$ (productivity of industrial labour), $0 < s < 1$ (a parameter that regulates the utility of consumers according to the combination of wheat and cars in their basket), $\gamma > 0$ (a parameter that regulates scale economies in industrial production), $0 < T < 1$ (magnitude of transport costs), $0 < r < 1$ (consumer

---

[21] See Venables (1996)

preference for a variety of industrial products), as well as other parameters indicating the importance of input-output links among firms.

A qualitative analysis of some results and of a few main trends is therefore preferable. The dynamics we examine are not the only economic development paths outlined by NEG; they are simply some of the possible trends that are of particular interest. As we have said, then, the overall effect on the profitability of a region of a new firm's entry is still unclear. It can, however, be shown that: i) a greater preference for variety in industrial products tends to reinforce the market effect and cost effect rather than the price effect, and tends to favour agglomeration[22]; ii) a higher proportion of industrial workers out of the total number of workers tends to reinforce market effect and cost effect rather than price effect, and tends to favour agglomeration; iii) stronger scale economies reinforce market effect rather than price effect but weaken cost effect against price effect, with an ambiguous consequences on agglomeration; iv) lower transport costs tend to reinforce the market effect and cost effect rather than the price effect, and tend to favour agglomeration.

NEG is thus in a position to explain industrial agglomeration as the result of an increasing preference for variety (owing to growth in per capita incomes), a greater proportion of workers in non-agricultural sectors, and lower transport costs.

There is empirical evidence to support this claim. Table 1 shows foreign trade as a share of industrial value added in several industrialised countries on average since the 1960s. There was a significant increase in foreign trade from the 1960s to the 1990s (first row of the table), which could be attributed to lower transport costs. However there are other causes to take into account: economic growth[23], vertical disintegration of productive processes[24], and a reduction in trade barriers, to name but a few. If we

---

[22] Where there is perfect substitution among variables (no preference for variety) there is perfect competition in which price effect always prevails.

[23] Consider two countries, a big country with a GDP of 100, and a small one with a GDP of 20. The highest level of exchange would be (20+20(/120=33% if the small country exported all its GDP and imported the same amount. If the two countries had GDPs respectively of 100 and 50, the highest level of exchange would be (50+50)/150=67% (Feenstra, 1998).

[24] Feenstra (1998) has documented a marked increase in imports and exports of intermediate goods owing to restructuring of integrated productive processes. In many industrial sectors an increasing number of firms no longer finalise production from raw materials; they produce goods that correspond to one productive stage or another. This leads to an increase in the ratio between foreign trade and industrial value added that does not depend on a reduction of transport costs but on new possibilities offered by technologies and organisational structures.

apply these parameters obtained in research on Europe, a very cautious estimate[25] of the share of international trade whose growth depends only on a reduction in transport costs can be made (second row of the table). The table shows that, while most of the increase depended on other causes, lower transport costs contributed significantly and regularly: increasing 1 percent every ten years.[26]

The effects of a country's economic growth, as well as the vertical disintegration of productive processes, could be attributed to increasing preference for variety. In this case, the fourth row of the table reveals a further indicator: another variable to which NEG models attribute agglomeration, alongside lower transport costs. It would seem that this factor has caused a significant increase in international trade.

As regards agglomeration, which could be considered an effect of these changes, row 5 of the table shows the Gini indices for the concentration of locations in manufacturing industry in Europe calculated at country level (an increase in the index corresponds to an increase in concentration) since the 1980s. It is evident that concentration has increased. Moreover, Figure 1 shows that spatial concentration, calculated at a country level for an 11-member European Union by sector (19 manufacturing sectors), tends to increase above all in sectors that used to be less concentrated, and does not diminish significantly in sectors that are no longer concentrated. In row 6 of Table 1, we can see the Gini indices for unemployment levels at a regional level that confirm a limited but significant[27] tendency to polarise in this respect.[28]

---

[25] In particular, the reduction of trade barriers has been highly relevant in Europe. The extension of this effect to other countries (such as Australia, Canada, Japan and the United States) probably produces an over-estimate, and therefore under-estimates the effects of reduced transport costs.

[26] A cross-country confirmation was obtained by Wilson, Mann and Otsuki (2004), who found significant effects on the volume of international trade in 75 countries (in 2000-2001) owing to the efficiency of ports, transport infrastructure and favourable trade policies. See the empirical work of Overman, Redding and Venables (2001).

[27] It is significant even though it grows only 3 percentage points from 28 to 31, taking account of the brief period of time taken into consideration, and of the well-known fact that unemployment rates (as a share of the active population) is effected by the impact of unemployment on activity rates. This means that when unemployment grows (or drops) activity rates diminish (or increase). Thus the unemployment rate increases (or diminishes) less than it would increase (or decrease) if activity rates were invariable.

[28] Agglomeration processes at the European regional level are confirmed by Overman and Puga (1999). They calculated the distribution of 150 European regions by unemployment level in 1986 and the transition matrix from 1986 to 1996. Using the data the distribution trend can be calculated (to 2012). It can be seen that unemployment rates lower than average are unvaried (43%), while regions with higher levels of unemployment rise from 29% to 37%. The number of regions with unemployment rates around the average, on the other hand, drops from 28% to 20%.

While empirical research is not yet able to claim beyond reasonable doubt that the NEG model is the best possible explanation for these stylised facts in comparison with other possible theories, the data we have just examined is encouraging. Increases in spatial concentration of industry, and consequent development, do appear to be correlated to a reduction in transport costs and diversification of production.

Tab. 1 – Growth in international trade in some industrialised countries ($_a$), spatial concentration of industry in European countries and unemployment rates in European regions

	1960s	1970s	1980s	1990s
1. International exchange of goods in % of industrial value added ($_b$)	31	38	52	56
2. Estimate of the share dependent only on reduced transport costs ($_c$)	31	32	33	34
3. Estimate of the effect of reduced tariffs and ceilings ($_d$)	-	2	6	6
4. Residual: effects of the growth of a country's economy and the disintegration of productive processes.	-	4	13	16
5. Spatial concentration of manufacturing industry by country (EU-11), Gini indices	na	na	0,156	0,188
6. Regional concentration of unemployment rates (168 European regions) Gini indices	na	na	0,284	0,302

Note: ($_a$) Australia, Canada, Denmark, France, Germany, Italy, Japan, Norway, Sweden, United Kingdom, United States, for data in rows 1 to 4; EU-11 row 5 (Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxemburg, Holland, Portugal, Spain). ($_b$) Simple averages of data for single countries with reference to the sum of their exports and imports divided by two ($_c$) ($_d$) Calcolated on the basis of Baier e Bergstrand'sresults (for Europe).
Sources: Feenstra (1998), Baier e Bergstrand (2001), Brülhart (1998), Eurostat, our elaboration.

Fig. 1 – Spatial concentration at country level (Gini indices) 1980-1990 (vertical axis on the right) and percentage variations (vertical axis on the left), manufacturing industry, EU-11
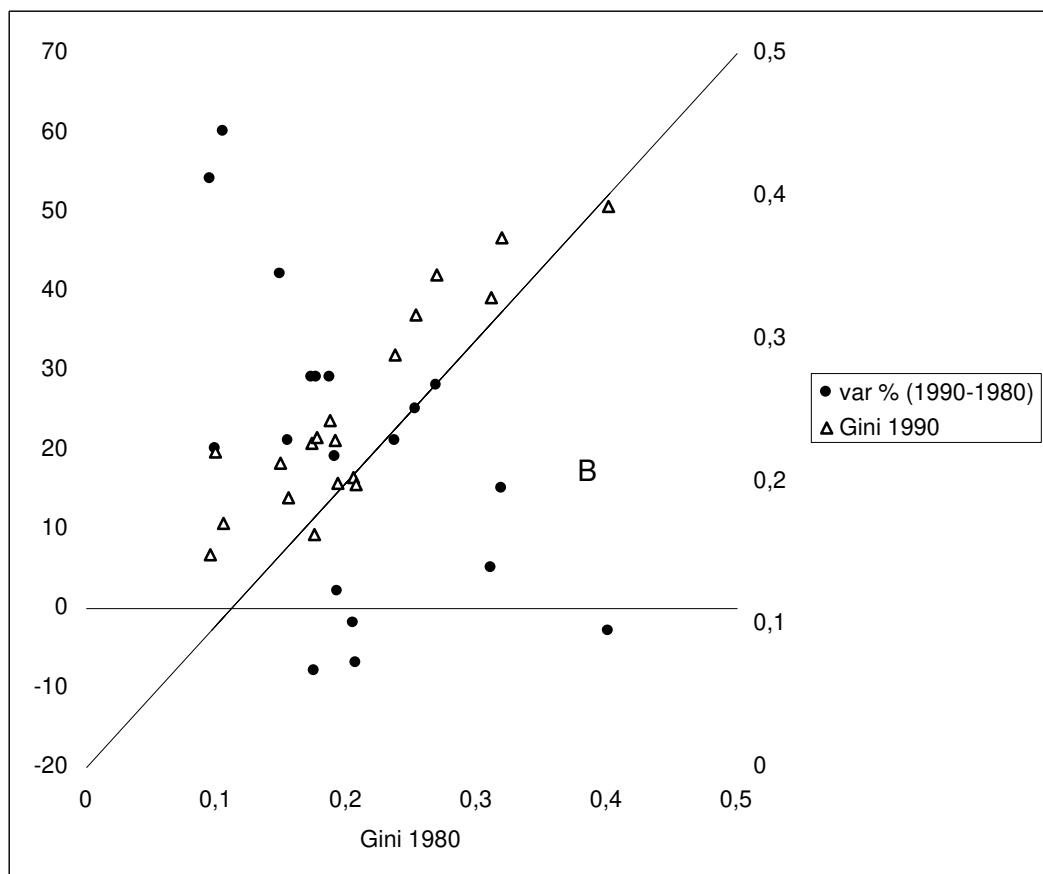
Source: Brülhart (1998), our elaboration.

Let us go on now to examine the link between transport costs and the location of firms. In the model we assume that the shifting of firms and workers from one area to another is not instantaneous but takes place incrementally over time. Figure 2 shows the link between transport costs and the location of firms.

Fig. 2 – Relation between transport costs and location in region A

if Firms in A
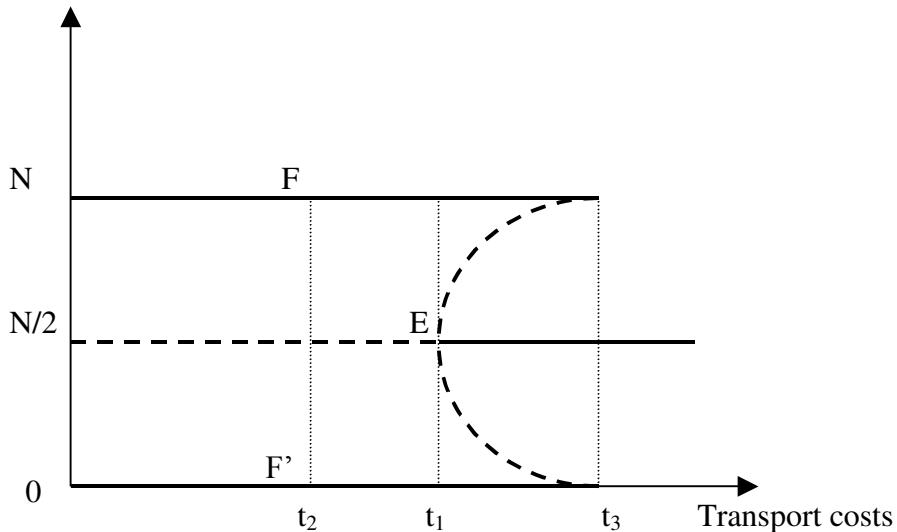
Let us suppose that N is the total number of firms. In one of the two economies (A represented in the figure) a point along one of the two continuous or dotted lines represents a possible equilibrium of a stationary state in which the system can stop. A point along the continuous lines represents a *stable* stationary state, towards which the system tends to move if it is a situation of disequilibrium. A point along the dotted line, on the other hand, represents a state of *instable* equilibrium, from which the system tends to distance itself if it is in a state of disequilibrium. The N firms could thus all be in equilibrium in region A (the top continuous horizontal line corresponding with N), or all in region B with none in A (the bottom continuous line corresponding with 0), or half in one region and half in another (the middle dotted line corresponding with N/2).[29] If the N firms are *not* along any of these lines then their location tends to change, shifting towards more stable equilibria, that is, in relation to transport costs, either all towards A or all towards B, or again half and half.

Let us suppose then that starting with a situation in which N firms are half in economy A and half in economy B. Figure 2 shows the results of the model. Starting on the left, if transport costs are less than t_1 and if we suppose, as we have done, that initially half the firms are localised in one region and half in another (the middle

---

[29] Economy B is part of the model and part of the argument, but it is not represented in the figure. Needless to say, the firms that are not in the figure, out of the total number of firms, would be in economy B.

dotted line), the system is in a state of instable equilibrium. All it would need is one firm to shift, and all the others would follow, moving production to either region A or region B (the top and bottom continuous horizontal lines). For example, if transport costs were equal to t_2, the two possible states of equilibrium are F or F'. This takes place, as we have seen, owing to market and cost effects prevailing over price effect. In other words, given that the two economies are equal to start with, and each have half the total number of industries, all it would need is for one firm to shift production from one economy to the other in order to determine an increase in demand and a reduction in production costs in the input-output link between firms. Since these effects are greater than the increase in competition, the profits of industrial firms increase and attract other companies. These companies increase local demand because of their re-location, and therefore increase profits, but they do not need to concern themselves with demand from the other region because transport costs are low. That is, they can incur transport costs without jeopardising the profit gained by re-locating.

With transport costs higher than t_1, by contrast, a sole firm (which, in the example above, by shifting production, made it economically advantageous for other firms to re-locate) can no longer initiate sustainable re-location. A whole block of firms would need to re-locate in order to set in motion effects of increased profitability sufficient to compensate higher transport costs. The higher the transport costs, the greater the number of firms would need to move, until – at level t_3 – re-location would be sustainable only if all the firms in one region moved. If transport costs fall, on the other hand, the movement would be in the opposite direction, from right to left in Figure 2. The size of the block that needs to re-locate diminishes, until it is sustainable for a sole firm to shift production. This would create equilibrium E, on the border between instability and stability.

Figure 2 thus helps us understand situations where relatively high transport costs have led to a distribution of firms that is not exactly divided between two regions. If to start with, there were slightly more firms in region A than in region B, a reduction in transport costs towards t_1 would cause a gradual, parallel shift of firms towards region A, until region B were completely empty. The opposite shift would take place if there had been more firms initially in region B than in region A.

To sum up, if transport costs are high, the forces that impel firms to re-locate tend to prevail and production is distributed over different regions. If transport costs are reduced, the forces that persuade firms to concentrate prevail, and production tends to

agglomerate. In the latter case, the fact that production is concentrated in region A or B depends, in the simplified structure of the model, on a single element: the initial distribution of the firms. Production is concentrated in the location where there were more firms to start with.

Some extensions

We have seen the implications of the mechanisms described by NEG in a simplified model using only two economies with identical structural features. We are now going to examine the effects of eliminating these simplifications in order to analyse the consequences of heterogeneousness among economies and with many different economies.

The assumption of structural homogeneousness could be eliminated in many different ways. One form of heterogeneousness that is particularly interesting is in the number of workers employed in the traditional sector.[30] Workers employed in this sector are generally considered immobile, and thus represent a share of demand for goods rooted in the location where they are first produced. Since the other share of demand is mobile (because industrial workers are mobile), the fact that one region has more agricultural workers than another means that this region has a greater potential demand for goods, in the sense that the demand (say, for cars) has the potential to reach a higher level than in the other region.[31] In this hypothetical economy, we have assumed production functions with labour as its single output, making population (labour supply) the only resource, and the agricultural population the only immobile resource. The agricultural population represents a share of demand for modern goods rooted in its territory because it is an immobile productive resource there.

The same holds for any other resource that is needed for production and that can be considered immobile. If we use a more complex production function that includes skills linked to the territory (such as those exploited by industrial districts) or natural

---

[30] See Krugman, 1992

[31] See the following numerical example. There are 200 workers, 100 in industry and 100 in agriculture. One hypothesis would be that economy A: 50 industrial and 50 agricultural; economy B: 50 industrial and 50 agricultural. In this homogeneous case, if all the industrial workers went to A (or to B) the maximum number of workers (and thus of potential demand for cars) would be the same: 50 agricultural and 100 industrial, equal to 150 workers. If the initial division had been economy A: 50 industiral, 10 agricultural; economy B: 50 industrial and 90 agricultural, and if all the industrial workers moved to economy A, A would have 100 industrial workers and 10 agricultural workers, making a total of 110, while economy B would still have 90. If all the industrial workers went to economy B, B would have 100 industrial and 90 agricultural workers, making a total of 190, while economy A would only have 10 workers.

resources (such as climate or landscape), these resources function in the same way as the agricultural population functions. When these resources are exploited, whoever possesses them earns an income that can subsequently be spent locally and that sustains demand for industrial goods. Since the logic is the same, we may as well simplify the production hypothesis with a single function (labour) and with the agricultural population as the only immobile resource.

We shall now go on to examine the effects on the location of firms of differences in agricultural sector populations as immobile resources. Let us imagine that transport costs are such that they guarantee forces for agglomeration. As figure 2 showed, industrial firms tend to concentrate in one of two areas, and if the two economies are homogeneous, firms will always locate where there were more firms to start with. Where there is heterogeneousness in the stock of agricultural workers, the results are different. In order to examine this new scenario, let us assume that there are more agricultural workers in region B than in region A. The simplest situation is that in which the initial number of modern firms is only slightly higher in one of the two economies, for example in region A. Let us suppose, however, that the difference is marginal and that the total number of workers is nevertheless greater in region B. In this case, the incentive for a modern firm to re-locate from B to A will be less pressing than its incentive to stay put, while the incentive for firms in region A to move to region B, where there is a greater number of agricultural workers. In region B, in fact, both present and potential demand for cars is higher. Agglomeration will thus take place in regions with a greater number of workers (including agricultural workers) even when this region had fewer industrial firms initially.

The conclusion would be very different and more interesting if there were a greater difference in the initial number of firms in the two economies. If, for historical reasons, there were far more industries operating in region A, it is possible that the present demand for cars would be greater in region A than in region B, even though there were fewer agricultural workers. Nevertheless, the potential demand for cars would be greater in region B where most of the immobile population lives. Thus, if the effect of higher present demand prevails in determining location then agglomeration will tend towards region A. If, however, potential demand prevails, then agglomeration will tend towards region B.

It is not hard to see what this result depends on. It depends on the conjectures made by each firm as to the behaviour of the 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 the most workers, given that there are more available agricultural workers. And yet, because this kind of collaboration is not possible[32], each firm would only decide to locate production in economy B if it thought most firms in economy A were going to move there too. If each firm, on the other hand, were convinced that the other firms were localising production in economy A, they would too. To sum up, 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 localise in many different ways and there is no equilibrium in their distribution.

The start up phase of a process of industrial growth (exogenous, triggered by an external event rather than by the model's variables) plays an important role. An agricultural area with little industrial development has greater potential demand for industrial goods but less present demand. We do not know how firms form their conjectures. We do know, however, that if a process of industrialisation is triggered in a traditional area, the forces that impel firms to agglomerate in the more industrialised economy (present demand) will be weakened, while the process in which firms re-locate to the more traditional economy will be strengthened. Once the localisation and agglomeration process has been triggered in economy B, it will tend to be self-reinforcing.

The second simplification that can be eliminated is the hypothesis of a system with only two economies where firms can locate production. The hypothesis can be substituted by the assumption that localisation can take place in any number of places. This more realistic assumption has been examined in detail by NEG,[33] but we cannot analyse it too closely here because it is highly complex. Nevertheless, some of its conclusions are easy to summarise. First, NEG has shown that, when transport costs are sufficiently low, agglomeration usually takes place in a subset of regions rather than in one region alone. Second, it has observed that the dimension and composition of this subset of regions changes when the structural parameters of the economies

---

[32] It is assumed that each firm makes its choices individually.
[33] See Krugman, 1993, 1994.

vary, as well as when the initial distribution of firms among the various locations is modified.

NEG thus provides us with a useful springboard for considering the development potential of, and obstacles for, lagging areas (the South) that entertain significant interaction with more developed areas (the North).[34] NEG makes it clear, for example, how wrong the idea commonly held in Italy is that development of the South is an automatic consequence of development in the North. According to NEG, development in the North detracts from development in the South, unless something is done to help the South to discover and valorise its immobile resources. Another commonly held idea is that development in the South depends on improving North-South transport infrastructure. The basic NEG model, as well as the example of the industrial sector in Europe, indicates that reduced transport costs produce greater divergence rather than convergence in regional disparities. And yet, NEG has explored more complex models showing how divergence can be avoided. These allow appropriate development policies to be established with a certain degree of success. If these policies were to intervene and modify decisively the conditions that normally worsen disparities, NEG would contribute to understanding how the forces unleashed by these policies can determine recovery in lagging areas.

In order to illustrate in greater detail the unpredictable processes that differ from the basic model and reflect more complex models, we are now going to examine two particularly interesting examples.

Two examples

The first example is a comparison between the territorial concentration of manufacturing industry in EU-11 and in eleven countries in southern Africa. The second is a study of local labour markets in unified Germany. In less advanced countries such as those in southern Africa, the fact that transport costs are higher than in industrialised, for example European, countries does not mean there is a greater spread in production; industry tends rather to agglomerate. It would seem that the heterogeneity of the single areas drives agglomeration towards wealthier regions. In the case of Germany, on the other hand, low transport costs do not trigger agglomerative processes. Agglomeration is held back by heterogeneity because the

---

[34] Both in the sense of under developed countries open to exchange with more industrialised countries and lagging regions in an industrialised country.

wealthier regions are also more congested, which can contribute to creating a greater spread of firms.

Let us take the first example of the comparison between Europe and southern Africa. For industrialised countries such as those in Europe, it can be safely assumed that transport costs (in the wider sense) are significantly lower because infrastructure is highly developed. It could be expected, then, that in less developed countries the little industry that there is would be more widely spread over the territory. And yet the opposite can also be the case, as Figure 3 shows. The graph compares the Gini indices for the concentration of manufacturing industries in eleven European countries and eleven southern African countries.[35]

Fig. 3 – Comparison between the Gini indices (Europe and southern Africa) for territorial concentration of manufacturing industry in the early 1990s
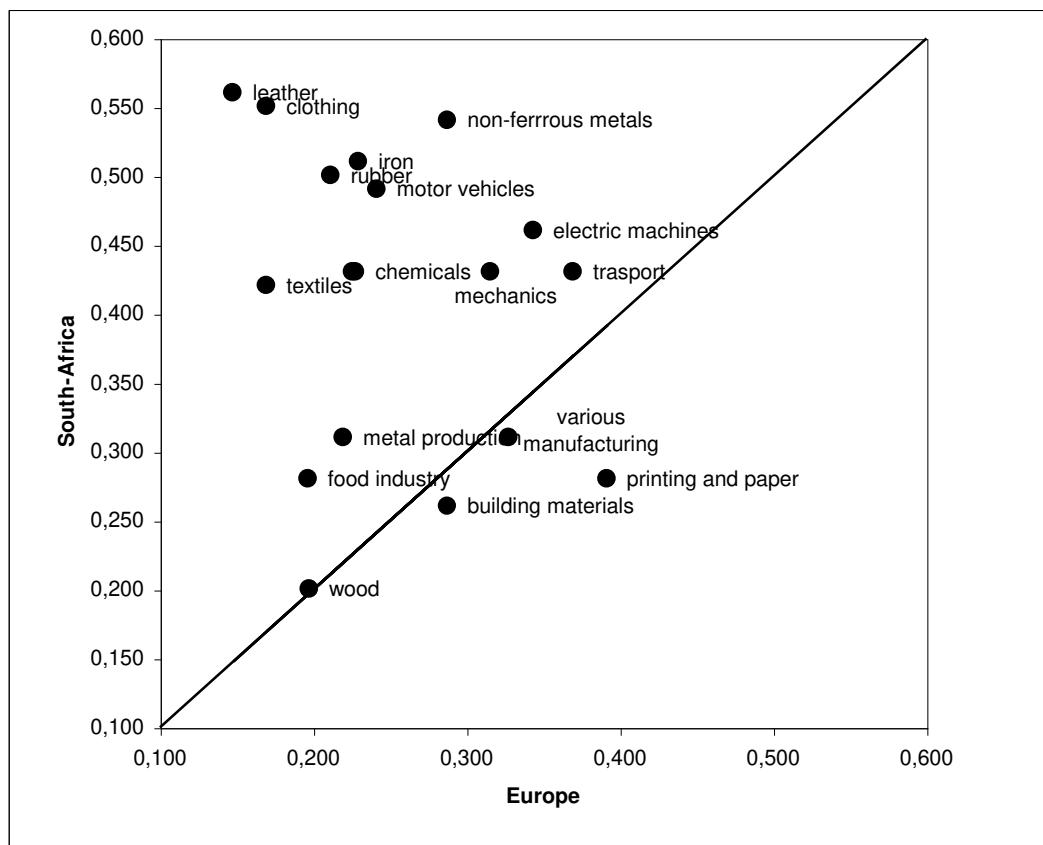
Source: Our elaborations on the basis of data from Petersson (2000) and Brülhart (1998).

It is clear that nearly all sectors are more concentrated territorially than in southern Africa. It can also be observed that the leather and shoemaking, clothing, metals and

---

[35] The eleven southern African countries are Tanzania, Namibia, Zambia, Malawi, Zimbabwe, Mozambique, Botswana, Lesotho, Mauritius, Swaziland, South-Africa (Petersson, 2000)

iron, rubber, motor vehicle, textile and chemical sectors are the most concentrated, as they are in Europe. The explanation for this evident anomaly, compared to the paradigm regarding only transport costs, must take into account many circumstances. The most relevant regard the marked differences in population distribution, urbanisation, per capita income and human development (correlated to the quantity and quality of labour supply for a given population). In southern African countries the population density per square kilometre varies from a maximum of about 600 people (Mauritius) to a minimum of 2 (Namibia), while in Europe the figures are 460 and 60 respectively. The urbanisation rate varies from 50% (South Africa) to 13% (Malawi), with European rates ranging from 97% to 58%. Per capita income in international dollars for (1997) ranges from 7,380 (South Africa) to 740 (Mozambique); in Europe the variation is between 30,800 and 14,300. The human development index (HDI) is 0.700 in South Africa and 0.341 in Mozambique, while in Europe the highest rate is 0.923 and the lowest is 0.858.[36]

These figures reflect some of the determining factors we have already talked about: a region's population and its early industrial development (urbanisation); present and potential demand (per capita income and population); labour supply (HDI). Evidently these factors played a far more important agglomerative role in southern Africa than the role played by greater transport costs, which in itself, as we have seen, would tend to drive firms to spread out over the territory.

Reunified Germany provides a case study of the opposite effect. If we complicate the model by assuming imperfect mobility of industrial labour (in as much as the competencies are linked to the territory and are therefore less productive elsewhere), as well as a certain mobility of agricultural labour, coupled with the possibility of increasing revenue from this sector, in addition to nominal salaries determined according to the local cost of land (as a proxy of housing costs), then the relationship between transport costs and localisation changes.

Brakman, Garretson and Shramm (2002) applied this model, known as the Helpman-Hanson model[37], to Germany after unification. There is general consensus that the re-unification of Germany, which raised both hopes and fears, reduced

---

[36] Petersson, 2000, p.4 and UNDP, Human Development Report
[37] The model was proposed by Hanson (2001) on the basis of Helpman's work (1998), which is why it is known as the Helpman-Hanson model. Hanson has used it recently with explicit reference to the economic regions of the US.

transport costs (in the wider sense) between East and West as well as within the East. It then became the subject of debate whether the advantages in the long run outweighed the disadvantages in the Eastern regions as well as in the West. On the basis of our simplified model, it could be claimed that, by accentuating the agglomeration effects, reduced transport costs would create disadvantages in the lagging regions of the East and advantages in the more developed regions of the West. On the basis of a model in which transport costs are less significant than heterogeneity (as in the case of southern Africa), by contrast, there would be advantages for the East and disadvantages for the West, owing to the delocalisation of production in search of lower labour costs.

The Helpman-Hanson model has made it possible to estimate empirically the relevance of transport costs compared to the relevance of territorial differences in wage levels, which in their turn are linked to differences in per capita income and in the cost of housing. In other words, this model allows us to verify empirically the combinations of forces that contribute to dispersion and hinder agglomeration. It emerges that in Germany, with its 37 districts, transport costs are relevant but forces for dispersion such as congestion costs are also significant. These costs are the result of a mechanism whereby high incomes and low availability of land drives up house prices that in their turn increase nominal salaries (and vice versa). The re-unification of Germany thus allowed Western firms to sell to the East more cheaply and therefore to maintain production in the West, but also to re-locate production to the East since in the West salaries were higher and would be driven even higher if the income gap between east and West continued to grow.

The correlation between revenues per square kilometre, the cost of land, and nominal salaries emerges clearly at a regional level, as Figure 4 shows.

Figure 4 - Regions of unified Germany ranked by nominal salary per hour in the industrial sector, trend in revenues per square kilometre and cost of land per square metre (1994). Percentage indices 100 = regions with the highest rates; hourly salaries on the right, revenues and land cost on the left (logarithmic scale)
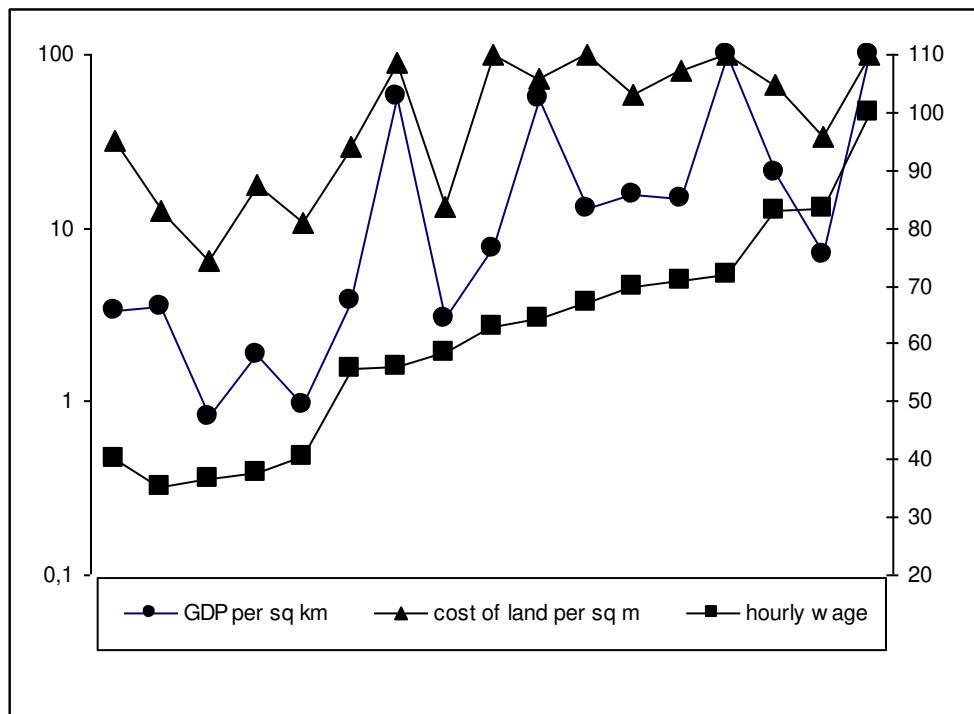
Source: our elaborations based on data from Brakman, Garretsen, Shramm 2002.

The two examples contribute to the North-South debate. The comparison between Europe and southern Africa makes it clear why policies based on incentives alone are ineffective. Since the development of the South is a challenge, it is easy to claim that transferring funds from the North would solve the problem. These transfers, it is claimed, would support local demand and provide new stimulus for the birth of new firms. As we have seen in this chapter, the claim is based on the hypothesis that reduced transport costs stop Northern firms from satisfying demand in the South. The example of southern Africa and the more complex NEG models, are a warning that dispersion, and thus re-equilibrium, may well not take place, even with higher transport costs. Such deeply seated structural differences between North and South can in fact make it more advantageous for modern firms to stay in the North even though they have to incur higher costs to satisfy the demands of the South.

The example of post-unified Germany also sheds light on the existence of agglomeration tendencies based on congestion costs. The idea that development in the North automatically contributes to closing the gap with the South could be defended

by claiming that scarce resources, such as land and housing, could keep labour costs higher in the North, thus impelling firms to re-locate to the South. The study of Germany, where the gap between East and West in terms of labour costs in relation to transport costs and externalities was historically very high, demonstrates that relying on the effects of congestion is also unrealistic. In short, congestion costs would only have an effect if they were very high, but in this case they would also be socially unsustainable.

The conclusions reached by NEG

So far we have examined 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. We have seen how forces that foster agglomeration and those that lead to dispersion often work together. On the subject of local development, NEG has shown that, all things being equal, improving infrastructure and transport services (both important for development in general) tends to accentuate the development gap between regions because these improvements support the spatial concentration of modern production. Concentration in one area rather than another, it has been shown, can depend entirely on the initial distribution of firms in the territory, thus signalling that local development could indeed depend exclusively on history.

And yet, it is important to stress that the initial distribution of firms is not the only significant element. The endowment of immobile resources is also relevant. A lagging economy with few modern firms can attract other firms if there are a great number of workers and in general if it has valuable immobile resources. This outcome is just one of the possible results, however, inasmuch as there are many different forms of heterogeneity that contribute to different location choices made by firms. By stressing the importance of regional heterogeneity, our analysis carves a middle path between the basic NEG model and international trade theory outlined at the beginning of this chapter. In the latter, regional differences are paramount and transport costs are irrelevant, while in the former only transport costs are relevant. In the real world, both factors are relevant, to a variable extent in each case. The agglomeration effects of increasing returns for internal and externalities, despite high transport costs, are also important; these effects are overlooked both by international trade theory and by the basic NEG model.

Keeping this complexity in mind, as the examples given in this chapter show, NEG invites us to consider the necessity for international interventions that set local development processes in motion in lagging areas. NEG also makes it clear that these interventions cannot be limited to income transfers, and must aim rather at valorising existing immobile resources. Finally, NEG warns that it would be possible in the abstract to support development in the North and wait for congestion costs to make it impellent to re-direct development effects towards the South, but that the strategy would be socially and politically costly.

Bibliography for Chapter 1

Baier, S.L. e J.H. Bergstrand (2001), The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity, *Journal of International Economics,* 53, pp. 1–27.

Brakman, S., H. Garretsen, M. Shramm (2002), *New Economic Geography in Germany: Testing the Helpman-Hanson Model*, Hamburgisches Welt-Wirtschafts-Archiv Discussion Paper 172.

Brülhart, M. (1998), *Economic Geography, Industry Location and Trade: The Evidence*, University of Manchester, mimeo, for: *The World Economy Special Issue*.

Caselli, F. (2004), The missing input: accounting for cross-country income differences, Cap. 8, P. Aghion e S. Durlauf (a cura di), *Handbook of Economic Growth*, North Holland, Amsterdam.

Christaller, W. (1933), *Die Zentralen Orte in Suddeutschland: Eine oekonomisch – geographische Untersuchung ueber die Gesetzmaigkeit der Verbreitung der Siedlungen mit staedtischen Funktionen*, Fischer, Jena.

Dixit, A.K. e Stiglitz J.E. (1977), Monopolistic Competition and Optimum Product Diversity, *American Economic Review*, 67, pp. 297-308.

Feenstra, R. (1998), Integration Trade and Disintegration of Production in the Global Economy, *Journal of Economic Perspectives*, 12, 4, pp. 31-50.

Gola, C. e A. Mori (2000), Concentrazione spaziale delle produzione e specializzazione internazionale dell'industria italiana, in: L.F. Signorini (ed.) *Lo sviluppo locale – un'indagine della Banca d'Italia sui distretti industriali*, pp. 67 – 100, Donzelli, Roma.

Hanson, G.H. (2001), Scale economies and the geographic concentration of industry, *Journal of Economic Geography*, 1, pp. 255-276.

Head, K. e T. Mayer (2003), *The Empirics of Agglomeration and Trade*, CEPII Working Paper, 2003-15.

Helpman, E. (1998), The Size of Regions, in D. Pines, E. Sadka, I. Zilcha (a cura di), *Topics in Public Economics*, Cambridge University Press, Cambridge.

Henderson, J. V. (1974), The Sizes and Types of Cities, *American Economic Review*, 64, pp. 640 – 656.

Hirschman, A. (1963), *The Strategy of Economic Development*, Yale University Press, New Haven.

Krugman, P. (1992), *A Dynamic Spatial Model*, NBER Working Paper 4219.

Krugman, P. (1993), On the Number of Location of Cities, *European Economic Review*, 37, pp. 293-298.

Krugman, P. (1994), Complex Landscape in Economic Geography, *American Economic Review*, 84 (Papers and Proceedings), pp. 412-416.

Krugman, P. (1991), *Geography and Trade*, MIT Press, Cambridge (Ma).

Lasagni, A (2006), *TFP, sviluppo economico e specializzazione: un'analisi empirica a livello internazionale*, Tesi di Dottorato Università di Pavia.

Marshall, A. (1920), *Principles of Economics*, MacMillan, London.

Neary, J.P. (2000), *Of hype and hyperbolias: Introducing the new economic geography*, University College Dublin and CEPR Working Paper.

Ohlin, B. (1933), *Interregional and International Trade*, Harvard University Press, Cambridge (Ma).

Overman, H.G. e D. Puga (1999), *Unemployment Clusters Across European Regions and Countries,* CEPR Discussion Papers 2255.

Overman, H.G., S. Redding, A.J. Venables (2001), *The Economic Geography of Trade, Production, and Income: A Survey of Empirics*, mimeo, London School of Economics and CEPR.

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.

Psacharopoulos, G. e H.A. Patrinos (2002), *Returns to investment in education a further update, World Bank* Policy Research Working Paper, n. 2881.

Schmutzler, A. (1999), The New Economic Geography, *Journal of Economic Surveys*, 13, 4, pp. 355-379.

Venables, A.J. (1996), Equilibrium Locations of Vertically Linked Industries, *International Economic Review*, 37, pp. 341-359.

Wilson, J.S., C.L. Mann, T. Otsuki (2004), *Assessing the Potential Benefit of Trade Facilitation: A Global Perspective*, World Bank Policy Research Working Paper 3224.

Chapter 1 Appendix 1 - Regional disparities in Europe

The construction of Europe was a complex operation during which political considerations, no doubt, were paramount. It was, however, also an operation that progressively created a multi-regional economic space within which highly diverse systems were integrated until a single currency was finally adopted. The process has been in motion for several decades[38], and it is interesting to appreciate how regional disparities have been reduced at the same time as local economic systems have gradually opened up. The following brief outline will show that European cohesion policy played an important role in building consensus for a local dimension in terms of the intentions it displayed and the results it obtained, but that it was much less relevant in terms of financial output.

---

[38] It might be useful to outline the main phases leading to the single currency. In September 1950, the EUP (European Union Payments) agreement was signed as part of the OEEC (Organisation for European Economic Cooperation) set up on April 16, 1948 mainly to manage US aid, after 1961 to become the OECD. This organisation established a system to compensate balance of payments in participating countries (Austria, Belgium, Denmark, France, Great Britain, Greece, Republic of Ireland, Island, Italy, Luxembourg, Norway, Netherlands, Portugal, Sweden, Switzerland and Turkey, as well as the three zones of Germany assigned to the Americans, the French and the British. In 1949 the Federal Republic of Germany, and in 1959 Spain, joined. In April 1951, the ECSC (European Coal and Steel Community) was established. In March 1957, the Treaty of Rome was signed (going into force on Jan 1, 1958) establishing the EEC (European Economic Community) or common market and EURATOM (European Atomic Energy Community). The original six signatories were Belgium, France, Italy, Luxembourg, the Netherlands and West Germany. In July 1969, European customs union was completed, removing custom tariffs between members and setting a Common External Tariff for imports from non-member states. In February 1971, the Council of Ministers, on the basis of the Werner report, resolved to establish economic and monetary union within ten years. In April 1972, the European Monetary Snake, was signed by EEC Council of Minsiters to reduce the fluctuation margin between the community currencies to +/- 2.5%, that is half of what had been agreed in Washington. The aim of the snake was to reduce the consequences in Europe of the instability created by the August 1971 declaration that the dollar would no longer be convertible. In January 1973, EU-9 is created with the addition of Great Britain, Republic of Ireland and Denmark to the original six members. In March 1979, EMS (European Monetary System) was created to stabilize exchange rates between members of the European Community. At the same time a currency that was never actually coined, the ECU (European Currency Unit) was introduced. This was a basket of currencies, preventing movements above 2.25% around parity in bilateral exchange rates other member countries. The EMS (European Monetary System) was a more sophisticated and complex version of the monetary snake devised to achieve monetary stability in Europe, but it too was strained by the differing economic policies of its member states, some of which withdrew from the system. January 1981, Greece joins. January 1986, Spain and Portugal join (EU-12). Februrary 1986, the SEA (Single European Act) made the first major revision of the Treaty of Rome to harmonise laws and resolve discrepancies, allowing for the free movement of goods, labour, services and capital. December 1991, the Maastricht Treaty was signed with three objectives: to develop a single market with economic and monetary union through the creation of a single currency by 1999, to define and consolidate cooperation in foreign policy and security and to find ways to for police and the judiciary to collaborate. In December 1995, the third phase was launched with the definition of the Euro that would enter circulation on January 1, 2002. In 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.

European integration and cohesion: theory

Although the question of regional disparities was not considered significant in the debate that accompanied the process of building the European Union[39], it has nevertheless been a source of real concern.[40] This corresponds to the nature of the definition of cohesion used: "the degree of regional disparity that is socially and politically tolerable within the Community" (Begg-Mayes, 1992) in which the word "tolerable" is ambiguous and cannot be measured *a priori*. Most attention was paid to other issues[41], mainly the costs and benefits of integration for member states and the feasibility and sustainability of monetary union. Experts and politicians soon realized that integration was an advantage for member states[42] from an economic point of view; the only condition was that the structural adjustments in economic and social systems were not too strenuous.[43] On one hand it appeared that integration could take place even if regional disparities worsened, or failed to be reduced; on the other there was real concern that if cohesion were to dip below the threshold of tolerability (a threshold that was obviously never reached), the opposition of civil society would foil political projects for integration.[44]

---

[39] "Despite the concern expressed in the Délors report about the dangers of regional imbalance (Committee for the Study of Economic and Monetary Union, 1989, p.18) little attention is given to the effects of monetary union on regional disparities"Begg and Mayes, 1992, p.220). The regional question was totally absent in the 1957 Treaty of Rome and only emerged recently and very gradually (Cuadraro-Roura, Parellada, 2002, pp. 1-2). It did not become a critical matter until the 2001 Cohesion Report (CEC, 2001) and was then again scaled down after the 2004 Sapir Report that shifted the question back to European growth as a whole (slower than in the US), rather than on growth in its lagging regions. In the Treaty of Rome priority was explicitly accorded to competition among and within member countries, leaving regional disparities to be dealt with by national policies with supranational supervision to avoid interventions that would distort the market. The SEA (1986) takes a step forward, followed by the Maastricht Treaty (1992), when cohesion becomes the second most important EU objective after the creation of an economic space with no borders.

[40] The EU Directorate for Regional Affairs was created in 1967 but only started to operate after the Thomas Report (CEC 1973) supported the UK on joining the common market by stressing the importance of addressing regional disparities, which were a significant problem in the UK at the time. Oddly enough, Italy played no role in this area until the recent "new programme" after the "special programme" had failed.

[41] In economics, the books that survey these problems we recommend are Baldwin, Wyplosz, 2004 and Altomonte, Nava (2005)

[42] 1970 is the date that is conventionally thought to have marked the beginning of EMU. This is the year the Werner Report was published as a supplement to the EC Bulletin in November. The Report agreed with the findings of the Commission and set the agenda for the process to begin. The formal start of the process was on February 9, 1971, when the Council of Ministers accepted the recommendations contained in the Werner Report.

[43] Dehem (1972) produced a survey of the costs and risks of EMU and covered the debate at the time the Werner Report was issued.

[44] Events demonstrated that concern was well founded. Referendums held in various member countries to ratify the European Constitution turned out resounding No's. Comments regarding the French rejection of the Constitution were based on similar considerations.

In the Single European Act signed twenty years ago (1986)[45] it was declared that: "In order to promote its overall harmonious development, the Community shall develop and pursue its actions leading to the strengthening of its economic and social cohesion. In particular the Community shall aim at reducing disparities between the various regions and the backwardness of the least-favoured regions. Member states shall conduct their economic policies, and shall co-ordinate them, in such a way as to attain the objectives […] The implementation of the common policies and of the internal market shall take into account the objectives […] and shall contribute to their achievement. The Community shall support the achievement of these objectives by the action it takes through the structural Funds (European Agricultural Guidance and Guarantee Fund, Guidance Section, European Social Fund, European Regional Development Fund), the European Investment Bank and the other existing financial instruments. The European Regional Development Fund is intended to help redress the principal regional imbalances in the Community through participating in the development and structural adjustment of regions whose development is lagging behind and in the conversion of declining industrial regions."[46]

At the beginning of the process that would lead to a single currency, a close link with development in lagging regions was forged. It has been established (Sapir et al., 2004, pp. 37-8) that promoting economic development in the less developed European regions had to be included in the Treaty (while previously the issue was neglected)[47] owing to the fact that in the Single European Act member states were committed to completing the process leading to a single market, a "space with no internal borders with free circulation of goods, services and capital" by 1992. The issue had to be included because establishing a single market that removed barriers to the free circulation of production factors and products meant that member states effectively renounced all forms of protectionism (both explicit and implicit, including state aid to firms and state monopolies) as well as relinquishing national currency. This new market made national and local productive systems more open to competition.

---

[45] The SEA was the first revision of the Treaty of Rome. Signed in Brussels on February 28, 1986 during the intergovernmental conference, it came into force on July 1, 1987 following ratification by single member states.

[46] SEA (1986) Title II, sub-section IV: Economic and social cohesion, articles 130a, b, and c.

[47] Article 23 of subsection IV states "A Title V shall be added to part Three of the EEC Treaty as follows: Title V Economic and Social Cohesion."

The Single European Act was not as unambiguous as it appeared, however. On one hand, the aim of reducing regional disparities was coupled to the achievement of the single market; on the other, it was linked to intentional "support" interventions. The theoretical backing for this position was uncertain about the possibility of reducing regional disparities. Simplifying a great deal, this uncertainty can be illustrated with two arguments: one is optimistic and the other is pessimistic about the chances of effectively reducing regional differences. Their relative force is difficult to evaluate *a priori*[48], making results uncertain by definition. The two arguments can be summarized with two very well known graphs.

The first regards the *advantages of trade*. In traditional theory of customs' unions, if you lift trade barriers in economic systems that are part of a union, trade increases inside, but decreases outside, the union. The benefits for single countries that participate in the union depend on the balance between these two opposing features (Johnson 1962). If we look at the matter, as we must in this context, from the point of view of regions - that is economic systems within countries, rather than of countries themselves - then Johnson's cautious theory does not hold in the same way. It is likely that many regions within a country participating in a customs' union would see a rise in the volume of their total external trade.[49]

From the point of view of regions, then, customs' unions simply open unions to trade. If this is the case, then the traditional analysis of comparative advantages originally proposed by Samuelson in 1939 (see Figure 5) holds. There are two goods, A and B; FF indicates the frontier of production capacity in a regional economic system. In isolation, equilibrium – in which the marginal transformation rate between A and B on the supply (technology) side is equal to the marginal transformation rate on the utility side – will be established in x, c_1 being the collective indifference curve. If trade is opened up, there will be benefits as long as the external price of good A

---

[48] That is according to common opinion among policy makers at the end of the 1980s, without taking into consideration new theoretical developments such as NEG and Endogenous Growth, both of which, it seems, have accentuated pessimism regarding the effects of integration on regional imbalances.

[49] Suppose, for example, that for a country the volume of trade created with countries within the union and the volume of trade with countries outside the union is the same and the balance is zero. This country would be indifferent and would gain nothing from joining the union. Imagine, however, that there are many regions in this country. It is likely, in this case, that some of them would increase their trade while others, perhaps because they are on the border with other countries outside the union, would lose trade. The result could be that the benefits for the former were such that they made up for the losses of the latter, so that the country in question (considered as the sum of its diverse regional economic systems) decided it was worthwhile joining the union despite its zero balance between volume of trade created and volume of trade lost.

compared to good B is different from the internal, autarchic price and cannot be influenced by domestic policy. For example if the relative external price of A, indicated by line p, is less than it would be in autarchy (q). The result will be a new equilibrium on the production side (y), with a reduction in the quantity of A and an increase in the quantity of B with regard to x. At the same time there will be a new level of equilibrium of wellbeing given by the availability of goods (z) obtained by selling B outside the union and buying A at a relatively lower price. Wellbeing will have grown, as the graph illustrates in Figure 5, with the indifference curve c_1 becoming c_2.

Fig. 5 – Trade Gains
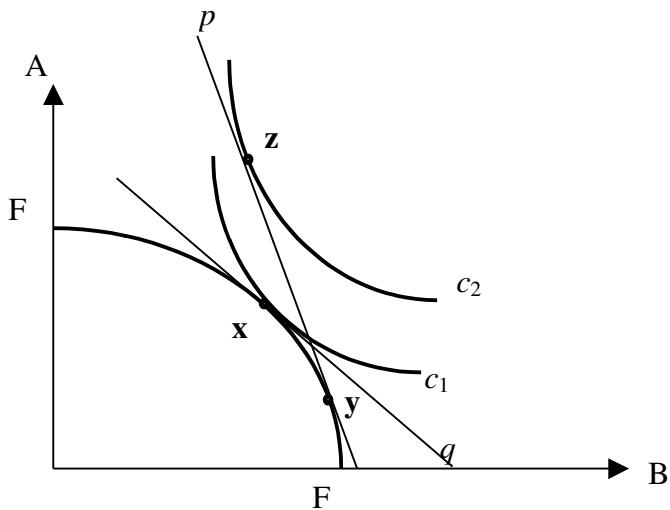
This exercise in *comparative statics* exemplifies an old idea: if people have to take care of their own needs on their own they are worse off than if they do something for others in exchange for a favour. What is needed is for people to be competent in different ways; a person who is more competent in one field will undertake that activity for all the others. As Hicks put it: "the gain from trade is the difference between the value of things that are got and the value of things that are given up" (Hicks, 1959, p.181, quoted by Thirlwall, 1972, p.360). There are also interesting consequences in terms of growth, though in the simplified graph in Figure 5 they are not made explicit. Imagine, for example, that the two goods A and B are actually two activities (as in the Grossman-Helpman model, 1991). The objective of one is to produce goods and services to satisfy immediate needs; the objective of the other is to generate "ideas", such as new technological or organisational solutions that can be

applied to the production of goods and services in order to increase their efficiency and effectiveness. In this case, the indifference curve from the consumer side would lie between satisfying immediate needs and satisfying them more effectively in the future, rather than between two goods. The division of labour according to comparative advantages would thus allow economic systems that are open to trade to gain greater benefits today without affecting growth, and vice versa. There remain the classic arguments, both of which were introduced by Adam Smith in his 1776 Treaty. These concern, first, the dynamic gains in trade, linked to increasing returns of scale of a dynamic and static nature, which an increase in demand would allow compared to the level permissible in autarchy, and, second, "vent for surplus".

The second argument, exemplified in Figure 6, regards *adjustment costs*. This argument concerns the effects of relinquishing economic policy tools when a country accepts joining monetary union. The previous exercise in comparative statics was highly simplified. In particular, it assumed that reducing internal production of A instantly brings about an increase in the production of B so that productive capacity is maintained to the full because imports of A are instantly balanced by exports of B. The exercise does not allow for construction time. Let us imagine, more realistically, that imports of A come in first because they are available at a lower price in the external competition. Owing to outside competition, there is an immediate drop in production of A in the domestic market but it takes time and money for resources that are not being used to find their way to the other sector, which must reorganise in order to increase its productive capacity and learn to sell on external markets. Let us imagine, however, that the gains in trade are positive even accounting for adjustment costs. The question is, since costs by definition precede gains, can a course of action that is efficient be implemented when the gains are greater in the long run but for the present the action is more costly? The traditional answer is yes, if all three tools of economic policy - fiscal, monetary and exchange rate policy - can be exercised. A reduction in interest rates can accelerate private investment and thus help reallocate resources and reduce unemployment caused by outside competition. This requires devaluating currency, which cannot be done in a regime of monetary union.

Mundell (1962, cited by Onida, 1974) first illustrated this point, exemplified in Figure 6. In the graph, G is public debt and *i* is the interest rate. Line XX indicates internal equilibrium with full employment and no inflation; line FF indicates external equilibrium of parity in the balance of payments. The initial position is at point *a*,

where, in order to have parity in the balance of payments, full employment and no inflation, it is necessary to exercise monetary and fiscal policies in order to obtain corresponding levels of public debt and interest rates. If it becomes necessary to reduce interest rates in order to accelerate investments in regions that are required to reorganise as a result of external competition, then, as the graph illustrates, the straight line FF must be shifted to F'F'. If this shift does not take place there is a risk of internal or external disequilibrium. Devaluating the national currency affects the shift, but in monetary union this cannot take place, as there is no longer a national currency.

Fig. 6 – Economic adjustment policy
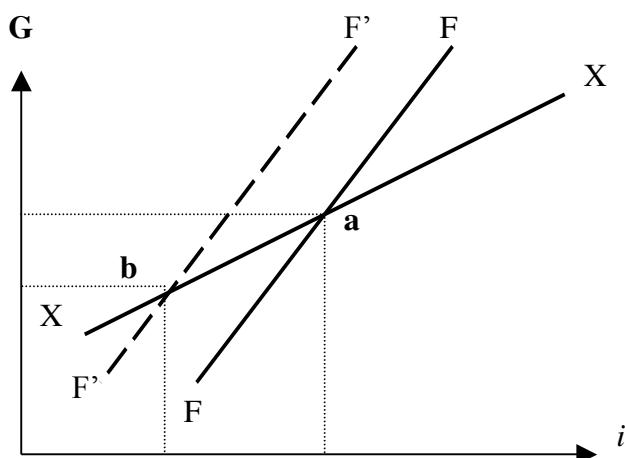
There are two alternatives, both of which are quite problematic. One can rely on the mobility of production factors, increasing mobility as the EU has sought to do. It can thus be hoped that the time and cost of adjusting is reduced, making specific policies, such as the reduction of interest rates, less necessary. Otherwise, one can count on external flows of capital (that would shift FF to F'F' on its own without devaluating), which would make a trade balance deficit sustainable. Capital flows, however, do not tend to find their own way to regions that need restructuring most. It is therefore a matter of implementing specific economic policies in order to sustain countries and regions that are in the throes of adjustment financially. This is precisely what the EU did, but already in the 1980s general opinion was that these two solutions were not the best possible. Increasing mobility in productive factors implied

reforms in the labour market[50] that were socially costly, and reforms in internal and international financial systems that met with great resistance. Any politician that supported these radical reforms was seriously at risk.

At the same time it was also clear that EU economic and spending policies could not be determined by needs. They were the result of the – mostly political – will of single member countries. As far as the prospects for regional imbalances after European integration were concerned, it was soon clear that two opposing elements had to be considered: potential gains in trade and adjustment costs. It was equally clear that it would be impossible to predict which of the two would prevail. Hence the idea that European regional cohesion policy would forestall future disparities and perhaps even support processes leading to convergence.

It could be added that, in the years that followed, concern about regional disparities was amplified by the findings of NEG[51] as well as by further arguments put forward by theories of modern growth (Armstrong, 2002)[52], specialisation processes (Krieger-Boden, 2002), and internal or external shocks of a political or institutional nature (Mazzola, Fazio, Lo Cascio, 2002)[53].

European integration and cohesion: policies and trends

Since the beginning of the 1990s, research has focussed on the entity and nature of regional disparities within the EU in order to shed light - implicitly or explicitly - on the scope of cohesion policy and to clarify whether there was a real need for it. While fears that regional imbalances would worsen as a result of EU liberalisation (and

---

[50] These reforms went in the direction of deregulation and met with great resistance. Major European players are mostly reconsidering them. The "regionalist" approach of the SEA until a few years ago accepted that resistance was a reality, while the Sapir Report criticizes "policies oriented towards maintaining what is socially regarded as an adequate degree of cohesion may reduce the efficiency and growth benefits arising from market liberalisation and integration. [… because] national labour market regulations which limit the effective degree of intra-country labour mobility across regions, sectors and/or firms in an attempt to reduce interpersonal inequalities tend to run counter to the requirements for a dynamic economy and thus reduce growth potential. (p.72)
[51] Used by Sapir in his Report, see p.106
[52] Armstrong discusses economic theory forecasts regarding the effects of convergence and polarization at a regional level in the light of: neoclassical theory of convergence (absolute and conditioned); the theory of endogenous growth; radical post-fordist theories; social capital theory; NEG models, models of growth encouraged by exports. His "eclectic" hypothesis is that the prevailing dynamic forces could be favourable to very slow convergence among regions, but could contribute to divergence at local sub-regional level.
[53] In reality, economic activity is never completely in "equilibrium". Apart from aggregate models that include cumulative processes, the literature that studies the effects of disturbances is relevant. These disturbances can be distributed in different ways in different local systems, and different local systems can react in different ways to them. Under enlargement and opening markets these disturbances could well become more frequent and more intense.

enlargement) were well founded, economists were divided as to the real effects, especially in the long term. The debate justifies the lively (but as we shall see as yet insufficient) production of studies and research on the topic. From a political point of view, moreover, the legitimation of cohesion policies was imperative, taking precedence over strictly economic arguments. The hypothesis that has gained ground is that the objective of cohesion was supported for reasons over and above the matter of developing lagging regions. If it is reasonable to claim that cohesion was a necessary element in the Single European Act establishing the single market, then it is not hard to appreciate the need to legitimate a European institutional system (a new axis between the European and the local level against the "conservative" resistance of national states) that would support its implementation. The system, it could even be said, was more significant for the future of the EU than the rapid reduction of national economic and social differences (Lebessis-Paterson, 1999).

This, together with the structural uncertainty regarding the tolerability of regional disparities, helps explain why there is often a gap between intentions and reality.[54] One important finding is how inconsequential EU policies for redressing imbalances were. In the US (in the 1990s) internal regional imbalances were two or three times smaller than in the EU where, after the 2004 enlargement, imbalances grew dramatically. Moreover, while in the EU imbalances grew throughout the 1980s and 1990s, in the US they diminished (Armstrong, 2002, p.247), just as faster growth in the US started to diverge from slower growth in the EU (Denis, McMorrow, Röger, 2004). This phenomenon has caused researchers and observers to claim that regional imbalances depend on overall growth; the gap closes when the EU enjoys sustained growth and widens when growth is slower. The role of "federal" redistributive policy, however, cannot be neglected[55]. As Espansa (2002) pointed out, in the US the federal budget represents 20% of the GDP and its redistributive capacity, in advanced regions as much as in lagging regions, is between 20 and 30%, while the EU budget represents 2% of its GDP and its redistributive capacity is estimated to be between 4

---

[54] It could, after all, have been considered important for the EU to establish direct relations with regional governments, whatever the economic resources available.
[55] This policy is neglected in the literature that examines the extent of, and reasons for, EU-US divergence in growth rates. See, for example, Denis, McMorrow, Röger (2004) in a report for the Commission that adopts an approach calculating growth, rather than considering the importance of local development. It may well be true that greater growth rates in the US (owing to capital accumulation and TFP) can be partly explained by the acceleration of growth in once backward regions, that this acceleration was encouraged by a federal redistributive policy, and that this policy was particularly successful in the last two decades.

and 6% [56]. This means that in the EU the problem is not that policies are not redistributive.[57] Their redistributive capacity can be expressed by the indicator 2-3 (4% and 6% divided by 2), while the same indicator is 1-1.5 in the US and 1-1.7 in Canada. The problem, rather, is that the EU budget is minimal, ten times lower than that of other "federal" countries. The comparison would seem to highlight little commitment on the part of the EU to redressing territorial disparities, even considering the fact that EU structural funds require recipients to co-finance regional policies and that the total resources available for regional policies are therefore the sum of EU and national funds. It could be observed, then, that EU regional imbalances are much greater than in the US, and that policies to redress these imbalances are less effective in the EU, though a lower budget[58] means that in absolute the difference is smaller.

Trends in regional disparities have been modified in the EU by regional policies, but probably less so than in other federalist countries. Let us now examine these trends, which paint a highly complex picture. Despite a lack of essential data that makes any conclusive results difficult to come by, a few clear tendencies can be outlined. Research[59] has unanimously stressed the fact that the gap between regions is greater than the gap between countries.[60] There is also agreement in documenting reductions in regional gaps in terms of product per employee during the 1960s and 70s, though the limited quantity and quality of available data has led to observations that these results are inconclusive. Research has shown that if convergence took place it was "conditioned". That is, regional economic systems that started off at significantly different stages of development did not converge towards a single level of product per employee, but towards various levels according to the varying natures of their production structures, human capital, infrastructure, and socio- institutional

---

[56] In Canada the federal budget is 18% of GDP. Its redistributive power is between 20 and 30%.

[57] Other researchers have made similar finds. For an updated bibliography, see Acconcia, Espansa, Leonida, Montolio (2003).

[58] In these calculations and comparisons territorial rebalancing policies in single member countries rather than over the whole of the EU are not usually taken into account. These policies are generally allowed and maintained by the EU, within certain guidelines and limits. Whether it is necessary to calculate them in order to make a correct comparative evaluation against other federalist countries remains to be seen. These national policies, however, could be considered independent of the integration process. Moreover, they are different in each country.

[59] Armstrong (2002) examines 23 from Barro and Sala –I –Martin (1991) to Cuadraro-Roura, Mancha-Navarro, Garrido-Yserte (2000). More have joined the ranks in the last few years. For an update, see Corrado, Martin, Weeks (2005).

[60] Hamburg is the richest region, with a per capita income 5 times higher than the poorest in Greece. Germany's per capita income as a whole is not even twice that of Greece.

environment.[61] Finally, more recent research has observed a tendency for regions to form groups, or "convergence clubs", some of which coalesce across national barriers. Non-parametric studies, which are more reliable in their findings for multiple polarising processes[62], have made similar observations.

The complexity of these findings makes interpretation highly challenging. One question is: was the early tendency to converge the real long-term trend, and its subsequent holding back a temporary deviation, or vice versa? As yet, research has not been able to provide an answer (Cuadrado, Roura, Parellada, 2002).

To conclude, the only alternative, as Corrado, Martin, Weeks (2005)[63] pointed out, is to valorise trends that on average converge - though very slowly and highly conditioned by their starting points - in clusters that are at least in part spatially contiguous. In short, it would seem that regional disparities persist (convergence in clusters does not impede divergence among clusters, even if it contributes to average convergence) even when a process of economic integration has progressively led to a single currency and to a certain extent has redressed imbalances.

Bibliography for Chapter 1 Appendix 1

Acconcia, A., M. Espansa, L. Leonida, D. Montolio (2003), *EU Policy, Regional Growth and Convergence across European Regions Evidence from Non-parametric and Semi-parametric Approaches*, mimeo University of Naples Federico II, University of Barcelona, University of Calabria, University of York.

Altomonte, C. e M. Nava (2005), *Economics and Policies of an enlarged European Union*, Edgar Allan Publisher, New York.

Armstrong, H.W. (2002), European Union Policy: Reconciling the Convergence and Evaluation Evidence, J.R. Cuadrado-Roura, M. Parellada (a cura di) *Regional*

---

[61] Considering per capita income (instead of productivity), the disparities seem to have increased owing to an increase in unemployment and absence from the labour market in lagging regions.

[62] Beaumont, Ertur and La Gallo (2003) discuss the ability of standard econometric methods to grasp polarising effects. They conclude, through an empirical analysis conducted with enhanced methods, that the formation of regional clusters cannot be excluded and that there are important spatial connections. See also Corrado, Martin, Weeks (2005), who find in favour of clusters.

[63] See Cheshire, Gordon (1998), Cheshire (1999), and Cheshire, Magrini (2002) for further theoretical conclusions. The statistics in this book are all from the 'Regio' statistics of Eurostat. Some have been integrated with data from the Crenos databank at the University of Cagliari (Paci, 2000; Cuadraro-Roura, Mancha-Narvarro, Garrido Yserte, 2002). The "Regio" department provides a wide range of information on the variegated aspects of European regional economies. However, further back in time the data is less exhaustive. It does not therefore afford the opportunity to calculate the variables that contributed to the 'initial' situation in the different regions, nor do the statistics go back far enough in time to appreciate the long-term effects, even for very few variables. The empirical exercises based on time series at various levels of territorial disaggregation reveal symptoms of clustering but are therefore unable to describe them in detail. To do this data regarding more variables for single regions for longer periods of time would be necessary.

Convergence in the European Union: Facts, Prospects and Policies, pp. 231-272; Springer-Verlag, Berlin, Heidelberg, New York.

Baldwin, R. e C. Wyplosz (2004), *The Economics of European Integration*, Mc. Graw Hill, London.

Barro, R.J. e X. Sala-i-Martin (1991), Convergence across states and regions, *Brookings Papers in Economic Activity*, 1, pp. 107-182.

Baumont, C., C. Ertur, J. Le Gallo (2003), Spatial Convergence Clubs and the European Regional Growth Process 1980-1995, B. Fingleton (a cura di), *European Regional Growth*, pp. 131-158, Springer-Verlag, Berlin, Heidelberg, New York.

Begg, I. e D. Mayes (1992), Cohesion as a Precondition for Monetary Union in Europe, in Ray Barrel (ed.) *Economic Convergence and Monetary Union in Europe*, pp. 220-240, Sage Publications London.

Capriati, M. (2003), Spesa pubblica e capacità innovativa delle regioni italiane, *Rivista economica del Mezzogiorno*, XVII, 4, pp. 735-772.

Capriati, M. (2005), *Spesa in R&S e sviluppo locale: prime evidenze empiriche*, mimeo, University of Bari, paper presented at the XXV Conference of Regional Science.

CEC - Commission of the European Communities (1973), *Report on the regional problems of the enlarged Community* (Thomas Report), Brussels.

CEC – Commission of the European Communities (2001), *European Unity, solidarity, diversity for Europe, its people and its territory; Second Report on Economic and Social Cohesion*, Brussels and Luxemburg.

Cheshire, P.C. (1999), Cities in Competition: Articulating the Gains from Integration, *Urban Studies*, 36, 5-6, pp. 843-864.

Cheshire, P.C. e I.R. Gordon (1998), Territorial Competition: some lessons for policy, *The Annals of Regional Science*, 32, 3, pp. 321-346.

Cheshire, P.C. e S. Magrini (2002), The Spatial Economic Impact of Euroland and the Implications for Policy, J.R. Cuadrado-Roura, M. Parellada (a cura di) *Regional Convergence in the European Union: Facts, Prospects and Policies*, pp. 211-230; Springer-Verlag, Berlin, Heidelberg, New York.

Corrado, L., R. Martin, M. Weeks (2005), Identifying and Interpreting Convergence Clusters across Europe, *The Economic Journal*, March 2005, 115, 502, pp. C133-C160(1).

Cuadrado-Roura, J.R. e M. Parellada (2002), Regional Disparities and Regional Convergence Problems in the EMU: Introduction, J.R. Cuadrado-Roura, M. Parellada (a cura di) *Regional Convergence in the European Union: Facts, Prospects and Policies*, pp. 1-22; Springer-Verlag, Berlin, Heidelberg, New York.

Cuadrado-Roura, J.R., T. Mancha-Navarro, R. Garrido-Yserte (2000), Regional productivity patterns in Europe: An alternative approach, *Annals of Regional Science*, 34, pp. 365-384.

Cuadrado-Roura, J.R., T. Mancha-Navarro, R. Garrido-Yserte (2002), Regional Dynamics in the European Union: Winners and Losers, J.R. Cuadrado-Roura, M. Parellada (ed.) *Regional Convergence in the European Union: Facts, Prospects and Policies*, pp. 23-52; Springer-Verlag, Berlin, Heidelberg, New York.

Dehem, R. (1972), Le mirage monétaire européen, son coût et ses aléas, *Recherches économiques de Louvain*, September.

Denis, C., K. McMorrow, W. Roger (2004), *European Economy: An analysis of EU and US productivity developments (a total economy and industry level perspective),* European Commission Economic Paper N° 208 July.

Espansa, M. (2002), The Territorial Redistributive Power of the EU Budget: Empirical Evidence at National and Regional Level, J.R. Cuadrado-Roura, M. Parellada (ed.) *Regional Convergence in the European Union: Facts, Prospects and Policies*, pp. 121-146, Springer-Verlag, Berlin, Heidelberg, New York.

Grossman, G.M. e E. Helpman (1991), *Innovation and Growth in the Global Economy*, Mit Press, Cambridge (Ma).

Greunz, L. (2003), The Technology Gap and European Regional Growth Dynamics, B. Fingleton (ed.), *European Regional Growth*, pp. 241-266, Springer-Verlag, Berlin, Heidelberg, New York.

Hicks, J. (1959), *Essays in World Economics*, Claredon Press, Oxford.

Johnson, H.G. (1962), The Economic Theory of Custom Unions, *Pakistan Economic Journal*, 10, 1, pp. 14-32.

Krieger-Boden, C. (2002), EMU and the Industrial Specialisation of European Regions, J.R. Cuadrado-Roura, M. Parellada (a cura di) *Regional Convergence in the European Union: Facts, Prospects and Policies*, pp. 77-94; Springer-Verlag, Berlin, Heidelberg, New York.

Le Gallo, J., C. Ertur, C. Baumont (2003), A Spatial Econometric Analysis of Covergence Across European Regions 1980-1995, B. Fingleton (a cura di), *European Regional Growth*, pp. 99-130, Springer-Verlag, Berlin, Heidelberg, New York.

Lebessis, N. e J. Paterson (1999), *Improving the Effectiveness and Legitimacy of EU Governance*, European Commission Forward Studies Unit Working Paper, Cdp (99) 750.

Mazzola, F., G. Fazio, I. Lo Cascio (2002), Regional Asymmetric Reactions to Shocks in EMU: an Assessment of Different Approaches, J.R. Cuadrado-Roura, M. Parellada (a cura di) *Regional Convergence in the European Union: Facts, Prospects and Policies*, pp. 95-120; Springer-Verlag, Berlin, Heidelberg, New York.

Mundell, R.A. (1962), *The Appropriate Use of Monetary and Fiscal Policy Under Fixed Exchange Rates*, IMF Staff Papers, march.

Onida, F. (1974), Aree monetarie ottimali e unificazione monetaria europea, in L. Giusso (ed.), *Teoria delle unioni monetarie e integrazione europea*, pp. 169-199, Edizioni Scientifiche Italiane, Napoli.

Paci, R. (2000), *Convergenza e divergenza tra le regioni europee. Implicazioni per lo sviluppo economico in Sardegna*, Osservatorio Economico del Banco di Sardegna.

Samuelson, P.A. (1939), The gains from international trade, *Canadian Journal of Economics and Political Science*, 5, pp. 195-205.

Sapir, A., P. Aghion, G. Bertola, M. Hellwig, J. Pisani-Ferry, D. Rosati, J. Vinals, H. Fallace, M. Buti, M. Nava, P.M. Smith (2004), *Europa un'agenda per la crescita*, Il Mulino, Bologna.

Thirlwall, A.P. (1972), *Growth and Development*, MacMillan, London.

Usai, S. e R. Paci (2003), Externalities and Local Economic Growth in Manufacturing Industries, B. Fingleton (a cura di), *European Regional Growth*, pp. 293-322, Springer-Verlag, Berlin, Heidelberg, New York .

Vaya, E. e J. Surinach (2003), The Spatial Distribution of Labour Productivity in the European Regions: A Study in Spatial Econometrics, B. Fingleton (ed.), *European Regional Growth*, pp. 183-214, Springer-Verlag, Berlin, Heidelberg, New York.

Chapter 1 Appendix 2 - Getting growth going: industrial districts

As we have seen, NEG does not provide answers to the question of how development in a lagging area is set in motion. It has demonstrated that development can be self-reinforcing once it has been set in motion, but the trigger has not been found. However NEG studies of agglomeration mechanisms are a good start. NEG claims that one of the factors that explain development disparities between different areas is the mobility of firms and their resources in the modern, industrial sector. This mobility implies that firms and resources - all other conditions being equal - tend to go where there is development and abandon areas where there is none, generating agglomeration processes. This implication leads to two considerations. First, a factor or mechanism that gets growth going (whatever it is) must be strong otherwise it will not be successful. NEG warns that where there is underdevelopment the level of wellbeing, however low, is not normally guaranteed. Rather, underdevelopment is seen as an evolving process that tends to deteriorate. Second, NEG sheds light on a way to promote development in lagging regions. If mobile resources tend to abandon the region, thus widening the gap with more developed regions, then the engine of change must lie in resources that cannot be shifted, that is, immobile resources.

The valorisation of immobile resources in underdeveloped economies becomes a feasible option especially when real life does not reflect the simplifications of the model -, the hypothesis of full employment, or that of perfect immobility of labour in the traditional sector and perfect mobility of labour in the modern sector, to name a few. A lagging area, in fact, usually has high levels of unemployment, and large numbers of workers abandon agricultural areas and traditional activities rooted in the land. At the same time many industrial sectors rely on specific skills equally rooted in the territory; these workers would be less productive if they were employed elsewhere.

If we introduce the concept that there is under-employment in the traditional sector and that some workers can move, then the lagging area is by definition impoverished unless it undergoes modernisation and increases its productivity. The same holds for the modern sector: a contributing factor to development would be to recognise that labour is not completely mobile. Development strategies should therefore aim to curb

losses in human resources by incrementing productive employment rooted in the local context.

At first sight, it would seem that the experience of Italian industrial districts, about which there is a vast literature reconstructing their history and describing how they work, might serve our purposes. Anticipating our conclusions, the answer, however, is that it does not. Neither NEG nor the literature on industrial districts solves the problem of identifying triggers for development in lagging areas.

Let us examine to start with the morphology of a successful local system as expressed in an "evolved industrial district". After years of research starting in the 1970s[64], Becattini (1978, 1979) analysed industrial districts as systems with three levels: "the productive apparatus in the strictest sense; the institutions linking the productive apparatus [...] to the district's community; the formation and transmission of values that are at the foundation of district-type behaviour." (Maccadelli, Sforzi, 1997, pp.259-62). At each of these three levels (interacting with one another) there are elements and mechanisms that are specific to districts. These are the distinctive resources that are used effectively in that district alone, any one of which, if it were taken away, would bring about the collapse of the whole district. In this sense, to adopt the concepts of NEG, the district's resources are immobile and defend the industrial district from agglomeration processes that may be taking place in other areas.

At the top of the list of specific resources there is one essential feature: districts operate as complex units. The distinctive element of their resources is that they interact at all three levels: at the level of productive apparatus, of the institutions linking the productive apparatus to the community, and of the core values of district-type behaviour produced and reproduced by the elementary institutions of the district (family, parish, infant schools).[65] This close and specific interaction between productive apparatus, institutions and values can be described (in our analytical bridge between a paradigm of industrial districts and new regional economics) as a

---

[64] Becattini's early systematic works (1978, 1979) are always quoted. In fact, Brusco (1989, p.461) refers that Becattini's first work, when he "for the first time in Italian economic literature since the war, claimed that systems of small firms could be vital and competitive and that it was a mistake to interpret them as residue's of productive systems destined to failure" was in 1969, in a volume published by Irpet (Tuscan Regional Institute for Economic Planning)

[65] This kind of interaction is present in all forms of economic organization because they are human constructs (Polany, 1957, pp.305, 307). In industrial districts it must be especially close and specific, as Becattini always underlines. It is therefore different in each district and different from other territorial economic organizations.

'set of immobile resources coherent with the system' or as the 'identity' of the district.

These resources include:

- Teams of firms belonging to the district organised around central firms that maintain links among districts and with outside demand continuously topped up by a pool of unconnected firms.[66] The presence of open, competitive markets for goods from different production phases proves that these resources are abundant. These could be described as 'diffuse immobile resources of specifically oriented productive organisations.'

- Price systems for diverse local markets regulated (and not completely free) between a minimum, to guarantee the social reproducibility of the district, and a maximum, to guarantee its external competitiveness. This regulation is undertaken by local organisations representing interest groups. These could be described as 'immobile resources of interested organisations to mediate between income distribution with social consensus and the competitiveness of the district.'[67]

- Tacit or contextual knowledge (of a productive or professional nature) and the capacity to renew it constantly through interaction with codified knowledge, which we describe as 'immobile resources including knowing how to do and knowing how to learn as mediation between tradition and technical progress' (Becattini, Rullani, 1993).

- A set of basic or elementary institutions (family, parish, infant schools) and a political and cultural context able to transmit values such an entrepreneurial spirit, a healthy desire to make money, an aptitude for competing but also for cooperating, a keen appreciation of technical and professional improvement for the new generations: 'immobile resources related to value systems that function as means of reproduction for districts.'[68]

---

[66] The use of the term 'firms' (rather than entrepreneurial resources) underlines the fact that they combine different productive factors. Firms are thus mobile resources, but immobile inasmuch as the specific combination is rooted in the territory.

[67] It could be said that the conditions for districts to be reproduced are not only linked to long-term profit rates, but also to long term 'consensus' rates.

[68] In the literature on districts, the social reproduction of these values, together with the close cohesion between the productive, institutional and cultural levels, are generally associated not only with the specific entrepreneurial supply, but also to the climate of trust that reduces transaction costs. See, for example, Dei Ottati, 1994.

With NEG models in mind, the concept that emerges is that in an evolved industrial district all these resources are immobile (in the sense that they are specific to the territory) and therefore safeguard evolved districts from losing their resources to self-reinforcing agglomeration processes taking place in other territories. If anything, they transform these districts into centres of attraction for other areas lacking these resources. Another concept that becomes evident is that these districts are case studies in dynamic scale economies (learning by doing) supported by external demand (competition). Immobile resources are thus exploited to the full: outside firms constantly top up teams of firms in the district; the price system achieves greater efficiency in allocating resources; choices involve investments and professional training; know-how is renewed, constantly interacting with codified knowledge; the political and social context is in symbiosis with a constantly evolving value system, which in its turn feeds the entrepreneurial spirit appropriate to and specifically sought by the district. A final observation is that the economic mechanism and its cultural dimension[69] are intimately linked, to the extent that culture comes first; without it the economic mechanism could not exist.

Having outlined the main features of evolved industrial districts, we can now go on to address the matter of transferring the district model as a possible solution to the problem of local development in a lagging area. Are immobile resources the secret?

The literature on industrial districts has spent many years[70] studying the question of whether the industrial district model can be replicated and what conditions (and resources) would allow the model to be transferred to lagging areas. As early as the 1980s, and until quite recently, experts in industrial districts were sceptical; the literature contains explicit (though rare) statements to the effect that districts cannot

---

[69] It is clear in Becattini's argument that the cultural dimension relevant for a district is closer to Braudel's idea (who stresses the idea of conceiving oneself and society, know-how, market and production relations, as well as group and family kinships. See Tenenti, 1988, p.12), than to culture, mentality and religious views as such.

[70] In Brusco's reconstruction (1989, pp.435, 436-7) it is necessary to go back to the early 1980s. "In the early 1980s the problem of designing interventions to encourage the growth of districts was raised for the first time. […] This issue could not be put off, because the literature on districts, both Italian and international, with exponents such as Sabel, Zeitlin and Piore, had expressed great interest – more abroad than in Italy it must be said – for the Emilia model (or the North-East-Centre model, NEC, the system area, the third Italy, whatever you want to call it). Experts abroad, researchers and policy makers, considered the problem and asked us questions concerning the transferability of the model. They asked us whether the model could be replicated in order to increase the productivity of, say, textile firms in Nottingham, or tiny knitting firms in Ridgewood, Queens, or small metal working firms in Detroit. Not to mention, though this has never been a concern for those responsible for public policy in Italy, whether the model could be used to encourage development in the backward Mezzogiono of Italy" (p.468).

be reproduced.[71] The highly specific features of the industrial district model lead us inevitably to the same conclusion, precisely because their cultural identity is so well defined and perceived as the result of a long history and tradition.[72]

In a 1997 interview, Giacomo Becattini said that the debate over whether industrial districts could be reproduced was still open, and he listed the resources that in his view would make it possible. In his view, a new aspect in the 1990s was globalisation, which provided two new directions. First, incomes continued to grow, creating greater and more differentiated demand for durables for personal and domestic use (and for the technologies used to produce them). Second, social integration[73] became a more important issue, perhaps as a reaction to a growing perception of fear and insecurity.

As far as resources were concerned (specific immobile resources in the backward area to be valorised), Becattini claimed, further interpretation was required, but an attempt was made to identify what was needed. Italian industrial districts developed in

---

[71] Perhaps, as Brusco remembers, the requests were too pressing.

[72] Brusco is again very clear on this point. In his introduction (1989) to a new edition of a previous essay (Brusco, 1989), he notes: "from research on domestic production and on the metal working sector in Bergamo, and above all, from all the studies that contributed to the piece on the Gilbrat law and the Emilia model, I was always very cautious about using the term 'culture'. I was convinced that everything could be explained in terms of mutual economic convenience between the parties involved in the exchange. […] The essay that follows, however, asks the question: what is the culture of an industrial district and where does it come from? And yet, it does not answer the more important question: can this blend of consensus and know-how be reproduced? My essay, after following a different path, ends up converging with Becattini's argument" (Brusco, 1989, pp.467-8). This convergence was contained at the end of the essay: "The development of industrial districts in the Veneto and Marches – where Christian Democrat administrators are certainly less efficient than in Emilia – is sufficient proof that this way of organizing productive processes in a territory is closely linked to the way basic forces operate, throughout the social fabric. […]These are the forces that […] have stimulated the spread of districts. Probably, in Emilia, local institutions have accelerated their development. But the main referents, if we want to understand and explain what has taken place, are not experts in industrial policy who plan incentives for depressed areas, but Hirschman and Braudel (Brusco, 1989, p.486). he reference to Hirschman cannot be only to his "linkages" (that belong strictly speaking to the productive apparatus). It must be to his more complex work, as an economic historian, development economist and economist-sociologist who "questioned the parsimonious paradigm of traditional economics: that of the isolated individual who is moved only by personal interest and who chooses freely and rationally between alternative courses of action having calculated that costs and benefits involved. […] Hirschman contested this approach and proposed as an alternative to incorporate gradually into the economic discourse a series of features and tensions that are part of human nature and have been forgotten until now" (Meldolesi, 1988, p.10). As for the reference to Braudel it is clearly to the great historian-economist who dwelt on the concept of large scale social change over long periods of historical time (see Kinser's 1981 essay), but also to his idea that "the engine of human development should be sought [not only] in terms of technological evolution [but also] in terms of the evolution of productive organization, trade and the way social groups aggregate" (Tenenti, 1988, p.12).

[73] By "growing demand for social integration" Becattini seems to mean a widespread need to reassert and defend (local) identities and sense of belongingness to reassure people in the new context of globalization (including factors such as immigration, loss of control over economic and social choices owing to super-national supervision, and cultural homologation).

the post-war period as a result of the region's entire history, territorial know-how, values and institutions. In many lagging areas in Italy, by contrast, there may have been a few factories (perhaps the result of previous failed national interventions such as the 'special programme' in the Italian Mezzogiorno[74]), a handful of craftsmen's workshops, self-production, domestic production and some conurbations. The values and cultural context were inadequate, but a surrogate factor was able to take over and allow districts to develop even in lagging areas. This factor was the growing demand for social integration.

There is no contradiction then between the claim that districts could be reproduced under certain circumstances and the statement that their value system could not be exported. The proliferation of Italian industrial districts in the post-war period was simply a first wave. A second 'spontaneous' wave is still underway today, stimulated by a 'world demand ' for industrial districts, and sustained by the same entrepreneurial resources and place-based know-how that Italy's countless cities and sub-regions are so full of.[75]

---

[74] The 'special programme' failed in its main objective, which was the industrialization of the Italian Mezzogiorno with industrial poles, but did leave behind various factories, each with long and tormented histories. In most cases, the big Italian firms have withdrawn. However, it should not be forgotten that in several successful industrial districts in the Centre and North-East of Italy, traditional big firms historically were important sources for the dissemination of technical and entrepreneurial know-how. An econometric study conducted at provincial level has shown that the variable 'employees' in industrial firms with more than 1,000 workers after 1937 was very significant in explaining variations in autonomous industrial labour between 1951 and 1971 (Seravalli, 1999, pp.76-77). Brusco made similar observations in his analysis: "Almost always at the origin of an industrial district there are big firms that have, even a long time before, been operational there. […] There are many examples, many of which are extremely clear. […] The logic of this pattern of development is evident: big and medium-sized firms, with their daily laboriousness, manage to introduce the necessary technical and professional competencies in a largely agricultural social context with few links with outside markets. Workers then learn how to manage the productive process, white-collar workers forge links with suppliers and other elements of the market for the product. Under certain circumstances, blue and white-collar workers tend to transform themselves into autonomous workers, transferring the skills they learned in the factory to their own productive units" (Brusco, 1986, pp.482-3).

[75] This is one step further than Brusco's three types of artisans (traditional craftsmen, dependent decentralization, independent decentralization) who co-exist but typically are dispersed. The first group are usually found in "in Puglia or Sardinia [the second group] in Lombardy and Piedmont, [the third group] around Bologna Prato, Venice an d Ancona. These should not be seen as three stages of evolution; in fact there is no point in hoping that aid given to small firms in general will transform artisans in the South into pillars of a new industrial organization. […] Without the necessary markets and experience, this aid can only make traditional craftsmen richer; it will never transform them into something different" (Brusco, Sabel, 1981, p.113). This point of view reflects a lack of trust in incentives, as well as, more importantly, Brusco's conviction that traditional craftsmanship is of a completely different nature and cannot be transformed into the independent artisan of industrial districts.

This kind of 'impotent optimism'[76]is troublesome. The problem is that the approach of this analysis, based on the genesis and development of Italian industrial districts, almost entirely excludes the influence of intentional policies, both national (central government) and local - especially if these policies were never part of the long-term history[77] of the territories involved. This is, in our view, unjustified. The literature of industrial districts has hardly ever conducted research into national or local policies supporting small firms. There are inevitable objective difficulties in terms of analysis and measurement[78], but that does not excuse a priori judgements.[79] A symptom of this is the literature's cautious, slow – but nor imperceptible – march towards including the role of policy. And yet, there has been very little research in recent years.[80] Although this marks a step forward compared to the reductive

---

[76] Optimism because – as, paradoxically, certain neoclassical analyses of growth pointed out – everything adjusts itself naturally; Impotent because there is really nothing important that can be done intentionally.

[77] Putnam, Leonardi, Nanetti (1993) had a great (and disastrous, in our view) influence with their 'demonstration' that the institutions of the Italian regions were highly diverse, especially between North and South, and that this diversity correlated positively with 'civic sense'. These findings became a formidable basis for arguing that, even if local institutions had done something to support development, what they did had in some way to be explained away by the long-term history of the area that left different heritage and cultural attitudes in different areas which have to do with basic institutions rather than intermediary institutions, with custom and tradition rather than with policy making. More detailed empirical research shows that the opposite can also be true. That is, politics and institutions can create civic sense (Arrighetti, Lasagni, Raimondi, 2000).

[78] A very basic fact that is often completely ignored, for example, is that these policies take variable lengths of time, and anyway a long time, to have an effect. This means that conclusions as to their effectiveness are often over-hasty. A case in point is the policy of infrastructured industrial areas, which in the 1980s was quickly considered a failure. Today the situation is highly diversified and research is still underway, but the experience can by no means be judged a failure. For many of these areas the problem was simply that results took ten or fifteen years to come about rather than two or three years.

[79] Obviously industrial policies in favour of districts (already in place) are another matter. At times there is a certain amount of confusion. Italian experts did not neglect the matter of industrial policy. Fuà focused on it it (Fuà, 1983), as did Brusco (Brusco, 1984, 1991, Brusco, Righi, 1985), and Becattini (Becattini, 1998, for example). These experts even inspired a national law. And yet this does not mean that they really believe that autonomous intentional actions of politicians and institutions are relevant factors among other things in the birth of industrial districts.

80 It might be useful to consider some of the contributions. We have already mentioned Becattini. For him, basic institutions were local markets, local organisations representing interest groups, family, parishes and nursery schools. In the year 2000 Becattini published a history of Prato, which offers an extraordinarily rich historical reconstruction of the city. While his dominant idea of development is that it is spontaneous, bottom-up and supported by a specific historical context, when he poses the question of where competition for products made in Italy's industrial districts is going to come from, and what form it will take, his answer is that it will come from: "some replica, spontaneous or constructed, identical or adapted, of their organisational formula and their evolutionary path. […] Their chances are limited by barriers that have been largely (and understandably) unexplored. Basically, their chances lie in the possibility that some of the more established districts (or some specialised intermediary at a multinational level) finds it advantageous to use the new productive unit's workmanship, 'exploiting' its skills, educating and providing its workers with the appropriate tools, as a master would his apprentices. Otherwise, its chances lie in the possibility that a government that has understood what is going on adopts the objective of promoting industrial districts and is in a position to implement its

traditional approach of the early district experts, the literature[81] has not been able to get beyond the idea that policy counts, but only in creating a favourable 'climate'.

---

policy. The government policy would need to avoid the usual traps of paternalism and aid-dependency and help firms inside the embryonic districts, supporting research, training, technical and market information, quality control, increasing access to credit, tax benefits and foreign exchange relief, and, last but not least, weighing in with the 'international influence of the country' as well as its political and diplomatic channels so that districts gain access to markets that have been closed and protected." (Becattini, 2000, p.227.). The obvious question is: did this ever take place? As we have said, optimism has faded, and there is a greater sense of impotence. Policies of this kind might well have been at work in the history of Italian districts and in the city of Prato without being trumpeted around or without being recognised. Again in 2000, Signorini wrote: "If a question of public policy appears right from the beginning of the literature, it's usually at a local government level rather than at central government level. The question seems to refer to cultural aspects, in the widest sense, rather than to specific intervention policies. Local administrations, parties and social formations are, in all the classic literature on districts, an important part of the 'climate' of the district. The interaction between administrations and entrepreneurial forces, based on a common identity and shared values can play a determinant role. In concrete terms, a favourable attitude on the part of local administrations can be translated into regulations and programmes that take into account the requirements of the district in order for the area to develop. It can also contribute to the creation of infrastructure (roads, transport, waste disposal etc), as well as support collective ventures in promotion and training. Firms on their own find it hard to achieve the same results in a coordinated, efficient and socially acceptable way. (Signorini, 2000, p.xxxiii). There is an apparent contradiction between the idea that the role of local politics seems to count only to the extent that it contributes to the cultural climate of the district but not for its specific interventions, and the brief but highly concrete list of possible interventions provided. Putnam might inspire an interpretation. What counts, it seems, are the shared values; all the rest follows. There is never any doubt that the opposite may also be true. In 2001, Dei Ottati wrote: "We have just seen how the collective capital of trust is essential for the formation and reproduction of districts as vital socio-economic systems. Even though it depends on its past, inasmuch as it is based on belonging to the same social culture, this kind of trust is not independent of human action. […] It is anyway the result of a 'political' action in a wider sense. This statement is of vital practical importance, because it opens the door to the possibility of drafting intervention policies to revitalise industrial districts that are failing and to help new successful districts to come about in regions that are still economically depressed" (Dei Ottati, 2001, p.177). Here, finally, there is a different idea, though the phrase 'political action in a wider sense' dilutes the strength of the claim. E believe it is 'political' in the closest sense (Arrighetti, Seravalli, 1997).

[81] Industrial districts in Italy, according to the ten-year data collected between 1951 and 1991 that can now be compared, absorbed a share of industrial workers of 35-40% of the total (Brusco, Paba, 1997, p.320). If we adopt a wider definition (that includes firms operating just outside the borders of the specialisation filière), "in 1996, 44.8% of workers in Italian manufacturing were employed in districts. […] Between 1991 and 1996 the weight of districts in terms of workers employed there increased by 1.5% (Iuzzolino, 2000, p.19). In 1996, in 17 provinces the weight of workers in districts was more than 80% of the total manufacturing sector, and in 34 provinces it was more than 60% (data in Iuzzolino, 2000, p.13). In 2001, the preliminary analyses of the Census data with reference to local labour systems were published, revealing that for the first time the relative weight of workers had shifted in favour of medium-sized firms, which do not share the same features as industrial districts. These data are important, though industrial districts were probably still doing better (and had done better) than the rest of the industrial system. "The firms that survive do so for longer […] The market is highly particular. Young workers enter the system relatively early, and consider the first stages of their employment as apprenticeship. Investments in human capital are successively repaid, both in terms of salaries and in terms of contractual and negotiating power" (Casavola, Pellegrini, Romagnano, 2000, p.65). As far as exports are concerned, although a "worrying weakness in the model of Italian specialisation - which mostly leans on traditional sectors, […] and which does not have the advantages of concentration; […] which suffers from high congestion and an inadequate ratio between circulating vehicles and the road network - is one of the factors that negatively influences trade balances, nevertheless, industrial districts have played an essential role in compensating all these weaknesses" (Gola, Mori, 2000, p. 81). "The profitability of firms in industrial districts is significantly higher than that of single firms.

The industrial district paradigm, then, is doubtless useful, but only partially so. Looking more in general at the local development of small firms in Italy, the usefulness of the model is limited both by the fact that it neglects other typologies of production and by the fact that it underestimates intentional policies. These two failings have been revealed by historical reconstructions of the features of Italian economic development. The limits of the district approach are the result of the idea that immobile resources are geographically and historically specific. In this view, they are the result of the cohesion and functional coherence of their territory; the product of the area's elementary institutions. All of these features call to mind cultural connotations rather than intentional actions. Thus, some important immobile resources that play a recognised role in economic development, such as public or collective goods in the widest sense, are absent. The limited role of intentional policies also conceals the starting and finishing point of development. The paradigm of industrial districts is more successful in tracing both the relationships among agents of the district and their origins, but it has not managed to solve the problem of their transferability.

Bibliography for Chapter 1 Appendix 2

Arrighetti, A. e G. Seravalli (1997), Istituzioni e dualismo dimensionale dell'industria italiana, in F. Barca (ed.), *Storia del capitalismo italiano dal dopoguerra a oggi*, pp. 335–388, Donzelli, Roma.

Arrighetti, A., A. Lasagni, M. Raimondi (2000), *Capitale Sociale, Istituzioni e Rendimento Istituzionale,* mimeo Sezione di Scienze Economiche Università di Parma.

Becattini, G. (1978), The development of light industry in Tuscany: An interpretation, *Economic Notes*, 2-3, pp. 107-123.

Becattini, G. (1979), Dal settore industriale al distretto industriale: Alcune considerazioni sull'unità d'indagine dell'economia industriale, *Rivista di economia e politica industriale*, 1: Sectors and/or districts, some remarks on the conceptual foundations of industrial economics, in E. Goodman et al. (ed.), *Small firms and industrial districts in Italy*, pp. 136–152, Routledge, London, 1989.

Becattini, G. (2000), *Il bruco e la farfalla – Prato: una storia esemplare dell'Italia dei distretti*, Le Monnier, Firenze.

Becattini, G. e E. Rullani (1993), La natura e la dinamica dell'impresa distrettuale, *Economia e Politica Industriale*, 80, pp. 25 - 48.

Beccattini, G. (1998), *Distretti industriali e made in Italy*, Bollati e Boringhieri, Torino.

Indicators are 2-4% higher - size, economic activity and time factors being equal" (Fabiani, Pellegrini, Romagnano, Signorini, 2000, p.29).

Brusco, S. (1984), Quale politica per i distretti industriali? *Politica ed Economia*, XV, 6, pp. 68 – 72.

Brusco, S. (1986), Small firms and industrial districts – the experience of Italy, in D. Keeble e E. Wever (a cura di) *New firms and regional development*, pp. 184–202, Croom Helm, London: Piccole imprese e distretti industriali – l'esperienza italiana, in S. Brusco, *Piccole imprese e distretti industriali – una raccolta di saggi*, pp. 469–488, Rosemberg & Sellier, Torino, 1989.

Brusco, S. (1989), *Piccole imprese e distretti industriali – una raccolta di saggi*, Rosemberg & Sellier, Torino.

Brusco, S. (1991), La genesi dell'idea di distretto industriale, in F. Pyke, G. Becattini e W. Segemberger (a cura di) *Distretti industriali e cooperazione fra imprese in Italia*, Banca Toscana, Studi e Informazioni, 34, pp. 25 – 34.

Brusco, S. e C. Sabel (1981), Artisan production and economic growth, in F. Wilkinson (ed.), *The dynamics of labour market segmentation*, pp. 99–113, Academic Press, London.

Brusco, S. e E. Righi (1985), Enti Locali, politica per l'industria e consenso sociale; ora in: S. Brusco, *Piccole imprese e distretti industriali – una raccolta di saggi*, pp. 433–460, Rosemberg & Sellier, Torino.

Brusco, S. e S. Paba (1997), Per una storia dei distretti industriali italiani dal secondo dopoguerra agli anni novanta, in F. Barca (ed.), *Storia del capitalismo italiano dal dopoguerra a oggi*, pp. 265–334, Donzelli, Roma.

Casavola, P., G. Pellegrini, E. Romagnano (2000), Imprese e mercato del lavoro nei distretti industriali, in: L. F. Signorini (ed.) *Lo sviluppo locale – un'indagine della Banca d'Italia sui distretti industriali*, , pp. 51-66, Donzelli, Roma.

Dei Ottati, G. (1994), Trust, Interlinking Transactions and Credit in the Industrial District, *Cambridge Journal of Economics*, 18, 6, pp. 529 – 546.

Dei Ottati, G. (2001), Concertazione e sviluppo nei distretti industriali, in: A. Ninni, F. Silva e S. Vaccà, *Evoluzione del lavoro, crisi del sindacato e sviluppo del paese*, pp. 163–196, Franco Angeli, Milano.

Fabiani, S., G. Pellegrini, E. Romagnano, L. F. Signorini (2000), L'efficienza delle imprese nei distretti industriali italiani, in: L. F. Signorini (ed.) *Lo sviluppo locale – un'indagine della Banca d'Italia sui distretti industriali*, pp. 21-50, Donzelli, Roma.

Fuà, G. (1983), L'industrializzazione nel Nord Est e nel Centro; in: G. Fuà e C. Zacchia (a cura di) *L'industrializzazione senza fratture*, Il Mulino, Bologna.

Iuzzolino, G. (2000), I distretti industriali nel censimento intermedio del 1996: dimensioni e caratteristiche strutturali, in: L. F. Signorini (ed.) *Lo sviluppo locale – un'indagine della Banca d'Italia sui distretti industriali*, pp. 3–20, Donzelli, Roma.

Kinser, S. (1981), Annalist Paradigme – The geohistorical structuralism of F. Braudel, *The American Historical Review*, LXXXVI, pp. 63 – 105.

Maccabelli, T. e F. Sforzi (1997), Totalità e cambiamento il paradigma dei distretti industriali intervista a Giacomo Becattini; in: C. M. Belfanti e T. Maccabelli (a cura di), *Un paradigma per i distretti industriali, radici storiche, attualità e sfide future,* pp. 259-276, Grafo, Brescia.

Meldolesi, L. (1988), Introduzione – cinque tesi ed un excursus; in A.O. Hirschman, *Come complicare l'economia,* Il Mulino, Bologna.

Polany, K. (1957), *Trade and Market in the Early Empires Economies in History and Theory*, The Free Press, New York.

Putnam, R. D., R. Leonardi e R. Y. Nanetti (1993), *Making democracy work – civic tradition in modern Italy*, Princeton University Press, Princeton;

Seravalli, G. (1999), *Teatro regio, teatro comunale – società, istituzioni e politica a Modena e a Parma*, Donzelli, Roma.

Signorini, L. F. (2000), Introduzione; in: L. F. Signorini (ed.) *Lo sviluppo locale – un'indagine della Banca d'Italia sui distretti industriali*, Donzelli, Roma.

Tenenti, A. (1988), Introduzione all'edizione italiana; in F. Braudel, *La dinamica del capitalismo*, Il Mulino, Bologna.

Trigilia, C. (1986), *Grandi partiti e piccole imprese*, Il Mulino, Bologna.

Trigilia, C. (1991), Il lavoro e la politica nei distretti industriali della Terza Italia; in F. Pyke, G. Becattini e W. Segemberger (a cura di) *Distretti industriali e cooperazione fra imprese in Italia*, Banca Toscana, Studi e Informazioni, 34, pp. 173 – 196.

Chapter Two - Underdevelopment Traps and Social Traps[82]

Underdevelopment is hard to combat; the only thing to do is fight the forces that tend to preserve or aggravate it. A community or territory has at its disposal immobile resources, which are potential features of - though not necessarily forces for - local development, while competition between economic agents on it own does not generate adequate incentives for development.

The aim of this chapter is to illustrate the reasons for this market "failure". We shall see how a single person or a sole firm cannot valorise resources or opportunities, because underdevelopment traps and social traps act together to create obstacles for a lagging area's development, or to drive an area that has been struggling to escape it back into underdevelopment.

The first section presents a simple graph that illustrates underdevelopment traps. The second section shows that this issue has occupied, and continues to occupy, an important place in economic literature. The third section deals with the objection that these traps can be overcome by means of reduced labour costs, and argues that it is precisely the reduced cost of labour that creates social traps, which in their turn curb the development of a significant modern sector. In the fourth section, combined social and underdevelopment traps are discussed using a classical model with endogenous techniques. The fifth and final section examines consequences for development policies. The appendix illustrates the mechanism of social traps by means of a simulation of an artificial, multi-agent economy.

Underdevelopment traps, the mechanism

The concept of underdevelopment traps can be used in a context of backwardness where it is assumed there are immobile resources (natural and human) with low or inexistent productivity that could be put to better use. In this mechanism, the transition from unoccupied to more productive immobile resources does not take place incrementally. The transition only takes place if a sizeable block of resources

---

shifts at the same time - but nobody knows how big the block needs to be before the transition takes place. The situation becomes a trap because a small number of immobile resources cannot exit low productivity, while individual holders of resources cannot come to spontaneous agreement to transfer a block of sufficient size, since nobody knows how many would need to move.

In advanced economic systems, these reallocations take place all the time, without any difficulty because the market stimulates them. Resources are disinvested from activities with low profit margins and reinvested in activities with higher margins. There is, therefore, a substantial difference between advanced economic systems and lagging ones. In advanced systems, the opportunity for resources to be used is wider and more articulated than in backward ones. A sole firm can disinvest from one activity and invest in another without incurring excessive costs. There are already many different operators in all fields and thus knowledge and skills can be easily acquired. In lagging areas, by contrast, economic activities are not very articulated. Traditional activities usually dominate the scene, together with some form of industrial activity in embryonic form. A firm that wants to leave a traditional activity and invest in the modern, industrial sector will find it difficult, if not impossible to acquire, without sustaining elevated costs, the knowledge and skills that are indispensable for his new activity. It would have been easier if others before him had already successfully made the transition from the traditional to the modern sector. In this case, one firm would benefit from the experience of others.[83] It could be said, in short, that in a lagging area there are direct costs associated with the transfer from traditional to modern activities as well as future externalities associated with modern activities.

The mechanism can be represented very simply in Figure 7. *V* represents marginal productivity (of a generic resource pertinent to a single agent); *Ltr* represents the share of agents in the traditional sector (with low marginal productivity), and *Lm* the share of agents operating in the modern or industrial sector (with high marginal productivity. Let us assume that differences in marginal productivity (*AB*) depend on the presence of externalities in the modern sector, as well as the transfer costs for shifting from the traditional to the modern sector. The two bold lines (to the left and

---

[83] In a recent study on development in emerging countries (Hausmann, Hwang, Rodrick 2005) the central theoretical hypothesis is that "the local cost discovery [that] generates knowledge spillovers." See pp.1-2.

right of the vertical line *V*) represent for a sole firm the alternative payoffs. Until a share of firms less or equal to *Lm** transfers to the modern sector, a sole firm will still find it advantageous to stay in the traditional sector. The threshold of *Lm** is not known to the agents, which means there cannot be collaboration.

It is thus clear that the actions of single economic agents deciding independently, even if they have the resources to sustain their actions, cannot get growth going. What is needed is a "Big Push" to reallocate resources in a sizeable block.

Underdevelopment traps in economic literature

That it is hard to get growth going and that development requires an intentional "Big Push" has long been recognised in economic literature. There are, however, some different paradigms – today the majority – that consider incremental approaches or look at markets. [84]

In a well-known essay, Paul Rosenstein-Rodan (1943) proposed a path for economic development based on intentional interventions[85] calling externalities into play. Rosenstein-Rodan claimed that this was the only path: in his view, development based on existing wealth was impossible because "Capital must be supplied internally at the expense of standard of life and consumption which are already at very low level."[86] Rosenstein-Rodan's contribution was influential mostly because of its argument for "balanced development" by means of public initiatives supporting different firms and sectors simultaneously. Hirschmann opposed his argument using his concept of linkages that favoured "unbalanced development" (with partial initiatives and spontaneous contagious effects). Rosenstein-Rodan's essay did contemplate this possibility, but his main emphasis was on externalities, "Marshallian economies external to the firm" [87] (see p.206, point 7) as well as other externalities associated with a drop in the perceived risk.[88]

---

[84] Among empirical research conducted by reducing growth rates to their component parts, works such as Young's (1995) have claimed that development in emerging countries should be attributed to the availability of workers. Other research has stressed the role of the "total productivity of factors" or of technological and organizational change. The idea of traps belongs to the second category because it claims that development is not possible only on the basis of vast availability of labour. A recent contribution, containing a survey of the debate, claims that with new methods it is possible to prove that technical and organizational change is dominant (Khan, 2006).

[85] He was anyway clearly thinking of small firms and of 'light' sectors, technologies employing a great number of workers, not large firms or intense capital sectors (see pp.203-4 and p. 206).

[86] Idem. See p.203

[87] These are indicated in pp. 259-62 of the Principles of Economics (Marshall, 1890) and regard common knowledge, the division and specialization of labour, a labour market with a strong

Basically, Rosenstein-Rodan's idea is that in lagging countries and regions nothing can change while the modern sector remains relatively small. He maintains that there is a threshold beyond which everything contributes to development and before which nothing can change. In this sense, then, it is clear that, in Rosenstein-Rodan's approach, the primary objective of policy was to focus on getting growth going, recognising that this cannot take place on its own and that neither the market nor incremental steps determine its outcome.

Economic development literature has always gone back to this point, from Leibenstein (1957) and Nurkse (1953) to the early formalisation provided by Murphy, Shleifer and Vishny (1989), and by Krugman (1993), up until the proliferation[89] of works published in the 1990s. Considering Hoff and Stiglitz's survey (2000), these works provide a detailed exploration of the various characteristics of complementarity between agents or processes and the risk of coordination failures with a stable underdevelopment equilibrium.

In the more recent literature on the subject, transfer costs, externalities and underdevelopment traps have been examined. According to these works, traps are caused by: low capita income in the population, which is therefore unable to sustain sufficient demand for new final goods; an insufficient demand for specialised equipment owing to the small number of firms interested in using it; an insufficient supply of qualified labour owing to the small number of firms demanding it; an insufficient level of professionalism in the labour force owing to widespread lack of knowledge and a low standard of living; an insufficient financial savings owing to the

---

professional and qualified component. A less known element is when Marshall quotes Hobson: "The mutual influences of the localization of industry, the growth of towns and habits of town life, and the development of machinery are well discussed in Hobson's *Evolution of Capitalism*" Marshall 1890, p. 262, who evidently knew Hobson's work before its publication after the 'Principles'. Hobson had in mind an evolutionary dynamic in which good salaries in local systems that grow allow a cumulative process that proceeds from the differentiation and qualification of labour, another aspect of externalities "The arts of production and consumption will [...] be found inseparable: not merely will they be seen to be organically related, but rather will appear as two aspects of the same fact [...]. Just in proportion as our tastes become so qualitative that we require to put our own spontaneity, our sense of beauty and fitness, our vital force, into what-ever work we do, and likewise require the same elements of spontaneity and individuality in all we enjoy, the economic conditions of a perfect society will be attained." (Hobson 1894, p. 380)

[88] "His subjective risk estimate is [in a depressed area] bound to be considerably higher than the objective risks. If the industrialization of international depressed areas were to rely entirely on the normal incentive of private entrepreneurs, the process would not only be much slower, the rate of investment smaller and (consequently) the national lower, but the whole economic structure of the region would be different [than that of substantial international investment]." Rosenstein, Rodan, 1943, p.206-7

[89] Hoff and Stiglitz's survey (2000) takes into account 43 works published since 1990.

insufficient productive use of financial resources; an insufficient capacity to combat corruption owing to the reduced number of firms that can effectively be controlled; an insufficient drive for change (of behaviour, attitude, beliefs, workplace, etc) owing to the minimal number of subjects that have already changed; an insufficient capacity for discriminated groups (including women) to take action, which perpetuates discrimination.

Underdeveloped areas are thus characterised by many traps,[90] which join forces to become powerful agencies for conservation: low per capita income, limited modern sector activities, lack of savings, unqualified labour, low level education and training, a static tradition hostile to change, corruption, criminality, discrimination. The very existence of these factors leads one to consider development strategies requiring highly decisive intentional policies.

Social traps

One possible objection to the concept of underdevelopment traps, from the point of view of valorising immobile resources, is that transferring them from less productive to more productive uses can in fact cost very little. In this case, there would be neither the high transfer costs assumed by the model, nor traps, nor the need for decisive intentional policies.

This objection has often been justified in economic development literature,[91] by observing that in backward areas the modern sector can make under-utilised or unoccupied resources in families or in society more productive. It is true that most artisan or commercial enterprises in rural societies are characterised by the fact that they are set up by individuals who were not employed efficiently in the agricultural sector and who as a consequence created their own work opportunities by establishing micro-enterprises in crafts or services (Haggblade, Hazel, Brown, 1989; Grosh, Somolekae, 1996; Mead, Liedholm, 1998; Weijland, 1999). The effect of this unproductive segment of the population on the overall labour supply is that opportunity costs are very limited and the reservation wage is very low.

---

[90] Empirical research based on multiple equilibria is still not very common (neither is work on traps). There have been some encouraging results, however, in Graham and Temple (2001).

[91] This is the main element, for example, of the dualistic model of Nobel prize winner Arthur Lewis (see Lewis, 1979).

It could be claimed, then, that low labour costs compensate for training costs and that this is simply a cheap way to reallocate under-utilised workers to new activities. If this were the case, a new and growing modern sector could gradually develop, creating enough efficient firms to push the area beyond the threshold indicated in the paradigm of underdevelopment traps.

All this is perfectly reasonable if looming over this optimistic picture there were not social traps. Social traps make linkages impossible between low initial transfer costs for resources from the traditional to the modern sector (owing to under-utilised resources) and the creation of a sufficiently strong modern sector. Going back to Figure 7, while underdevelopment traps can be seen as impeding the shift of resources from traditional to modern sectors, social traps can be seen as an obstacle to resources staying in the modern sector where they have already been transferred. Put in another way, the former are mechanisms that do not allow exit processes to take place, while the latter are mechanisms that do not allow resources to stay away from underdeveloped territories.

The reason for the existence of social traps is that pursuing a limited advantage in the short term causes more serious harm in the long term.[92] Imagine a closed local market with a multiplicity of enterprises working in the traditional sector (Ltr=1 in Figure 7). Subsequently, a generic innovation that is perfectly reproducible[93] takes place in various stages, and this opens up the opportunity for a new sector to work alongside traditional activities. Given the fact that the income received is additional in

---

[92] A single member of the community exploiting natural resources is the most common example of social traps. The over-harvesting (beyond the natural replenishment level) of fish species or forest reserves, or over-grazing, allows short-term benefits for the individual, but in the long term it leads to impoverishment and finally to the depletion of the common resource. Positive reinforcement in the present soon becomes long-term negative loss to society if other members of the community imitate the behaviour of the individual. On the other hand, rational agents have no choice but to do so. If there are no regulations or obstacles to this behaviour, the fear of no longer having access to a resource in the future will persuade all the agents to accelerate their exploitation in order to accumulate advantage before it is too late. The result is the over-exploitation of a resource and, in the long run, its depletion. In a context of social traps, individually rational behaviour can lead to collectively irrational behaviour. The literature on social traps is vast. Essential references are Hardin, 1968, 1982 and Ostrom, 1990.

[93] Imagine that in a backward area where there is only traditional activity, an innovation (of product or process) brings about the early stages of modern sector activity, and thus becomes the first of a series of innovations. Let us also imagine that the first step is easy to achieve, without sustaining additional costs, while further steps are not as easy, and require hefty investments. This hypothesis is what actually happens in real life. Backward economies have the advantage of being able to imitate the technical and organizational solutions adopted by more advanced economies. However, owing to barriers of tacit knowledge and the need to adapt to circumstances, assimilation of innovations takes place over time and with experience, each step preparing the way for the one that follows. See Boggio, Seravalli, 2003.Ch.4

nature, and given natural risk aversion when a new enterprise has not been experimented and when its outcome is uncertain, previously under-utilised resources will be employed in the early stages. Imagine, however, that to start with there is only one firm starting this new activity, temporarily in a monopolistic capacity, operating at average marginal costs equal to C', and making maximum profits equal to π_1. This could be represented by the grey area $(P_1-C')Q_1$ in Figure 8.

Costs are made up of variable exogenous costs C" and the prevailing wage according to current local parameters (C'-C")=w. The absence of barriers to entry (investment in fixed capital is limited) and the perfect reproducibility of the innovation increase production and reduce prices, causing an erosion of π_1. The level of production Q_2 corresponds to a level attainable in perfect competition where prices are the same as costs and profits are zero. For a single firm, demand shifts from $d1$ to $d2$ because total demand $d1$ is sub-divided among all firms. If no other firms came in beyond the level of perfect competition, the firms present would pay the prevailing wage for the labour they employed. Since this labour supply is assumed to be unoccupied, the prevailing wage paid represents a flow of additional income that could sustain in part the demand for products in the same modern sector. There would thus be a two-fold effect on demand. The entry of new firms would reduce demand for each producer. Distributed additional income would, on the other hand, increase it. Taking both these effects into account, one could imagine that, for each producer, demand would be situated somewhere between $d1$ and $d2$. In this case, profit margins would still be positive (π_1 would not be annulled completely). With positive profit margins, firms in the modern sector would persevere in their innovations.

What actually happens is that new firms continue to arrive, pushing the threshold beyond the level of perfect competition because they are not forced to pay the prevailing wage. A wage below the prevailing level, however modest, is still a source of increased family income for people who have been unemployed and the labour opportunity cost is almost zero. The entry process for new firms could continue (in theory) as long as the curve of the nth firm's residual demand ($d3$) has average costs net of labour costs (C") in a hypothesis of zero reservation wage. There would thus be a higher than optimal entry of productive units, with a limit level of Q_3 of total production. Demand in the modern sector is not therefore sustained significantly by income distributed within the same sector, and firms do not make enough profits to

allow them to go onto the next stage of innovation and to undertake a development path,

This is true because a sole producer tends not to consider general welfare, but only its own profits. Overall demand in the modern sector is at the end of the process practically unaltered; it is simply distributed among the new entries. The new firms do not distribute additional income, or, if they do, they do so to a very small extent.

In development models with unlimited labour supplies (Lewis, 1979) this effect is not considered, and the unit productivity of labour in the modern sector (value added per employee) stays above the standard wage (C'). In these models a constant reservation wage is a factor that helps increase profit, accumulation and development. The difference compared to the social traps situation depends on differing assumptions about two features. First, Lewis's models do not allow the reservation wage to be below the prevailing wage. Second, they take as given that barriers to entry are high, which means that it is not so easy for new firms to come in. These models reflect a context of dualism, where economic and social systems have a 'community based' traditional sector. In this context, the whole social system is based on the distributive principle "from each according to his ability, to each according to his need". The unit wage in the traditional sector is thus equal to the average productivity of labour, even for those who are unoccupied or under-utilised. The reservation wage is fixed at that level even when the productivity of these workers is lower, or inexistent. Lewis's models do not consider that capitalistic market approaches are predominant even in regions and in backward areas (indeed, traditional agriculture is obliged to buy seeds, fertilisers, transformation services, storage and transport on the market), and that communitarian rules and practices can only survive marginally. In the modern sector, these models assume large and medium sized firms and neglect the fact that often in the real world firms are small, even very small. Dualistic models envisage a world with two completely different social and economic systems: one genuinely communitarian and one based on advanced industrial capitalism. It is more realistic to imagine that hybrid systems co-exist in lagging regions: the traditional sector could be modernising and moving towards the market, while the modern sector could still be primitive and unstructured.

The mechanism of social traps can be seen in many case studies,[94] and in the Appendix there is a good example, backed up with a multi-agent simulation on the Swarm platform. What both this exercise and real case studies show is that there is a link between the macro dimension (distributed income) and productivity at a micro level. Incremental innovation (a significant and stable increase of productivity in the modern sector) is held back by low levels of distributed income, owing to the absence of barriers to entry and the lack of rigidity (verso il basso) in the cost of labour. As early as 1894, Hobson discovered the same phenomenon, as Marshall then pointed out.

Combined traps: a model inspired by classical economics

This problem brings us back to the central issue in classical economics. It is possible to build a very simple dynamic model, in the spirit of classical development economics, which evidences the nature of traps at a level of entire economic systems.[95] Underdevelopment and social traps combined create a threshold under which economic development cannot take place and beyond which it gets going very quickly.[96]

---

[94] McCormick (1998) has studies the cluster of mechanical firms in Kamukunji (Kenya). The local system comprises 2,000 metal workers and manufacturers. Its position in an industrial area and near the regional bus station gives access to markets for factors and products. Low-income residents in Nairobi and neighbouring cities are their main consumers. Most firms use productive methods based on manual ability and basic technology; only 10% has enough electricity to run more technologically advanced machines. Consequently, productive processes are slow, the products are low quality and the few firms that have electricity work on commissions rather than in a market. The low technological level of the craftsmen does not generate spillover and does not create an accumulation of knowledge and implies low barriers to entry, which, together with low skills, translate into a multiplication of mini units that does not allow any firm to grow enough to become efficient.

[95] The classical school is relevant here because it flourished during the Industrial Revolution. The beginning of modern economic development is central to the social context of the time as well as to its thinking.

[96] The processes of economic development that were successful (both after the second world war and between the eighteenth and nineteenth century) all had an initial phase of acceleration when GDP growth rates are particularly high. It is possible now to see why this is plausible, by means of a recent sample of a number of countries. Taking into account the yearly series (1995-2000) of real per capita income and equal purchasing power of 45 backward countries, for each one a different base year is selected so that they are aligned according to their initial "development stage" defined in correspondence to a quota of 70% of the rural population. From the base year, the growth rates in per capita income for every country are calculated together with the differences from average world growth. On the basis of the results (for every country it is possible to have growth rates for ten years from the base year) the 45 countries are divided into groups. The first group includes all countries that since the base year have always registered a positive difference from world growth rates; the second group comprises countries that have always registered a negative difference; he third group embraces countries that have changed trend. The last group includes only 8 countries out of 45, and describes cases where a weakly positive initial stage of development is soon aborted. Most countries either grow well (especially in the early years after the base year) or continue to decline.

As regards the theoretical construction of the model, the question that the classical school of economics considered crucial was where surplus ended up. Surplus means the part of total output that does not need to be reinvested in the productive process in order to keep production at the same level. Surplus can be consumed completely without compromising the firm's future output (or the surplus itself). If, however, a part of the surplus is reinvested in the firm, the volume of production can be increased. The question of surplus – how much there is and what happens to it - is in fact crucial in the take-off phase of development. With no surplus, growth simply does not get going. Second, in the classical approach, goods consumed by workers to survive and reproduce are not considered surplus; they are seen as necessary to maintain pre-existing levels of production. This consideration is relevant for those dealing with the problem of getting growth going in a lagging area. The systemic presence of vast swathes of poverty means that unsatisfied subsistence needs cannot be neglected. A third peculiarity of the classical 'model' is that, as in Ricardo's essay on the price of grain, it only envisages circulating capital, while aggregate models in subsequent literature foresaw only fixed capital. The consequences on the crucial matter of surplus in the classical model are not particularly relevant. However, in contemporary macroeconomic models, which envisage only fixed capital with no depreciation or proportional depreciation, the problem of surplus is not so simple.

The basic model: constant techniques

We consider that the traditional sector is dominant in a lagging area. All the other variables refer to the traditional sector. The embryo of modern activity is considered external to the productive system described in the model.

In the first version of this model (constant techniques), unlike the model that ensued, it is supposed that in order to obtain one unit of traditional output in time $t+1$ you need (at time t) a_M (constant) output units as means of production, and a_L (constant) output units to pay workers. Let $a=a_L+a_M$. If this is the case, then a is the physical capital that must be advanced to time t in order to obtain an output unit after a year, that is at time $(t + 1)$. This is the unit cost. To produce in time $(t + 1)$, the quantity q_{t+1}, the capital to advance to time t will thus be aq_{t+1}.

Let us call K_t the capital invested in traditional output. Thus:

$$K_t = aq_{t+1}$$

After a year, this capital no longer exists because it has been completely used up by means of production totally absorbed by the annual production cycle and by means of subsistence that have been consumed. Capital is thus only circulating capital, meaning that it doesn't last longer than the production period, in this case one year.

If $a=1$, in order to reproduce in time $(t+1)$ the quantity of output produced a year before (in time t) the output must be re-employed completely ($K_t=q_t$). The system just manages to achieve simple reproduction; that is, to reproduce itself at pre-existing levels, without any possibility of growth.

If, on the other hand, $a<1$, then to reproduce in time $(t+1)$ the quantity of output produced in time t all that is needed is to re-employ some of the production. In this case the system is able to generate a surplus; that is, to produce more than is necessary to achieve simple reproduction.

If we indicate surplus with S_t we have:

$$S_{t+1}=q_{t+1}-aq_{t+1}=(1-a)q_{t+1} \qquad [10]$$

Thus the rate of surplus over capital is:

$$\frac{S_{t+1}}{K_t}=\frac{(1-a)q_{t+1}}{aq_{t+1}}=\frac{1-a}{a} \qquad [11]$$

Let us call the growth over time of K_t (equal to the output growth rate, since it is proportional to the capital invested) 'capital accumulation'. The capital necessary to set a new productive cycle in motion can be taken from output revenue. If this capital is higher than in the previous cycle, 'capital accumulation' takes place.

Let $g(t)$ be the rate of growth capital between time t and $(t+1)$, defined as:

$$g(t)=\frac{K_{t+1}-K_t}{K_t}=\frac{K_{t+1}}{K_t}-1 \qquad [12]$$

If the output at time $(t+1)$ were completely reinvested in production, that is if $K_{t+1}=q_{t+1}$, then the situation would be:

$$g(t)=\frac{1-a}{a} \qquad [13]$$

Which is equal to the surplus rate.

This is another way to express the condition we have already indicated. The necessary minimum condition for a system to be able to reproduce itself, that is, in order for its growth rate not to be negative, is that $0<a\leq1$. If this condition is satisfied, both the surplus and the output growth rate are positive.

Let us now suppose that only part of the surplus is invested for capital accumulation; that is, the given share χ, with $(1>\chi>0)$. The other part can be employed in the 'external' modern sector.

$$K_{t+1}=K_t+\chi S_{t+1}=K_t+\chi(1-a)q_{t+1}= K_t+\chi(1-a)K_t/a$$

$$K_{t+1} = K_t\left[1-\chi+\frac{\chi}{a}\right] \qquad\qquad [14]$$

Formula [14] is illustrated in Figure 9.

Figure 9: The time path of movement of capital
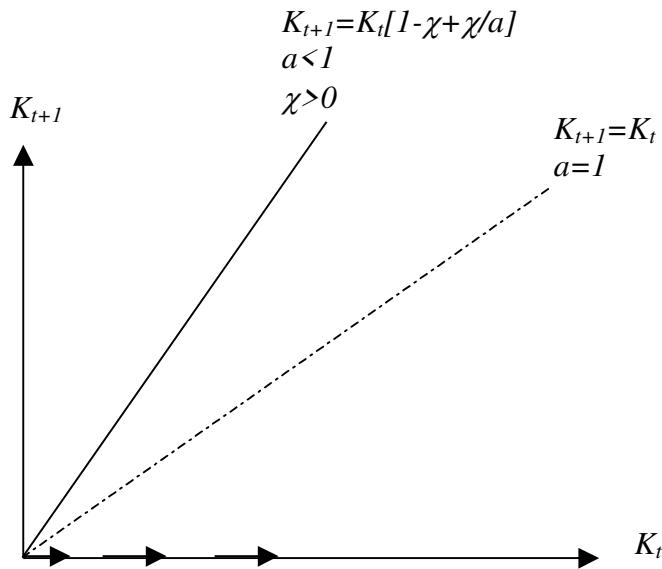
As long as a is less than 1, the capital will continue to grow (whatever the starting level was and whatever the value of $a<1$) if χ is positive. In this case the contribution

of the traditional sector to the modern sector in terms of continuous flow of resources will also be positive.

The rate of capital growth and of the output will thus be:

$$g(t) = \frac{K_{t+1}}{K_t} - 1 = \frac{\chi(1-a)q_{t+1}}{aq_{t+1}} = \chi \frac{1-a}{a} \qquad [15]$$

$$g(t) = \chi \left(\frac{1}{a} - 1 \right) \qquad [16]$$

Formula [16] is illustrated in Figure 10.

Figure 10: Growth rate in the basic model
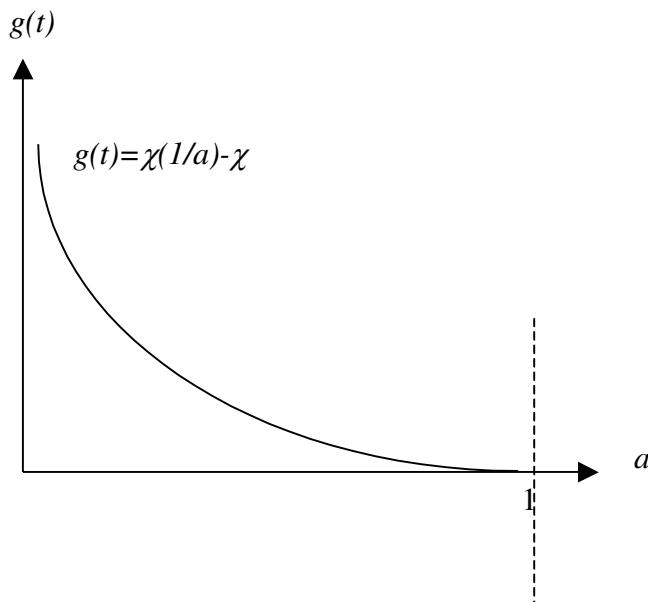
If $a<1$ then the growth rate will be positive, for a value of χ that is not zero. Whatever the rate is, even if it is very low, it will be proportionate to the parameter (1-*a/a*). Remember that this parameter represents the surplus rate on the capital. For every given technology, then, there will be a given constant growth rate.

If we now abandon the hypothesis of a constant unit cost per product of the advances in circulating capital that we assumed in the first version of the model, and if we suppose, by contrast, that the unit cost increases, then the surplus – even when it existed at a given time – could disappear. A hypothesis of this kind is incorporated in Riccardo's model in which less fertile land, where unit costs depend on quantities

produced, leads to a decreasing unit surplus on land. Riccardo's model in this version stresses that new growth can be held back, or rather that growth rates can diminish as new growth and development takes off owing to the presence of indispensable resources with a given availability whose supply costs increase as demand increases. In this case we can talk about a brake on growth, opposed by a force that sustains growth, as Adam Smith pointed out, and as history has proved: the division of labour which favours technical progress.

Our hypothesis reflects the idea that the forces that sustain growth prevail, creating a self-reinforcing mechanism. Let us assume more precisely that, wherever there are the conditions for surplus and growth in the traditional sector (producing wheat), a part of the surplus is spent on the modern sector and thus continues to sustain demand. The modern sector can thus escape the trap created by the absence of barriers to entry and is able to be more innovative. If the modern sector produces technologies used by the traditional sector (ploughs and fertilisers), one consequence is a continuous improvement of productivity in wheat production.

Second version: endogenous costs

Let us suppose then that the traditional sector has endogenous unit costs, rather than given ones as in the previous model, and that these are subject to reduction if the accumulation rate is positive.

Let

$a_{Lt}=a_L(K_t)$ with a negative derivative of : $a_L'<0$

$a_{Mt}=a_M(K_t)$ with a negative derivative of: $a_M'<0$

For example, indicating with A a scale factor of ($A>K_t$):

$$a_{Lt} = \frac{a_L A}{K_t}$$

$$a_{Mt} = \frac{a_M A}{K_t}$$

$$a_t = \frac{a_L + a_M}{K_t} A = \frac{aA}{K_t}$$

This implies that:

$$q_{t+1} = \frac{K_t}{a_t} = \frac{K_t^2}{aA}$$

$$S_{t+1} = (1 - a_t)q_{t+1} = \left(1 - \frac{aA}{K_t}\right)\frac{K_t^2}{aA} = \frac{K_t^2}{aA} - K_t$$

With the result that:

$$K_{t+1} = K_t + \chi\left(\frac{K_t^2}{aA} - K_t\right) = (1 - \chi)K_t + \frac{\chi}{aA}K_t^2 \qquad [17]$$

The instantaneous growth rate is thus:

$$g(t) = \frac{K_{t+1}}{K_t} - 1 = \chi\left(\frac{K_t}{aA} - 1\right) = \chi\left(\frac{k_t}{a} - 1\right) \qquad [18]$$

where:

$$k_t = \frac{K_t}{A}$$

Figure 11 shows the comparison between equations [18] and [16]

Figure 11 - Growth rate in a model with endogenous production costs
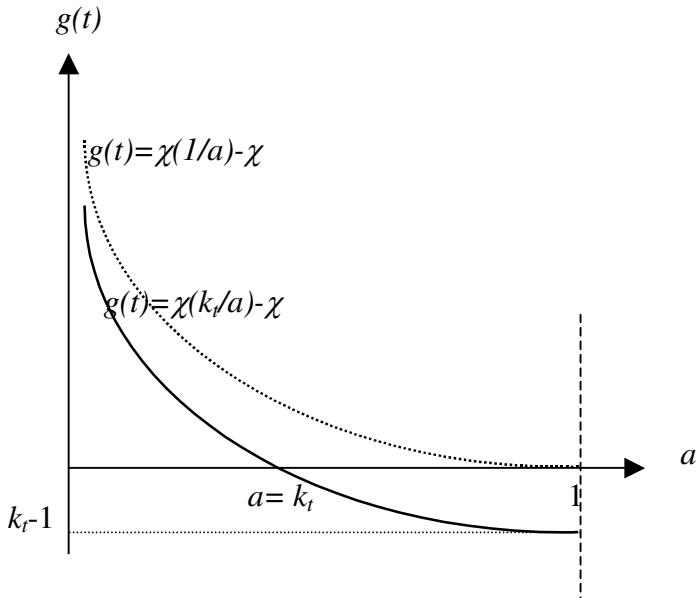
A positive growth rate is only possible where $\chi>0$ if the potential technology is at least as efficient as the level that leads to units costs lower than a certain threshold not just lower than one. Moreover, the instantaneous growth rate for every technology must be lower than the rate in the basic model. It will be the same as that in the basic model only asyntotically when $a=0$.

The development process can begin (with capital growth) if a further necessary condition is fulfilled in addition to the conditions $a<1$, $\chi>0$ in the previous model. That is:

$a\leq k_t$; or rather $a\leq K_t/A$.

The potential unit cost of production must be lower than a positive value greater than zero. However, equation [17] could be written:

$$K_{t+1} = K_t + \chi\left(\frac{K_t^2}{aA} - K_t\right) = (1-\chi)K_t + \frac{\chi}{aA}K_t^2 = K_t\left(1-\chi+\chi\frac{K_t}{aA}\right)$$

This shows a performance of capital as described in Figure 12.

Fig. 12 - The temporal path of movements of capital with endogenous production costs.
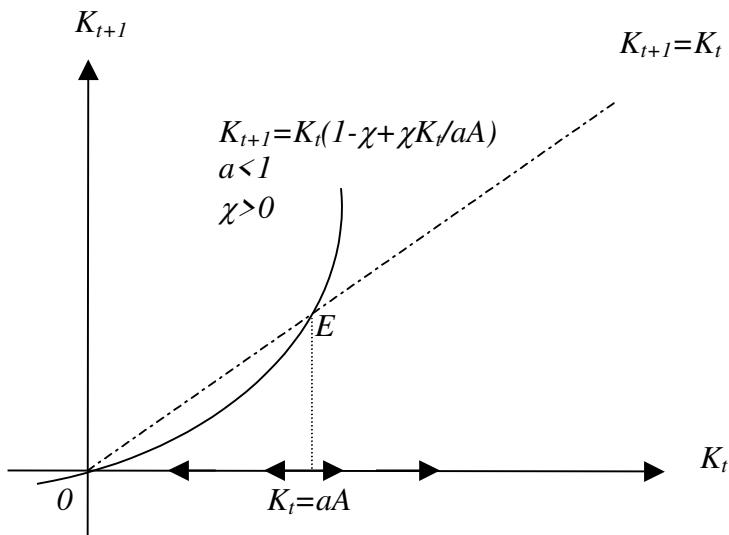
We can thus confirm that capital growth can only take place above the technology threshold $a=K_t/A$, but that, beyond that threshold, the growth rate will increase (point E represents an unstable stationary equilibrium. If the threshold condition is fulfilled, there is not only the possibility of growth but growth is accelerated. Under the threshold, on the other hand, growth tends to slow down and go into decline.

Development policies

Analyzing combined traps, as we have said, allows us to shed light on specific policies that might help avoid them. We have seen that these policies must aim to support competitiveness by sustaining labour productivity.

Creating access to external markets (national or international) for traditional products such as transformed agricultural products, by activating trade channels that are stable over time, could help trigger virtuous growth patterns.[97] A significant increase in demand for these products would rapidly lead to a depletion of labour stocks with low opportunity costs. Industrial activity would then express a demand for labour that could be satisfied only by people who were previously employed in the traditional sector. Their inclusion in the modern sector would lead average costs to increase to a level above the average subsistence costs. The most relevant consequences would be: the number of small firms would decrease relatively; the unit

---

[97] A similar approach is developed in Humphrey and Schmitz (1996).

volumes of production would increase and therefore more firms would be closer to the optimum level of activity; there would be higher wages. Foreign demand, however, can only be sustained if local firms are competitive.

We shall now see that competitiveness can only be achieved with the help of pubic and collective goods. A combination of underdevelopment traps and social traps presents us with the problem of lack of information. Resources cannot be reallocated from less to more productive uses when there is no information available. If the agents involved know enough, they can solve their own problems of reallocation. What they need to know is quite evident. They need to know how many resources they must transfer from the traditional to the modern sector in order to repay the transfer costs with the returns from externalities. As visualised in Figure 7, they need to know what Lm^* is; following the classical model in the previous section, they need to know what A is. If all the agents involved have this information, they are able to avoid the two consequences of these traps: one of these is that no resources are reallocated; the other is that too many resources are reallocated too fast. It would not be difficult, in fact, for the agents to stipulate an agreement among themselves because they would be able to measure the advantages.

With imperfect information, the critical threshold is unknown, which means that the agents involved cannot make an agreement or pursue a fixed path to reach it. One solution seems fairly easy, or at least logical. This solution is to consider imperfect information a missing good (information) that can be produced or bought by sustaining the cost of producing it or buying it, just like any other useful good. We would then simply add the cost of information to the cost of transferring resources from the traditional to the modern sector and then adjust the threshold of resources that need to be transferred in order to reach equilibrium between costs and benefits. With inadequate information the threshold cannot be known, but once the necessary cost of acquiring information has been sustained, the threshold can be predicted. On the basis of this shared knowledge it would not be difficult for the agents involved to come to an agreement among themselves and thus avoid the traps.

And yet, unfortunately, this is not possible. As Stiglitz (and much of the literature) has pointed out, information is not a good like any other because it is 'relational' in nature. The most relevant information is transmitted, or created, by the very behaviour of the agents involved; it modifies their behaviour and everybody knows this. "A simple lesson emerges: some individuals wish to convey information; some

individuals wish not to have information conveyed (ether because such information might lead others to think less well of them, or because conveying information may interfere with their ability to appropriate rents). In either case, the fact that actions convey information leads people to alter their behaviour, and changes how market function. This is why information imperfections have such profound effects." (Stiglitz, 2003, p.19)

Let us now look at how the information mechanism behaves in the case of traps. Imperfect information implies that: a) not all the agents involved know the threshold level, perhaps none do; b) each agent has different information of conjectures about what the level is. If we assume the most favourable case where one agent knows precisely what it is and the others have no idea, a solution could only be reached if the agent in the know reveals his information and if the others believe him. What really happens, however?

The fact that most agents do not know what the threshold level is will make them doubt whether the agent who claims to know is telling the truth. What will the agent in the know do? He could transfer his resources from the traditional to the modern sector, thus signalling to the others that he is credible. The others could then follow his example, but the agent in the know could never be sure whether they will do so, or whether they will do with the right timing. Knowing this, his actions will be very risky. The risk is inherent in the fact that if the others do not follow his example he will not enjoy the benefits of externalities that would be able to equal his transfer costs; if they follow him too fast or with too many resources his profits as a temporary monopolist will rapidly be reduced and he will not be able to stay in the modern sector and introduce innovations. Alternatively, he could decide to do nothing to avoid the risk, actively hiding information, or he could transfer his resources but be prepared to be the first to go back on his steps, leaving the others to learn from their own mistakes. Everybody knows that this is the way things work, so information becomes something to interpret. The result is that a mechanism is established that tends to lead to the exact opposite of what is needed to create an efficient equilibrium requiring neither too many nor too few resources to be transferred. Faced with the actions of one agent, the others will conclude either that he has grabbed an opportunity and that they should hastily follow in order to exploit it themselves (social traps), or that he is taking a risk but that he knows how to get out of it and will

leave all the damage in his wake, and that at this point it would be better to avoid the whole venture (underdevelopment traps).

An efficient equilibrium would require an agent with no information to be sure that an agent with information acts without taking into account the fact that he has information and the others do not. In real life, an agent with no information is absolutely sure that the actions of the other agent are dictated above all by the fact that he is in the know while the others are not.

The conclusion is that there is no way to establish an effective exchange between greater transfer costs that are certain today and benefits that are uncertain in the future. The only way is to create one or more goods of collective utility,[98] which must give rise to an increase in productivity for all resources and for all the agents involved so that each party is able to sustain the transfer costs today, rather than when a transfer of a block of resources equal to the critical threshold has already taken place.

These considerations focusing on the central role of public or collective goods contrast with the dominant, mainstream approach that still thinks it is the market that guides economic growth.

A multi-agent simulation

An artificial economy such as a "Swarm" studies the interactions between micro-behaviour and macro-phenomena involving a large but circumscribed set of agents. The virtual agents behave and make decisions on the basis of imposed routines. They are subject to cognitive and computational limitations, but they have the capacity to learn over time owing to their memory of the results of the interactions. The simulation is based on adaptations of the procedures proposed by Bruun and Luna (2000). The system comprises initially self-employed agents. These agents can create innovative firms and coordinate heterogeneous workers; they produce outputs designed to satisfy the demand for consumer goods. The entrepreneur's remuneration

---

[98] With reference to backward contexts, a significant role is often attributed to sustaining the productivity and competitiveness of public goods of a physical nature, such as infrastructure, telecommunications, transport etc. Immaterial public goods, however, should not be overlooked. Looking at the development stages that precede the successful consolidation of European district systems, it soon becomes clear that basic education and techniques - as well as contexts where the division of labour between firms and collaboration among firms was encouraged - were essential ingredients of their success. These resources are at times the result of explicit programmes, but more often they are the unconscious result of intuition. Even more often they are the result of a slow accumulation of institutional interventions that have become public goods. In the majority of cases it has been the non-excludability of these resources that has most contributed to the enrichment and reinforcement of the districts and to the accelerated take-off of local systems.

is what is left of his earnings after labour costs. Profits and wages represent the GDP of the economy. GDP in its turn generates demand for consumer goods that are exchanged. At the end of a pre-established period the firm can go bankrupt. In this case, its losses are redistributed among the community and thus reduce the income available for consumption. The existence of a negative differential between supply and demand and of qualitative mismatches between production and consumption increase the bankruptcy rate. Firms learn over time to adapt the quality and quantity of supply to demand. If innovation is profitable it generates a flow of new entries. The flow of new entries is a function of the costs sustained to set up new firms (imitation of innovation). Setting up a new firm takes place in two stages: in the first, the stage of a potential firm, the entrepreneur employs labour and sustains the relative costs but is not sure whether he will succeed in placing his output on the market; in the second, once he has observed the match between supply and the features of the demand, the firm is established and becomes effective.

The hypothesis that needs to be verified is whether or not a high level of imitation of innovation gives rise to an excess of new entries with the result of creating an aggregate supply greater than the actual demand. The mismatch between quantities supplied and demanded triggers a process whereby firms exit the sector. Selection remains inefficient. Firms learn that profitability is overall inexistent, but they also realise that the failure of a pre-existing firm can give rise to temporary profits in a sub-local consumer niche. A part of the demand satisfied by the failed firm is in fact redistributed among potential firms in the vicinity. The absence of entry costs and the positive probability of making profits mean that firms try repeatedly to enter. The effervescence of these attempts reduces the availability of workers for the incumbent firms, holds back their growth and increases the probability of their failing. Over time, and through learning, the propensity to enter diminishes, but it stays high enough to influence negatively the economic result of the firms already on the market.

The results of the simulation seem to confirm the hypothesis. On the basis of a relatively limited number of periods (2,000), the trends are quite explicit. The system's GDP tends rapidly to contract, and at the same time the initially very high number of successful firms decreases. In the initial phase, the high failure rate in the system does not dissuade potential firms from trying to enter; rather, it stimulates them. The presence of an increasing number of attempts to create new firms contributes to stressing the precariousness of incumbent firms and accelerates their

exit. Over time, with fewer successful firms, the propensity of new entries to set up firms in the system is still very high. The selection process is incomplete because achieving an optimal number of firms is hindered by the fact that new firms repeatedly try to come into the system. Finally, there are alarming tendencies even in terms of income distribution. The Gini indices of inequality display a tendency to increase, although there are marked oscillations.

Bibliography for Chapter 2

Boggio, L. e G. Seravalli (2003), *Lo sviluppo economico: fatti, teorie, politiche*, Il Mulino, Bologna.

Bruun, C. e F. Luna (2000), Endogenous Growth with Cycles, in a Swarm Economy, in F. Luna e B. Stefansson (ed.) *Economic Simulations in Swarm: Agent-Based Modelling and Object Oriented Programming*, Kluwer (Advances in Computational Economics Series), New York.

Graham, B.S. e J. Temple (2001), *Rich nations, poor nations: how much can multiple equilibria explain?*, CID Working Paper 76, Center for International Development University of Harvard.

Grosh, B. e Somolekae, G. (1996), Mighty Oaks from Little Acorns: Can Microenterprise serves as Seedbed of Industrialization, *World Development*, 24, p. 1879-1890.

Haggblade, S., P. Hazel, J. Brown (1989), Farm-nonfarm Linkages in Rural Sub-Saharan Africa, *World Development*, 17, p. 1173-1201.

Hardin, G. (1968), The Tragedy of the Commons, *Science*, 162, pp. 1243-1248.

Hardin, R. (1982), *Collective Action*, The Johns Hopkins University Press, Baltimore.

Hausmann, R., J. Hwang, D. Rodrick (2005), *What You Export Matters*, Harvard University Working Paper, December.

Hobson, J.A. (1894), *The Evolution of Capitalism – A Study of Machine Production*, Walter Scott, London.

Hoff, K, e J.Stiglitz (2000), Modern Economic Theory and Development, in G. Meier and J.Stiglitz (ed.), *Pioneers in Development*, Oxford University Press, New York.

Humphrey, J. e H. Schmitz (1996), The triple C approach to local industrial policy, *World Development*, 24, 12, pp. 1859-1877.

Khan, F.A. (2006), *New Product Technology, Accumulation, and Growth*, World Bank Policy Research Working Paper 3846.

Krugman, P. (1993), Toward a Counter-Counterrevolution in Development Theory, *Proceedings of the World Bank Annual Conference on Development Economics 1992*, pp. 15-38.

Leibenstein, H. (1957), *Economic Backwardness and Economic Growth*, Wiley, New York.

Lewis, W A. (1979), The Dual Economy Revisited, *The Manchester School of Economics and Social Studies*, 47 (3), pp. 211-29.

Marshall, A. (1890), *Principles of Economics. An Introductory Volume*;

McCormick, D. (1998), *Enterprise cluster in Africa: on the way to industrialization?*, Discussion Paper, n 366, Institute for Development Studies.

Mead, D. C. e C. Liedholm (1998), The Dynamics of Micro and Small Enterprises in Developing Countries, *World Development*, 26, p.61-74.

Murphy, K., A. Shleifer e R. Vishny (1989), Industrialisation and the Big-Push, *Journal of Political Economy*, 97, 5, pp. 1003-1026.

Nurkse, R. (1953), *Problems of Capital Formation in Underdeveloped Countries*, Oxford University Press, Oxford.

Ostrom, E. (1990*), Governing the Commons. The Evolution of Institutions for Collective Action*, Cambride University Press, Cambridge.

Rosenstein-Rodan, P.N. (1943), Problems of industrialization of Eastern and South-Eastern Europe, *Economic Journal*, 52, pp.202-211;

Stiglitz, J. (2003-04), Information and the Change in the Paradigm in Economics, *The American Economist*, 47, 2, pp. 6-26; 48, 1, pp. 17-49.

Weijland, H. (1999), Microenterprises Clusters in Rural Indonesia: Industrial Seedbed and Policy Target, *World Development*, 27, pp. 1515-1530.

Chapter Three - Collective Goods[99]

In order for economic development to take place, economic, social and institutional transformations must first be tackled. These include reallocating resources to more productive uses, widening markets, introducing new technologies, accumulating fixed and human capital, and changing forms of governance and organisation. Bringing about these changes is a challenge because they break with consolidated practice, and because the transformations need to take place more or less simultaneously.[100]

Each change seems to have its own specific set of problems, but there is a common element: people must accept that a portion of their own satisfaction relies on others, and that they an give up "autarchy" in exchange for improving their wellbeing.[101] It is, after all, a matter of reallocating resources from traditional to more modern activities, where the vertically integrated production process has more steps.[102] These resources must be allocated over time, sacrificing current needs in the hope of accruing greater satisfaction in the future. Because there is a threshold of resources to be invested, below which results are nil, actors must come to an agreement. Moreover, it is not a question of reallocating either resources or time, but of reallocating both, since, as we have said, transformations must be simultaneous. The relevance of thresholds and the lengthening of productive processes show that the root

---

[99] Section 1, and the sub-section Informative and coordination costs and Appendix 2 are written with Alessandro Arrighetti, and Appendix 1 with Lidia Seravalli.

[100] Documented economic history has shown that often changes have taken place in one of the areas without lasting development taking place. It then becomes clear *a posteriori* that changes in complementary areas did not take place. There is a very clear example in the long history of extraordinary interventions in the Italian Mezzogiorno, where large firms brought new technologies and new capital to various areas with the effect of causing a transferal of labour from agriculture to industry, and the opening of new markets. This was not sufficient, however, to set in motion development processes that were sustainable in the long term because social and institutional changes did not take place to accompany these transformations. The new large firms were isolated, dependent on powers that were far away from the firm's location, while society and local institutions were dependent on aid and clientelism. Del Monte and Giannola (1997) write: "Our hypothesis is that, starting in the second half of the 1970s, the weight of the indirect negative effects of public intervention determine an accentuation of the malfunctioning of institutions, which, in addition to other negative exogenous factors accentuates the loss of efficiency in the economic system of the Mezzogiorno." (p.131)

[101] Barrington Moore Jr (1969) notes that the first and foremost problem of modernization was how to give bread to those who no longer produced grain.

[102] We can include in this category: the re-allocation between hidden unemployment and employment, between agricultural and industrial employment, between material transformation activities and creative, organizational or governmental activities, between productive activities for domestic or international demand, etc.

problem is relying on others[103]: while the disadvantages are clear and immediate, the advantages are less tangible, further in the future and anyway dependent on others. It is thus impossible, or at least highly improbable, for development to come about as a result of autonomous, decentralised, or spontaneous decisions. Actors should work together as a team, but each one has a specific vision of the future, which is very different from the others', especially in view of the fact that there are no institutional institutions or support to help them come to an agreement.

The solution lies in collective goods and services. It could be said that the modern age spelt the demise of *the* collective (community) good and that in its place we now have a series of collective goods. When community ties within small social circles are no longer solid, collective goods and services play three important roles. First, they reduce risk; then, they provide technical intermediation; finally, they help a group of people come to an agreement. There are three immediate reasons why decentralised and autonomous decisions tend not to contribute to development. To start with, trusting others is risky in terms of personal wellbeing. The higher the provision of collective goods (given the level of private resources that are needed to produce them), the greater the opportunities for trust. Secondly, in order to be able to rely on others you have to provide and gain information about, communicate, and transport whatever the division of labour technically requires. In this sense, collective goods and services are a key factor in ensuring that the division of labour actually takes place. Finally, a common outlook is essential. Collective goods and services contribute to bringing expectations and behaviour into line simply because they are collective; everybody uses the same services and are equally advantaged or conditioned by them.[104]

---

[103] Even in traditional community life people had to rely on others. However, relying on others in modernity is very different in two ways. First, our wellbeing now depends on many others outside the tightly-knit community where ties are broken. Second, we have to rely on others despite the fact that the prevailing individualism has broken the ties of solidarity; we do not rely on others out of solidarity but out of convenience. In Bell (1993) there is a brilliant explanation of communitarianism, in contrast to the "neo-Kantian liberalism of the Left" attributed to Rawls, which highlights these aspects.

[104] An example from Italian economic history: the Autostrade del Sole. It is a commonplace to hear that the Italian model of development was predicated on private transport, and that the construction of this motorway made an important contribution. Another smaller example: technical training in Emilia Romagna. It has been shown that the development of mechanics in central Emilia Romagna (which was predominantly agricultural until the beginning of World War II) was significantly driven by the promotion of technical training by local administrations, who were far ahead of their time in terms of demand.

This chapter focuses on what kind of collective services are useful for development. It is divided into two sections. The first section examines the proposition that in conditions of imperfect information and therefore of incomplete markets, the goods and services essential for development that are lacking cannot be supplied through spontaneous collective action in a backward area. In other words, we show that essential goods and services that have not come about as a result of individual initiative cannot successfully be provided by collective action without the support of state intervention. The second section illustrates the limits and inefficiencies of state intervention, given the premise that the role of the state in the provision of public goods and services is essential. These limits and inefficiencies are thus an inevitable corollary. The chapter concludes that it is not a matter of achieving the impossible. It is more a question of recognising that state intervention is necessary - despite its limits and inefficiencies - because single agents cannot carry out individual initiative, and because bottom-up initiatives lead to sub-quality collective services.

The limits of collective action

A backward area lacks many of the goods and services that would be useful for developing and increasing the productivity of its existing resources, which are often unused or misused. Providing these missing goods and services could help the lagging area become more productive. Likewise, land is often little or badly cultivated, and natural resources are underused or inaccessible. Again, providing access to wider markets, as well as to internal and external financial resources that an inefficient credit system does not allow them to tap, could valorise these resources. Similarly, endangered cultural values and traditions could be re-evaluated in order to provide a valid defence against the tyranny of exported models and sharpen competition. Valuable and fast-vanishing skills honed through traditional activities could be saved, not to mention customs - already disappearing with the advent of corrupt local administrations in cahoots with organised crime - such as honesty, wisdom and respect for those who undertake to administer public services. The appendix presents a case study of how a potentially relevant resource was not valorised, causing the majority of the population to live in poverty. In this case study it was easy to identify

the collective services that encourage development: cooperative firms or networks of firms that reduce the fragmentation of the productive process.[105]

In a lagging area each of the following fields - and the list could be much longer - lacks essential structures and activities: schools, professional training institutes, hospitals and clinics, water supply and drinking water, waste collection, services providing technical, administrative and promotional assistance, communication, logistics, material infrastructures and networks. Similarly, entrepreneurial, organisational and co-operative skills are hard to come by.

Often this endless list of seemingly inexistent services is reduced and prioritised at a local level. Some of these deficiencies are particularly heart-felt or especially long-standing; and yet, nothing gets done. In all these cases the market is incomplete. However real a need is, even when the need has been perceived and expressed, nothing happens. Or rather, nobody makes a big enough offer to fulfil the need. In many lagging areas these goods or services are not altogether absent; they are simply prohibitively expensive or, rather, access to them is restricted to a small minority, perhaps through familistic, clan or criminal ties. Indeed, the small circle of those who benefit from whatever goods and services exist strenuously defends the system, which is based on a monopolistic provision of services at exorbitant prices entailing significant rent positions. Incomplete markets do not occur in lagging areas for "innocent" reasons; they are common and deeply-rooted in backward areas because landed interests there wish to preserve the status quo.

Thus incomplete information, which some individuals or groups have an interest in maintaining, has the effect of obscuring the fact that everyone could pay a price to get goods or services that do not yet exist on the basis of the greater income that would accrue if those goods and services were to be created.[106] There is no such thing as an "effective demand", while "potential demand" exists and is often explicit. In a

---

[105] It could be considered surprising that a cooperative firm is considered part of the category of collective goods and services. Despite many schools of thought that consider a cooperative a firm just like any other firm, there are good reasons to believe that, in a capitalistic market system, they are above all a 'public good' that cannot be realized or sustained without support that is external to the mechanisms of the capitalistic system. These supports include: the active support of the community, the contribution of participants beyond the call of short-term convenience, public support. I think the considerations made in Seravalli (1988) are still valid, as Fiorentino and Scarpa (1998) confirm.

[106] Incompleteness also regards the credit and insurance market. In an evolved economic system there is widespread physiological effective demand for something that will provide future revenues. The financial system intervenes by providing credit, and the insurance system provides guarantees. In backward areas, money-lending is common and the only guarantees are in kind (at worst one's life), or belonging to criminal organizations.

situation of this kind, collective initiatives are needed; if they are to be efficient they must also have the following requisites.

Minimum endowment – a slow increase in the endowment of goods and services is not efficient. A minimum level must be guaranteed to start with; otherwise the benefits are gradually eroded by an increased demand that outstrips the new supply.

Critical mass – the presence of strong interest groups requires goods and services to benefit a good number of agents, otherwise they rapidly decline.

External domination – trade channels outside the community must be sought, infrastructure (transport and communication) must also be provided. At the same time, many collective goods and services require the collaboration of actors outside the community (governmental organisations, national and multinational firms, international organisms). This brings about heterogeneous rules of conduct, patterns of behaviour and restrictions that force agents to collaborate.

It is important to take these characteristics into account when evaluating the possibility of implementing either of the following solutions: bottom-up collective action to achieve collective goods, or delegated action to build public goods.

Bottom-up collective action: building cooperative norms

There is a vast literature on the subject of collective action, and it would be hard to summarise it here. The literature, however, delineates two main paradigms: Axelrod's evolutionary approach to the formation of cooperative behavioural norms (Axelrod, 1997, especially Ch.3), and Buchanan's economic club theory (Buchanan, 1989). Taken singly, these paradigms have their limits[107] but together they provide a useful tool.

In order to appreciate the nature of the two solutions for providing bottom-up collective goods without state intervention, it is important to remind ourselves of the specific difficulty arising from imperfect information. As Stiglitz pointed out (1988, 2003-2004), and as we have also already said, imperfect information does not only lead to information costs, as if information were a good like any other that can be produced by sustaining costs which are at the margin equal to their utility.

---

[107] The evolutionary paradigm regarding cooperative norms leaves the social environment undefined. The club goods paradigm assumes a representative agent. The emergence of cooperative norms can provide an operative basis for the fiction of the representative agent (when we make a commitment we all reveal our preferences). Defining on this basis the dimension of the group and of the collective good can also help define the social environment.

Information, on the contrary, depends on agents' behaviour (that others will have to interpret), which in its turn gives rise to uncontrollable externalities, except in particular contexts. There are two negative effects: one is adverse selection (who acts) and the other is moral hazard (how one acts).

When the issue is creating collective goods or services (infrastructure), adverse selection means that - given the uncertainty of individual agents regarding the behaviour of others, and given their uncertainty regarding the final outcome (technical and financial choices that are often anything but obligatory, utility of the enterprise that only becomes clear after the event, unsure future maintenance costs) - if anyone declares that they want to participate in a collective enterprise, or shows signs of wanting to join, they are likely to be considered superficial, lightweight or untrustworthy by others, especially by those who are most interested. Those interested parties are likely to think that the other parties acted only in order to draw them out, dump the responsibility for the enterprise on them, and then reap the benefits. The result is that the most interested and reliable parties, for whom the 'infrastructure' is most necessary, withdraw from the enterprise. Moral hazard then takes place when participating members have been selected and the parties that want to obtain immediate benefits behave inconsiderately (over-demanding or excessively cautious), causing the overall cost of the infrastructure to rise. Such different mechanisms as adverse selection and moral hazard thus work hand in hand to throw obstacles in the path of collective action.

The evolutionary paradigm presents a solution. Without leaning on the paradigm of rational expectations[108], it relies on the hypothesis that "That what works well for a player is more likely to be used again, whereas what turns out poorly is more likely to be discarded. As in the game theory, the players use their strategies with each other to achieve a payoff based upon their own choice and the choices of others. In an evolutionary approach, however, there is no need to assume a rational calculation to identify the best strategy. Instead, the analysis of what is chosen at any specific time

---

[108] Game theory assumes the players are fully rational and choose the strategy that gives the highest expected utility over time, given their expectations about what the other players will do. Recent work by economists has shown great sophistication in dealing with problems of defining credible threats and of showing the consequences of requiring actors' expectations about each other to be consistent with experience that will be generated by resulting actions. Although deductions about what fully rational actors will do are valuable for their own sake, empirical examples of changing norms suggest that real people are more likely to use trial-and-error behaviour than detailed calculations based on accurate beliefs about the future. Therefore, I have chosen not to study the dynamics of norms using an approach that depends on the assumption of rationality. (Axelrod, 1997, p.47)

is based upon an operationalization of the idea that effective strategies are more likely to be retained than ineffective strategies. Moreover, the evolutionary approach allows the introduction of new strategies as occasional random mutations of old strategies." (Axelrod 1997, p. 47-48). Once a norm has been defined such as: "A norm exists in a given social setting to the extent that individuals usually act in a certain way and are often punished when seen not to be acting in this way" (idem, p. 47), if in a local context a norm is established by evolutionary means (trial and error handed down through the generations) this could give rise to the complete internalisation (within the local context) of the externalities stemming from incomplete information, and thus avoid the mechanisms of adverse selection and moral hazard.

Let us suppose that this norm dictates that one must keep a promise until all the people to whom the promise was made recognise that it is impossible to keep any longer, or that the commitments made have all been fulfilled. If a norm of this kind is established in a social context where a collective good needs to be realised, nobody will make a commitment light-heartedly and therefore everybody can relax. Moreover, nobody, once committed, will withdraw from the agreement, making everybody equally reliable. This situation is extremely beneficial for the collective good, as well as analytically interesting. There is still imperfect information regarding technical and financial choices, the utility of the enterprise that only becomes evident after the event, and future maintenance costs; there also continues to be asymmetry between agents. This imperfect information, however, could be dealt with as if it really were a good like any other that can be produced according to its cost in relation to its utility. In other words, while a reduction in the cost of imperfect information is not possible in general, it becomes possible in a context where a norm that avoids the risks of adverse selection and moral hazard has been introduced. In this context, then, we must identify: 1) what obstacles in the way of collective action are created by the very process that established the norm, and 2) what information and coordination costs need to be accounted for in order to calculate whether collective action is worthwhile.

There is thus a clear analytical path to follow, with the following steps:

1) Examine the process that leads to a norm being adopted whereby a promise is a promise;

2) Identify what conditions must be satisfied in order to establish this norm as well as the consequences in terms of constraints to collective action;

3) Calculate the information and coordination costs needed to sustain the achievement of collective action.

A promise is a promise

Following Axelrod (1997), we define behaviour as 'vindictive' or 'eccentric' when it is based on a cost-benefit analysis. You are vindictive if you punish behaviour that does not correspond to a socially recognised norm, and eccentric if your behaviour does not take social norms into account at all.[109] In the social context we are assuming here, the distribution of vindictive and eccentric characters is random. Let us suppose, then, that by chance somebody decides to behave according to the idea that a promise is a promise, and that this behaviour turns out to be advantageous. Axelrod demonstrates that the norm 'a promise is a promise' will at first spread rapidly, but that it will not become an established code of behaviour on the basis of three assumptions (an evolutionary mechanism, a random distribution of vindictive and eccentric behaviour, the greater benefits of a kept promise). What happens, he shows, is that, to start with, eccentric behaviour declines rapidly because it is very costly. Once the number of eccentrics has been reduced, they drag the vindictives down with them, because their behaviour is also relatively costly. At that point eccentric behaviour makes a comeback and becomes dominant, creating a stable equilibrium that does not allow the norm to become established.

The result is very different when a meta-norm is introduced along similarly evolutionary lines, according to which 'you must punish whoever fails to punish violations of the norm.' In this case the norm 'a promise is a promise' can be established. We shall now examine the conditions that allow this meta-norm to be introduced.

Conditions and consequent constraints

There are two basic requirements for a meta-norm to be introduced: dominance and reputation. By 'dominance' we mean the presence in a social context of a group that counts more, either in terms of numbers or of power, than other social groups or than the group formed by the rest of society. By 'reputation' we mean that a way of

---

[109] The social context is by implicit hypothesis based on identity. The cost sustained by those who punish people who do not respect the norm is the disapproval of the eccentrics, while the benefits are the approval of the vindictives.

behaving adopted by an individual in a certain context transmits signals that are interpreted as belonging to his or her 'nature' and that can therefore be used to define within that context his or her expected behaviour in other contexts.

Where there is dominance, a meta-norm is more likely to come about as the members of a dominant group can more easily punish other members of the same group (even if they do not punish those who do not respect a norm). Reputation reinforces dominance because networks and relationships among individuals are stronger within their own social group; it is advantageous for individuals to build a reputation in order to facilitate these multiform relationships.

And yet, these two mechanisms tend to establish norms that benefit the dominant group rather than the social context as a whole. In our case, for example, it is more likely that a norm is established whereby 'a promise is a promise only for the members of one's own group until… etc.' than a norm dictating that 'a promise is a promise for everyone, period'. In this instance, the norm is established by restricting its sphere of application, unless the mechanism of reputation succeeds in widening it (that is, if the network of relations were stronger among groups rather than within groups – but that is highly unlikely).

It can be concluded, then, that to establish a norm that internalises the externalities deriving from imperfect information in view of collective action as well as neutralising the effects of adverse selection and moral hazard, implies accepting constraints in terms of size and quality of the collective good. It becomes clear, in other words, that it is realistic to imagine successful bottom-up collective action when it is a matter of providing 'infrastructure' that is useful above all for the dominant group and for which that group has made a commitment. It is far less likely that a collective action is successful if it is needed by the whole social context and everybody's commitment is needed. A good example of this observation can be seen in Appendix 1.

Information and coordination costs

The problem of coordination regards the ability of agents to identify a single solution (the precise nature, size and location of the equipment) that mutually satisfies the requirements and plans of each of the individuals involved in the decision. This is a huge problem when, owing to the complementarity of individual actions and therefore increasing returns, there is no such thing as an optimal solution. In this case

what usually takes place is called 'coordination failure' (see, for example, Leijonhufvud, 1968; Cooper, John, 1987; Colander, 1992; Boitani, Delli Gatti, Mezzomo, 1992; Van Es, Garretsen, 1992). Economic agents who fail to harmonise their decisions as individuals simply do not make choices that have the potential to increase their wellbeing.[110] In more analytical terms, the problem of coordination imposes specific information costs once the commitment has been made to avoid making discordant choices. These costs vary according to the number of agents and the characteristics of the group involved in the choice.[111] One can conclude then that:

1) Agents can undertake collective action after defining the harmonisation of individual behaviours and selecting which of the many possible actions to undertake;

2) Synchronisation and selection costs affect the value of the coordination payoff and thus significantly condition whether the collective action is undertaken or abandoned.

Given the relevance of these synchronisation and selection costs it is therefore necessary to examine the factors that determine them.

Synchronisation costs

Synchronising the courses of action of individuals is subject to constraints (costs) that vary according to the nature of the equipment and the characteristics of the agents. Independently of the solution that has been found to be most effective, the synchronisation of decisions is itself conditioned by the presence of technological as well as information constraints (distorted communication, ambiguous signals, physical distance). The variety of contributions and heterogeneous time scales deriving from division of labour or role specialisation makes it very difficult to make cooperative decisions. These difficulties are amplified as participation widens. One result is that, except in the case of groups with a high degree of overlapping (Hardin, 1982), and cases where the same good is reproduced over time in an identical form

---

[110] It is important to stress that this problem is separate from opportunism. The reason why agents are unable to harmonise their individual decisions is not seen here as linked to the fear of the losses that could be inflicted by opportunistic behaviour on the other side. This is excluded by the assumed consolidation of the norm 'a promise is a promise'. One interpretation is simply ignorance of real opportunities for success and risk aversion.
[111] See Bulow, Geanakoplos, Klemperer, 1985, Artus, 1993.

(Lewis, 1969), agents are required to engage increasing resources as the group gets bigger and more heterogeneous. The consequence is that, already at this preliminary stage, incentives for collective action – meaning the difference between the payoffs for coordination and the payoffs for non-coordination – diminish.

Selection costs

Identifying a cooperative solution is conditioned to start with by the cost of acquiring and transferring the information that is relevant for the agents involved. Each agent must communicate whatever information might influence the outcome of the coordination solution (Arrow, 1974). Plans can only be drafted after enough information has passed hands regarding each individual's contribution, the features of the technology used, and the resources required for realising the infrastructure. The more homogeneous the information, resources, and technology are, the lower the transferral costs; the more heterogeneous, the higher the costs. Becker and Murphy (1992) have demonstrated that coordination costs increase alongside increases in division of labour and specialisation. Specialisation increases productivity but reduces the overall knowledge level, making communication and interaction between holders of differentiated knowledge more difficult. Moreover, the kind of specific knowledge possessed by individual agents must provide appropriate skills for drafting and evaluating plans. Finally, practical alternatives need to be identified and prioritised, with resources invested in research and development for alternative technologies with which to realise the required infrastructure. Only a complete picture of all possible alternatives allows for plans to be drawn up and decisions to be made.[112] Selection costs, moreover, are not usually limited to setting up projects; they generally extend to their implementation and efficient renewal (maintenance). Finally, widening the field to n agents[113] will determine an increase in average benefits and will reduce the costs of setting up the enterprise, but at the same time it will amplify the coordination problems. Any increase in participants requires additional information to be gathered and transferred to each agent, as well as a re-prioritisation of the possible alternatives. Widening the number of participants also tends to increase their heterogeneousness in terms of information and skills in evaluating projects. The risk is that the benefits of

---

[112] Schelling, 1980, p.84) pointed out the role of difficult communication and scarce information in conducting coordination games.

[113] The absence of opportunism continues to be assumed.

collective action in realising an infrastructure are annulled by the rising selection costs as more participants join the enterprise. This theoretical conclusion is well illustrated in Appendix 1.

Bottom-up collective action: club goods

Solving coordination problems by means of a decentralised and cooperative decision-making mechanism depends on two factors: a) the probability that the agreement, once reached, is subject to sanctions (we looked at this with reference to the emergence of evolutionary norms); and b) the possibility that the level of information, knowledge and contractual costs required to draft a project is not so high that it annuls the payoffs for cooperation.

We have seen that for both these factors, the difficulties multiply with an increase in number of participants, in heterogeneousness and in the number of alternative technologies available for setting up the equipment. The case study in Appendix 1 clearly shows this conclusion. It could be said, then, that when the infrastructure is simple (or well established) and useful for a restricted group - especially a dominant group - of participants in a social context, then cumulative circular causative links can be established that allow club goods to be established by means of bottom-up collective action without institutional intervention. When this happens, synchronisation and selection costs are reduced, the nature, cost and expected utility of the infrastructure are defined early on, the norms regarding the value of promises are promptly established and propagated. All these factors clearly contribute to the success of collective action.

Unfortunately, however, club goods are not what lagging local systems need the most. As we have seen, club goods guarantee minimum endowments, critical mass, nor external domination – features needed for infrastructure to make a difference to development. These features and the mechanisms we have described, if anything, help explain why there are collective goods and services in lagging areas, and why these are usually minimal and reserved for a restricted circle of people (those who are rich or influential, belong to a privileged family or clan, or are affiliated to criminal organisations).

By contrast, a number of collective goods are produced by means of bottom-up collective action in more advanced areas. It would seem that some of the constraints to collective action become less stringent in a well-developed or advanced context.

Networks of relationships among agents, as well as appropriate specialisation and division of labour, extend in all directions not just within a given social context. This gives the mechanism regarding reputation a greater chance to combat that of dominance, and makes it more likely that a meta-norm leading to greater respect for keeping promises is adopted throughout the community. Technological alternatives and private knowledge will also be better known and/or more easily communicated, while selection and synchronisation costs will be lower.

But in this case, the causative links are reversed. It is not collective action promoting development but the other way round. In order to make this clear, a graph such as that in Figure 13 could be useful. In this graph the dimension-effectiveness of the collective good in relation to the level of development is seen both from the 'demand' side of cooperative norms (required in relation to the dimension and technical nature of the infrastructure needed) and from the 'supply' side representing the establishment of such norms.

Figure 13 - Collective bottom-up action with opportunism
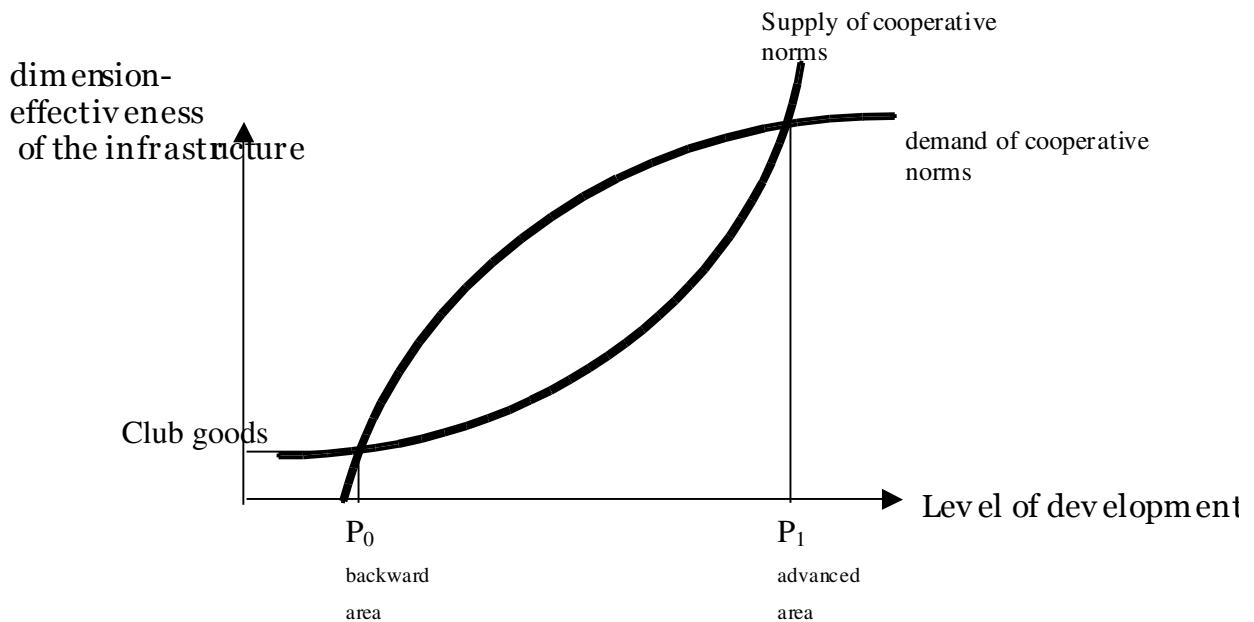
There are two conditions of equilibrium in which the two mechanisms – evolutionary and 'technical' – are complementary and reinforce one another. These are the equilibrium of backward areas and the equilibrium of advanced areas. The former (P_0) only allows bottom-up collective action to set up small or ineffective collective goods, while the latter (P_1) allows for more compelling action. In the area

between the two equilibriums, supply and demand for norms favouring cooperative behaviour do not meet. In order to emerge from underdevelopment, collective action would have to lie in this area, which, as we have argued, is highly unlikely. In backward areas, in fact, if norms of cooperative behaviour are established, they are likely to be among a restricted group of agents that is too small to create the necessary infrastructure and therefore make a difference for development. Setting up the necessary infrastructure at that point would be subject to adverse selection and moral hazard.

It is important to highlight the fact that even when these norms are established among the entire community and are therefore potentially powerful enough, there are no guarantees that coordination failure will not take place. Excluding uncertainty about the cooperative attitude of the participants, there would still be difficulties inherent in the high information costs involved in setting up the project. In this case, the absence of opportunism does not annul the problem of coordination, a point that the literature has largely ignored. If we assume that participating agents are perfectly trustworthy then there is no need for enforcement, but information costs are still high. A decentralised solution does not simply entail neutralising the effects of not keeping promises, or, more in general, the risk of free riding.[114] Considering the characteristics of collective action, non-opportunistic agents could come across many obstacles on their path to achieving collective action.

As Figure 14 shows, a solution that allows opportunism to be reduced by shifting the curve of cooperative norm supply upwards does not solve the problem of collective action in backward areas. Once coordination costs and opportunism have been introduced, the dimension-effectiveness of the infrastructure no longer depends only on the supply and demand of cooperative norms. It could be put forward that these norms determine the size of the collective that can take action, while the dimensions of the infrastructure also depend on information costs. In Figure 14 we imagine (fourth quadrant) that the greater the collective that potentially takes action, the higher the coordination costs. What could happen, in this case, is that the dimensions of the infrastructure are anyway no greater than those of club goods even when far-reaching cooperative norms are established. If the dimension of the

---

[114] Rizzo and Sindelar (1996) reach similar conclusions, showing that perfectly rational agents and contracts that are able to eliminate moral hazard do not necessarily avoid coordination failure.

infrastructure depends positively on the size of the collective but negatively on coordination costs it can be imagined that the ratio between dimension of infrastructure and coordination costs rises only very slightly (third quadrant). The result is that in any case the dimension of the infrastructure is reduced.

Many generous development efforts predicated on the basis of a sense of solidarity among local populations can be seen as aiming to shift the curve of cooperative norm supply upwards. Often those who support these efforts, however, neglect the information and coordination costs function. The result – as we often see –is that local actions can be successful on a small scale, but often fail with large-scale projects.

Institutional authority

There are quite severe limits to bottom-up collective action, then. It is unsuccessful when it is needed and has some chances of success when development is already underway. Setting up collective goods or services requires institutional support or substitution. The only way to achieve this is to channel the decision-making process into one decisional authority (concentrated decisional mechanisms), which acts on behalf of individual agents. This delegated action can take the form of incentives and direct support for collective action that reduce coordination costs.[115] It can also, more in general, produce public goods directly. Whatever the case, public intervention, unlike private action, is by definition 'for everyone' and 'for ever'. The pre-requisite for centralised decision-making is non-excludability: people would not transfer decision-making power to an authority that might then exclude them from the potential benefits of the decision. Similarly, an institutional action is obliged to provide continuity over time. This is clearly a 'functionalist' assumption, but it does not take away from the fact that in many backward areas - and even in some more advanced ones - institutions (or the State) are captured by interest groups and by no means guarantee continuity over time in their provision of public goods. This pathological state of affairs seriously impedes development.

---

[115] See Arrighetti, Raimondi (2000) for an empirical survey that sheds light on how collective action between firms (consortiums or associations of firms in the manufacturing sector and services in Italy), especially in the case of small and medium-sized firms, is successful when – in addition to previous collective action – local institutions intervened by means of incentives and direct participation in the consortiums of associations. If this is true in an industrialised country, it is all the more true in a developing country. See the examples in Appendix 2.

It is worth noting, in conclusion, that the institutional product is not only made up of regulatory goods. Current economic literature, which considers as being uniquely to "constrain and regularize behaviour" (Scott, 1995, p.35) as well as containing the consequences of failures in collective action brought about by free riding, restricts the sphere of influence of institutional action to an unjustifiably limited area. In the present approach, institutional action is extended to include a vast set of (material and immaterial) goods, with constraints of continuity and non-excludability, which are subject to collective action failure owing to coordination costs. In addition to norms and legal codes, and the administration of justice and defence, other fields that can be classified as public goods useful for economic development include interventions in the field of professional development, education, scientific research, infrastructure, the collection and dissemination of economic information, and the provision of social and cultural services.[116]

The limits of public intervention

So far we have said that to encourage development in a lagging area State interventions play a useful role, alongside markets and firms, and can become important examples of the way to go forward. The kind of institutional bodies we are talking about should be built 'for ever', involve everyone, not just those that want to take part, and be endowed with coercive power. From this point of view, institutional bodies are very different from markets and private organisations that are by definition temporary, based on voluntary participation, and lacking in coercive power.

Using this basic difference as his starting point, Stiglitz (1991) presented an analysis of public institutions (from now on, for simplicity's sake, we shall refer to them as 'State') that highlighted their limits compared to the greater efficiency of private organisations. He then went on, however, to claim that private organisations are unable to satisfy the demand for essential public and collective goods. The limits

---

[116] While this approach has had little attention in economic literature, it is widely adopted in other disciplines. Karl Wittfogel (1957), for example, when analysing the solutions adopted in Imperial China for its water projects, provides an example of a public good that required vast collective mobilization of resources . Referring to the link between cooperation and ex ante coordination, he points out that: "Orderly cooperation involves planned integration. […] Most writers who mention the cooperative aspect of hydraulic agriculture think in the main of digging, dredging, and damming; and the organizational tasks involved in these labours is certainly considerable. But the planners of a major hydraulic enterprise are confronted with problems of a much more complex kind. […] Even in its simplest form, agro hydraulic operations necessitate substantial integrative action. In their more elaborate variations, they involve extensive and complex organizational planning." (Harris, 1977).

of the State must be examined in detail not because we wish to conclude that private action is preferable to public, but in order to propose that development policies should be conceived hand in hand with policies that reinforce the administrative capacity of the State, thus reducing – as far as possible – its inefficiency.

The very nature of the State makes three aspects important: the choice of managers and political leaders, attending to fiduciary duties, and both planning and control.

Choice of managers

In conditions of imperfect information, the choice of managers is an 'agency problem' in all organisations. The interests of a 'principle' are carried out by 'agents' who are placed in control over resources that are not their own and who, for this reason, possess information the principle does not. Agents with delegated authority can therefore exploit their 'private' knowledge to feather their own nest at the expense of those whose interests they are supposed to be serving. There are many examples: a shipbuilder and a captain of his ship; a producer and a director of a film; partners and managers of a firm. In private organisations (which are voluntary and temporary) there are two circumstances or mechanisms that help reduce the agency problem in choosing managers, though they do not solve the problem altogether. For organisations this means focussing (or limiting) their final aims and subsequently the tasks of those in charge; for partners, it means acquiring the same – or at least similar - skills and knowledge as their managers in order to control them. In this way managers are selected to carry out a limited number of well-defined tasks, and they are accountable to a higher level that shares their knowledge. Private organisations are easier to focus and control because they are temporary and have a limited number of participants. If partners are unhappy with their managers they can leave the company, and managers would lose their job.

The State ('for everyone' and 'for ever') is much harder to focus and control because it serves an unlimited number of purposes and no group of citizens can know enough to supervise its functioning. The Public Good is a public good, and there is no interest in spending in order to raise awareness. Moreover, a vote carries little weight. Internal incentives, such as those used in private organisations, are less effective in recruiting and keeping good administrators and political leaders, which means that external conditioning and pressure take over. At that point a situation well described by Hirschman (1970) is established: in a State, where participation is universal and

involuntary, there is a way of 'defecting' that replaces Exit, which is not allowed, or very costly. In a democratic State it is formally sanctioned by the possibility of abstaining from voting in elections. In any case 'defection' is expressed by disinterest, disenchantment, a low consideration of politics and politicians, and, ultimately, by boycotting the system. Defecting thus lends strength to loyalty; and a mechanism similar to Axelrod's is established. As we have said, positive attitudes to establishing a norm generally survive alongside eccentric positions. Except that the State, being an organisation 'outside' civil society, provides no guarantee that the evolutionary mechanism will kick in. Those interested in politics can actively construct attitudes that valorise or devalue it, thus creating self-fulfilling prophecies. If the attitude cultivated places value on loyalty or approves of a policy, then competent, incorrupt managers (inasmuch as there are any) will be selected, and they, in their turn, will make politics more attractive. If, on the other hand, the prevailing attitude is one of disaffection that socially devalues politics, then incompetent and corrupt managers will be chosen, their policies will be mediocre or even reprehensible, and this will confirm the prevailing opinion that politics is by definition corrupt and incompetent.

In backward areas, then, there can be obstacles to development – which are very had to avoid - when dominant and socially privileged social groups, whose aim is to protect their own interests, on the one hand do not let behavioural norms that would benefit everyone be established, and, on the other, cultivate and encourage negative attitudes towards politics, and thus manipulate its outcomes. And yet, members of so-called 'civil society' can (and often do) commit the error of trying to be a foil to politics, cultivating the idea that where they are involved they are doing good and wherever politics is involved there is only rot. In this way, civil society deprives itself of the chance to select good politicians and good managers, and therefore to have the public goods and services that civil society is unable to achieve.

Figure 15 represents this situation. The horizontal axis is the level of inclusion, which is lower for private firms than for State firms. The vertical axis represents managers' competence. The graph shows the hypothesis that demand for competent managers grows according to the level of inclusion (dimension of the organisation and therefore complexity of their managerial tasks) when the organisational formulas are the same, but that it is higher in the private sector than in the public. This is because when we say 'competent managers' we mean competent in an economic sense. If we shift our attention to the State, a manager's skills are less focussed and

they are required to take on a wider variety of tasks. The graph therefore indicates 'supply', which presents a bifurcation that widens as inclusion increases. The hypothesis that what counts is the perceived quality of politics is the most convincing: a high social opinion of politics allows better managers to be selected. By contrast, when the common view is that politics is degraded, it is much harder to find competent managers (the best managers are not willing to work for the State). As the graph shows, there are three states of supply-demand equilibrium: one in the case of private firms and two in that of the State. As we have said, the prophecies are self-fulfilling: when politics is perceived as valuable (greater loyalty), better managers are recruited. The quality of political life will improve, there will be greater inclusion, and the functions of the State will be amplified. When the opposite is true, worse managers will come forward; politics will be worse, there will be more disaffection, less inclusion and a 'minimal state'.

Fig. 15 - Supply and demand of competent managers

Managers' competence

"demand"
in the private sector

"demand"
in the case of the state

"supply"

inclusion

Equilibrium
in the private
sector

low
value

high
value

Indeterminate equilibriums in the case of
the state
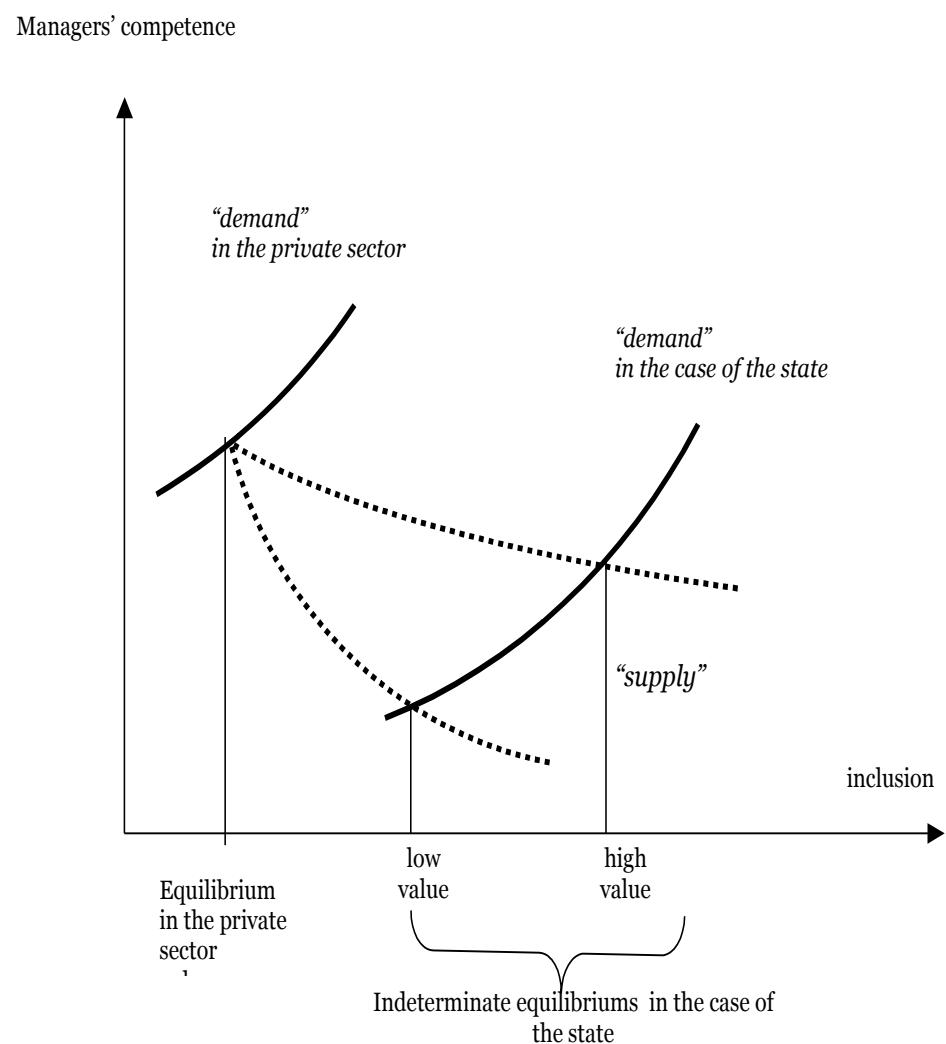
Fiduciary duties

In a condition of uncertainty, the legitimation of the authority of the State in the production of public goods and services cannot come about simply on the basis of results, which are still unavailable when the State must take action. The State, and institutions in general, must take on fiduciary duties. Responsible institutions demand, and citizens call for, constraints. As far as economic activity is concerned, there are employment constraints and constraints of equity.

Employment constraints

Every organisation functions on the basis of leadership (managers, agency problems) and the executive apparatus (hierarchy). The executive apparatus of a private organisation usually has incentives to carry out orders efficiently on the basis of a relationship whereby salary depends on performance and there is a credible threat of losing the job (or of the firm going bankrupt). In the State there are constraints on public spending and tax revenues that make it impossible to pay wages that are commensurate with performance, making the wage relationship very different from that in the private sector. The State pays less but guarantees a job (generally States don't go bankrupt). Wages in the State sector play an arbitrary role, and the threat of losing your job is less credible. The executive apparatus thus has fewer incentives to create efficiency.

Equity constraints

Constraints on public spending and tax revenues in the State are not only quantitative but also qualitative: public spending and tax revenues must be distributed fairly. This is the other side of fiduciary duties. Establishing what is considered a fair distribution of spending and revenues in a way that is accepted by the entire community is impossible and everyone knows it. The State therefore exploits the circumstance that what it does must be perceived as fair, even if it is not. This inevitably leads to institutional hypocrisy, greater spending for communication and the manipulation of consensus. In its turn, this leads to hierarchical inefficiency: managers (politics) could be tempted to place greater value on semblance than action.

Efficient hierarchical relationships

Once we have identified the variable 'efficient hierarchical relationships', we can represent in a graph the consequences of State fiduciary duties as employment and equity constraints. What we mean by efficient hierarchical relationships is the extent to which managers are able to obtain active collaboration from the people they are in charge of, given the characteristics of the organisation. In private organisations, hierarchical relationships are e more efficient than in the State, which, because of the constraints on public spending and tax revenues, cannot pay its employees according to productivity and tends to value appearance more than (or at least as much as) action. We can therefore imagine that there is an inverse ratio between the variable we

use to characterise organisations – inclusion – and the variable of hierarchical relationships. Moreover, because in the case of the State (and in general for inclusive institutions) employment and equity constraints are both important, we can imagine that this inverse ratio increases with inclusion, as in Figure 16.

Fig. 16 - Efficiency of the hierarchical relationship
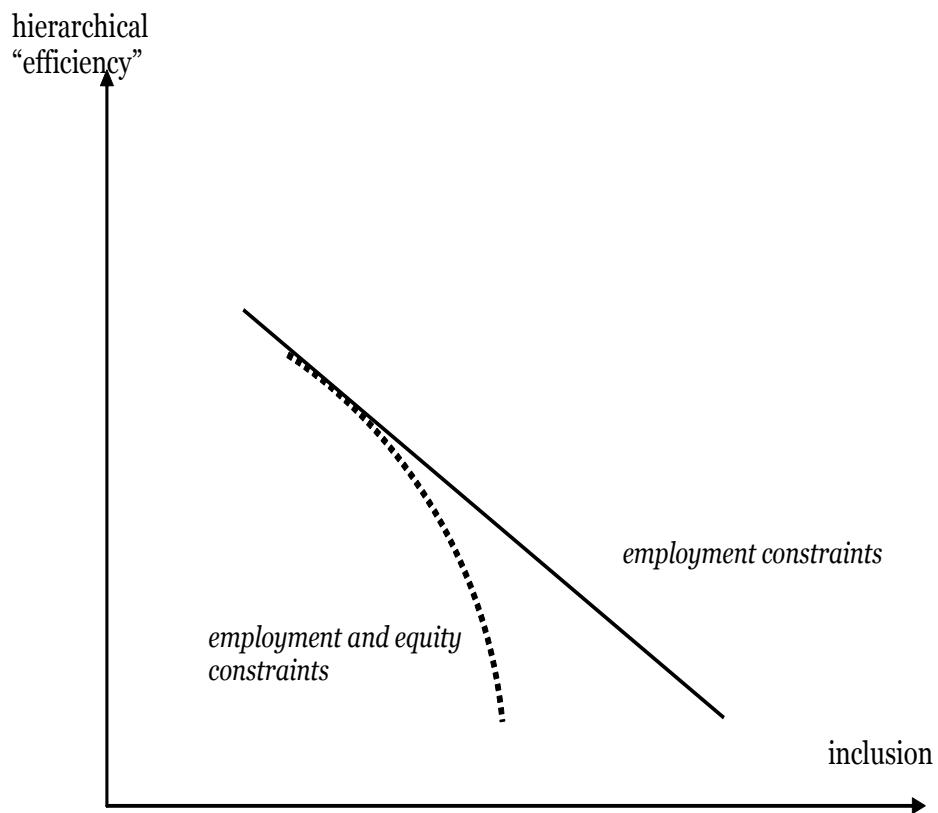
Let us now put Figures 15 and 16 together. In Figure 15 we saw a decreasing ratio between inclusion and management skills with a bifurcation that identified three different levels of competency in correspondence with the State, the 'minimum State' (when politics is devalued), and private organisations. In Figure 16 we saw another decreasing ratio between hierarchical relationship and inclusion. If we put the two graphs together, as in Figure 17, there is an association between the three competency levels (the State, the 'minimum State', and private organisations) and three distinct levels of efficiency in the hierarchical relationship. The result is very interesting. While it is predictable (given the assumptions we have made, which are evident *a priori*) that private organisations can attract better managers and at the same time enjoy greater hierarchical efficiency, what is surprising is that the State has the lowest level of hierarchical efficiency but an intermediate level of competence in its

managers, higher than in the 'minimum State'. This result is not immediately evident *a priori* and is the specific consequence of the cumulative effects of the appreciation or depreciation of politics.

Fig. 17 - Quality of managers and hierarchical relationships
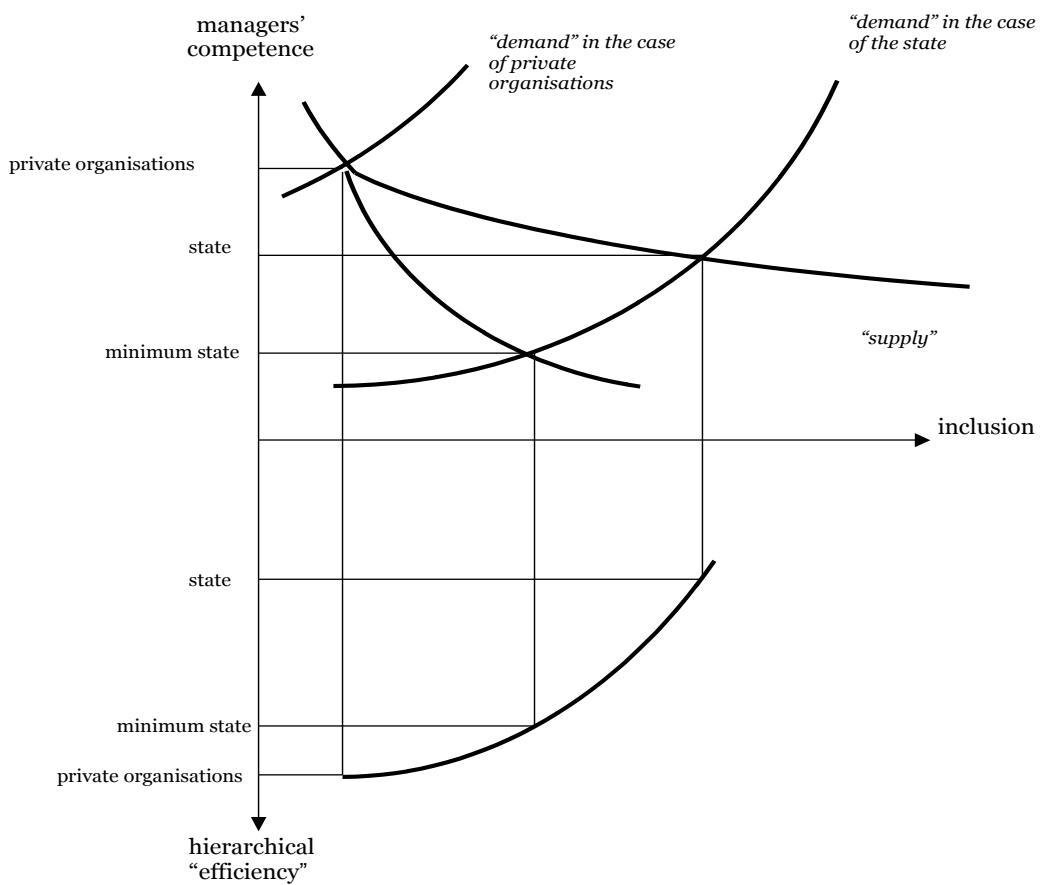
The graph suggests an outline for a micro-economic theory of the State. We have so far established three variables: competence of managers, inclusion, and hierarchical efficiency. We have also defined two relations: between competence and inclusion, and between inclusion and hierarchical efficiency. We need two more variables: the ultimate aim (for example, efficiency-effectiveness, or the extent to which what is needed for the public good is obtained at the lowest possible cost), and an intermediate measure of whether results are being obtained (planning and control). We also need two more relations: between hierarchical efficiency and a results-

oriented approach, and between different combinations of results-orientation, competence and the ultimate aim of efficiency-effectiveness (the problem of public interest).

## The measurement of results and the problem of planning and control

Difficult dilemmas about how to allocate resources without adequate information, and especially about results, arise in every organisation; decisions can turn out to be wrong when it is too late to do anything about them. That is why planning and control are important. Centralised planning and control requires rigid organisations within which specialised organisms are responsible for each and every function according to a sectorial strategy. Every sector of State activity is organised vertically. The centre administers its periphery by means of separate commands with no links between them. There are three disadvantages to this logic. First, rigid organisations tend to reduce or hide available information rather than reveal it. Second, they allow for less diversification of risk, and, finally, they cramp the learning process because experimentation is not encouraged. At the same time, there are advantages for 'degenerate' and decentralised organisations.[117] In these bodies, the strategy is horizontal, and sectorial interventions take place within a framework that provides a common outlook for every area, and involves local administrations to a greater degree. This kind of organisation in public institutions is apparently more chaotic and less specialised, thus more 'degenerate'.

Where information is incomplete, we can only talk about a results-oriented approach, while in rigid organisations there is - by their very nature - a great deal of resistance to this approach. We could even say that by their nature they are oriented towards the procedure or the process rather than results – in short, that they are bureaucratised. Maintaining a sectorial logic, we can envisage an inverse ratio between a results-oriented and a procedure-oriented organisation. The inefficiency of a wage relationship (you are paid less in exchange for a guaranteed job and you are paid to seem you are doing something rather than actually doing it) requires strong

---

[117] 'Degenerate' organizations are those where every department or organ is potentially able to undertake many tasks (in contrast to 'redundant' organisations). In Lane and Maxfield's approach (2005), degenerate firms are better able to innovate and are less prisoners to routine. The results of complexity theories (Van der Leeuw, Aschan-Leygonie, 2000) and of biology (Edelman, Sporns, Tononi, 1999; Edelman, Gally, 2001) take on new value in this light.

incentives for a procedure-oriented rather than a results-oriented approach. An inefficient hierarchical relationship in terms of incentives is defended (in the joint interest of the managers and the organisation) by means of a highly formalised set of norms – a fixed procedure – that can easily be manipulated by the organisation for its own interests. This can be represented in a graph, where the right side of the Figure is the same as in Figure 17, and the new context is added in the bottom left-hand corner. In Figure 18, the two variables are results-orientation and hierarchical efficiency, with a direct relation between them.

Fig. 18 – Results - orientation
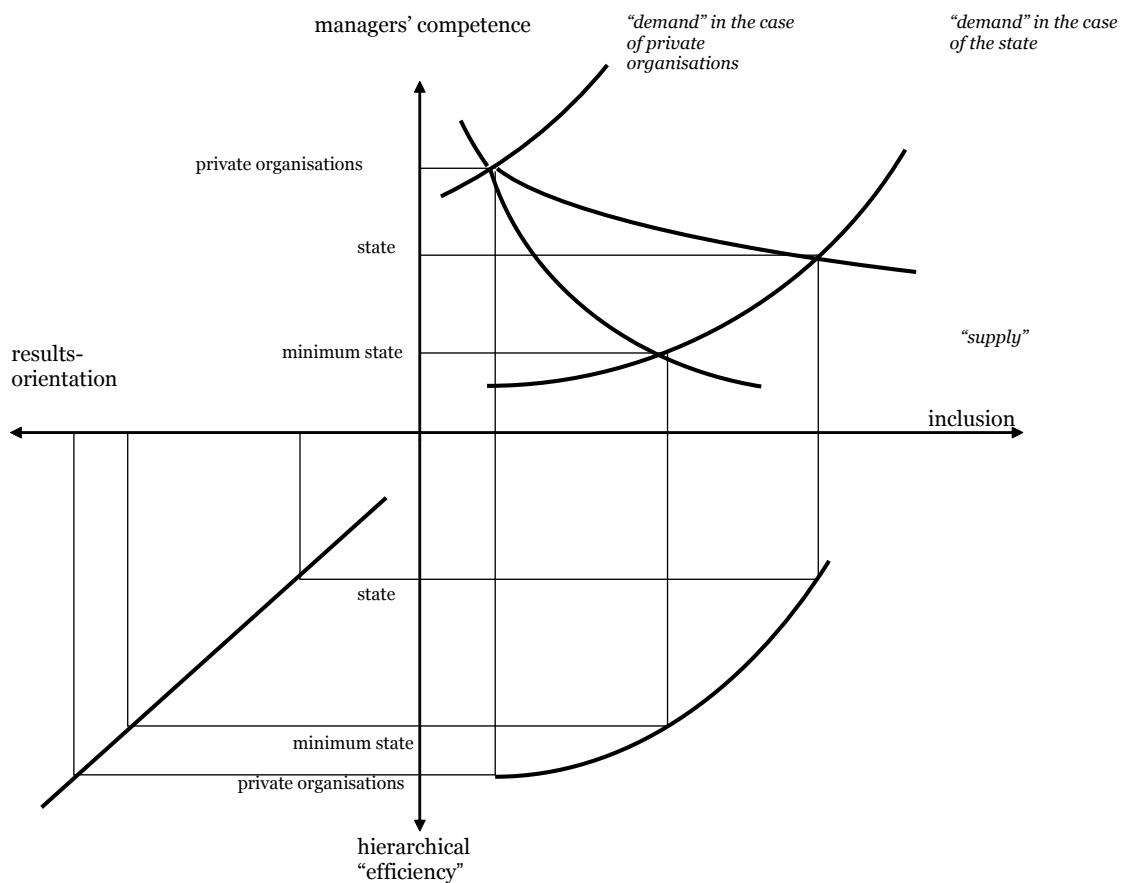
As we can see, results-orientation rises as procedure-orientation diminishes, and procedure-orientation diminishes as the hierarchical relationship becomes more efficient. Our three focal points (private organisations, State and 'minimal State') are

thus transformed. Until now these three points were characterised by different levels of managerial competence and hierarchical efficiency; we now have three different levels of results-orientation. The order is the same as before, but private organisations are the most results-oriented, followed by the 'minimal State' while the State is the least results-oriented.

In order to complete the model we need to add a relation between combinations of managerial competence, results-orientation and efficiency-effectiveness (the problem of public interest). It can be observed that incompetent managers and procedure-orientation give rise to public organisations that act according to private interests and are therefore against public interest. Similarly detrimental to public interest, however, are competent managers in a procedure-oriented organisation (prisoners to the apparatus), and results-oriented organisations with incompetent managers (absence of strategy).

The last piece of the mosaic needs to be added. Figure 19 completes Figure 18 adding the top left-hand corner, which illustrates the function of efficiency-effectiveness, and which, in its turn, depends on two variables: the competence of managers and results-orientation. We thus have $E=E(C,O)$ where E stands for efficiency-effectiveness (ultimate aim), C for managerial competence, and O for orientation to results.

Fig. 19 – Efficiency-effectiveness (a)
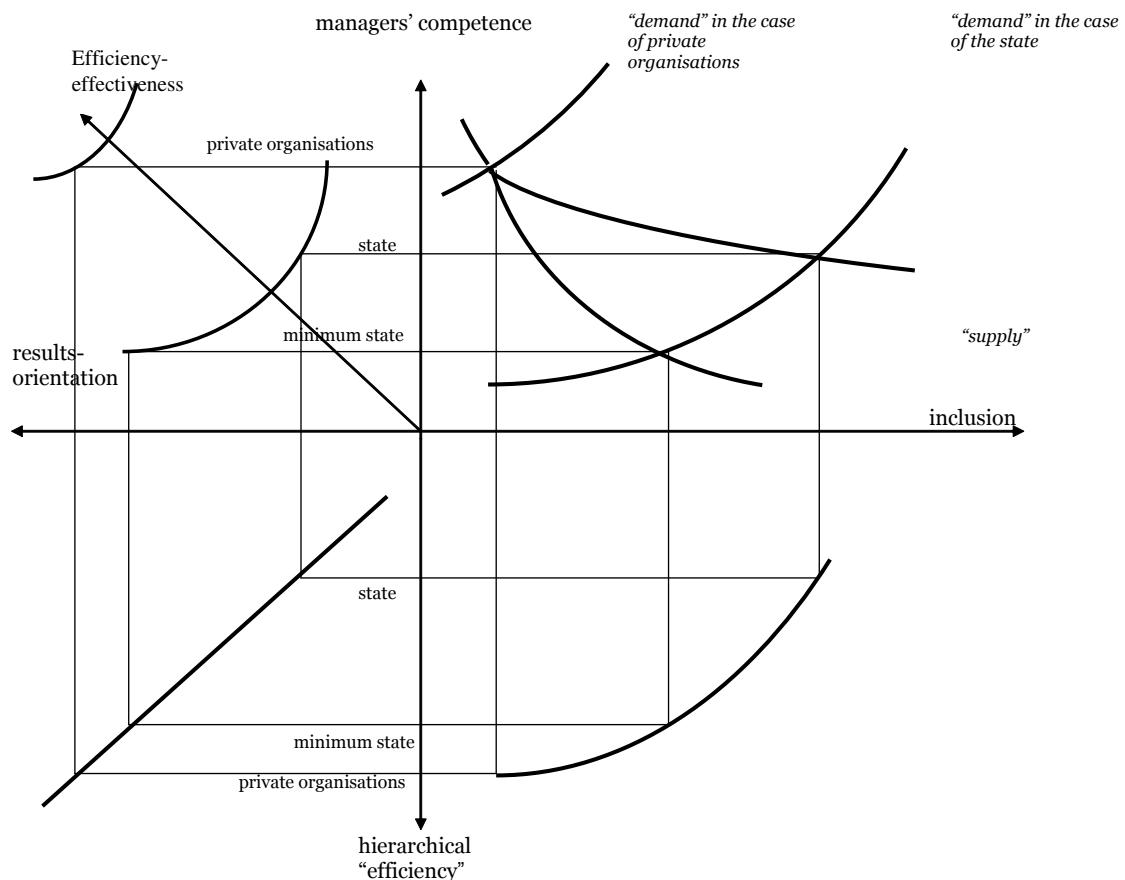
In this graph we obtain that private organizations achieve the best results in terms of the ultimate aim of efficiency-effectiveness when they have competent managers and when they are more results-oriented. The most interesting result, however, regards the State. There is, in fact, a chance that State and 'minimum State' achieve the same objectives: the State by means of more competent managers, and the 'minimum State' by increasing their results-orientation. To conclude, when institutions and, in particular, public institutions intervene in the same fields and with the same aims as private ones, in a centralised and sectorial manner, the State is bound to fail. This failure is equally serious in the case of a 'minimum State', which is usually lacking in strategy. As we have seen in the early sections of this chapter, however, there are goods and services that private companies are not able to provide

and that cannot be achieved even through bottom-up collective action. Figure 19 is therefore not quite complete.

One way to indicate that the State can achieve what private organisations cannot achieve is to modify the ratio between hierarchical efficiency and results-orientation in the third quadrant, identifying two ratios – one for private organisations and one for the state. In Figure 20 the graph is finally complete.

Fig. 20 – Efficiency-effectiveness (b)
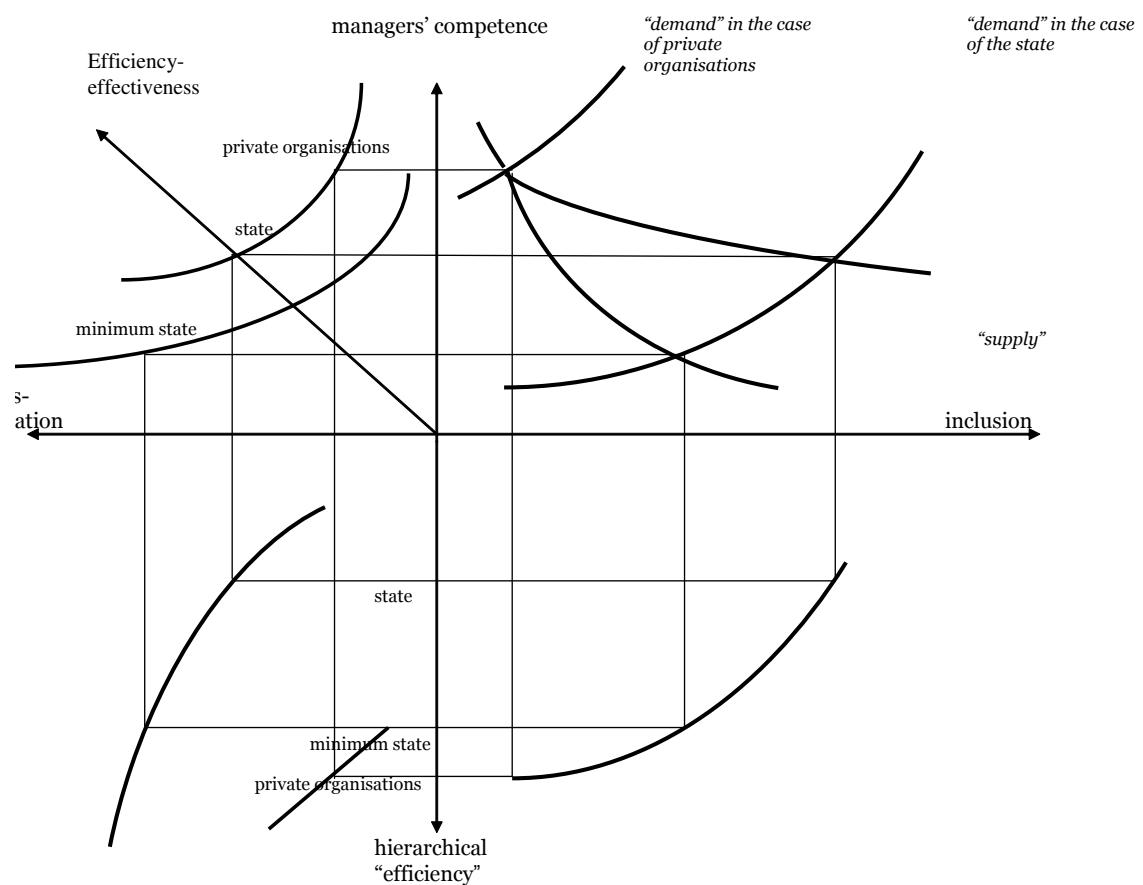
A new line has been drawn in the third quadrant indicating a ratio between results-orientation and efficiency in the hierarchical relationship shifted upwards in the case of the State. The idea that the State can achieve what private organisations cannot achieve is thus depicted graphically. As a result, the fourth quadrant shows a possible consequence: that the State can reach the same objectives at the same level as those

reached by private organisations. In this case, the inefficiency of the State is compensated by the fact that it can intervene and act in environments in which private organisations cannot function. It is important to stress, however, that this compensation does not mean that the micro-economic inefficiency of the State is cancelled out. In short, the 'minimum State' appears to be the worst of both worlds: more inefficient than private organisations in fields where the private sphere can achieve results, and less effective than the State in fields where the private sphere is less compelling.

Conclusions

In the presence of incomplete markets owing to incomplete information, which are particularly serious in lagging areas, an increase in productivity of resources requires public goods to be set up. Generally speaking these public goods cannot be realised by means of bottom-up collective action, because of the obstacles in the path of collective action, which include adverse selection, moral hazard, and coordination costs. A centralised State, on the other hand, is highly inefficient compared to private organisations. It is hard for the State to recruit competent managers, the executive apparatus is less results-oriented, and managers and apparatus can collude to feather their own nests against public interest. And yet, public intervention is inevitable. We thus need to take into account that, in addition to development policies that actively realise public goods, specific action must be taken to address this inefficiency. The analysis we have undertaken suggests two ways. First, politics should be sustained and valorised, and those who predicate its devaluation should be countered. Second, an institutional framework and examples of best practice should contribute to making the State less rigid and centralised, and more flexible, 'degenerate' and decentralised.

Bibliography for chapter 3

Arrighetti, A. e M. Raimondi (2002), *Fattori influenti sull'azione collettiva tra imprese*, WP Dipartimento di Economia, Serie Economia e Politica Economica Università di Parma.

Arrow, K.J. (1974), *The Limits of Organization*, W.W. Norton & Company, New York and London.

Artus, P. (1993), Défauts de coordination des activités, *Revue économique*, 44, pp.551-568.

Axelrod, R. (1997), *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton Studies in Complexity – Princeton University Press, Princeton.

Barrington Moore jr (1969), *Le origini sociali della dittatura e della democrazia*, Einaudi, Torino.

Becker, G.S. e K.M. Murphy (1992), The Division of Labor, Coordination Costs, and Knowledge, *Quarterly Journal of Economics*, CVII, pp.1137-1160.

Bell, D. (1993), *Communitarianism and its critics*, Claredon Press, Oxford.

Boitani, A., D. Delli Gatti e L. Mezzomo (1992), Concorrenza imperfetta, esternalità e spiegazioni delle rigidità nella Nuova Economia Keynesiana, *Economia Politica*, IX, pp.299-361.

Buchanan, J.M (1989), Una teoria economica dei club, in Id, *Stato mercato e libertà*, Il Mulino, Bologna.

Bulow, J.I., J. Geneakoplos. e P. Klemperer (1985), Multimatket Oligopoly: Strategic Substitutes and Complements, *Journal of Political Economy,* June, pp.448-511.

Cooper, R. (1987), Dynamic Behavior of Imperfectly Competitve Economies with Multiple Equilibria, NBER Working Paper, n.2388.

Cooper, R. e A. John (1988), Coordinating Coordination Failures in Keynesian Models*, Quarterly Journal of Economics*, August, pp.441-465.

Del Monte, A. e A. Giannola (1997), *Istituzioni economiche e Mezzogiorno*, Nuova Italia Scientifica, Roma.

Edelman, G.M. , J.Ae Gally (2001), Degeneracy and complexity in biological systems, *Proceedings of the National Academy of Sciences PNAS*, 98, 24, pp. 13763-13768.

Edelman, G.M., O. Sporns, G. Tononi (1999), Measures of degeneracy and redundancy in biological networks, *Proceedings of the National Academy of Sciences PNAS*, 96, pp. 3257-3262.

Fiorentini, G. e C. Scarpa (1998) (a cura di) *Cooperative e mercato*, Carocci, Roma

Hardin, R. (1982), *Collective Action*, The Johns Hopkins University Press, Baltimore.

Harris M. (1977), *Cannibals and Kings. The Origins of Culture*, Random House, New York; trad. it. *Cannibali e re. Le origini delle culture,* Feltrinelli, Milano 1994.

Hirschman, A.O. (1970), *Exit, Voice and Loyalty: responses to decline of firms, organizations and states*, Harvard University Press, Harvard; trad. it. *Lealtà, Defezione, Protesta. Rimedi alla crisi delle imprese, dei partiti e dello Stato*, Bompiani Milano 1982.

Lane A.D., e R. Maxfield (2005), Ontological uncertainty and innovation, *Journal of Evolutionary Economics*, 15, pp. 3–50.

Leijonhufvud, A. (1968), *On Keynesian Economics and the Economics of Keynes*, Oxford University Press, Oxford.

Lewis, D.K. (1969), *Convention. A Philosophical Study*, Harvard University Press, Cambrudge (Ma); trad. it. *La convenzione. Studio filosofico*, Bompiani, Milano 1974.

Rizzo, J.A. e J.L. Sindelar (1996), Optimal Regulation of Multiply-Regulated Industries: the Case of Physician Services, *Southern Economic Journal*, 62, pp.966-978.

Schelling, T.C. (1980), *The Strategy of Conflict*, Harvard University Press, Cambridge (Ma).

Scott, R.S. (1995*), Institutions and Organizations*, Sage Publications, London.

Seravalli, G (1988), *Cooperazione e autogestione nel settore industriale*, Nuova Italia Scientifica, Roma.

Stiglitz, J. (1991), The Economic Role of the State: Efficiency and Effectiveness, in T.P. Hardiman and M. Mulreany (a cura di) *Efficiency and Effectiveness in the Public Domain, The Economic Role of the State*, Dublin Institute of Public Administration, pp. 37-59; trad. it. *Il ruolo economico dello Stato*, Il Mulino Bologna 1997.

Stiglitz, J. (1988), Economic Organization, Information, and Development, in H. Chenery e T. N. Srinivasan, *Handbook of Development Economics*, vol 1, pp. 93-160, North Holland, Amsterdam, New York, Oxford.

Stiglitz, J. (2003-04), Information and the Change in the Paradigm in Economics, *The American Economist*, 47, 2, pp. 6-26; 48,1, pp. 17-49.

Taylor, M. (1987), *The possibility of Cooperation*, Cambridge University Press, Cambridge.

Van der Leeuw S. E., Chr. Aschan-Leygonie (2000), *A Long-Term Perspective on Resilience in Socio-Natural Systems*, Santa Fe Institute, Paper presented at the workshop on System shocks and system resilience, held in Abisko, Sweden, May 22-26

Van Ees, H. e H. Garretsen (1992), On the Contribution of New Keynesian Economics, *Eastern Economic Journal*, 18, pp.465-477.

Wittfogel, K. (1957), *Oriental Despotism. A Comparative Study of Total Power*, Yale University Press, New Haven; trad. it. *Il dispotismo orientale*, Vallecchi, Firenze 1968.

Chapter 3 Appendix 1 - Misiones: a case study of collective action failure

Yerba mate is not just a plant and exotic hot drink few people know in Europe; it is a symbol of identity for Argentina and neighbouring Latin American countries such as Brazil, Paraguay and Uruguay, where it was drunk in pre-Columbian times. The plant and its brewed infusion (*mate*) maintained and even increased its importance after the conquest, becoming a profitable economic activity, in particular for the Jesuit missions who earned significant revenues from the trade. Its spread, and the creation of industries to transform the leaves coincided with the birth of Argentina as a nation. A new nation with a mixed population needed unifying symbols to represent the spirit and life-style of the emerging State. Drinking *mate* with friends is a common and time-consuming social ritual that represents just such a symbol. The brewed *yerba mate* is steeped in a decorated hollow gourd (also called a *mate*), and the metal straw (*bombilla*) is often preciously carved. Initially the brew was consumed in the countryside as a social pastime, but the habit soon caught on in the cities and

throughout all social classes. Candidates for political elections can be seen campaigning with a *mate* in their hands, and historical figures are depicted after their heroic conquests sipping the drink with their comrades.

From the point of view of local development studies, it is hard to find a product that incarnates as effectively the idea that valorising a relevant immobile resource can set development processes in motion. Owing to the particular nature of the climate and soil, the *Yerba mate* plant can only grow in a limited area, the northernmost province of Argentina called Misiones. The product thus has a guaranteed market, anchored in the creation of a national identity, and at the same time a territorial monopoly. It is interesting, therefore, to investigate why these factors have not led to development in the region, which, on the contrary, is one of the poorest in Argentina. The marked contrast between potential and results has led to a certain amount of research and documentation undertaken, in particular by academics in the University of Posadas, the capital of the Misiones region. Together with field research and on-site interviews[118], the reasons for this failure can be delineated fairly clearly.

What is clear from the start is that *Yerba mate* could have been valorised to the extent that it could have become a factor of development; for a while it was even known as 'green gold'. There are also two very clear tendencies: first, the production process became increasingly vertical (without a significant or growing presence of cooperatives, which would have sustained the producers' income); second, large distribution companies slowly bought out independent industries (*molinos*) with their own brand name. These twin processes have contributed to worsening the living conditions of vast swathes of the population. A third factor is that the social actors involved in the production process have never been able to come to an agreement over how to valorise the resource, even though benefits would accrue to all of them if they could. Even with a recent opportunity, afforded by a new negotiation tool envisaged by a new law, which was passed in order to obviate the serious damage inflicted on the sector by deregulation during the last decade, agreement cannot be reached. Owing in part to a fall in prices, relations between different actors in the production process have soured even further, and the new measures fail to live up to

---

[118] Apart from written material, this study is based on interviews with small producers, managers of organizations defending the category of producers, aid workers involved in development projects, and academics. Direct contact with these participants was essential for understanding the context and mentality of the actors involved, as well as appreciating their point of view and attitudes.

their expectations and do not present solutions to the underlying structural problems. The very subsistence of small producers – the majority of the population - is at stake, but each component of the production process clings to the short-sighted view. Small producers consider the renewed promise of a minimum price level an important victory, even though the level does not guarantee them a dignified standard of living. The middle-sized firms and transformers defend their profit margins and try to keep prices down, at the cost of damaging the prospects of the small producers and salaried farm hands, by employing illegal, unprotected workers. None of the parties realise that they need to organise in order to keep large distribution companies at bay. Finally, cooperatives fail to appear although this form of organisation could be the winning card in the context.

Cooperation represents the only possible solution for small producers. Cooperative management of funds would provide access to inputs that individual farmers cannot even strive for. Similarly, the creation of integrated cooperatives to oversee all the transformation stages would create greater transparency and justice in labour and production relations. Moreover, it would be the only way to hold back the advance of large distribution, because it would allow producers to compete by means of quality and diversification rather than through price reduction. There are a few integrated cooperatives in the region, the Santa Pipa, is a good example. It soon became clear, however, in the interviews conducted, that cooperatives were a last resort. Small producers are harried by life's daily necessities and struggle to sell their products at the best possible price. Managers of their organisations are committed to the battle to obtain an acceptable basic price, and do not get round to debating the structural position of the *molineros* or the commercialisation of *mate*. Both groups support the idea that wide participation is important because increased funds and large quantities of raw materials are necessary. In view of these requirements, agreements should include the *molinos*, and extend to packaging and distribution firms. And yet the financial risks of entering a cooperative project of this kind are so great -, especially for the packaging and distribution sectors - that no agreement is possible without a significant contribution from the government, which would need to create the necessary structures. The government, of course, has no intention of intervening, and limits its scope to regulating markets.

The situation clearly illustrates a textbook failure of bottom-up collective action. Vertically integrated cooperatives, or networks of firms, need to be set up with the

following characteristics: a minimum endowment (big enough or numerous enough to involve a significant number of small producers and farm hands); critical mass (*molinos* and packagers must be involved); external dominance (the cooperatives must be able to compete against large distribution and run significant risks, with the support of outside contributions. Although the region has had experience with cooperatives, there has been no evolution in this direction, despite the fact that all parties know that they represent perhaps the only structural solution to their recurrent problems. As Axelrod described so well, the reason why this evolution has not come about is that dominant groups (*molineros*) are convinced they do not need them. The textbook problems of information and coordination costs are also present in this example: none of the parties involved in production appreciates the overall process, and nobody commits to making it more efficient.

It is also interesting to note that this failure of bottom-up collective failure appears to be caused mainly by one unassailable factor: the short-term profit mentality that is widespread at all levels and all social classes and that has been ascertained not to be attributable only to a state of emergency. This would seem to call for a cultural explanation, but in this case the explanation is not cultural but institutional. The failure of collective action was determined by the lack of state interventions to promote and support production. State initiatives have always been limited to regulating quantities by restricting harvests; they have never dealt with the structural problems behind production instability and over-supply. There is another historical, political and institutional factor to consider. At the crucial time when the province of Misiones was established within the State of Argentina, land was divided up following the logic of exploitation rather than of long-term valorisation. This created a deep-rooted "robber" mentality, whereby *mate* production was not perceived as an opportunity, and certainly not as a collective opportunity to preserve and develop. Just as collective action theory dictates, a strong heterogeneity of agents, owing to an unbalanced distribution of resources (as we shall see when we look at the pyramid structure of *Yerba mate* production), together with a context lacking in institutional interventions, has created mechanisms regulating economic and social activity which are far from efficient but nonetheless surprisingly tenacious.

The following sections will argue this synthesis. The first section examines salient moments in the history of the province that led to the fragmentation of production activity and interest groups. The second section outlines briefly the current situation,

stressing the tragic level of poverty in the majority of the population. The third section illustrates the production cycle of *Yerba mate* and the relationships that have been established among the various actors, evidencing the reasons why so many small producers are unable to obtain fair dues for their labour. The fourth section describes the new giant that has become part of the production process: large distribution companies that supply the market under their own brand-names and have huge bargaining power, allowing them to cut producer prices. The fifth section analyses the role of decades of regulation. The sixth and final section looks at the political ad social actors participating in the debate over the recent law and its implementation, and shows that its claims and its potential results cannot be considered sufficient to bring about a significant change in the economic outlook of small producers.

A little history

The territory included in today's Misiones, as well as neighbouring regions, were settled in pre-Columbian times by tribes belonging to the Tupì-Guaranì linguistic family. The Guaranì tribe achieved considerable social and production levels of development, and their cultivation techniques were well suited to the climate and soil. Colonial occupation changed this picture completely: the pattern of domination in the Latin American subcontinent considered conquered terrains as exclusively in the service of the conquering power. Colonisation in the region of Misiones was organised with one aim in mind – exploiting the rich seams of precious metal already mined in Peru. This required absolute control over the territory in order to guarantee the safe road transport of valuables to and from Potosì. The agricultural system also needed to be reorganised as a satellite economy for mining. The region was thus inevitably subordinated to the interests of the city of Potosì, though it was still regarded as highly strategic. The region became a breadbasket for the mining district, which meant a lot of ready labour was needed. The system of *encomiendas*[119] established with the first land distribution in 1555 considered the natives free citizens, but they were obliged to work in construction, or in defending colonial infrastructures. They were also forced to share the benefits of producing *Yerba*

---

[119] This is translated literally as a 'commission' or 'task', but in the spoken Spanish of Latin America it means 'the assignation of a group of Indios to a colonial master.'

mate[120] with their colonial masters. Thirty years later, the system was partly replaced by the more modern system known as *reducciones*[121]. First introduced by the Franciscans in 1585, it was improved by the Jesuits in the following century. Surprisingly advanced for the time, the system proved to be the most efficient way of exploiting native labour in exchange for protection from marauding Brazilian slave traders and from the exhausting obligations imposed by the conquistadores. The drop in population after the early conquests caused by wars, exposure to new illnesses, arbitrary killing and inhuman working conditions was thus stemmed. The enclosure of Indians in *reducciones*, or stable villages, and the exploitation of their labour, created what could be called local development. By the year 1707, a century after the royal decree bestowing the Company of Jesus with the opportunity for peaceful colonisation, the system counted 30 villages, 8 in Paraguay, 15 in Argentina and the rest in Brazil. These *reducciones* constituted the Christian province of Misiones, under the name "misiones del Paranà, Uruguay e Tapè", with a government headed by a Father Superior resident in the village of Candeleria (today Posadas). The missions' most important product was Yerba mate, at first gathered in the wild then cultivated, but they also grew wheat and reared cattle.

As far as mate was concerned, the Jesuits introduced an important innovation. Gathering the plant had always been a complicated, uncertain and dangerous activity, because in the wild it grew in the most impervious regions. Expeditions to search for the shrub and harvest its leaves and twigs were exhausting trials of resistance lasting months and decimating its native Indian participants. When the Jesuits took over the Franciscans' activities, they were sufficiently powerful and self-sufficient to protect the Indians within their *reducciones* from the obligations imposed by the *encomiendas*. This meant that the Jesuits had enough labourers to begin cultivating the shrub by planting seeds near the villages. A better quality *Yerba mate* could thus be harvested with less effort and a lower mortality rate than before, and the leaves could be sold on the market at a higher price. Since the *reducciones* were all linked, a harmonious system of complementary activities was created, allowing for self-sufficiency and export of excess production. In this case, up to 30-40% of the

---

[120] According to Garavaglia (1983), the natives worked for their Spanish masters six days a week, with only Sunday and holidays to work in their own fields where they grew corn, manioc, and beans for their own subsistence.
[121] Reduccion is translated literally as "a village of Indians converted to Christianity."

production of *Yerba mate* was sold, making the Jesuits strong competitors against the *encomienderos*, and explaining both the success of the Jesuit missions, compared to the Franciscans who were more subservient to the system of *encomienderos*, and the constant pressure exerted by the authorities on the Jesuits to hand over the natives living in their protected villages. It was precisely this battle for control over the workforce that led to Charles III's expulsion of the Jesuit order on October 27, 1767, which was carried out without resistance the year after. The Jesuit fathers were replaced by Franciscans or by private administrators representing the interests of the *encomienderos* and traders who were more interested in turning a quick profit than supporting the village organisations. While the Guaranì Indians conferred almost magic powers and prestige on the Jesuits, and were extremely grateful for their protection, the new managers were only able to exploit the workforce by means of the whip. Without the Jesuits, the exaggeratedly exploited Indios began to abandon the villages, and the production system built and maintained by the Company of Jesus fell into ruin.

From then onwards, a long process that led to a complete re-mixing of the ethnic, cultural and social makeup of the population and to the fragmentation of the agricultural structure began. From the beginning of the nineteenth century, with the formation of Nation States, former colonies radically changed their systems of relations and ties. Regions lost their autonomy and were required to answer to new, centralised powers. A national bourgeoisie was not yet influential in economic terms, but eager to make the most of the newly liberated territories, above all to control them in order to ward off the territorial claims of neighbouring powers. The area of Misiones was appetising because it was strategically important, for its wealth of natural resources, and for its abundant Creole workforce.[122] The battles that took place to gain control in the area were between members of the dominant elites in the cities of Asunciòn, Buenos Aires, Corrientes and in Brazil.

---

[122] Between the eighteenth and nineteenth centuries, a long process revising social and production relations brought about a simultaneous process of racial mixing, with marriages between Spaniards and Indios. At the same time there was a process of destruction and re-construction of society: Indians who fled from the encomiendas or from the villages allied themselves with a powerful Spaniard who protected them in exchange first for almost slave labour, then for salaried work. Changes that started with the expulsion of the Jesuits culminated in the arrival of a new social figure: a Creole of Guaranì culture, who made up the basic labour force of the new production system. See Garavaglia, 1973. In addition there was immigration from Paraguay and Brazil, escaping from wars in their respective countries, and later by Argentinians.

The political and military situation was complex: the region was to start with relatively autonomous, but was then partly or wholly occupied by Brazil, Argentina then Paraguay. In 1882, the province of Misiones was split from that of Corrientes and was administered from the central government based in Buenos Aires. From an economic point of view, there was no long-term attempt to re-launch the province after the colonial administration, not even from the point of view of exploiting the workforce. The territory was highly contended but economically backward, which made it easy for the ruling class to appropriate its resources (*wood* and *Yerba mate*) and consider it an "extractive front" (Unam, 1993). The system that characterised the region from the middle of the nineteenth century onwards was different from previous systems for exploiting the region's resources, and the centre of economic power was Buenos Aires not Asunciòn. The region was no longer considered a breadbasket to feed the mining industry, but a source in its own right of wealth. Invested capital provided quick returns. The production processes adopted a hunter-gatherer approach, relying above all on an abundant workforce and basic technology. With the collapse of the Jesuit *reducciones*, *mate* was no longer cultivated near the villages and had to be collected in the wild. Through a system of permits and regulations, the government guaranteed selected entrepreneurs legal harvesting rights for the shrub. Villages were not allowed to be established, and cattle, sheep and goats could not be raised, so that the free exploitation of the workforce could be maintained (in the regions of Misiones, northern Paranà and northern Uruguay). A tax calculated on the quantities of *yerba* harvested was then levied on every impresario.

There were many disadvantages for the region. Conflicts over concessions continued until the Buenos Aires government established its project of land distribution. Natural resources were impoverished, and the region underwent a gradual process of desertification. The only resource left was land, for which property rights had never been established; the region thus became a natural destination for repopulation in the early decades of the twentieth century, in particular for settlers from Europe. The government of Buenos Aires approved of this redistribution of immigrants, as the damp Pampas areas were filling up fast, despite local government disapproval. The absentee landlord system of latifundia[123] was partly replaced by

---

[123] All the land in Misiones was sold to a restricted number of buyers: 38 landowners divided up a total of 750 square leagues (in Argentina a league is about 5 kms) and created vast unproductive latifundia. The price they paid was minimal, and this superficial division of land was the first step towards the

successive waves of settlers establishing small farmsteads that more or less guaranteed them a living[124]. Land rights, however, were never well defined.

In 1937, 60% of cultivated lands were in the hands of families with no official title to the land. Land distribution created many different kinds of production units: there were foreign settlers, whose families farmed small or medium-sized lots; Creoles from neighbouring regions who settled on tiny strips of land, just survived off the land and worked as labourers during the harvest; a few land-owners and impresarios who mostly raised cattle in the southern region; and, finally, absentee landlords, mostly from Corrientes and Buenos Aires, who owned the huge latifundia in the north and either cultivated it extensively , or left the land largely uncultivated. In the 1940s, the re-population process was terminated, and all the productive land was assigned. As a result of this variegated process, combining official colonisation, private settlers, and unofficial squatters, the demographic and production structure of the province today began to take shape. Its main features are: the high percentage of rural population; extreme cultural and ethnic heterogeneity[125]; a predominance of small family homesteads over other forms of production organisation; and the presence of large latifundia especially in the area of northern Paranà.

Poverty

The province of Misiones extends over 29,801 square kilometres, with a population in 2001 of almost a million. The population growth rate, owing in part to the colonisation policies enacted, is one of the highest in Argentina: 2.1% average growth per year from 1961 to 2001. Only forty years before, the population was little over 400,000. While the population in the province represents 2.6% of the national population, only 1.6% of Argentina's GDP is devoted to it. The per capita income of he province is just over half the national average. In 1993, the Senate commission for ecology and development calculated the human development index (HDI) of the province of Misiones – which can range from zero to one, with one being the lowest

---

creation of the enormous private agricultural and animal husbandry businesses still present in the Misiones area. The sale was the last act by the government of Corrientes before the nationalization of Misiones, and took place so precipitately that often the land was not even measured, and land rights were granted with a great number of irregularities.

[124] After nationalization, the irregularities committed in the bills of sale allowed the new government to regain much of the land in the middle of the province, which set off a process of colonization of public land starting with the area known as 'del campo' and then extending to the whole of the central region.

[125] In the Oberà region alone there are more than 40 different religions.

possible level of development, and which takes into account key factors such as life expectancy, literacy rates, school attendance, and per capita income according to purchasing parity - as 0.842, one of the worst levels in all the provinces of Argentina. In 1991, 80% of families were unable to satisfy their basic needs (Ipec, 1999) and suffered variously from overcrowding, precarious housing conditions, lack of basic services, truancy and food shortage. In the years that followed, life conditions worsened further. In 1996, nearly 30% of the population in the province lived under the absolute poverty line, double the percentage of ten years before (Lopez, Romeo, 1999). Research undertaken at the University of Misiones, which focussed on income distribution, showed that inequality has increased. While in 1985 in the city of Posadas 40% of the poorest inhabitants had access to 15.5% of the total income, in 1996 they only had 14.6%. At the same time, 20% of the wealthiest citizens enjoyed 50.9% of the total income, three more percentage points than in 1985. The labour market has also deteriorated. In 1980 there was still a protectionist regime, in 1991 liberalisation started to take effect, and in 1996, the system underwent total deregulation, privatisation and opening to outside markets. This process increased both unemployment and underemployment, reaching the double figures.

The main economic activities in the region are agro-industries. This creates four different sets of consequences.

I. A great number of small producers supplying raw materials have to contend with a concentrated demand, which means they have little bargaining power, made even weaker by the absence of both union support and solid cooperatives;

II. Perennial or semi-perennial cultivation needs long-term investments which take longer to yield results; in the case of *Yerba mate*, it takes four years before the shrub can be harvested;

III. Alternative products to *Yerba mate* are mostly destined to foreign markets (almost all the Tung oil, 80% of the tea, more than half of the tobacco, and about 30% of the cellulose paste) creating strong pressure on prices and instability;

IV. Most of the products are cultivated or transformed only in the province of Misiones, or at most in the region, meaning that their relevance in the national economy is low, as is the attention of central government.

The first two characteristics – a buyers' market and long-term investments that penalise small producers and farm labourers in particular – need specific counterweights such as employment alternatives or support policies. The dominance of agro-industries depends on the fact that the manufacturing sector has shrunk (in terms of value added from 27% of total production to 21%), as well as becoming more concentrated, with a reduction in the number of production units, and a 37% drop in employment share. Tourism could provide a valid alternative, but it is as yet unable to employ enough people. The vast majority of the local population depends on producing Yerba mate, tea, tobacco, wood and its by-products. In this context, Yerba mate represents the most important opportunity, because the other products are negatively affected by points III and IV (dominance of foreign markets and irrelevance in the national economy).

Yerba mate

The area where the *Yerba mate* plant grows is limited to the East by the Atlantic Ocean, the West by the River Paraguay, and to the north and south by the parallels 18° and 30° in the Southern latitude. The sub-tropical shrub requires the high temperatures, damp soil and heavy rainfall, sand and clay soils rich in potassium, iron and phosphorous acid offered by the red soil of Misiones. Gortari (1997) calculated that the total world production of Yerba mate is approximately 500,000 metric tonnes per year, 260,000 of which in Argentina (57%), 180,000 in Brazil (38%) and the remaining 30,000 in Paraguay (5%). In Argentina 200,000 tonnes of *Yerba Mate* are consumed per year, and the rest of the production is exported; 70% goes to countries belonging to Mercosur, and the rest to Chile and Bolivia. Outside Latin America, *Yerba mate* is becoming popular in Syria, Lebanon, Israel, Saudi Arabia, Germany and the US. Ninety percent of Argentina's production is from the province of Misiones, while the rest is harvested in the north-eastern province of Corrientes. In Misiones, 24,000 producers cultivate the shrub, of which 16,000 are small producers, called *colonos*, with plots no bigger than ten hectares (1 ha = 10,000 msq). At harvest time (*cosecha*), the activity employs about 4,000 extra labourers, or about 22% of the agricultural workforce. The industrial transformation of the leaves and twigs employs about 15% of the workforce in the manufacturing sector. The production structure comprises 24,000 producers, 300 *secaderos*, 200 *acopiadores*, and 124 *molinos*.

The land

Data on the quantity of land currently under cultivation for *Yerba mate* is hard to come by owing to large discrepancies between satellite pictures, declared quantities and the findings of surveys. The main culprit is no doubt the black market.[126] The most reliable figures derive from satellite imaging, most recently undertaken between March and April 2001. Fifty percent of the Agricultural Surface Used in the province of Misiones is cultivated with *mate*. The images are divided according to density of cultivation into three categories: high (more than 1,800 plants per hectare), medium (between 1,000 and 1,800) and low (less than 1,000). The intensity of cultivation also reveals information about the technologies adopted, the relevance of the cultivation for family economies, and the workers' standard of living. In order to absorb the working capacity of two adult members of a family, taking into account alternative production that is also needed, 10-15 hectares of land are necessary. Surveys into the number and size of smallholdings of this kind, known as *chacras*[127], provide the following data: 21,000 are less than ten hectares (with 116,000 ha cultivated all together), 2,400 are between 10 and 25 hectares (38,980 ha total), 597 are 25-50 hectares (21,163 ha total), and 179 are bigger than 50 hectares (20,542 ha total).

The graph in Figure 21 shows that the distribution of land in the province of Misiones is less unbalanced than average in the country. This is important because it shows that, while the very survival of small producers is at risk, this is not only owing to the enormous imbalance between the limited amounts of land available for small farmers and the vast latifundia in the rest of the region. In the production of mate, as we have seen, less than 10 hectares of land cannot be considered sufficient to feed a family. At the same time, the amount of land available for smallholders is not irrelevant: more land means potential cooperative behaviour, a chance of increasing productivity and thus sustaining the population, product diversification and economies of scale.

---

[126] Satellite images produced the figure of 24,000 producers, while official surveys only counted 17,000.

[127] A *chacra* is a unit of mixed farmland, typical of the province of Misiones. One *chacra* of 25ha usually mixes perennials (mate, tea, tung), annuals or biennials (corn, soy, tobacco, manioc, sugar cane or pineapple) and cattle or pig raising.

Fig. 21 – Distribution of land cultivated with mate in the province of Misiones and Corrientes, compared to the distribution of cultivated land on average in Argentina – "Lorenz curve"
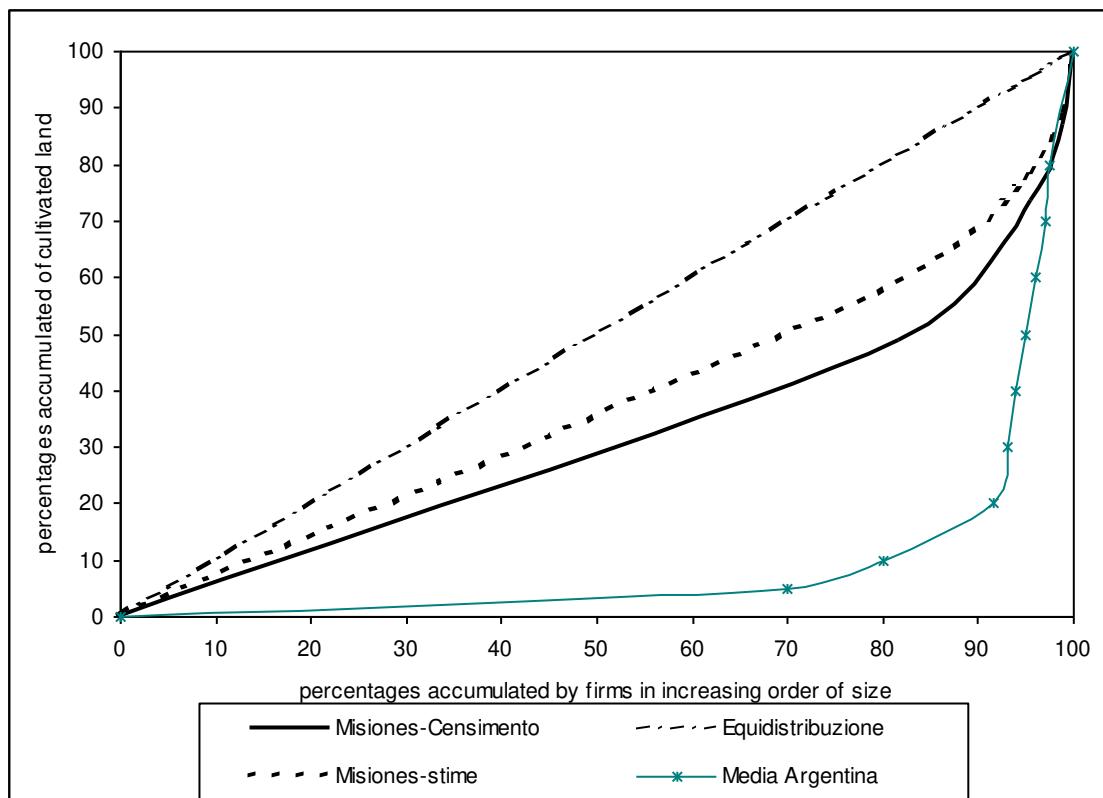
Sources: My graph using data from "El productivo" in Misiones e Corrientes, and www.gtz.de/orboden/tenure/te3 for Argentina, with reference to the year 2000

The production process

The food code established by Argentina's Ministry of Agriculture (MAG, 1971) defines *Yerba mate* as a product made out of the dried and lightly toasted leaves and twigs of the *Ilex paraguariensis*[128], which are then chopped or ground. In order to be fit for consumption, the industry regulations require it to: contain less than 11% humidity, be well conserved and free of harmful or alien substances; contain less than 9% ash, 1.5% of which should be insoluble in chloridric acid; have a minimum of 85% of broken, dried or ground leaves with no more than 15% twigs in the blend. The food code also considers it necessary for the product to contain more than 0.6% of caffeine, and for the watery extract to be more than 25%. Over thirty years, when the tree reaches full maturity, it can grow up to 15 metres tall, but cultivation techniques

---

[128] This is a species of holly native to Latin America. *Ilex aquifolium* was used a great deal in the seventeenth century for medicinal purposes. The earliest records go back to the sixteenth century, which mention its use in an infusion.

introduced over the years keep the shrubs down to the more practical height of 3-6 metres.

The first operation in the production process is the *cosecha*, or harvesting of leaves and twigs. This is generally carried out with scissors, machetes or little saws, except in the large, integrated firms with wider areas cultivated and greater capital investment, which have undertaken research and bought machinery. The first picking, with a low harvest, takes place four or five years after planting; from then on the shrubs can be picked annually, and in some cases twice a year. Under favourable conditions, a plant can yield up to 20-25 kilos of green leaves. The first cleaning takes place on the spot, leaves are separated from unsuitable twigs and the product is then wrapped in jute cloths for transport. Until several decades ago, it was the *tarefero*[129] (seasonal or daily farmhand) who transported the leaves and twigs to the place where phase two – the sale - begins. Nowadays the *yerba mate* is loaded onto trucks, which generally belong to the buyer. The protagonists of phase one are farmhands on a medium-sized *chacra*, the families of smallholders on small farms, or employees in large, integrated firms; once the harvest is over, about 73% ends up on the market. Only about 27% belongs to the few operational cooperatives, which manage both the harvest and the successive stages of transformation.

The first stage of commercialisation is highly fragmented: 24,000 producers enter into contact with 300 *secaderos* (who buy about 53% of total production). Added to this, there are 200 intermediaries: the *acopiadores*[130], who buy about 13% of the product, and the *molinos*, who buy the remaining 7%.[131] It is at this point that small producers are at a disadvantage because they cannot impose their own conditions, being at the base of a pryamid where the summit has much higher bargaining power. In addition to fragmentation and the number of producers, there is also the problem of the precarious financial condition of an average family of *colonos*. With the pressure created by the bottleneck that keeps demand and supply apart, it is easy to see why the basic price of the raw material is insufficient to inject capital into the company. The

---

[129] There are several versions of the meaning of this word, which means a farmhand employed by the hour or day by the plantation owner during the harvest. One origin of the word could be the Spanish for 'tariff', meaning the agreement between the labourer and his employer. Another etymology is from the Portuguese, meaning 'task' or 'job'.

[130] The word *acopiar* means literally 'to pile up'; his task is to buy and amass leaves and twigs from different producers in order to sell them to the *secadero*.

[131] Ieral, 1998. This study regards a very limited area but is representative of the whole production sector in the province of Misiones. The figures are based on a survey.

higher levels of the pyramid, the *secadores* and *acopiadores*, similarly have little leverage; they too have to negotiate with the far more powerful distribution companies. This pyramid structure is one of the causes of the poor standard of living of small producers.

Fig 22 – The pyramid structure of the production process of yerba mate
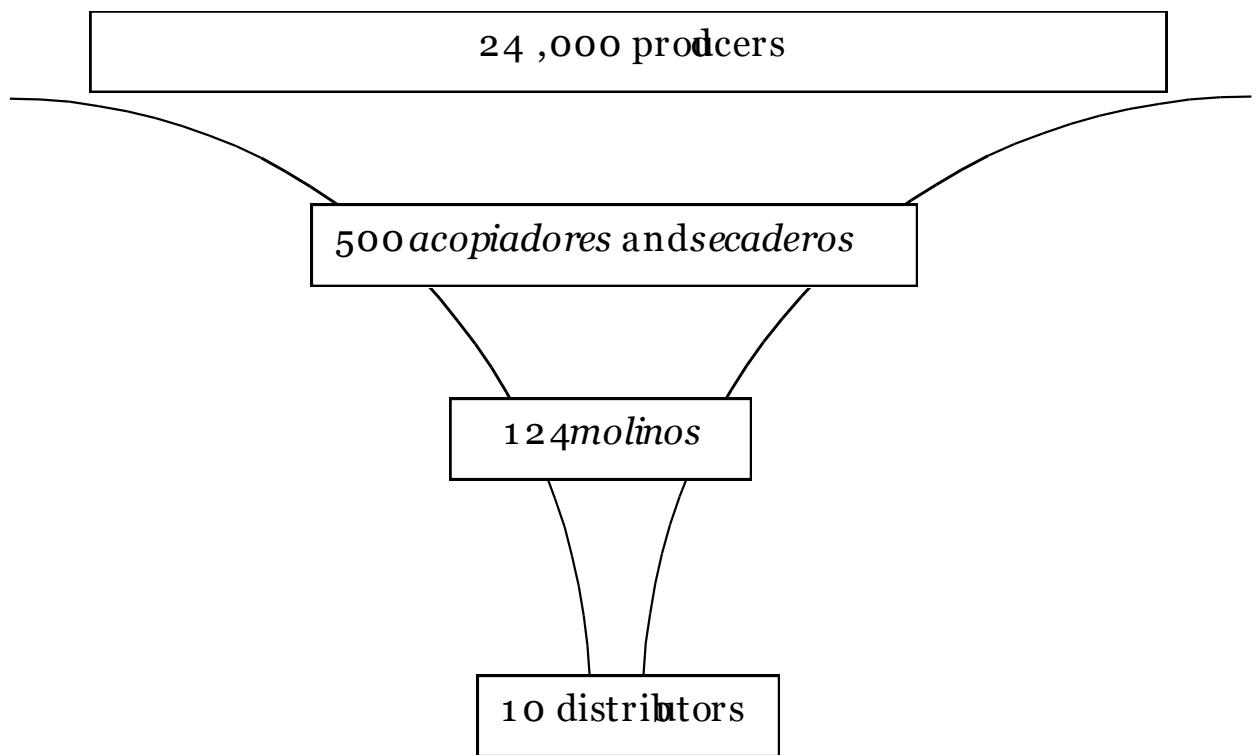
The next stage of production is the drying process (*sapecado*[132], *secado* or *secanza*, and *canchado*). This process must be undertaken within 24 hours of the harvest in order to avoid fermentation and the consequent waste of the substance. This technical requirement is relevant because it explains why the *secaderos* play an important intermediary role between producers and *molinos*, thus creating two sets of exchanges: between producers and *secadores*, and between *secadores* and *molinos*. The lack of adequate infrastructures, transport, and preservation techniques means that the *secadores* must be physically located near the harvesting areas. Straight after the harvest, the product is subject to fermentation and weighs a great deal, whereas

---

[132] The word is Guaranì for 'exposing leaves to a source of heat'. Generally this is a live flame, where the leaves are exposed for a few seconds. This stabilises the leaves chemically.

after the drying process it is stabilised – and can therefore be stored - and weighs about one third. After the live flame, which destroys the ferments, stops oxygenation of the tannin, and ensures that the leaves stay green, the product is further dried and lightly (*secado* or *secanza*) in order to reduce humidity to 5 or 6%, and finally ground down (*canchado*) in order to facilitate transportation. The *yerba* is then left to season, originally for one year, but nowadays for 9 months, or even less with the more moder machinery. At this stage the product is called *yerba mate canchada* and provides the raw material for the *molinos* to transform.

At the Molino, the product is selected according to its provenance and its features, and cleaned, eliminating some of the excessive twigs. It is then further ground, at varying levels, sieved and then sorted into two types of mate: *con palos* and *sin palos*, according to the percentage of twigs in the blend. At this point it is called *yerba molida*, and it is ready to go to the 124 *molinos*, who are high up on the pyramid and take advantage of the bottleneck. Only 10% of production goes to cooperatives, 30% to independent firms and the remaining 60% to twenty large *molinos*. Moreover, their location reflects the concentration of the production regions: there are 95 in the province of Misiones, 11 in the neighbouring Corrientes, 6 in the capital, Buenos Aires and another 10 in other provinces (Freaza, 2002). This is true despite the fact that the *molineros* need technologies subject to scale economies and that the seasoning period separates the drying stage from the *molinero* stage. The fact is that there is no need for the factories to be physically located near production. Historically, between the two alternatives – integrating commercialisation with necessary investments in their distribution networks, or integrating, or interacting with the *secaderos* – the *molineros* have always chosen the latter. It is easier for them to exploit the highly scattered and small-scale producers. The consequence, especially today, is a further shrinking of the summit of the pyramid during distribution.

Large distribution

While traditionally the *molineros* –who sold the product under their own label – were at the top of the pyramid, since the end of the 1980s owing to the liberalisation process underway in the country, their privileged position has been undermined, and supermarkets and large distribution chains have become dominant. The increase in competition without an agro-industrial national policy has had several consequences. There has been a concentration of the commercialisation of the product: 80% ends up

in the hands of only six firms, and 20% in four others, who compete by means of massive advertising campaigns. Many traditional labels linked to the *molinos* have disappeared, and the product has been standardised. The production process is thus pyramidal, with an extremely small summit, and many different qualities and types of product can no longer be found.

Demand

The demand for *yerba mate* is inelastic: increased consumption depends neither on a reduction in prices nor on an increase in income in the population. And yet some properties of the plant are as yet unexplored: its energetic, nutritive and preventive qualities, as well as its potential as a preservative, colorant, deodorant, antioxidant and bactericide. In short, while currently demand is static, there is nothing to say that its potential could not be tapped, with the support of appropriate policies, especially in research and development.

Supply: the end of the line

On the supply side, the most important phenomenon has been the increased dominance of large distribution. Its disproportionate economic force, however, does not invest in research to improve the quality of the product: its main aim is to promote *yerba mate* commercially. Profit margins are mostly sought by squeezing the production process, thus leading to even greater concentration of the *molinos*, only the biggest and best equipped logistically are able to compete and stay on the market. The result for the weakest links of the production process is a deterioration of their bargaining power and therefore of their standard of living.

Supply: the beginning of the line

By contrast, producers of *yerba mate* have always had problems in satisfying demand. Because it is a perennial, and can only be harvested after a cycle of four years, when higher prices would suggest increasing supplies, it is hard to increase production accordingly. By the same token, it is hard to contract the supply when prices drop. To leave the market is very costly, both in terms of loss of investment and in terms of expenses incurred. Producers often leave the plantation and resign themselves to low prices. The result is a structural regime of excess production and lack of demand, diminishing producers' bargaining power even further. This situation

justified State control, when excess production was stockpiled according to a government policy – that was ineffective but doubtless better than no policy at all - called *mercado consignatario*. In short, in the last decades, producers have seen their profit margins diminish owing to the arrival of large distribution and the abandonment of government policies that tried to control and stabilise supply.

*Fig. 23 – The traditional structure of the sector (until the mid eighties)*

133
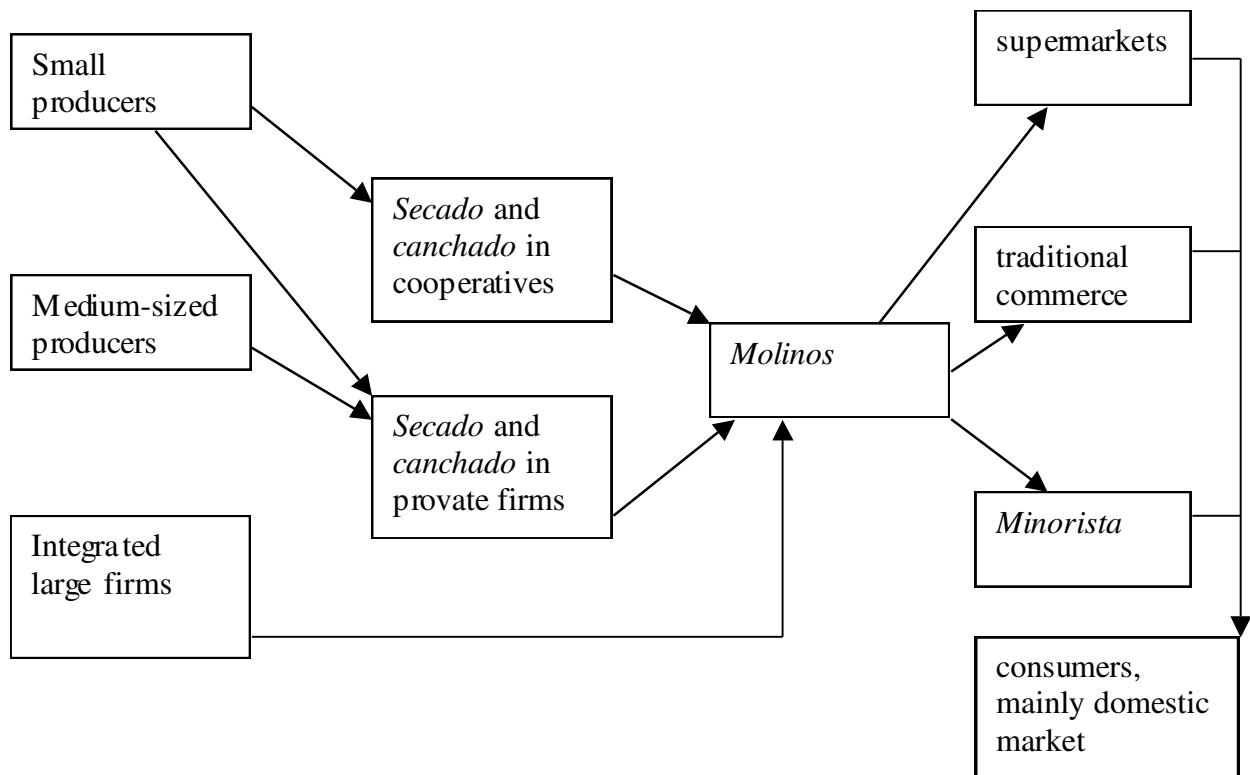
Fig. 24 – The structure of the sector, current trends
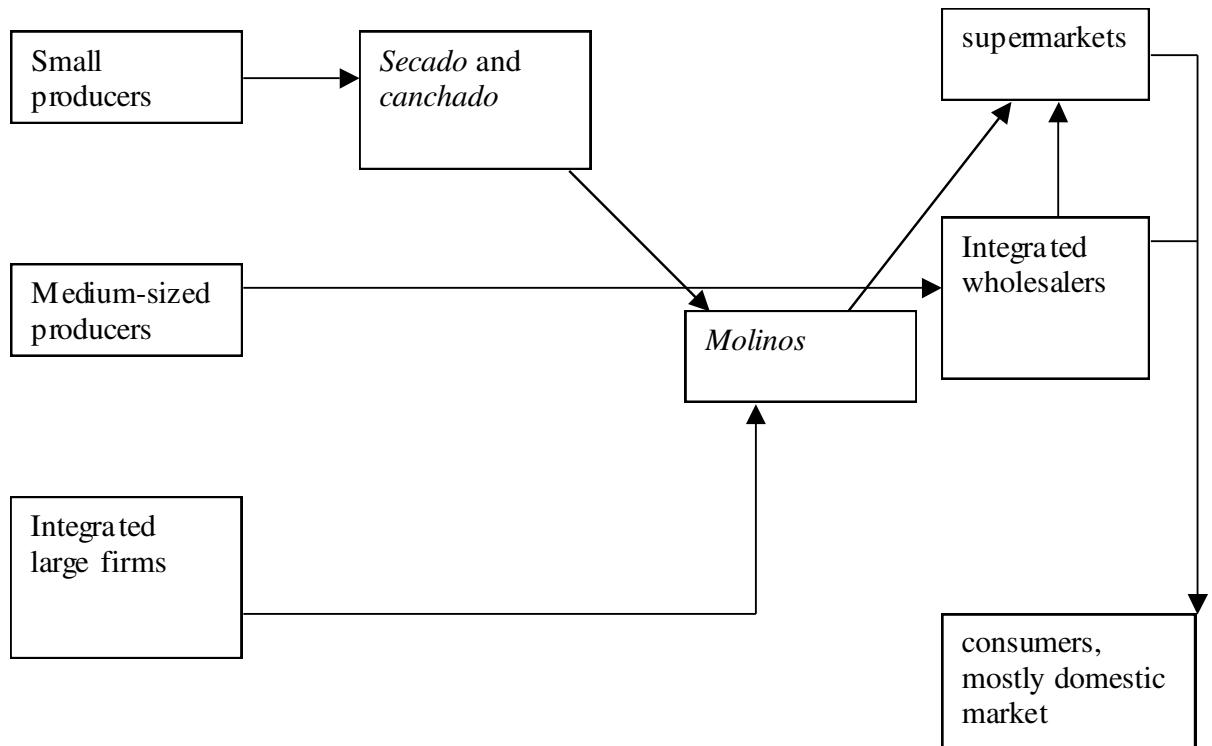
A vicious circle

This situation has aggravated another significant phenomenon: illegal labour. As their profit margins shrink, many producers, *secaderos* and *molinos* have begun to evade taxes, as well as administrative and contractual obligations, contributing to the black market. This makes it impossible to draw up a clear map of who works and sells what and where, which at the same time makes it difficult to draft a law that supports certification of the product, or of the production process, in order to improve quality and diversification. At the same time the black market reduces income for small producers, as informal agreements, especially between producers on one side and *secadores* and *acopiadores* on the other, replace legislation. Agreements whereby *yerba* is exchanged under a barter system are not rare, and price variations, as well as credit extensions, are equally common. Producers have no defence against this lack of transparency: there are no detailed, reliable facts available about prices, the extent of cultivations, production or stocks. Decisions regarding investments, harvests, or sales, are made on the basis of limited information, especially since liberalisation (Ieral,

134

1998). Without adequate interventions, it is easy to imagine the sector evolving towards a progressive domination of large distribution over the entire production process, with subsequent loss of employment, small producers being reduced to becoming farmhands, and small firms that could resist with the support of cooperatives, dying out.

State interventions

Public interventions have followed several paths. At one time, the government tried limiting supply by fixing quotas and restricting harvests in order to sustain prices. The result was often an increase in the final sales price rather than in the price paid to small producers for raw materials. Moreover, price-sustaining policies led to higher administrative costs, which, in the long run were unsustainable given the budget constraints of such an economically instable country. Excess production was then absorbed by government funds (Mercado Consignatario). The next step was to subsidise the product, but on the one hand this was difficult to sustain financially, and, on the other, it encouraged farmers to increase production and thus aggravate the situation of over-supply. Another attempt was to support producer cooperatives, but with little continuity and determination.[133]

Following Argentina's general economic approach - and unlike experiences in some countries that succeeded in emerging from underdevelopment by abandoning protectionism for open markets - institutional interventions in support of *mate* production were structurally instable and in continual flux between liberalism and protectionism through state intervention. Setting up state production and commercialisation companies was a solution that worked well from the mid-1970s to the late 1980s because they worked in a protected market in a regime of import duties. When the markets were opened, however, large, private firms started to compete, the state companies were forced to privatise. In recent years, round tables were set up to negotiate agreement among the various interest groups. In order for this bargaining tool to be effective, state functionaries as a rule need to be completely impartial referees, and their prime aim must be to distribute gains fairly. The obvious problem

---

[133] The reform that took place at the beginning of the 1990s eliminated the preference that had been accorded to cooperatives by imposing a capital tax of 2% on cooperative funds and creating more intransigent regulations than for private firms, which can always resort to double accounting. The few cooperatives that there were at that point started to fail. Today there are only three significant cooperatives producing yerba mate: the Santa Pipà, Montecarlo and Colonia Liebig.

is that the state did not absolve this function with absolute transparency, thus generating a lack of trust and a failure of the negotiations: the state representative at the round table was in fact a Senator, owner of one of the biggest *molinos*. A recent attempt at re-negotiating agreement through a law that created the Istituto Nazional del la Yerba Mate (INYM) needs further examination.

The 2001 law aimed to "promote and reinforce development of the production, transformation, industrialisation, commercialisation and consumption of yerba mate and its derivatives for different uses, stimulating the sustainability of the various sectors involved in the activity" (Scdra, 2002), and therefore improve and stabilise producer prices. The law envisages a directive organ comprising 11 members: one representative appointed by the central government resident in the province of Misiones for at least five years; one representative of local government from Misiones and one from Corrientes; one representative from industry, three from producers, two from agricultural cooperatives, one from the category of farmhands, and one from that of the *secaderos*. This composition aims to represent all the parties involved in the production process, but clearly the central and local government representatives are simply two more votes for the *molineros* and industry. The substance of the law is no different from previous government initiatives: it is an attempt to propose a method to negotiate price stabilisation.

For decades, structural changes have been mentioned but never implemented because of the failure of collective action[134]. The closest the government has got to taking concrete action in support of producers is to fix guaranteed minimum prices, estimate quantities of the plant that are produced and, if necessary, restrict harvests. Other actions that are potentially envisaged by the law include certifying quality and supporting cooperatives. Since 2001, conflicts between producers and industrialists

---

[134] The functions of the Institute were very similar to those of the previous Crym: control the implementation of regulations; support producers, transformers, industries and shopkeepers with incentives; identify and draft strategies to support competition in the sector; plan, organize and participate in promotional activities for *yerba mate* and its derivatives both at home and abroad; conduct research and innovation in order to diversify its uses and increase consumption; coordinate improvement plans with competent food organisms; create a data bank for disseminating information about hygiene laws and quality requisites; disseminate statistics and surveys in order to sustain demand and supply, and, if necessary, collaborate with the Ministry of Agriculture, Food and Fishing to establish norms to restrict production; create registers to identify all participants in the production process; provide technical assistance; facilitate training and updating of technical and professional personnel with conferences and scholarships in Italian and foreign universities; promote forms of networking between producers and in particular cooperatives in the production area; include those who work in the sector among the beneficiaries of the national unemployment fund; attempt to find a common agreement on the price of the raw material.

have made its implementation very difficult. Strikes, debates and demonstrations have characterised the political climate since then.

The debate

Small producers are, as we have seen, at the wide base of the production pyramid. Their standard of living has never been high, but if during their lifetime they had earned enough to set aside a few savings, the whole economy of the region would probably be wealthier and more dynamic. The return of democracy in 1983 also marked the return of trade unions, as well as of the *Movimento Agrario de Misiones* (Mam[135]), though the political climate was still fraught with tension and fear. After the precipitate price drops owing to over-production, in the early 1990s workers' and producers' associations again became active. Despite the climate of fear, there was optimism in the air, as union leaders realised that the sheer fact of meeting again meant a new opportunity to rebuild the country, starting with their own history, and their individual and collective identity. This enthusiasm, however, blinded some activists to the fact that the context had changed and that perhaps a new form of organisation was called for (Goldsberg, 1999). Apart from the absence of the smallest producers, too frightened and mistrustful after long years of repression, one of the reasons collective action did not take off immediately was that the provincial government supported the creation of sectorial associations for different products, with the final aim of atomising the producers' demands. In this period, by way of example, this policy led to the creation of the Aptm, an association of small tobacco producers, and the Atém, an association of tea produces from Misiones.

---

[135] Born in 1971, this trade union of Obertà small producers shared the usual aim of stemming a new price slump. It was very influenced by the Movimento Cristiano de la Acciòn Catolica, a Catholic organization that provided leaders for the Mam and imposed its ideology. The structure was made up of 20 producers who elected one representative every 100 farms who participated in the general assembly. The movement was funded by membership fees and by contribution from the Catholic church. Farms with over 25 hectares, considered medium-sized farms , formed the majority of its members, meaning that they were not on the lowest rung of the social scale but were anyway unable to accumulate capital. During the 1973 elections, Mam decided not to support any particular candidate, and this decision sparked disagreement within the movement. Some members accused the union leaders of being Peronist and subsequently split off, forming the union *Agricultores Misioneros Agremiadas* (Ama), which did not aim to subvert social order in any way. Another splinter group was the *Ligas Agrarias Misioneras* (Lam). No agreement was reached between the groups and in the 1975 elections Mam supported the Frejuli party, Lam the Partido Autentico of the Montoneros, and Ama the U.Ce.De. Colonos who protested were violently repressed after the 1975 crisis, as the interests of the agro-industrial groups were threatened. The government declared Mam and Lam subversive. In 1976, with the military coup the army occupied the province, and many union activists died or disappeared mysteriously, effectively suffocating union activity.

The important demonstrations at the end of the eighties, such as the tea strike, gradually led both sides to consider a less conflictual approach and negotiate solutions between producer associations and government institutions. Starting in August 1991, Mam proposed that the State should once again play a role in exploiting natural resources, financing commercial activities, and experimenting and developing new products, within the existing production structure. In 1999, when prices once again precipitated, there were renewed protests. The associations active at the time, in addition to Mam were: the *Union de Trabajadores Rurales y Estibadores* (Uatre); the *Federacion de Cooperativas Agricolas de Misiones* (Fedecoop); the *Asociacion de Productores Agrarios de Misiones* (Apam); and, for the industrialists, the *Camara de Molineros deYerba Mate de la zona Productora* (Cmyzp). All these organisations participated in the discussions leading to the new law concerning *Yerba mate*.

The new law that instituted the national institute governing *Yerba mate* production (Inym) was considered indispensable after the liberalist wave had hit producers and left havoc in its wake not only for the weakest links in the production line, but also for medium-sized producers and – perhaps decisively - even for the traditionally strongest subjects, the *molinos*, who were increasingly threatened by supermarkets and large distribution chains. The various associations have differing views regarding how well the new (old) system of regulations works.

Apam, which represents medium-sized farms and which is represented in the bargaining organ of the national institute, supported the logic behind the law: the only thing that really counts is fixing a guaranteed minimum price for raw material. Despite obvious differences between the various approaches, this point of view was - either tacitly or openly - a lowest common denominator for all parties involved. The mobilisation, including a six-week peaceful occupation of the main square in the city by the families of small producers and tareferos, to implement the law - at least in these minimum terms – finally brought about an agreement for a guaranteed minimum price of 13.5 cents per kilo for the raw product, and 50 cents per kilo for ground *yerba canchada*

Table 2 – Production costs for Yerba mate per hectare in pesos in the province of Misiones

Factors	Medium Technology	Low Technology
Manual work	6	6
Tractor	16	16
Ploughing	7	7
Raking	22	22
Grinding	12	6
Pesticides	18	-
Herbicides	24	24
Fertilisers	90	-
Covering	21	-
Total	216	81

Table 3 – Producer spending per hectare in 2001 in the province of Misones

Factors	Spending per hectare (excluding harvest*) In pesos	Spending per Kilo (excluding harvest*) In cents	Spending per Kilo (including harvest) In cents
Total direct costs	*220*	*5,5*	*8,0*
Ammortisation machinery	114	2,9	2,9
Ammortisation plantations	142	3,6	3,6
Interest on loans	77	1,9	1,9
Total without margins	*553*	*13,6*	*16,4*
Earnings 10%	55	1,4	1,6
Total	**608**	**15,0**	18,0

Note (*): Our evaluation comparing various sources.
Source: MAM (2001).

Fig. 25 – Ups and downs in production of Yerba mate in tonnes
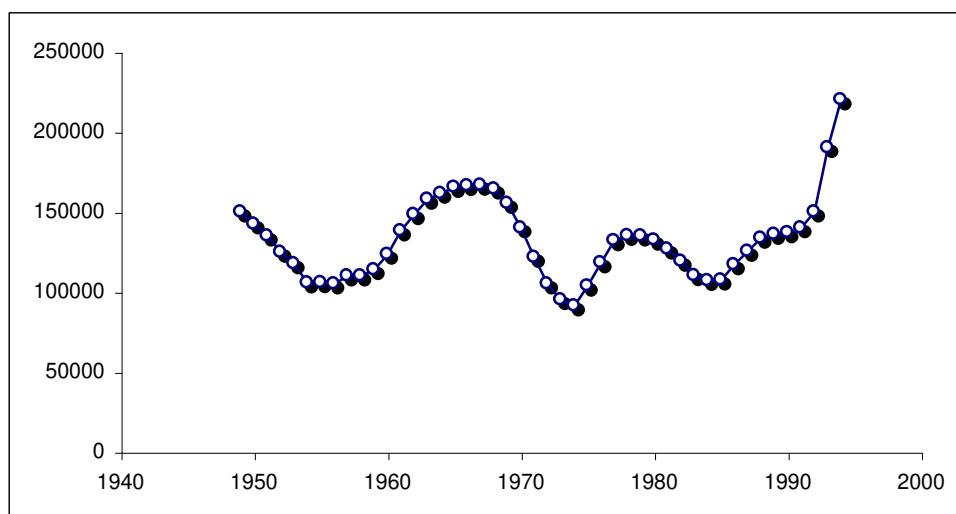
Sources: MAG (1971), Ieral (1998), our interpolations.

Bibliography for chapter 3 appendix 1

Freaza, M.A. (2002), *Economia de Misiones,* Editorial Universitaria de Misiones, Posadas, 2002.

Garavaglia, J.C. (1973), *Un modo de produccion subsidiario: la organizacion economica de las comunidades guaranizadas durante los siglos XVII, XVIII en la formacion regional altoperuana rioplantense,* Modos de Produccion en America Latina, "Pasado y presente", Cuaderno N°40, Cordoba.

Garavaglia, J.C. (1983), *Mercado interno y economia colonial*, Grijalbo editores, Mexico.

Golsberg, C. (1999), *El movimento agrario de Misiones en un escenario en trasformacion*, Trabajo para optar al titulo de Ingeniero Agronomo (tesi di laurea), Facoltà di Agronomia, Universidad de Buenos Aires.

Gortari, J. (1997), *El Mercosur y la economia yerbatera, una aproximacion al impacto en la pequeña produccion regional*, "Realidad Economica", n. 154.

Ieral (1998), Fundacion Mediterranea, *El sector Yerba Mate*, "Documentos de estudio", Posadas.

Inta, (1998), *Misiones*, in "Sociologia rural",1, 0.

IPEC (1999), Insituto Provincial de Estadistica y Censos, *Producto bruto Geografico. Serie 1991-1998*, Gobierno de la Provincia de Misiones, Posadas.

Lopez, A. e M. Romeo (1999), *Desarrollo regional y exclusion social. Transformacion y crisis en la Argentina Contemporanea,* Amorrortu Editores, Buenos Aires.

MAG (1971), Ministerio de Agricoltura y granaderia, Comision Reguladora de la produccion y comercio de la yerba mate, Comite de propaganda del consumo de la yerba mate, *La* yerba *mate*, Republica Argentina.

MAM, (2001), *Costos de produccion de la yerba mate*, Oberà.

Rau V. (2000), *Yerba mate: el "Paro verde"*, "Realidad Economica", n. 185, pp. 122-143.

Sarasola R. (1998), *Las voces de la yerba mate, Breve vocabulario ilustrado,* Editorial universitaria Universidad de Misiones, Posadas.

SCDRA (2002), Senado y Camara de diputados de la Republica Argentina, *Ley n. 25564.*

Schamber P.F. (2000), *Bajar y dar de nuevo. Consecuencias de la desregulacion en el sector yerbatero*, "Realidad Economica", n.169.

Unam (1993), *Notas sobre la historia de Misiones*, Documentos de trabajo del projecto Pobur, Edicion secretaria de investigacion facultad de humanidades y ciencias sociales, Posadas, 1993.

Chapter Four - Institutional Action

In lagging regions, resources have the potential to set a process of development in motion. The process is difficult to trigger, however, and it cannot be left to entirely to private economic agents; public goods and services must be established by means of institutional action. Institutions, on the other hand, especially state institutions, are not particularly efficient. Even when they are not prey to specific interest groups, whose aims are often anti-development, they suffer from any number of inefficiencies: in recruiting managers, in being procedure-oriented rather than results-oriented, in their tendency to privilege apparatuses over the common good. It is often said that decentralisation, which relies on local institutions, resources and skills, is a way to minimise the inefficiencies that are so much more evident at a central government level. A State that has many different levels of government is seen as being more flexible, more innovative; at the same time, local institutions are considered closer to the people and easier to control.

And yet, when this approach is adopted it is easy to lose sight of an important element. Precisely because they are closer to people's concrete interests, local institutions are subject to consensus. The considerations made in this chapter stem from the conviction that, even when institutions behave virtuously and in a non-predatory manner[136], there is always a 'consensus constraint' that long-term future results cannot satisfy. Institutional action, under these uncertain circumstances, does not necessarily weigh the whole package of features that make up the history of any given economic and social context – its resources, social structure, competence of the managerial class, political regime, collective experience and traditions. These historical conditions count, but they are not sufficient to determine what actions will be undertaken; there is always a margin of indeterminacy. Institutions must be seen as being 'entrepreneurial': their behaviour as 'intentional'. Why in some conditions institutions take the right steps in the right direction for development, and why – even

---

[136] 'Virtuous' in the sense that they act for their own good by means of the collective good. Predatory institutions aim to satisfy their own interests to the detriment of the common good. The emergence and consolidation of predatory institutions that are anti-development has been analyzed by Acemoglu (2006). The model he presents is very interesting: he stresses that institutions against development are seen as persistent and hard to dismantle once they have become established, and that they tend to establish themselves when powerful social groups are convinced that their personal advantages would be endangered by changes that would distribute economic benefits more widely. Acemoglu's model therefore lends support to our argument.

in similar circumstances – they either do not act, or the steps they do take are in the wrong direction, or ineffective, must have something to do with the quality of the managerial class (their competence and experience). It also has to do, however, with their ability or desire to take risks, or their inability or reticence to do so, given the fact that future or uncertain results fail to legitimise their decisions.

The first section analyses briefly the issue of political and administrative decentralisation, and finds that the available literature does not come to a definitive conclusion that decentralisation provides a solution to the problem of consensus. The second and third sections present the main focus of the chapter: local institutions are in closer contact with the interests of different groups, and whatever they do, even when it is justified by the common good, their actions generate differentiated advantages and disadvantages and maintain a high level of uncertainty. This means that local governments have to sustain high costs to draft complex development projects as well as high costs related to gaining consensus. Their actions are conditioned by the history of the territory, by the experience and competence of the managerial class, but there is always an element of indeterminacy. The chapter closes with a conclusion and two appendices. The first appendix describes the experiences of local development in Italy and shows that local action did play a role but achieved highly differentiated results even in the provinces that started out in the 1950s with similar per capita income levels. The second appendix provides a critique of a widespread idea that managers and administrators are as impregnated with the cultural attitudes of their territories as their subjects are, and therefore have no real autonomy.

Development policies and decentralisation

The issue of institutional decentralisation is a complex one. Simply adopting the paradigm of an inefficient central State does not necessarily lead one to believe in the virtues of policies that are drafted and decided locally. The limits of local governments need to be taken into account, as well as the difficulties they face.

Let us examine, to start with, some of the arguments in favour of decentralisation that add something to the basic paradigm of an inefficient central State. The first indirect argument is supported for the same reasons that two typical options of a centralised State approach to local development - automatic incentives and large-sale public works – are considered inadequate. As regards incentives, the argument is based on the concept of 'temporal inconsistency'. That is, in order to be effective any

form of incentive must be conditional and temporary; automatic incentives tend to be unconditional and become permanent (Middleton, 2001). Their dubious effectiveness is well documented by experience, both in industrialised countries (Bradshow, Blakely, 1999; Middleton, 2001) and in developing countries (Amsden, 2001; Amsden, Wan-Wen, 2003). There is less literature on the perplexities regarding large-scale public works as a solution for local development (Flyvbierg, Bruzelius, Rothengatter, 2003), but there is ample documentation of the systematically vast gap between estimated costs at the planning stage and final costs[137] and it has been shown that this gap hides all manner of interests, outside the sphere of legitimacy or functionality (Premchand, 2003, p.15).

The second argument in favour of local policies is that they are often able to set in motion positive expectations and behaviour that set development in motion. That is, active and capable local institutions indirectly create and support the kind of attitudes that are more common in private economic agents, as well as in the population at large, such as respecting common and collective goods and being willing to take risks and innovate. It is also important to stress, moreover, that the idea that the dynamic capacity of a given territory is dictated by its history, or its 'civic' heritage, is inaccurate, if not downright ideological. This culturally dominated approach generally cites Putnam (1995) with regard to Italy; but, as the appendix will show, there are many doubts as to the reliability of his results, not to mention his underlying hypothesis.

These two arguments are powerful, but not sufficiently persuasive if we are to discuss the actual realisation of local development policies. Azfar, Kahkonen, Lanyi, Meagher, and Rutherford (1999) demonstrate that available empirical research is inconclusive because it reveals highly variegated outcomes - in terms of efficiency and development - of decentralising reforms in different countries (Jutting, Kauffmann, McDonnell, Osterrieder, Pinaud, Wegner, 2004). The emphasis must be shifted at this point to the conditions under which the reforms were implemented, comparing them to the conditions that at least in theory[138] call for decentralisation.

---

[137] The Suez Canal, for example, cost 1900 times the estimated price. This is an extreme example, but there are many others: the new EU offices (400 times), the Hem motorway tunnel in Holland (444 times), Denver Airport (253 times). See http://home.planet.nl/viss1197/underest.htm

[138] Especially public finance theory, for example from Hayek, 1945; Musgrave, 1959; and Tiebout, 1956 to Oates, 1968, and Ostrom, Schroeder, Wynne, 1993.

The authors mentioned above underline that decentralisation increases efficiency in allocating resources, the responsibility of institutions (reducing corruption), and the capacity for initiative if the following conditions are satisfied. If: I) the constitutional framework is stable and the various levels of government are distinct; II) the political and electoral system guarantees real participation for its electorate as well as effective sanctions; III) there is effective coordination between different levels of government, avoiding the twin risks of extreme vertical control and anarchic decentralisation; IV) there is a sophisticated system of decentralisation in taxing and spending; V) information regarding spending and results is transparent; VI) civil society is mature and human capital abundant.

Looking at this list an obvious objection is that decentralisation is a good idea in situations where development has already taken place, or is at least well underway; not in areas where development must be set in motion. This objection can be attenuated[139], but not entirely rejected. In the model we are going to examine, the objection holds, to a certain extent. The local institution called upon to set development in motion faces a dilemma: to act or not to act? The ever-present risk is losing consensus. This dilemma is the result of two causes, both related to the list of conditions above: first, spending and results are uncertain and not very transparent; second, local society is not mature enough to allow 'losers' to see through the promises of social groups representing the 'winners'.

The basic model

As we have already said, one explanation for the marked differences in the success rates of local development policies is that local institutions come into closer contact with highly differentiated interest groups. Under uncertain conditions, there could be a particularly delicate problem of consensus.

A useful starting point for examining this issue is provided by Fernandez and Rodrick (1991). Let us assume that there is a clearly identified institutional action with a guaranteed positive impact worth a total of B (benefit) and with a guaranteed cost totalling L (loss). Let us also say that this action is economically justifiable from the point of view of the common good, in that $(B/L)=b>1$. The action, however, in this

---

[139] Shah (1998), for example, laments that the debate has been too schematic; while Seddon (1999) stresses that the list contains features with different levels of importance, and that what counts is the overall design of decentralization.

hypothetical situation, can only be undertaken is there is consensus inside the majority interest group, while the distribution of net benefits among the various groups is uncertain. It is assumed that institutions are virtuous; that is, it is assumed that if an action is collectively useful it will be undertaken as long as there is ex ante social consensus for undertaking it. We could call this a regime of 'deliberative democracy' (Cersosimo, Wolleb, 2006). A decentralised government, in more direct contact with the real interests of different interest groups, however, does not generally fall into this category. Usually, consensus is gained in order to implement specific policies rather than to uphold the rules of representation.[140]

To continue the hypothesis, let us assume that in the system we are describing there are three interest groups S_1, S_2, S_3, whose relative weight is $d_1=d_2=d$ for S_1 e S_2, and $(1-2d)$ for S_3, with $0.5>d>0.25$ and then $0<(1-2d)<0,5$. In this case, none of the three groups constitutes a majority on its own, but two groups together are quite consistent. Let us assume that the benefits and losses (costs) are distributed as follows: I) guaranteed B benefits for group S_1; II) guaranteed L losses for group S_3; III) uncertain benefits and losses for group S_2. Let us also assume that the uncertainty of the benefits or losses for the social group S_2, derives from the fact that ex post a quota p of S_2 will have a benefit B, while a quota of $1-p$ will have a loss of L. Thus individual members of the group have the probability p of gaining something, and the probability $1-p$ of losing something, but they do not know what these probabilities are. Taken all together, these assumptions concerning the social composition and distribution of the gains and losses distributed by means of public action are the simplest way of representing the problem of uncertainty in the distribution of benefits and therefore the problem of consensus. To put it simply, the future results of today's public action are uncertain ex ante for a significant segment of local society.

Regarding the institutional action that has been identified, group S_1 plus pS_2 (the share of the second group that stands to gain) should be interested, but ex ante only S_1 is interested because individually each member of S_2 has no idea who stands to gain or to lose. There is no chance, then, of gaining consensus or dissent before the action is implemented. Group S_1, which has guaranteed benefits, is not a majority, but neither is group S_3, which certainly stands to lose. Ex ante consensus or dissent

---

[140] Sartori (1993) remarks that in a representative democracy consensus is by definition guaranteed by rules so that governments are not subject to the problem of gaining consensus over specific issues because they gained it (and must continue to earn it) through elections.

depends on the influence that group S_2 decides to exert by joining the ranks respectively of group S_1 or group S_3.

Since $0 < p < 1$, if the institutional action is implemented, ex ante consensus will be higher than ex post consensus; if the action is not implemented, the opposite will be true. This means that there could be actions that are justifiable both economically and socially and that are implemented on the basis of social consensus but which will accrue advantages to a smaller number of beneficiaries than the number that supported the action in the first place. There could also be actions, by contrast, that are not implemented even though they would have given benefits to the majority.

A hypothetical solution to this institutional dilemma would be if the social group that stood to gain guaranteed benefits (the winners) succeeded in acquiring the support of a sufficient share of those who were uncertain as to their benefits by promising to cover possible losses. The total benefits in this hypothetical solution would outweigh total losses. The group of those with uncertain outcomes would thus be divided up and ex ante consensus would realign with ex post consensus. However, this solution is not realistic, since a promise to cover another group's losses would not be credible. Those who were uncertain would surely think that the ex post winners would form a majority and would thus be unlikely to keep the promises they made when they were a minority.[141]

This institutional dilemma is illustrated in Figure 26.

---

[141] If the winners' promises were considered credible by the losers there would be no dilemma and the institutional action would automatically follow from the ex post effectiveness of the project (even though it was only imagined by many). A great deal of the literature (especially among sociologists) claims that these promises are credible on the basis of the idea of 'social capital': credible promises made by single agents (Bourdieu, 1983), made by groups of winners or social components (Coleman, 1988, 1990), promises considered credible because a diffuse social group believes this (Putnam, 1993, Fukuyama, 1999). The problem is that social capital is insufficient in a context of backwardness, while it becomes relevant after development has taken place.

Figure 26 - Institutional dilemma when an action is socially beneficial but there is a substantial gap between ex ante and ex post consensus.
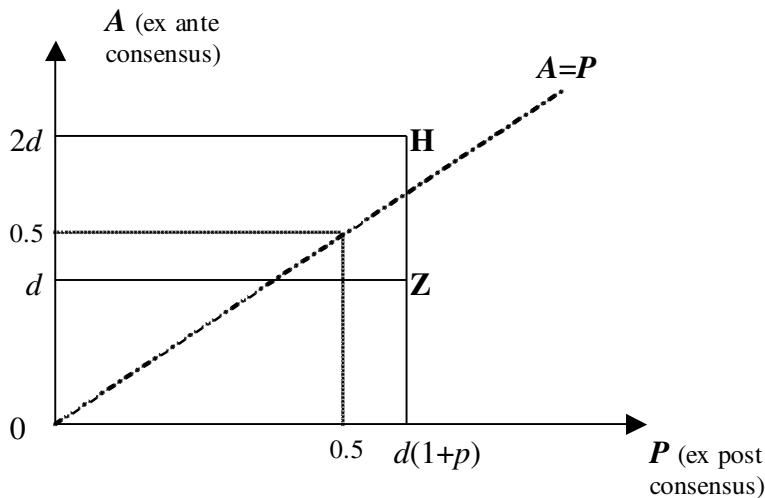
The number of votes that can be obtained when a decision to implement an action or not (ex ante consensus A) is indicated on the vertical axis. Ex post consensus P, or the number of people that would vote in favour of the action if they knew its result, is shown on the horizontal axis. Ex post consensus measures the entity of the group that would benefit if the action were implemented. The bisector $A=P$ is the optimal position, where ex post and ex ante consensus are equal. The points along this line show the various equilibriums that would arise if there were rational choice and perfect information; that is, if those who voted ex ante for taking action were the ones to benefit from it ex post.

Given that we assumed that the action was beneficial for a majority of the social body, ex post consensus is indicated in correspondence with the value $d(1+p)>0.5$. Since this utility is unknown ex ante, and only the utility for interest group S_1, which is $d > 0.5$, is known, one possibility is that ex ante consensus for implementing the action has a value of d, voting only for group S_1. This corresponds to point Z, which identifies a situation in which ex ante consensus is among a group smaller than the majority and therefore the action will not be implemented. The institutional dilemma arises because there is only one alternative, on point H, which indicates a situation where the action is implemented to the benefit of a majority that is even greater than the number of those who will have ex post benefits. In this case, ex post, the action

will leave many dissatisfied: all those who voted in favour of the action but gained nothing from it. Point H corresponds to the ex ante consensus that interest group S_1 (with guaranteed advantages) gains by convincing group S_2 to vote in its favour even though its members do not know whether they will gain benefits or not.

The graph shows that if the action under consideration gains ex ante consensus to a value of $d<0.5$ it is not implemented, while if it gains $2d>0.5$ it will be implemented even though it is such that it could gain ex post consensus equal to $d(1+p)>0.5$, which is a majority but still less than $2d$. Ex post consensus is thus smaller than ex ante consensus if the action is implemented (point H in the graph) and greater if it is not (point Z).

At this point we can take another step forward as suggested by Dur and Swank (1997). This consists in assuming that different actions are identified, each of which presents in isolation the same institutional dilemma (for each of them the condition of ex ante consensus is violated, and the condition of ex post consensus is verified). When these different actions are proposed together in the form of a project, they can align ex ante and ex post consensus if the guaranteed benefits and losses are distributed among different interest groups. Dur and Swank demonstrate that when this happens the condition of ex ante majority consensus is the more probably satisfied the greater number of actions that comprise the project and the greater equilibrium there is in the distribution of overall net utilities.

This last point is illustrated in Figure 27.

Figure 27a - Solution to the institutional dilemma where different actions (n) combined (project) each of which is socially convenient but taken singly, present a substantial gap between ex ante and ex post consensus.
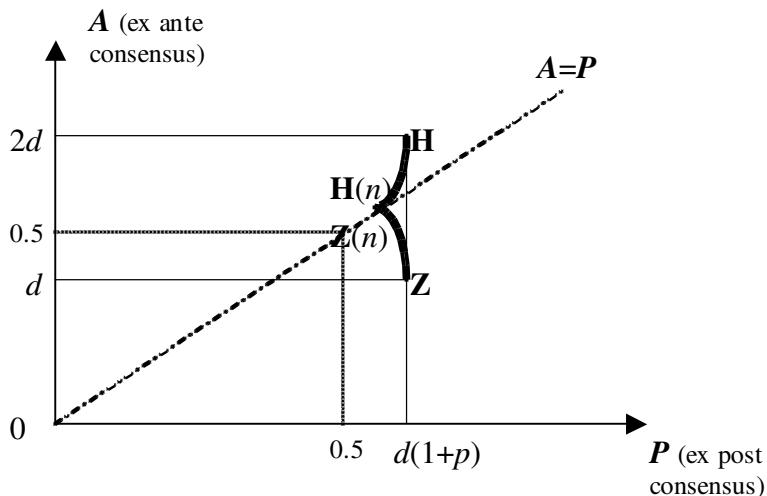

Solutions H (implementation) and Z (non implementation) in the hypothetical situation with only one institutional action, have been abandoned. Where in the previous graph there were two points (H and Z) indicating only two alternatives, in Figure 27 there are two curves, shown as functions of the number of actions that comprise the project (n): H(n), Z(n). The hypothesis is that the more complementary actions comprise a complex, articulated project (the higher the value of n), the more the solutions will approach equality between ex ante and ex post consensus along the curves, and the closer they will lie to the bisector $A=P$.[142]

The two curves are oriented towards the Cartesian origin. The hypothesis is thus that the project comprising different actions acquires ex ante consensus but loses it ex post, and therefore becomes less effective overall. This assumption makes sense, though it is not strictly necessary. It makes sense because a project that satisfies an ex ante majority is likely to be less innovative and/or more expensive. It is not strictly necessary, however, because if the curves were oriented towards the right there would be the same result with converging consensus. In this case the hypothesis would be

---

[142] It is thus assumed that, while interest groups cooperate when they make decisions concerning institutional actions in conditions of uncertainty, they tend to express themselves separately when the benefits likely to accrue to them are clear in their minds. We have seen that group S_2 votes all together for or against an action that gives the group uncertain results, but in this context it is assumed that the group splits and individuals vote for or against different actions comprising the project according to the benefits they are certain to obtain. This is another example of deliberative democracy, and illustrates well why it is supported.

that different actions comprising a complex, articulated project give rise to greater ex post collective utilities than actions taken singly (see Figure 27b).

Fig. 27b - Solution to the institutional dilemma where different actions (n) combined (project) each of which is socially convenient but taken singly, present a substantial gap between ex ante and ex post consensus. A case where the project is overall more effective than the single actions it comprises.
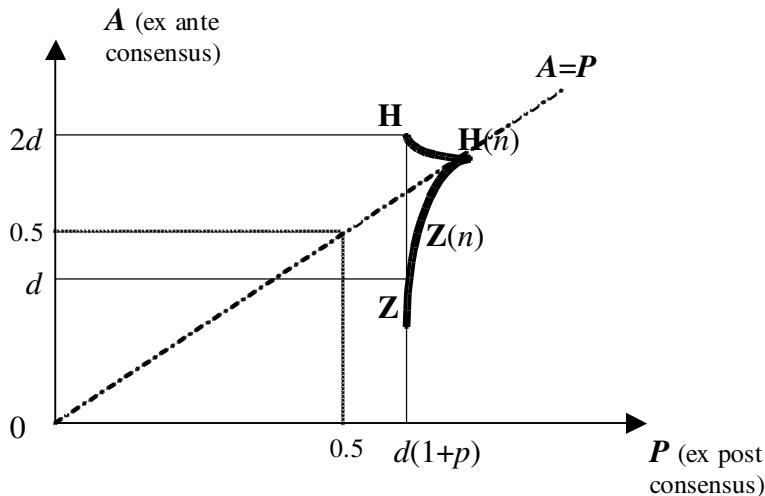
Lagging regions

We have until now assumed two hypotheses that evidently have no basis in real life, especially in lagging regions. The first was that institutional action couldn't be implemented without ex ante consensus that was precisely defined as belonging to a majority (more than 50%) and that was a direct consequence of the distribution of benefits among the various social groups. The second was that the institutions called upon to take action were capable of identifying, formulating and implementing a project without sustaining costs.

In real life, and especially in backward regions, however, forms of government are not particularly democratic, nor are they deliberative. Certain groups, even when they represent a significant portion of the population, continue to count for little and decisions are based on anything but majority consensus. On the other hand, there are few dictatorial regimes where social consensus can be done anyway with altogether. More realistically, in most cases, either by manipulating consensus and/or owing to undemocratic constitutions – be they formal or informal – institutions can implement projects even when there is no majority consensus. What this means, however, is that

governments need to support the costs in order to manipulate consensus or produce propaganda. Moreover, identifying and formulating actions to set development in motion, especially if these are articulated in a complex project, require local bureaucrats and politicians to have superior technical and administrative skills, which is rare in backward areas. This again means that governments need to sustain the significant project costs entailed in a complex project.

Government institutions are thus called upon to support the costs of identifying and formulating projects, the dimensions of which regulate the possibility of envisaging a complex project and therefore guide 'spontaneous' ex ante consensus. At the same time, they are called upon to support the costs of gaining consensus, which regulates 'induced' ex ante consensus. Consequently, the hypothesis implicit in the basic model examined in section 2, according to which institutions do not have intentions of their own, is no longer valid. In the basic model, institutions automatically translate opportunities into projects, with socially distributed gains and losses, and automatically gain consensus for doing so. That is, they do not have their own utility function. And yet, if there are costs for institutions which, as we have been reminded, are neither perfectly democratic nor perfectly competent, it is not up to the interest groups to decide whether, and to what extent, to sustain these costs, especially if they do not know how much they will be called upon to pay. It is the institutions that must sustain these costs, making it necessary at this point to ascribe an objective function to them.

If institutions are required to spend resources, it can no longer be assumed that the perfect alignment between ex ante and ex post consensus illustrated in the model can be achieved. Resources are always limited.[143] Let us assume then that ex ante and ex post consensus can be close but not aligned. In this case, given their stock of available resources, institutions decide how much to spend in order to implement the project (to cover both propaganda and project-building costs), and by doing so they also decide how innovative the project will be. Let us assume, in conclusion, that the project maintains significant margins of uncertainty regarding the outcome, which is anyway a long-term outcome.

---

[143] Assume that the stock of available institutional resources is given and does not depend on the results of the project.

We can now divide the analysis into two stages. In the first, we assume that the distribution of institutional resources between propaganda and project-building costs is given. In this stage, then, institutions only have to decide how much to spend and what projects to adopt. Let us assume, as we did in the basic model, that institutions are virtuous, but now let us specify that for any given project they will try to maximize its level of innovativeness. We could thus define a project drafted by institutions on the basis of an already existing social demand as 'conservative', while a project drafted to satisfy a demand that has not yet been formulated could be defined as 'innovative'. Let us assume, again, that this feature can be placed on a continuum, so that we can talk about a 'degree of innovativeness'. And let us assume, finally, that institutions, again as in the basic model, are required to submit to a consensus constraint, but now let us specify that this constraint is the result of the difference between two terms: one that depends directly on how many resources are used in what way, and that depends inversely on the degree of innovativeness of the project; the other that depends inversely on the resources used. The hypothesis we are testing is that by spending resources, institutes promote consensus, which depends negatively and to an unknown extent, however, on how many resources are used – precisely because these resources will be taken away from other projects.[144] Uncertainty affects both these terms. It affects the first because institutions know and can manipulate ex ante consensus in order to maximize it, but they cannot predict the degree of consensus they will obtain. It affects the second because institutions cannot know what consensus they are losing (owing to the subtraction of resources from other projects) because this is only revealed ex post, when it becomes possible to compare the net benefits of the project that was implemented with the losses attributable to the implementation of that project rather than any another within the economic system as a whole.

If we indicate consensus with the letter *A* then, we could say that:

---

[144] Notice that these losses are not the same as the ones accounted for in the project (which already envisages socially distributed gains and losses). This is a matter of taking into account the social losses owing to the subtraction of resources from other alternative uses. For example, building a bridge benefits those who have to cross a river, but causes losses for the ferryboat owners. The bridge, however, could also be a source of additional losses, for example the recipient of subsidies if the construction subtracted resources from subsidies, etc. Giving incentives to organic farm producers, to provide another example, damages those producers who cannot farm organically, but also takes resources away from other possible beneficiaries.

$$A = \gamma \frac{\beta}{T} a - \beta\varphi(\beta) \qquad\qquad [19]$$

with:

$\gamma > 0$ = Scale parameter (unknown) that regulates the level of ex ante consensus,

$0 < \beta < 1$ = Variable indicating the share of (given) resources used;

$T > 0$ = Variable indicating innovativeness (*1/T* degree of conservatism) of the project;

$a > 0$ = Resulting variable in terms of consensus in the distribution of resources to cover propaganda and/or project complexity;

$\varphi(\beta) > 0$ = Scale value (unknown) that regulates the subtraction of consensus, as a function (with a positive first derivative) of the share of resources used and subtracted from other uses.

We can assume, for simplicity's sake, that:

$$\varphi(\beta) = \varphi\beta$$

Thus:

$$A = \gamma \frac{\beta}{T} a - \beta^2\varphi \qquad\qquad [20]$$

The ex ante consensus constraint, for example, $A \geq \overline{A} > 0$, will oblige institutions, when they do not know the value of γ and φ, first to maximize *a*, and then to take inspiration from the experience around them to decide on which project *T* to choose.

Let us imagine that collective experience derives from the outcomes of projects of the same kind as these we have schematized, and that they can be observed as values. Let us also imagine that *T* is associated with *A=0*, that is with projects whose ex post outcome has not been considered positively in terms of net consensus gained by subtracting from ex ante consensus the loss of consensus ex post owing to the change of destination for resources from previous to new uses.

From [20] we thus get to:

$$A \geq 0; for: \frac{a}{\beta} \geq T\frac{\varphi}{\gamma} \qquad\qquad [21]$$

Collective experience thus shows that, for every φ/γ, the more innovative a project is (increasing T), and the greater the ex ante consensus a is in relation to the resources used, the more sustainable it is from the consensus side because [21] indicates a directly proportional ratio between T and a/β along the line which separates sustainable from unsustainable projects.

In reality, however, this consideration does not help institutions decide. What they would be interested in would be an indication of the relationship between T and new tools for consensus. Taking as given that these new generations of consensus depend on the features of the institutions (their history and structure), collective experience could be relied upon if it provided an indication as to what kind of innovations to introduce under what circumstances. This would almost certainly take place if the relation between what determines a and T were monotonous. In this case, it is easy to imagine that collective experience would soon indicate the correct path to follow: a project that is as innovative as the number of factors that determine a, which a specific institution would easily be able to identify. And yet, if this were the case, they would no longer have any freedom of choice; their intentionality, their very entrepreneurial spirit, would be lost.

It is a matter, then, of discussing whether the value of a can be such that it progresses monotonously whatever variety of factors determine it. If, for example, we should conclude that first it decreases then it increases as the factors that determine it vary, hypothesis [21] would be verified for both low and high values of these factors. Collective experience would then be hard to interpret. This is precisely what happens, as we shall now see.

Let us go on to the second step. We have said that consensus can be promoted in two different ways that we assume are not strictly complementary and can thus to a certain extent be substituted: by supporting the cost of manipulating consensus, and /or by supporting the costs related to making the project complex enough to be useful for a wide number of beneficiaries. We can thus imagine an output that depends on two inputs, propaganda and project building, whose costs are related respectively to the costs of manipulating consensus and the costs of building a complex project. The

hypothesis we are testing is that institutions are capable of making a choice that maximizes the joint result, combining propaganda and project building.

We thus have:

$$M = \psi^\alpha F^{1-\alpha}; con : 0 < \alpha < 1 \qquad\qquad [22]$$

The result (M) for institutions is the geometric mean[145] for propaganda (Ψ) and project building (F)[146], with a weight of α and 1-α that is positive and less than 1. The parameter α indicates the weight of propaganda costs compared to project building costs from the side of consensus. It could be claimed that this value will be high if the managerial class is 'new' and eager to show what it is capable of, and low if the managers have already been tried and tested.

Now the function of institutional cost becomes:

$$C = c_\psi \psi + c_F F = \beta K \text{ ; with } 0 < \beta < 1 \qquad\qquad [23]$$

In this formula c_Ψ e c_F are unit costs, C the overall resources that institutions decide to use, a share β of available stock (K).

It is hard to define what the unit cost of propaganda (c_Ψ) depends on; it could depend on the extent to which the media are controlled. Let us assume that it is a given, and that it is the same in all cases. The project building cost (c_F) rises as the competence of institutional apparatuses drops.

The problem of what the maximum constraint is can be resolved given the identity:

$$\frac{\psi}{F} = \frac{\alpha c_F}{(1-\alpha)c_\psi} \qquad\qquad [24]$$

---

[145] The geometric mean is a value that substitutes all the terms that comprise it but leave the product unvaried. The geometric average is zero if any of the terms are zero. Because of these properties, it represents very well the average intensity of values that compete as factors that determine a result.
[146] Notice that F indicates a different character of the project from T. While F indicates its complexity, according to the number of actions that comprise it, T indicates its innovativeness, meaning the extent to which (whatever its complexity) the actions have yet been demanded by interest groups. The project thus anticipates explicit social demand.

Institutions will tend to opt for a ratio between propaganda and project-building equal to the ratio between their returns multiplied by the inverse of their costs. This proportion will be high with a new but incompetent managerial class (α high, c_F high), and low with tried and tested, competent managers who will use their experience to spend more on complex project-building and less on propaganda.

Equations [22, 23, 24] can be resolved for M, obtaining:

$$M = \beta \left[K \frac{\alpha^\alpha (1-\alpha)^{1-\alpha}}{c_\psi^\alpha c_F^{1-\alpha}} \right] = \beta a \qquad [25]$$

The term between the square brackets is a as in equation [21].

It assumes positive values that however decrease as α drops and c_F rises, or to put it another way, going from new, competent managers to tried and tested but incompetent managers.[147] And yet, for every given value of c_F (for every given level of competence), the relation between a and α is not monotonous. Its progress forms a parabola with its high points corresponding to very new or very experienced managers and its nadir corresponding to $\alpha = 0.5$.

Fig.28 Ratio between a and α, for every given K, c$_\psi$ and c$_F$

---
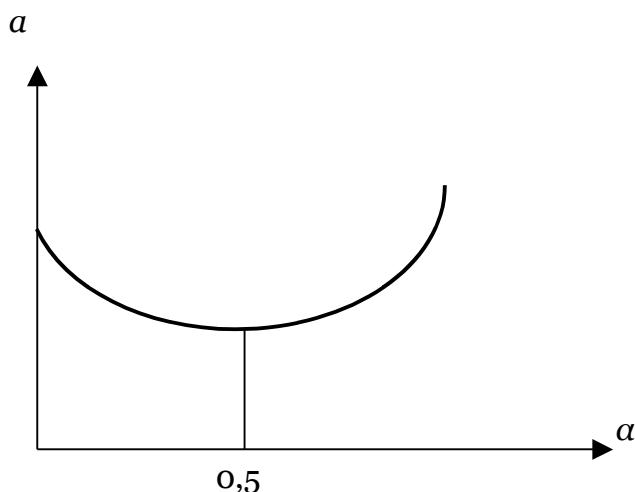
[147] Notice the distinction between the 'experience' of the managerial class and their 'competence'. We assume here that these two features are independent, i.e. that there can be competent new managers with no experience at all, and incompetent but extremely experienced ones.

What we find is that among the factors that determine a, there is at least one (α, or inexperience/experience of managers) that does not have a monotonous relation with a. This factor is also probably the most recognizable (compared to that of the competence of institutional apparatuses, which is harder to evaluate from outside) and it therefore plays an important role in interpreting collective experience. Thus collective experience is anyway hard to interpret because it shows that consensus can be won by innovative projects promoted by institutions whose managers are either experienced or inexperienced.

In conclusion, institutions should, as a rule, be capable of making decisions and risking unpopularity even in conditions of uncertainty. While their decisions are naturally influenced by historical and structural factors (K institutional resources; c_ψ media control, c_F apparatuses' competence; α managers' experience), their results depend on choices that have nothing to do with history, economic or social structures, or collective experience. These decisions are genuinely intentional and they are the result of an institutional and entrepreneurial attitude,[148] which guides decisions regarding how much to spend on what projects.

---

[148] This approach, envisaging intentionality that is not predetermined and therefore exogenous, is akin to Keynes' animal spirits. These mysterious animal spirits explain investments that are not an endogenous part of neoclassical economics. Keynes' aim was methodologically important. He wanted to build an economic theory with a beginning and an end, rather than a circular structure, interpreted along human lines. In Keynes' view there was a role for free will in the economy; it was not a self sufficient natural phenomenon with its own rules like the inanimate physical world.

# Chapter 4 Appendix 1 - The Historical Role of Local Institutions in Italy.

We have said that decentralizing development policies is not necessarily a solution in itself; what is more important is the local institution's real capacity to govern. This conclusion is borne out by the history of local development in Italy from the 1950s to the 1990s. On the basis of the per capita value added growth rate in the Italian provinces between 1950 and 1996[149], as well as of 18 other indicators that have influenced this main one[150], the provinces can be classified into groups[151] as in the map in Figure 29.

---

[149] The data stops at 1996. However, since 1996 there has been a new cycle and therefore the post-war cycle could be said to come to an end at about that time. The new cycle is also due to new countries coming onto the international market, the gap created by new technologies, and reforms implemented in Italy (Barca, 2006).

[150] See Arrighetti, Seravalli, 1999 and Arrighetti, Lasagni, Raimondi, 2000). The classification is more complex (by about 5 groups), but the essential message is clear in this simplified version.

[151] These variables include, as well as those indicated: an index of completeness of the original production structure, productive sectors and services to firms (1951), obtained on the basis of Utton's diversification index, modified taking into account the effect of sectorial specialization, number of firms with more than 50 employees (1961); an index of infrastructure endowments (1980s); a factorial index for cultural and economic association building (1982); a factorial index of social capital endowments (1960s-1970s); indices of civic maturity, measured by literacy, political activism, participation in referendums (1970s); a factorial index of activities supporting the economy on behalf of the chambers of commerce, local banks, professional training, territorial administration (1950s); a factorial index of infrastructure endowment regarding cultural services and social services (1950s); distance as the crow flies from Milan; proportion of mountainous territory; state incentives for agriculture (1972-81); state incentives for industry (1972-81). Using regressions and semi-parametric methods (tree-like classifications), the most significant independent variables were in order: distance from Milan, number of firms with more than 50 employees in 1961, social infrastructure, local institutions, completeness of the industrial structure, professional training, and civic sense.

Fig. 29 - Geography of development per province in Italy (1950-1996) The darker grey tones indicate provinces in group 1; light grey group 2; white group 3.
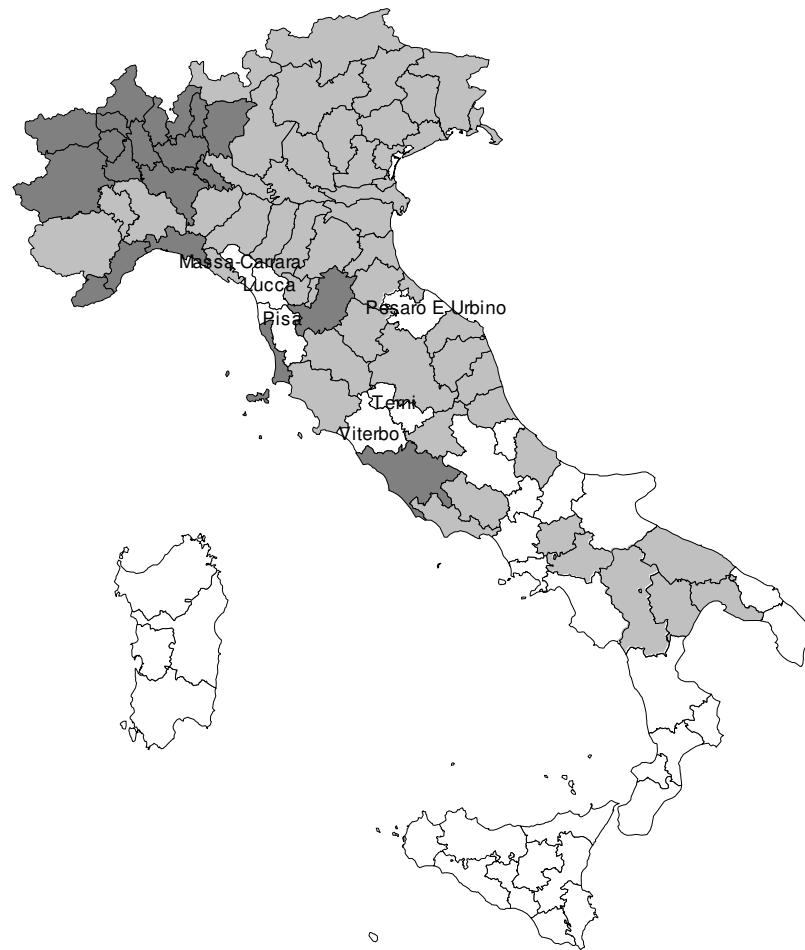
Source: Our elaboration using data from ISL (Databank Department of Economics, University of Parma).

Let us examine some of the data for these three macro-areas.

Table 4 - Italian provinces by group, per capita value added growth rate, role of local institutions, and maximum growth of neighbouring provinces.

	Growth	Index of per capita income to start with	Local institutions (professional training)	Maximum growth in neighbouring provinces
Gruppo 1 (Industrial Triangle)	2.6%	0.41	0.23	3.46%
Gruppo 2 (Centre-North-East)	3.5%	0.26	5.27	3.82%
Gruppo 3 (South)	3.1%	0.20	-7.97	3.52%

Source: Our elaborations with ISL data.

159

The provinces in the first group belong to the industrial triangle with the addition of Florence, Livorno and Rome. All together, these provinces had a higher per capita income in the 1950s, while the other two groups had a lower per capita income that was quite similar. In the fifty years that followed, there was greater growth in group 2 and lower growth in group 1. It is interesting to observe that this difference between growth and starting conditions corresponds to a marked difference – above all between groups 2 and 3, spread throughout the country but prevalently in the South – in the local institutions and professional training available to them. This variable was constructed by considering (in the 1950s) the proportion of the population of a similar age that attended schools, technical institutes or professional training centres, and by using the value of this variable that exceeds the share that in the regression is explained by the industrialisation rate at the time. It is therefore a proxy of the activism of local institutions in the field of professional training, which goes beyond what firms would have called for. Entrepreneurial spirit appears to be an important factor in the development of these provinces.

The last variable in the table (maximum growth of neighbouring provinces) was constructed by calculating the average per group of the maximum growth rate in neighbouring provinces for each of the provinces in the group. It therefore picks up the spill-over effect of development and also evidences the extent that agglomeration takes place. The highest values are to be found in the provinces of group 2 that are quite well distributed over the country but tend to be mostly in the central and north-eastern provinces.

The most relevant consideration for the analysis we have conducted in this chapter regards the partial but significant dissemination of group 2 towards the south. This shows that even in some backward southern regions there have been high rates of growth, such as in Avellino (3.9%), Potenza (3.6%), Matera (3.3%), and that this growth corresponds to positive values in the column that measures the activism of local institutions in professional training (Avellino +1.5; Potenza +1.8; Matera +6.2). By contrast, in other southern provinces with lower growth rates, this variable is negative, such as Naples (growth 2.2%, activism -5.9), Brindisi (growth 2.5%, activism -7.9) and Massa Carrara (growth 2.6%, activism – 8.7)

This data confirms that the hypothesis we have presented is plausible. Local institutions have a role in economic development but this role does not necessarily

depend on the existing economic conditions. It would seem, in fact, that we must consider their role as autonomous, and their policies as decidedly intentional.

Chapter 4 Appendix 2 - A Critique of the Cultural Approach.

In his observations of cross-national data, Dani Rodrik (2000) has shown that "democracy has no systematic effect on long-run growth rates."[152] Today, as in the past, undemocratic regimes have often enjoyed extraordinary economic success. For one case where democracy and growth correlate (Ireland), there are many others where growth and dictatorship seem to go hand in hand, such as many of the countries in South-East Asia, and China. In the past, development took place in a democracy in the cases of Holland, Britain, France and the US, but Germany, Japan and Italy underwent development under dictatorships[153]. These findings are in line with a long tradition[154], which thought the causes of economic development lay with individuals

---

[152] He highlights, however, that there is a statistically significant positive relationship between democracy and growth stability, and that there are good reasons to believe there is a causal relationship between democracy and stability rather than vice versa.

[153] Historians claim that Italy began to develop in the 1920s.

[154] Max Weber (1904-05) observed that "Protestants (especially certain branches of the movement to be fully discussed later) both as ruling classes and as ruled, both as majority and as minority, have shown a special tendency to develop economic rationalism which cannot be observed to the same extent among Catholics either in the one situation or in the other. Thus the principal explanation of this difference must be sought in the permanent intrinsic character of their religious beliefs, and not only in their temporary external historico-political" He then explains this tendency with an elaborate analysis based on the doctrine of predestination: "Thus, however useless good works might be as a means of attaining salvation, for even the elect remain beings of the flesh, and everything they do falls infinitely short of divine standards, nevertheless, they are indispensable as a sign of election They are the technical means, not of purchasing salvation, but of getting rid of the fear of damnation. In this sense they are occasionally referred to as directly necessary for salvation or the *possessio salutis* is made conditional on them." (p.115-6). He adds, with real anxiety: God's grace is, since His decrees cannot change, as impossible for those to whom He has granted it to lose as it is unattainable for those to whom He has denied it… In its extreme inhumanity this doctrine must above all have had one consequence for the life of a generation, which surrendered to its magnificent consistency. That was a feeling of unprecedented inner loneliness of the single individual. In what was for the man of the age of the Reformation the most important thing in life, his eternal salvation, he was forced to follow his path alone to meet a destiny, which had been decreed for him from eternity. No one could help him." Deepak Lal (1998) presents a reconstruction according to which "The instrumental rationalism that individualism bred was responsible for the scientific and industrial revolutions […] But individualism has paradoxically undermined the very cement of the prosperous societies it has created. […] this cement has been provided by the emotions of "guilt" or "shame" engendered in the process of socialization. The triumph of science that individualism has wrought has led to the death of God in the West, and thus of the guilt that underwrote its personal morality, whereas the democratization promoted by individualism has undermined the traditional hierarchical bases of these societies and erode the set of "manners" based on deference, which evinced shame in these societies in their process of socialization. […] By contrast, as Japan has showed, and as India e China are beginning to show […] it is possible for these non-Western societies to adopt the West's means to attain prosperity without giving up their souls. Their social cement is unlikely to be undermined by secularization, insofar as it does not depend on a belief in God. In these hierarchically ordered societies, the social cement was provided by ancient processes of socialization based on shame. "(p. 174, 176, 178).

and their inclinations rather than with relations between people and interests. If relationships are a cornerstone of an approach, then regimes are important, but if what counts in your approach is entrepreneurial spirit, creativity, hard work, saving, studying, loyalty, commitment and discipline, then these inclinations could be considered compatible with any regime since they depend on the history and culture of a given area. This 'culturalist' paradigm also adopts the approach that the competence of institutional apparatuses has to do with the general culture of a people and can therefore in no way be considered a factor in local development. In their view, development can take place even when political institutions are incompetent if society is all in all ready for it.

Criticism of this approach has been directed against Max Weber (Hirschman, 1977), and, more recently, against Deepak Lal (1998)[155], claiming that causal relationships between culture, tradition, religion and virtuous behaviour for development have often been invented outright.[156] The opposite approach would be that relatively competent institutional apparatuses (technically equipped and capable of decision making) are able to conceive and conduct sophisticated and complex development projects, which even contemplate the ability to exploit local culture intentionally and influence individual attitudes.

Wherever the truth lies, there is no doubt that the competence of institutional apparatuses was a decisive factor in the economic success of South-East Asia (Amsden, 2001). It is true that in the better-known experiences constructing and implementing projects was made easier by a centralised relationship between the State and big firms. There are nonetheless other experiences, such as in Taiwan, and South Korea after the mid-eighties (Seravalli, 1992), in which local political institutions and productive systems were the protagonists of development. In general, it can be claimed, as we have said, that competence in institutional apparatuses is required above all when complex projects need to be drawn up (Rodrik, 1995).

As regards the importance of local institutions being able to conceive and construct complex projects, the example of Italy could be a case in point. Table 8 shows the results for the various indices elaborated from ISL data.

Calvinism, at the end of the seventeenth century, for Weber, and Confuciansim today for Lal, are the true motors of development.

[155] In particular by Jones (2002)

[156] The fact that the managerial class has always tried to make their general behaviour appear 'traditional' and therefore 'morally necessary' has often been observed (Hobsbawm, Ranger, 1983).

Table 5 - Cross-section regressions at provincial level, dependent variable growth of per capita value added 1951-1996, ordinary list square (OLS)[157]

Independent variables	1	2	3	4	5
Constant	0.44	0.44	0.38	0.40	0.40
	32.5	*35.9*	*8.2*	*8.4*	*8.4*
Distance from Milan			-7E-6	-10E-6	-1E-5
			-2.9	*-4.3*	*-4.4*
Starting income (1951-63)	-0.07	-0.06	-0.06	-0.06	-0.06
	-10.2	*-10.4*	*-8.3*	*-8.2*	*-8.2*
Index of disarticulation (1960s – 1980s)			-0.0007	-0.001	-0.01
			-2.3	*-3.5*	*-3.5*
Spill-over (1951-1996)			0.2	0.2	0.2
			2.3	*2.1*	*2.1*
Professional training excluding local institutions (1950s – 1960s)					7.8E-5
					1.7
Professional training (1950s)				8.1E-5	
				1.9	
Local institutions (1950s – 1960s)		0.002	0.001		
		5.7	*2.6*		
Civic sense (1960s – 1970s)	0.003	0.002	-0.0002	-0.0005	-0.005
	6.7	*2.9*	*-0.3*	*-0.7*	*-0.7*
R adjusted	0.55	0.67	0.73	0.71	0.71

It would seem from column 1 that cultural conditions (civic sense) have had a significant impact on subsequent growth. This appearance is attenuated, however in column 2, even if you consider only local institutions; it disappears altogether if you consider the other variables, though local institutions continue to play a role. Development has been more affected by local initiatives that are inter-connected, rather than concentrated, even in a field as influential as professional training (see columns 4-5).[158]

---

[157] The independent variables were obtained as follows. Distance from Milan: as the crow flies kilometers from each provincial capital to Milan; Starting income: average per capita value added 1951-1963; Index of disarticulation: factor built by factorial analysis with high points for motorways and train lines, low and negative points for council and provincial roads and water supply (1965-1980); Spill-over: maximum rate of growth of per capita value added 1951-1996 in neighbouring provinces; Professional training: residual of specific rate of attendance for technical and professional schools regressed for rate of industrialization (1950-51); Local institutions: factor from principal components such as loans from local banks, cooperatives, savings banks and other local banking institutions as a share of loans from all credit institutions in 1960 by level of specialisation of manufacturing activity (1951), factor relative to the Chambers of Commerce, presence and activity (1951-53); Professional training , average 1961-63 of council payments for public works and education over ordinary spending; Civic sense: factor (Putnam, 1995) with high points for the share of population that habitually reads (1965) and voters for the referendum on divorce (1974) out of total number on electoral roll, low or negative points for preference votes out of total number on electoral roll in political elections 1953, 1958, 1963. Source: ISL data.

[158] Di Giacinto and Nuzzo (2004) in an econometric study based similarly on a provincial study come to the conclusion that: "There was no significant intervention on the part of local bodies in promoting local development." And yet, even though they had the same database as we had (ISL), they use only one variable to test the role of local institutions: city council payments for public works and education on the ordinary budget. Our results are different (with a significant role for local institutions in

Bibliography for chapter 4

Acemoglu, D. e F. Zilibotti (1997), Was Prometheus Unbound by Chance? Risk, Diversification and Growth, *Journal of Political Economy*, 105, 4, pp. 709-751.

Acemoglu, D. (2006), *Modeling inefficient institutions*, NBER Working Paper 11940.

Amsden, A. (2001), *The Rise of "the Rest": Challenges to the West from Late-Industrializing Countries*, Oxford University Press, New York.

Amsden, A. e C. Wan-Wen (2003), *Beyond Late Development: Taiwan's Upgrading Policies*, The MIT Press, Cambridge (Ma).

Arrighetti, A. e G. Seravalli (1999), Sviluppo economico, convergenza e istituzioni intermedie; in A. Arrighetti e G. Seravalli (ed.), *Istituzioni intermedie e sviluppo locale*, pp.25-50, Donzelli, Roma.

Arrighetti, A., A. Lasagni, M. Raimondi (2000), *Capitale Sociale, Istituzioni e Rendimento Istituzionale,* mimeo Sezione di Scienze Economiche Università di Parma, 2000.

Azfar, O., S. Kähkönen, A. Lanyi, P. Meagher, D. Rutherford (1999), *Decentralization, Governance, and Public Services: the Impact of Institutional Arrangements – A Review of the Literature*, IRIS Centre, University of Maryland, College Park, September.

Barca, F (2006), *Italia frenata – Paradossi e lezioni della politica per lo sviluppo*, Donzelli, Roma.

Bourdieu, P. (1983), Forms of capital, in J. C. Richards (ed.) *Handbook of Theory and Research for the Sociology of Education*, Greenwood Press, New York.

Bradshaw, T.K. e E.J. Blakely (1999), What Are "Third-Wave" State Economic Development Efforts? From Incentives to Industrial Policy, *Economic Development Quarterly*, 13, 3, pp. 229-244.

Cersosimo, D. e G. Wolleb (2006), *Economie dal basso*, Donzelli, Roma.

Coleman, J. C. (1988), Social capital in the creation of human capital, *American Journal of Sociology*, 94, pp. S95-S120.

Coleman, J. C. (1990), *Foundations of Social Theory*, Harvard University Press, Cambridge (Ma).

Di Giacinto, V. e G. Nuzzo (2004), Il ruolo dei fattori istituzionali nello sviluppo dei distretti industriali in Italia, in Banca d'Italia: *Economie locali, modelli di agglomerazione e apertura internazionale: nuove ricerche della Banca d'Italia sullo sviluppo territoriale*, Conference paper, Bologna November 20, 2003, pp. 251-278.

Dur, A.J. e O.H. Swank (1997), *The Role of Governmental Agreements in Breaking Political Deadlock*, Research Memorandum 9707, Erasmus University Rotterdam, Timbergen Institute and ECFEB.

Fernandez, R. e D. Rodrik (1991), Resistence to reform: status quo bias in the presence of individual-specific uncertainty, *American Economic Review*, 81, pp. 1146-1155.

Flyvbierg, B., N. Bruzelius, W. Rothengatter (2003), *Megaprojects and Risk: An Anatomy of Ambition*, Cambridge University Press, London.

Fukuyama, F. (1999), *The Great Disruption. Human nature and the reconstitution of social order*, Profile Books, London.

---

promoting development) because we use a factorial reduction of this, as well as considering the role of local banks, Chambers of Commerce and professional training.

Hayek, F. (1945), The Use of Knowledge in Society, *American Economic Review*, 35, pp. 519-530.

Hirschman, A.O. (1977), *The Passions and the Interests: Political Arguments for Capitalism Before its Triumph*, Princeton University Press, Princeton.

Hobsbawm, E.J. e T. Ranger (1983), (a cura di) *The Invention of Tradition*, Cambridge University Press, Cambridge; tr. it.: *L'invenzione della tradizione*, Einaudi, Torino, 1987.

Jones, E. (2002), *The Record of Global Economic Development*, Edward Elgar, Cheltenham (UK) e Northampton (MA).

Jütting, J., C. Kauffmann, I. Mc Donnell, H. Osterrieder, N. Pinaud, L. Wegner (2004), *Decentralisation and Poverty in Developing Countries: Exploring the Impact*, Oecd Development Centre, Working Paper n. 236.

Lal, D. (1998), Unintended Consequences: The Impact of Factor Endowments, Culture, and Politics on Long-Run Economic Performance (Ohlin Lectures), MIT Press, Cambridge (Ma).

Middleton, L. (2001), *Literature Review: Tax Abatements and Economic Development Incentives,* The University of Kansas Centre for Economic and Business Analysis Policy Research Institute, Technical Report Series, Report n.49.

Musgrave, R.A. (1959), *The Theory of Public Finance*, McGraw-Hill, New York.

Oates, W. (1968), The Theory of Public Finance in a Federal System, *Canadian Journal of Economics*, I, 1, pp. 37-54.

Ostrom, E., L. Schroeder, S. Wynne (1993), *Institutional Incentives and Sustainable Development: Infrastructure Policies in Perspective*, Westview Press, Boulder Colorado.

Premchand, A. (2003), *Ethical Dimensions of Public Expenditures Management*, Georgia State University, Andrew Joung School of Policy Studies, Working Paper 03-14.

Putnam, R. D. (1993), The prosperous community: social capital and public life, *American Prospect*, 4, 13.

Putnam, R. (1995), *Economic growth and social capital in Italy*, in idem J.F. Helliwell and R.D. Putnam (a cura di), Fondazione ENI Enrico Mattei, Milano.

Rodrik, D. (2000), *Participatory Politics, Social Cooperation, and Economic Stability*, Paper for the American Economic Association meetings in Boston, January 7-9.

Sartori, G. (1993), *Democrazia*, Rizzoli, Milano.

Seddon, J. (1999), Decentralization and Economic Growth, in J. Litvack, J. Seddon (a cura di) *Decentralization Briefing Notes*, World Bank Institute, WBI Working Papers in collaboration with PREM network.

Seravalli, G. (1992), *Profile for development policies in newly industrial countries of South-East Asia,* Quaderni del Laboratorio Sviluppo e Impresa, Università di Parma.

Shah, A. (1998), *Balance, Accountability, and Responsiveness: Lessons about Decentralization*, Policy Research Working Paper Series, n. 2021, World Bank, Washington, DC, USA.

Tiebout, C. (1956), A Pure Theory of Local Expenditures, *Journal of Political Economy*, 64, pp. 416-424.

Weber, M. (1904-05), Die Protestantisches Ethik und der Geist des Kapitalismus, Archiv für Sozialwissenschaft und Sozialpolik, XX-XXI, J.C.B. Mohor, Tübingen.

Weber, M. (1989), *L'etica protestante e lo spirito del capitalismo*, Sansoni, Firenze.

Chapter Five - Institutional Frameworks[159]

Local institutions, called upon to provide public goods and services in order to trigger and sustain development, would be better equipped to do so if they had more resources than the ones they can obtain from local taxation. Faced with the problem of uncertain results and the probable asymmetry of advantages and disadvantages for different interest groups in the community, their task would also be less arduous if the resources needed for a complex project were guaranteed not to be redirected to other more consolidated uses. Surveys and case studies have shown this to be true, and the international literature has accepted the principle that when decentralising development policies, it is a good thing if there is some external funding from central governments or international bodies, and that this funding should not cover the entire costs (in order to avoid dependence) but should be quite sizeable. It has also been agreed that the way funds are spent should be at the discretion of local governments, but that the principle of 'additionality'[160] should be maintained.

The resulting institutional frameworks, however, can be quite cumbersome. Central governments find it difficult to come up with criteria that have been agreed upon for transferring resources directly and distributing them among highly variegated localities. Generally, intermediate levels of administration or institution step in, with the result that in almost all experiences of decentralisation there is an institutional framework comprising at least three levels: national, regional and infra-regional local. This structure, in its turn, generates specific problems. The structures providing funds are not able to control directly the nature and effectiveness of the policies that have been adopted, which leaves significant margins of freedom at a local level. This

---

[159] Many points made in this chapter were inspired by a seminar held at the Department for Development and Cohesion (DPS) at the Italian Ministry of the Economy in February 2005, and in particular by a paper given by Paola Casavola on the effectiveness of simple rules in development policy making, and by the contributions of Paolo Bertoletti from the University of Pavia and of other DPS researchers, including Giorgio Pugliese and Paolo Praticò. Other ideas were generated in discussions with Paola Casavola, Giampiero Marchesi, Anna Natali and Tito Bianchi. For parts of section 3 and the introduction to section 4 I am indebted to Domenico Cersosimo and Guglielmo Wolleb.

[160] A system of transfers from central to local powers based on decentralized fiscal structures. Shah (2003) recommends this system "because information requirements and transaction costs are lower at lower levels of government and the government can be more responsive to the citizenry." Shah and others (see, for example, Von Braun, Grote, 2002) stress that local governments must be able to decide how to use resources, but at the same time they must avoid using them for ordinary administrative purposes, maintaining the idea of their 'additionality'.

creates a need for rules to be imposed on the beneficiaries of funding in order to 'protect' the resources from misuse. Often these rules are dictated from the centre, but accepted ex ante by local governments in order to receive the funding. Because they must follow the rules, however, local governments could easily become prey to incentives for implementing projects that are not particularly effective in their case, receiving indiscriminate funding for superficial projects with no particular priorities.

Institutional arrangements can provide a solution to these problems; in particular by ensuring that tasks and responsibilities are clearly identified at all levels. A good starting point for development, we have found throughout our analysis, is to design institutional frameworks so that these tasks and responsibilities are clearly identified and attributed. This is a necessary, but insufficient, condition. Even more important are: explicit and implicit rules, premiality systems, minimum standards for public goods and services, and, above all, valorising high level functionaries and institutional leaders so that they are capable of exerting authority in a coordinated way at all levels. All these mechanisms present a contrast with the 'naïve' idea of development, which believes all you need to do is bring resources and skills to the citizenry. Local development policies, in fact, require an authoritative and competent executive directorship.

This chapter is divided into five sections. In the first section we examine the topic of the availability and protection of resources for development policies in the context of institutional decentralisation. It emerges that explicit rules agreed upon by both donors and beneficiaries are necessary in order to avoid the risk of inactivity at a local level. In the second section, we show that simple rules are advantageous. In the third section, we underline, however, that in a multi-level institutional arrangement, even simple rules tend to generate significant costs for project building and implementation. If learning costs are relevant in relation to the possible benefits, projects tend to be insipid. If learning costs, on the other hand, could be contained, with a more rapid learning process, there would be very different final outcomes. As we show in section four, some local governments would spend on development and use their resources wisely, while others would not spend anything, or use their resources wastefully. In section five, in conclusion, we provide a few indications as to how to design a 'sophisticated' decentralised institutional framework, and we show that this requires strong commitment on the part of central government.

Decentralisation and resources for local policies

Many developing and emerging countries have in recent years adopted decentralised development policies, and numerous case studies have studied the phenomenon. An OECD synthesis regarding 19 countries[161] sketches a highly varied picture. There are, as yet, it seems, few examples of outright success[162] (these include Bolivia, the Philippines, and West Bengal), especially when public goods and services, essential for development, are taken into account. The most relevant observation for our purposes, however, is that a vital factor in explaining failure or success is the intentionality with which reforms are enacted by central and local governments. Apart from cases where there is no intentionality at all, and reforms are merely apparent (or, worse, designed to fragment discontent by distributing despotism and authoritarianism throughout the territory), in other cases success seems to depend on the transparency and effectiveness of the rules adopted by central governments for assigning resources, and on the capability of local governments to put them to good use. Looking at the last eight years[163] of 'regional' development policies, the Italian experience seems to confirm this observation[164].

---

[161] The countries are: Bolivia, Philippines, India (W. Bengal), China, S.Africa, Mexico, Ghana, Paraguay, Brazil, Nepal, Vietnam, Egypt, Sri Lanka, Ethiopia, Burkina Faso, Uganda, Guinea, Mozambique, Malesia, India (Andrah Pradesh and Madyar Pradesh). The study was conducted by the OECD (Jutting, Kauffman, McDonnel, Osterrieder, Pinaud, Wagner, 2004).

[162] The countries selected in the survey have either introduced in the last decade institutional decentralization reforms, or have reinforced existing mechanisms. The evidence shows that these reforms take a long time to bring results.

[163] Italy set up its 'New Programming' after the end of the 'extraordinary intervention' in 1988. In 1999 the new territorial pacts came into force, together with the new Strategic Project for the Mezzogiorno. At the same time the procedure to include the new strategy in the Community Support Framework (CSF) 2000-2006 was set in motion.

[164].The preliminary national strategic document (DSPN) produced by the DPS in November 2005 provides a survey of this experience. The document was compiled on behalf of the government economic planning committee (CIPE). "On the basis of guidelines established by the central government, regions and local authorities in agreement with the unified conference of Feb 2005 (according to Art.8 of Law 5/June 2003, n.131), Italy must establish before September 2006a National Strategic Framework (QSN) for the 2007-2013 period in order to programme spending for social and economic cohesion (EU Treaty, Art. 158-162) both for national funds and for regional policies (Fund for underutilized areas, Constitutional Charter Art. 119.5). Once the document is approved, the Community component must implement the indications established for 2007-13 cohesion policies and their relative strategic orientations established by the Council. In order to start writing the document, the CIPE deliberation n- 77/2005 implementing the guidelines established a coordination committee made up of 12 central government administrations coordinated by the DPS at the Ministry of the Economy in order to publish a technical and administrative report. This mandate was discharged in partnership with institutional representatives from local authorities and with the trades unions. The economic and social partnership, representing all the associations that had participated in the 2000-2006 round of programming and implementing the Community Support Framework for Objective 1 (Mezzogiorno) and Objective 3 (Social Fund for the areas of the Centre and the North), also drafted documents, discussed specific parts and participated in seminars on topics related to the DSPN. They

What we need to do, then, is to examine the incentives and behaviour of the various institutional levels, starting – for simplicity's sake – with only two levels: the centre (fund donors) and the periphery (local decentralised institutions). The aim of the central level in donating resources could be assumed to be development and social cohesion. And yet, if this is the ultimate aim, when there is tension between budget stringency and spending objectives (as there generally has been in most developing and advanced countries in recent years) there is urgent need for certainty regarding when and how much is to be spent. Limited financial resources and the pressure on public spending exerted by various interest groups make credible agreements necessary. Programming must therefore be well anchored in reality. A gap of months (or years) between budgeting and spending is not acceptable. A reasonable guarantee of certainty regarding when and how much is to be spent is needed in order to defend the resources destined for regional policies. If there is no such guarantee, temporal instability and quantitative uncertainty regarding transfers to regions and local institutions increase. In the political market, in fact, interest groups scattered over the various territories can easily take second place compared to concentrated interest groups, such as those in central administration or in 'vertical' associations and unions.

It is in the interest of donors to be certain when resources can be used, but also of beneficiaries because uncertain timing can jeopardise both the regularity of payments, and the overall amount available. The kind of joint payoff that can be gained through institutional cooperation, however, contrasts with the desire to keep local institutions

---

also took part in four collegiate meetings devoted to discussing the chapters organized jointly by the Ministry of the Economy and the Ministry of Labour and Social Policies (holder of the two funds, FESR and FSE). At the conclusion of these sessions, the resulting document was agreed upon, both in terms of content and in terms of methodology. Together with regional strategic documents drafted by individual regions, and the strategic document for the Mezzogiorno, written by the eight southern regions and the DPS, this document will be useful for establishing contact between the various levels of government, trades unions and associations which, according to the guide lines, will allow in the early months of 2006 an initial 'plan' for the national strategic framework to be draw up. "(DPS, 2005, p.II). "In Italy, in the current language of official documents, and therefore even in this document (DSPN), regional policies mean those development policies aimed at cohesion, redressing economic and social imbalances, and competitiveness in specific territories. In line with the indications of the EU Treaty and the Italian Constitution, the emphasis is thus on the intentionality and additionality of funds used for specific areas. In other European countries (UK, Austria, France, Germany, etc) where there is a similar conceptual and operative framework, this more modern form of policy is called territorial policy. Regional policy, therefore, does not coincide either with overall development policies (most of which are aimed at achieving national objectives, not for specific territories) or with the set of interventions for development implemented by Regions. Regions implement on their territory ordinary interventions which are not additional, wohle part of 'regional policy' can be implemented by the Centre. Regional policy in this document, to conclude, means those development policies that are added to ordinary action undertaken both by the Centre and by the Regions." (DPS, 2005, pp.II-III).

inactive. Later on we will illustrate some cases where explicit and cogent rules to accelerate spending were introduced after it became clear that the gap between earmarked funds and actual spending was getting wider.[165] The fact that the rules introduced were relatively effective shows that inactivity was not the result of local institutions being unable to spend, but rather of low motivation. We could call this an implicit incentive for inactivity on the part of local institutions receiving funds earmarked for development policies from the centre. This incentive could derive from the fact that local administrations have always distributed benefits to local interest groups, and any new policies aiming to bring about change and propel development in a backward area are likely to challenge this traditional form of cake sharing. On the other hand, donors generally impose programming and monitoring procedures on fund beneficiaries, and require funds to be spent on the basis of dialogue and consensus among the various interest groups.[166] Local interest groups, then, should prefer incisive policies that bring about change, but, at the same time, they are tied down by the need for the consensus of those who do not want things to change. Local governments often find it preferable to wait. The payoff of cooperation thus goes hand in hand with a payoff of defection. Local authorities can easily gain advantage by obtaining funds transferred from the centre, not paying for the organisation or learning costs required by the project, and waiting for the various local interest groups to come to some sort of agreement so that the benefits of the resources just made available can be shared out without conflict.

It is like a game of strategy, in a context where reiteration[167] is expected, but where, without explicit rules, there is a great deal of 'noise'- or rather, the constant

---

[165] Both in the EU choices regarding 1997-98 Structural Funds, and in those of the government regarding 2002 FAS spending (Funding for Under-Utilised Areas), the imperative soon became clear to move on from programming on the basis of financial commitments but on the basis of cash flow. The Commission stated this aim because it had always got the balance wrong between earmarked funds and actual spending, with the result that the European Parliament repeatedly made more resources available than were actually used. When the Italian government chose to contribute to the compliance of the Maastricht rules, by holding the level of capital expenditure through the use of EU resources, it had to find out from the General Accounting Office the precise distinction between domestic and EU sources.

[166] Not only is there a consensus constraint for local politics because they are more visible for single interest groups, but there is also a request for cooperation on the part of central institutions donating funds. The reason for this request is that central institutions do not want to take the blame for interventions that have not been agreed upon at a local level. The result of this blame would be gradual but widespread delegitimisation.

[167] In time but also in different territories, assuming realistically that local institutions watch and take into consideration the experience of others. This is thus more like a networked game of strategy. Lippert and Spagnolo (2004) show that cooperation among agents can be easy, but it is precisely the

risk of interpretative errors. These can easily come about when donors and beneficiaries do not know the conditions that would induce one party or the other to defect; for one party defection could be unavoidable, but the other party might interpret it as opportunistic. Both sides are perfectly aware that this can happen. In a similar context, Axelrod (1997, ch.1-2) shows that on the basis of adaptive behaviour, the traditionally efficient result (when a cooperative act is followed by a tit for tat action – cooperation for cooperation, and defection for defection) can be replaced by another strategy. This consists in repeated acts of defection in order to find out what the implicit rules are; or, rather, to see how far they can be forgiven, or how much they can delay action, without losing the funding. If various local institutions behave this way, to start with there is a widespread tendency to wait and see (local institutions do not spend, donors forgive them). Soon, however, the system of transfers begins to feel the crisis (as EU Structural Funds did, and as the Fund for Under-Utilised Areas did in Italy).

There are good reasons, then, for introducing explicit rules, such as the rule of 'automatic disengagement', $n+2$[168], in the case of EU Structural Funds, and, in the Italian case, the rules defined by the inter-ministerial programming body, CIPE, concerning the earmarking and use of resources for the Fund for Under-Utilised Areas (FAS).[169] These rules are reinforced with penalties, which generally eliminate those

---

conditions that allow this to take place give rise to a form of cooperation that does not always end up on the same side as the aims of the donors (Picci, 2004).

[168] In order to receive Structural Funds the Commission must certify beneficiaries' spending (Regulation 1260/99, art.9). This certification consists in a series of verifications (admissibility, conformity, timing) conducted by the Certification Authority. The $n+2$ rule intervenes in the sense that at the end of this process, the member States ask for resources which are financed to the extent that the temporal limit between earmarking funds and actual spending does not exceed 2 years and 8 months (funds earmarked April 30, 2001, are automatically disengaged if they are not spent by December 31, 2003, etc). Resources that do not comply with this procedure are automatically withdrawn. They are not transferred from one year to the next; they are taken out of the budget and result in a lower EU budget and a lower level of contributions by the States to the overall EU budget.

[169] With reference to the CIPE deliberation n.20/2004 (the system was changed after the introduction of a disengagement clause for the first time in deliberation 36/2004), the budget assigned every year a certain amount to FAS, which substitutes law 208/1998, in order to finance material and immaterial infrastructures and to use as incentives. CIPE allocated these funds to various objectives and different tools (incentives, infrastructure, regional programmes, agreements and specific projects). With a later deliberation, the government body distributed regional funds, according to consolidated practice, but this time it spelt out clearly the way the funds should be used (General and Operative Agreements) as well as establishing rules. These rules include the fact that FAS resources must be engaged year by year, by means of legally binding acts, by the final beneficiaries within a certain period of time. The government programming committee, CIPE, will reprogram any resources not used. The rule, established every year, has so far had the following deadlines: resources financed by deliberation 36/2003 must be spent by December 31, 2004; by 17/2003 by December 31, 2005; for 20/2004 by December 31, 2007.

who have disregarded them from their right to use resources that have already been assigned. Results have been consistent. In the case of FAS, for example, from 1999-2001 the share of resources[170] actually spent out of those available was about 30%, while from 2002 onwards the share increased dramatically (together with a hefty increase in total resources made available): 48% in 2002, 77% in 2003, 95% in 2004, 97% in 2005 (DPS data).

Simple rules

Explicit rules introduced in order to avoid inactivity in decentralised institutions must be simple, like the ones we have given as examples. This consideration is important because, as we shall see, even though they are necessarily simple, there are still difficulties in programming, spending and implementing effective development policies. What happens is that, in order to avert one risk – local inactivity that jeopardises the regularity of funding as well as the overall quantity of resources to be made available - further risks are generated owing to the poor quality of the interventions. On the other hand, as we have seen, leaving the distribution mechanisms without any rules does not avert these risks either. Consequently, it becomes necessary to identify specific measures to avert or reduce the effects of these risks.

First, however, let us examine why rules, such as 'automatic disengagement', need to be simple. Simple rules are easy to communicate; beneficiaries know what donors require of them. They also know the time-line that indicates when a certain action must be taken. Similarly, they can appreciate the kind of behaviour that is expected of them. Rules are simple if, and only if, those who have to follow them can answer the following questions without thinking: 'What do they expect us to do?' and 'What reaction will they have if we behave differently?' Simple rules, which must also be inflexible, are required when there is contractual incompleteness; that is, when it is not possible to know and specify in advance all the possible contingencies that may take place after the event. In a situation of this kind, rules that were complex, and therefore complete, would be inapplicable because they could never specify every possible case and the rules' enforcers would be lost if any situations came up that had not been envisaged. Put in another way, implementing complex rules in circumstances

---

[170] The 'final beneficiaries' by means of legally binding acts programmed this share.

of contractual incompleteness would give rise to exorbitant costs as well as to constant legal challenges. Simple rules – in the sense that they are not complex and not rigid – can be written down with a limited number of phrases that contain 'principles' and refer to figures that can be identified and quantified in an unambiguous manner (Glaeser, Shleifer, 2001). In the real examples we are citing, for example, concepts such as 'the state of advancement of works', or 'legally binding commitment' are devised and the rule is then written, with precision but without aiming for completeness. If the CIPE deliberation had tried to define the concept of 'legally binding commitment' in all its possible specifications, it would have been obliged to write down n of the $n+m$ possible cases, and it would therefore have been impossible to enforce.

Another requisite of simple rules becomes clear when we consider not just contractual incompleteness but also asymmetric information. People who set and enforce rules generally lack the information that those who have to obey the rules hold. If we imagine a situation where rules are not simple (without an aggregate financial objective), but where they aim to list specific results and create incentives for achieving them. If some results have incentives (the ones that can be identified and checked), other objectives will be sacrificed (usually the ones that are less easily identified and checked by donors). Take a more general example: when quantity and quality becomes the object of incentives. If quality is hard to measure, and quantity is the object of incentives, quantity can easily override quality. What is needed is a system of equal incentives in order to avoid the danger of certain objectives becoming more important than others. If it is not possible to include all the objectives in the rules, then it is best not to try and put in just a few. If it is impossible to measure all the objectives, if, indeed, it is hard even to define them, but they are considered anyway important aims, then a system of articulated rules risks becoming counter-productive (Milgrom, Roberts, 1992). Simple rules are preferable, in the sense that they should have a composite objective. Nor are simple rules naïve because they limit themselves, for example, to financial measures and deadlines.

To sum up so far: simple rules aiming to avert inactivity and to maintain the flow of resources, should be rigid and non-complex, as far as contractual incompleteness of concerned; non-articulated and composite, as far as asymmetry of information is concerned. The merit of this simplicity is that rules that observe these conditions are credible and appreciated. They are credible because they are linked to an objective in

a transparent fashion. They are appreciated because their mechanisms are easy to recognise and comprehend. Simple rules are also precise. That is, they do not allow room for interpretation; or rather, as we have seen, they do require interpretation but they are less likely to be challenged than complex rules with a list of circumstances that is by definition incomplete.

If we wanted to be more analytical about the credibility of simple rules we could introduce a further specification. We have seen that inactivity itself can be considered advantageous by those who should take action, and therefore disadvantageous to those who provide funds. We have also seen that in circumstances where there is asymmetry of information, those who hold information could have a stake in taking action against those who do not. Under these circumstances, there is a potential conflict of interest among those who should cooperate in order to achieve results that are useful for all sides. Where there is potential conflict, then, there is all the more need for credibility so that the rules create incentives to cooperate. In this case, credibility takes on the specific meaning of being capable of actually influencing behaviour. This connotation of credibility becomes strategically important because the realignment of plans that are potentially unaligned depends on it. To finish the analysis, then, the capacity to actually influence behaviour generally requires certain kinds of behaviour to be pre-determined (Kreps, 1990, pp.90-143), with the result that a certain amount of discretionary action is sacrificed. With this in mind, we can now add that rules must be non-renegotiable, and for this reason all the more simple.

In conclusion, we have made an exhaustive list of reasons why rules must be simple, and why rules must be: I) easy to communicate and clear as to when they must be fulfilled; II) non-complex and inflexible (rigid); III) non-articulated according to objectives but based on a composite aim; IV) credible in the sense of being non-renegotiable, and therefore particularly simple.

Simple rules with difficult final results

Rules of this kind, however, for the same reasons as above (contractual incompleteness and asymmetry of information) may be able to counteract inactivity on the part of beneficiaries of funds, but they can play no role in guaranteeing that resources are well spent.

For example, with reference to Italy, the experience of territorial development policies has been very mixed. The idea, generally shared by all those involved in local

development, is that economic development requires the intentional mobilisation of forces disseminated throughout the territory, and of a great number of people, institutions and organisation that share the firm intention of acting for the collective good. Mobilisation on this scale does not take place spontaneously either on the market or as a result of central policies.[171] In order to come up with a definition of local policies that encourage development, we must thus include three considerations: decentralisation, diffuse involvement, and an inclination towards public and collective goods. Taking these three points into account, ten years of local development experience in Italy have been successful as far as decentralisation and participation is concerned, but less successful and decidedly more problematic regarding involvement and inclination towards public and collective goods (Cersosimo, Wolleb, 2006).

Development policies contained in the 'New Programming' adopted decision-making procedures based on negotiation that increased dramatically the number of participants compared to previous experiments with central government policies, especially those concerning the Mezzogiorno.[172] It cannot be denied that in certain circumstances there have been manipulations, partial inclusions, collusive coalitions, and negative externalities. In general, nonetheless, the procedures adopted have allowed different positions and interests to be aired, with a gradual alignment of the preferences of individual actors in the territory, leading to widely shared decisions. Negotiation has not been excessively manipulated and has been relatively inclusive, leading to collective decisions without overt conflict and free of negative externalities. Negotiation has also brought about marked improvement in the social and institutional quality of a territory. It has allowed tighter networks to be formed among institutional bodies, and created a more fruitful relationship between

---

[171] This idea is at the heart of the article by Fabrizio Barca (2005) who was DPS Department head. The article is an important testimony to an experience of local development policy in Italy. The idea is also supported by numerous theoretical and empirical results, in particular by those that stress the obstacles to development presented by market incompleteness and by agglomeration processes in mobile resources, as well as by those underlining the complementarity between accumulation of human capital, innovation and accessibility of immobile resources that requires public and collective goods to be earmarked and disseminated.

[172] There have been a number of analyses of previous development policies. An overall view would have to take into account five essential features: I) substitute and non-additional interventions compared to ordinary territorial policies; II)uncertain rules that change with little transparency; III) central policies with no responsibilities at regional level; IV) corruption and little administrative capability; V) only one dominant theory – development poles based on external investments. The 'New Programming' envisaged, by contrast, symmetry in: I) additional interventions compared to ordinary investments; II) certainty of rules; III) responsibility for local governments; IV) administrative capacity building; V) attention to endogenous development mechanisms. This list was suggested by Paola Casavola.

institutions and civil society organisations. It has, moreover, encouraged capability building and passed on skills in the field of development policy making, generating new leadership and at least a partial renewal of the managerial class. These quite impressive results, of course, are not the same everywhere, but all in all, the negotiation method embodied in the local development policies has had the effect of increasing participation, as well as social and institutional cohesion.

One problem, however, was the actors involved. Local institutions were always more dynamic than civil society, which either played a very partial role, or limited itself to consultancy. Another problem was that the rules laid the necessary groundwork, but did not go far enough to ensure success. The idea that there was a direct link between greater democracy and the quality of institutions on which development depends through the production of public goods has turned out to be simplistic.[173] If we overlook the procedural and methodological aspects of negotiation, it is clear that this approach has not brought about the expected results; in short, not enough local public goods have been established. Moreover, various analyses of Territorial Pacts, and then later of Integrated Territorial Plans, have shown that most local projects were small, lacklustre, fragmented, and based excessively on aid for firms (Cersosimo, Wolleb, 2006).

The next step in examining the way simple rules work, and analysing why results are not entirely satisfactory, is to introduce an element of realism into our argument. As we have already said, in most experiences of decentralised institutional frameworks, there are more than two levels. Very rarely is there a simple dualistic structure pitting the centre against the periphery, however. There is nearly always an intermediate level, which we could call 'regional', that interacts both with the centre and with the periphery. Rules regarding spending deadlines are typically negotiated

---

[173] The *Strategic Document for the Mezzogiorno 2006-13* reads: "The expectation of achieving early results in terms of increased growth, that the Strategic Framework (QCS) certainly envisaged, was linked to an idea that the policies (acting with determination in a general context in which citizens live and firms work and promoting direct actions to valorize existing comparative advantages) were so solid and credible as to persuade the private sector – that at the time seemed to be on an upward curve – to anticipate development (and thus 'break' with the unsatisfactory economic tradition). Recent history has shown that this virtuous circle did not take place. During mid-term revisions for the Strategic Framework (QCS), which made a similar diagnosis, underlined that the failed achievement of this 'break' with tradition was owing to three factors in particular. First, the potential of the private sector in such disadvantaged conditions was probably over-estimated. Second, to start with it was impossible to gauge how unprepared the administrations were to implement the policies. Third, the cultural and political confrontation at different levels was not strong enough to uphold the overall framework and apply the policies." DPS

between the centre ad regions, while their implementation depends on the concrete actions of the various local sub-regional governments. Ideally, then, what we have said about the advantages of simple rules would be confirmed only if there were perfect information between regional and local governments, and if regional governments thus considered themselves part and parcel with local administrations in their dealings with central government. If this were so, there would indeed be a dualistic structure.

If, on the other hand, there is asymmetric or no information in central- regional dealings, compounded by asymmetric or no information in regional-local dealings, then things get more complicated. Regional governments can find themselves in a position where they have to form an 'alliance' either with central government or with local administrations.[174] This rarely takes the form of a rational choice on the basis of long-term payoffs. With imperfect information, the de facto coalitions that emerge will do so on the basis of other considerations, having to do with short-term advantages and relying on a playing field which is anything but level and where the proximity of relationships counts more than anything.[175] Axelrod (1997, ch.4) shows that the choice of alliance is usually determined in correspondence with a minimum local level of "energy" (the sum of the frustration of the actors involved, who are unhappy about having chosen certain allies rather than others). There can be any number of these minimum local levels, and different alliances do not necessarily establish themselves in a global minimum. Moreover, a minimum local level that depends on its starting conditions is not easy to escape from.

This technical language is easy to translate: a choice of alliance underlines where you want to be considered friendly or obsequious and where you want to be seen as intransigent and unforgiving. In a concrete case, every institutional act in the three levels of the institutional framework always has two sides to deal with. Central government can decide to ally itself with the regions 'against' local authorities, or vice versa; regions can decide to ally themselves with central or local governments;

---

[174] The hypothesis we have assumed is that it is frequently the case that one institutional level finds it hard to agree spontaneously and easily with the other levels. This is coherent with the assumption regarding the scarcity of resources and incentives for inactivity, both of which create conflicting interests among different levels.

[175] By relational proximity we mean, among other things, a proximity in size (which is why central governments find it easier to deal with regions than with individual local realities, and vice versa), organizational proximity (again to the advantage probably of central-regional dealings), the frequency of dealings, similarity of rhetoric, similarity in decentralization process.

local authorities can choose an alliance with central or regional governments. There are a multiplicity of combinations, taking into account that central government comprises many different ministries and departments, and that there are in Italy 21 regions, 110 provinces and 8,100 local councils. Some of these combinations are stable, others are instable, depending on the level of discontent that the institutional actors express, either because their overtures have not been reciprocated, or because the alliance they have formed has not been particularly fruitful. Axelrod acknowledges that there could be many stable structures (expressing minimum general dissatisfaction), and that an optimal structure (with the lowest levels of dissatisfaction) may not even exist. Many unsatisfactory structures are surrounded by even less satisfactory ones, making them anyway the 'best' local structure, from which it is hard to escape.

Technically, regional governments can decide an alliance with central administrations and negotiate rules concerning spending deadlines without any discussion, and without taking on board the difficulties posed by these rules for local authorities. The outcome of these rules can only be their formal appliance, with regions leaning on local authorities to ensure they spend resources within the deadlines without considering quality. Adopting behavioural rules that have been agreed upon ex ante does not eliminate the root causes for which the rules were written. Rules usually transform these root causes into costs that need to be sustained. Costs, in this case, fall on local governments, forcing them to make the organisational investments needed in order to spend the resources both well and within the time limits. In view of the size of these costs, then, implicit or explicit agreements can come about between regions and local governments accelerating spending at the cost of sacrificing the quality of the interventions that could be useful for development.

On the other hand, different alliances produce different costs, which are distributed differently. For example, a hypothetical alliance between regions and local authorities would presumably lead to less simple rules, with more controversial applications, which are more subject to re-negotiation. In this case, central government would be liable to have to cover the costs, and it would consequently offer regions an explicit or implicit pact allowing them to make less regular or lower payments, in particular by means of a higher regional co-financing.

Transforming difficulties deriving from imperfect or asymmetric information into costs by means of rules presents another advantage, other than simply allowing things

that seemed impossible to get done, even if it is costly. This additional advantage can be evidenced more clearly if we adopt a more dynamic approach than the static one we have followed so far. From a dynamic point of view, supporting costs that are produced by rules allows learning to take place. In the institutional framework there is ample room and a great need for different players at different levels to learn as they go along.

Learning processes would be supported if accurate monitoring and evaluation systems were able to pinpoint who or what is responsible for achieving good results when the rules are followed (as well as who or what is responsible for bad results). A system of ex post controls on the achievement of final objectives would also be a good idea, not so much to condition the assignation of resources (making the rules less simple), but more to feed into the learning process. This would require not only an efficient evaluation system (based on the identification of minimum standards and premiality indicators), but also a clear idea of who is responsible for what tasks.

If learning processes were indeed encouraged by means of these mechanisms, there could be the risk that they become cumulative and therefore increase variation in results. Specific measures are needed to level the playing field all over the territory. In this respect, collaboration between the different levels of the institutional framework is particularly important, as is the presence of institutional leaders who are able to operate at different levels. All these elements (a clear institutional design with clearly identified tasks and responsibilities, monitoring and evaluation supported by minimum standards[176] as well as premiality indicators, leaders who can operate at all levels) would allow decisive steps to be taken towards the ultimate aim of rules: to make sure resources are available, and that they are protected not only from inactivity but also from being spent badly.

Costs and benefits of learning

Simple rules in order to guarantee regular spending within well-defined time frames, guaranteed costs and a facilitated learning curve are three ingredients of a

---

[176] Particular attention must be paid here because fixing minimum standards (for example for the functioning of public services financed with transfers) can transform simple rules into complex ones. There is a certain amount of experience in this field that seems to show that it is possible, though the path is narrow, to preserve both simple rules and minimum standards as long as there are very few, they are agreed upon, precisely delineated and come with a clearly identified verification method.

package that can be represented by a simple model. The model should help us understand why it is important to have a clear institutional design, monitoring and evaluation, minimum standards, and professional figures who are able to work at all levels. But first we must define three types of local intervention that could schematically represent a whole range of cases.

The first type (A) is an intervention that encourages development and that relies to a significant extent on public and collective goods. For this very reason, its results fall into one of the following four patterns. I) the benefits of A are distributed asymmetrically among local interest groups but while the collective benefits of A are higher than the collective benefits of B (the second type of intervention), the benefits of A for each specific interest group are not greater than those of B − or rather, the benefits of A are higher than those of B for the whole collectivity but to a different extent for each group; II) the benefits of A are greater than those of B but they are postponed in time; III) the benefits of A are less easy to appropriate, and therefore less easy to spend politically than those of B (with intervention B the link between action undertaken, good produced and benefit enjoyed is easier for citizens to perceive); IV) the benefits of A are greater than those of B but in the eyes of local actors they seem particularly uncertain − or rather, the benefits of A are greater than those of B but most local actors do not know it.

The second type (B) is an intervention that does not contribute much to development prospects in the territory, does not rely to a significant extent on public and collective goods. Its results fall into the following four patterns: I) the collective benefits of B are lower than those of A, but individual benefits are distributed fairly symmetrically; II) the benefits of B are lower than those of A, but they are immediate or almost immediate; the benefits of B are easier to appropriate, and therefore the link between an intervention and its utility is easier to perceive; IV) the benefits of B are seen as being less uncertain than those of A.

The third type (C) is an intervention that fails to be effective, wastes resources or leaves resources unused. This kind of intervention produces greater collective damage than good.

As we have seen, a decentralised institutional framework, with rules in place to protect transferred resources and impose guaranteed time frames on spending, usually makes it possible to avoid the results of type C, but often at the cost of limited results of type A. Let us assume that a decentralised level of administration formulates

'programmes' comprising many different interventions. These programmes are often required by the planning procedures incorporated at the central level. It is also advisable for local administrations to formulate them, given the lack of public goods and services. In addition, a programme comprising many different interventions can also be met with greater consensus by the different local interest groups.

Let:

c = "standard" degree of efficiency;

b = parameter b>0;

e = degree of average efficiency in development for as *set of local interventions* assuming average interventions of type C in correspondence with values (c-r)>e≥0, of type B with values (c+r)>e≥(c-r), of type A with values c+b≥e≥(c+r), with a small r anyway definitely b>r>0. We also assume that type C interventions cannot be funded, and that the actual operative interval is c+b≥e≥(c-r);

q = net project cost;

q* = net cost of standard project (defined as one with efficiency e=c), as given average cost considered.

s = variable indicating relative net project cost, as ratio q/q*;

The gross cost of the project includes the whole package of costs sustained by the territory in terms of organisation, administration, information, control for every single intervention, its approval, its funding and its implementation. By net cost we mean the gross cost minus the lower costs owing to the learning process. By learning process we mean the learning of skills and capabilities able to formulate and conduct interventions that allow administrations to avoid, at least in part, the total burden of costs for each intervention.

$$e = c + b(1-s) \qquad\qquad [26]$$

Equation [26] indicates the assumption that the degree of average efficiency in interventions, considering a continuous variable[177], is decided so that:

$$e - c = b(1-s) \qquad [26b]$$

It is, therefore, equal to c when s =1, better $[(e - c)>0]$ when s < 1, and worse $[(e - c)<0]$ when s > 1, since s is the net cost of relative project building.

Equation [26] thus assumes that interventions will on average be of the standard type, when the project costs of those who decide are the same as the costs of standard interventions, and that interventions will be chosen with greater efficiency for development when project costs are lower and with lower efficiency for development when project costs are higher than the costs of standard interventions.

Let us assume that formulating, presenting, funding and implementing a set of interventions (a programme) takes time, which can only be defined as a 'time span for project building'. Along this time scale, we can now define the following dynamic variables:

Let:

a = the learning rate, or the rhythm at which skills and capabilities are accumulated;

f, d = parameter >0

$$a = f + de \qquad [27]$$

Equation [27] indicates the assumption, for linear simplicity, that the learning rate is an increasing function of the degree of efficiency of the intervention. The hypothesis is that every single intervention of type A gives rise to a significant amount of learning along the time span for project building, and therefore generates a high learning rate when the different interventions are on average of type A. This means that it is more difficult to reconcile the standard rules and procedures with the effectiveness of interventions for development, and therefore that more rapid learning of skills and capabilities takes place when interventions are more difficult. With

---

[177] From equation [26] and its discussion it is clear that the average efficiency depends on the cost of the project as well as on other parameters that are here considered non- variable.

simpler and less efficient interventions, less learning takes place for each intervention, and the learning rate is thus lower.

Finally, let:

p = rate of variation of the gross cost of the project on the time span of project building;

m = parameter >0

n = parameter >0.

Parameter n is important in interpreting results, so let us anticipate that it is the parameter on which the progress of the gross cost of project building in proportion to the learning rate depends;

q = rate of variation of the net cost of the project (derived logarithmically from the variable q) defined as the difference between p and a, or $q = p - a$

$q*$ = rate of variation of the net cost of a standard project.

$$p = \text{m} + \text{n}a \qquad\qquad [28]$$

Equation [28] indicates the assumption that the gross cost of project building rises according to the learning rate, but that learning reduces the net costs of the project, which is equal to the increase in gross cost minus the gross cost of learning.

A solution for the system of equations could be given for ds/dt=$\dot{s}$.

$$\dot{s} = \text{bd}(1\text{-n})(s^2\text{-s}) \qquad\qquad [29]^{178}$$

This indicates the dynamics for the net project cost in relation to the standard net cost as the various interventions are identified and defined along the time span of project building.

From equation [26] we can also obtain:

$$\dot{s} = -\frac{\dot{e}}{b} \qquad\qquad [30]$$

---

[178] Assuming that the proportional variation of the net cost of average planning for a standard intervention takes place according to the same mechanism as the one indicated in equation [28] for every given a.

Finally:

$$\dot{e} = d(n-1)(c-e)(c-e+b) \qquad\qquad [31]$$

Equation [31] indicates the dynamics for the average degree of efficiency of interventions along the time span of project building. To the right of the equals sign, this is a parabola that possesses zeros in correspondence with values e=c and e=c+b, the direction of whose concavity is regulated by the value of (n), according to whether it is greater or less than one.

Remembering that (n) is the parameter on which the movement of gross costs of project building in proportion to the learning rate depends, for a limited learning cost (n<1) the parabola cuts the horizontal axis of zeros from above.

Figure 30 - The dynamics of average efficiency (e), with limited learning costs.
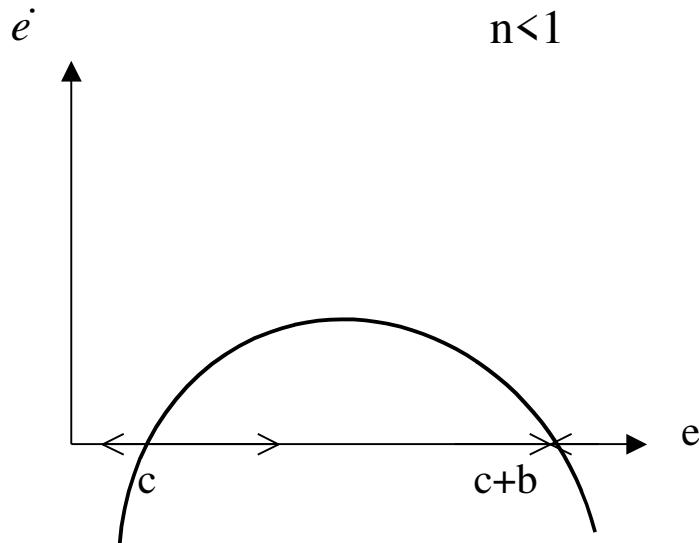
In this case the point of equilibrium e=c is unstable, while the point of equilibrium e=c+b.

In the other case, with high learning costs (n>1), the parabola faces the other way.

Figure 31. The dynamics of average efficiency (e) with high learning costs
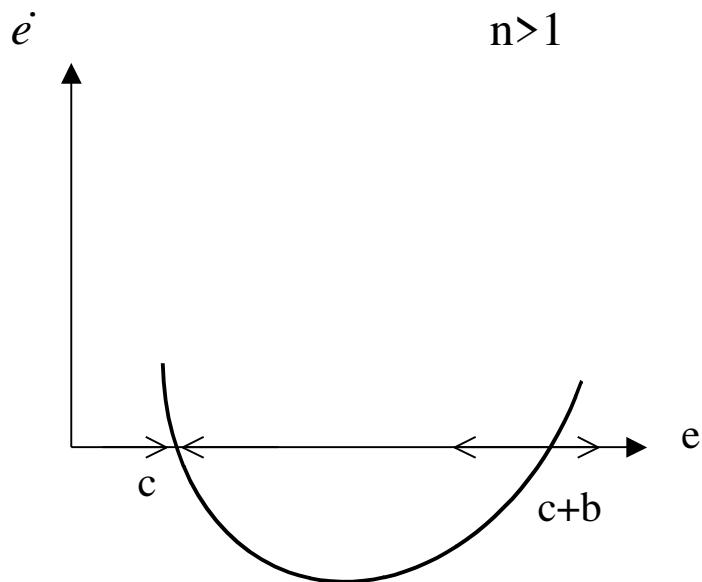
Now the solution with a stable equilibrium lies in e=c. This result leads to the consequences we shall now examine.

Limited learning costs

If learning costs are limited within the institutional system, the project that comes about comprises interventions mostly of type A in areas where project-building capacity is already higher than the norm. The learning rate is high, overcoming learning costs and allowing constant improvements to be made. In areas, by contrast, where project-building capacity is lower than the norm (even when they are in the admissible interval in correspondence with c+b≥e≥c-r), the learning rate is slower and costs are also be reduced, but the losses owing to low learning levels are greater than the gains arising from reduced costs. The result is continuous worsening of quality. The movement of e (in the case included between c and c+b) is towards improvement.

High learning costs

If learning costs are high, in areas with superior project building capacities, interventions are mostly of type B, unlike in the previous case, because in the first interventions considered (of type A) learning is fast but learning costs are even higher, with the result that there is a worsening of quality of interventions (except in

cases that can realistically be excluded where project-building capacity is very high to start with). Even in areas where project-building capacity is lower than the norm, the movement is towards improvement, or towards interventions of type B. In these circumstances, there is not very much learning in the first interventions considered but learning costs are even lower, with the result that the quality of interventions can be improved. The movement of e (in the case included between c and c+b) is towards worsening.

Learning, as we have seen, means accumulating skills and capabilities. Learning costs are high when obstacles and inertias that require spending to overcome hinder this kind of accumulation. Usually these hindrances are generated by a lack of information, which becomes more damaging, the greater the difference or distance between available information at each institutional level. We can therefore imagine that learning costs depend on a lack of transparency at every institutional level and at every stage of the process, from budgeting, to distributing, to actually implementing resources for local development. When there is no information, interpretation is necessary, as is trial and error, creating obstacles and inertias that hinder the accumulation of skills and capabilities. The model produces a result that allows us to explain why most local projects are oriented towards interventions of type B if the institutional framework, rules and practice bring about high learning costs or a lack of transparency. With high costs, there are incentives to homogenize behaviour by adopting interventions of type B, and thus to reduce variance at the cost of a possible lack of efficiency. In this case, local governments have to submit to strict rules regarding time limits for spending, which are established by central governments in agreement with regions, but which do not take into account local difficulties. These local difficulties, in other words, remain unexpressed and are not known by regions, not to mention by central government (denoting a lack of transparency). The difficulties arise from the political sustainability, or lack of same, of specific initiatives. If these projects are good, they bring results that are uncertain and postponed in time. Moreover, their benefits are distributed unevenly among local interest groups in contrast with consolidated patterns of distribution. Under these circumstances, local governments, whether they possess good project-building capabilities or not, risk adopting projects that scatter funds indiscriminately to please everyone just a little, rather than realizing effective development strategies. They also risk being captured by dominant interest groups resisting change. Regional

governments, on the other hand, are 'allied' with the centre, as we have seen, and tend to keep their distance, encouraging this kind of behaviour.

If rules and procedures are more transparent, on the other hand, if they are designed to take account of difficulties at every level, and if there are limited learning costs, then the results are more articulated: better in some areas (with interventions of type A) and much worse in other areas. In this case, local governments can more easily follow the rules regarding time limits for spending because the shared programming process comprises mechanisms that help them achieve their aims (if they want to). What counts, at this point, is the starting point and the first steps. If local governments possess good project-building and administrative capabilities, and if they set off down the path undertaking interventions that are useful for development, they create a virtuous circle in which the learning process allows increasingly efficient interventions to take place. By contrast, incapable local and regional governments that take on inefficient interventions generate a vicious circle in which the opposite takes place.

There is thus a trade-off between unexceptional but homogeneous results and enhanced results that are more heterogeneous, which depends on the assumption that it is important to learn by doing. As far as the question of what measures to adopt is concerned, the suggestion is to contain learning costs by making the institutional framework more transparent, and to ensure that there is a more level playing field for all parties, especially in terms of project-building capabilities. We could say, in conclusion, that in order to avoid local policies that tend towards interventions of type B, which are not very useful for development, it is necessary to reduce learning costs by increasing the transparency of the institutional framework. However, if this is achieved, there will be a tendency towards a great variety of behaviour, only some of which is virtuous. At the same time, then, it is necessary to promote and sustain the project-building and organisational capabilities of the weaker local institutions.

Designing institutional arrangements

Let us go on to analyse how to reduce learning costs by introducing transparency in the institutional framework. In the international literature on decentralising development policies, the focus is on the design of the institutional framework, within which the tasks and responsibilities should be clearly delineated at every level following the principle of subsidiarity (see, for example, Shah, 2003). Our analysis

confirms this view, as do the available case studies (Jütting, Kauffman, McDonnel, Osterreider, Pinaud, Wagner, 2004, pp.17-29)- Clear tasks and responsibilities encourage equally clear indications from monitoring and evaluation systems; these, in their turn, encourage learning at all levels (Shah, 1998). Since the institutional design of decentralisation is necessarily complex[179], each institutional level should have specific tasks, which should be kept distinct from other levels. For each level, the responsibilities should be clearly assigned. The central level should have the overall tasks of programming and finding the necessary resources, and be responsible for ensuring these resources are both guaranteed and supplied regularly. The regional level should have the task of territorial programming and should guarantee that the rules dictated by the centre are kept to by interacting with the local level. The local level should have the task of building projects and implementing them within the time limits and according to established guidelines; they should also ensure these projects are effective and useful for development in agreement with the regional level.

From this analysis, it soon becomes clear that there are elements that go beyond this idea of specialisation at different institutional levels. In conditions of incomplete and asymmetric information at all levels, simple rules, which should be dictated by the central level in order to ensure regular resources within guaranteed spending time limits, there is a risk that learning costs will be high even in an institutional framework where tasks and responsibilities are clearly identified. The crucial question is thus one of inter-institutional collaboration. It is essential that skills and roles at one institutional level do not conflict or overlap with those of other levels, and that who has to decide what and how is established before the projects are assigned, and with great precision. It is equally important, however, that different levels coordinate their activities and integrate their functions. Different institutions should bring about forms of cooperation that encourage knowledge to be shared, administrations to adapt, and each level to be a little less self-referential (Barca, 2005; Sheuerman, 2004; Stato e Mercato. 2005) The problem is that protocols, agreements and negotiation tables are never enough. Not only is knowledge often concentrated in some areas or institutions, but also the kind of knowledge we are talking about – whether it is codified or

---

[179] For the following reasons: I) it is better if funding for local development policies is partly from within the local system and mostly from outside; II) local authorities should be able to exercise discretion when spending resources, and yet be guided; III) collective goods and services are being established, and these are complex projects to build and implement.

uncodified – is the personal heritage of individuals. Taking this consideration hand in hand with the fact that designing development policies at all levels requires the exercise of margins of discretionality - if not the expression of 'institutional entrepreneurship[180]'- we can translate the necessity for inter-institutional collaboration into a recommendation to train and support experienced institutional leaders who are able to take on authoritative roles at different levels simultaneously.[181]

A third recommendation (in addition to clear institutional design and multi-level leaders) regards an evaluation system. Evaluating local development policies is difficult because many different actors with very different missions are involved and results are not easy to measure. And yet evaluation is critical. Evaluation *ex ante* allows alternative development programmes to be selected, with a consequent reduction in the failure rate. Evaluation *in itinere* allows shifts from the original plan to be verified, objectives to be re-assessed and actors to re-align themselves. Evaluation *ex post* allows the gap between aims and results to be measured, so that important lessons can be learnt for the future of the policy or similar policies and the results of different ways of implementing the same policy in dissimilar territories can be compared.

In order to reinforce the evaluation system - once the institutional framework is clear and multi-level institutional leaders are successfully operating - premiality systems can be put in place after establishing standards for public goods and services. This issue is particularly sensitive since, as we said in section 2 (Simple rules), with imperfect information it is counter-productive to complicate rules for assigning resources. Premiality systems that link additional resources for local authorities to the achievement of targets regarding the availability of goods and services[182] for the

---

[180] In the Italian experience of local development policies, especially in successful cases, out of a group of local actors a particular individual or group emerges with greater power than others. These individuals or groups were able to make decisions against the will of others, solve conflicts even without consensus, and pursue innovative solutions before they had been identified by the majority (Cersosimo, Wolleb, 2006).

[181] This consideration, suggested by Anna Natali in the light of her experience at the DPS, is highly relevant for us. There are a number of theoretical works that evidence that the problems of inter-institutional cooperation dealt with by means of game theory (Dahlby, 1996; Dixit, Londregan, 1998; Broadway, Marchand, Vigneault, 1998) either end up with assumptions that are distant from reality or achieve sub-optimal results (Bardhan, Mookherjee, 1999; Azfar, Kähkönen, Lanyi, Meagher, Rutherford, 1999) because they do not perceive the presence or role of a multi-level actor.

[182] For example, drinking water or a regular domestic electricity supply, functioning hospitals or day clinics, good teaching in schools, etc.

citizenry could have the negative effect of further complicating rules and thus making them inapplicable and less credible. If, on the other hand, there is a clear institutional framework, or one that is well on its way to being defined, as well as a functioning evaluation system and effective multi-level institutional leaders, then the lack of information is likely to be reduced, or on its way to being reduced. If this were the case, rules could evolve on the basis of experience, becoming gradually less complicated (Ahmad, Martinez, 2004).

Standards for goods and services should preferably be identified by means of inter-institutional agreements, rather than being top-down, so that there would only be a limited number of them, and priorities would be clear (for example, water, health, education, etc). Another advantage would be that the established standards would be high enough to create incentives for local authorities to commit to, but at the same time low enough to be credible and achievable (Martinez, Vazquez, 2001; Alm, Martinez, Vazquez, 2002). Finally, premiality systems and established standards are preferable to making the programming guidelines stricter for regional and local governments, for example by tightening up the list of sectors or compartments that are eligible for funds, or making spending more difficult (Barca, 2004). Empirical research bears this out. Majumdar, Marcus, 2001, stress that it is useful to adopt rules that indicate objectives, leaving decisions as to how to achieve these objectives to the protagonists, rather than imposing behaviour on them, because this reduces their innovative effort.

Finally, the project-building capabilities of local institutions should be reinforced and their project implementation sustained, especially in lagging areas. Aside from wide experience in technical assistance, both internationally and abroad, which needs evaluating, the UN has also had some interesting experience.[183] The UN Fund of 40 million dollars a year created to encourage microfinance is less than a thousandth of the Official Development Assistance (ODA) funds of 50 billion dollars per annum (UNCDF, 2002). With this limited budget, the Fund's administrators had a simple but

---

[183] There is not room here to list all the examples of institutional capacity building, including in the World Bank. Concrete experience, however, has not really been verified (Gillespie, 2003); there is a well-founded suspicion that in this field the significant amount of money that has been spent has benefited the donors rather than the recipients of funds. The case of UNCDF is different, mostly because of low level of funds.

perceptive idea: to devote most of their work to replicating 'best practice'.[184] The UN Capital Development Fund (UNCDF) compiled a manual to help its functionaries share their experience, and to encourage other institutions involved in international cooperation to describe the technical assistance they were able to give. Once these promising, clear and replicable experiences have been identified, the UNCDF uses its funds to spread the word and disseminate best practice. This method needs further analysis to see whether this is indeed the right path to encourage development.

Bibliography for chapter 5

Ahmad, E., L. Martinez (2004), *On the Design and Effectiveness of Targeted Expenditure Programs*, Georgia State University, Andrew Joung School of Policy Studies, Working Paper 04-18.

Alm, J. e J. Martinez-Vazquez (2002), *On the Use of Budgetary Norms as a Tool for Fiscal Management*, Georgia State University, Andrew Joung School of Policy Studies, Working Paper 02-15.

Azfar, O., S. Kähkönen, A. Lanyi, P. Meagher, D. Rutherford (1999), *Decentralization, Governance, and Public Services: the Impact of Institutional Arrangements – A Review of the Literature*, IRIS Centre, University of Maryland, College Park, September.

Axelrod, R. (1997), *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton Studies in Complexity – Princeton University Press.

Barca, F (2004), *Measuring for Decision Making: Soft and Hard Use of Indicators in Regional Development Policies*, Paper presented at "Statistics Knowledge and Policies - OCDE World Forum on Key Indicators" – Palermo, Italy, 10-13 November 2004.

Barca, F. (2005), *Spunti in tema di ritardo di competitività e politica di sviluppo nelle diverse Italie*, DPS, mimeo, July; now revised in F. Barca, *Italia frenata – Paradossi e lezioni della politica per lo sviluppo*, Donzelli, Roma 2006.

Bardhan, P. e D. Mookherjee (1998), *Expenditures Decentralization and the Delivery of Public Services in Developing Countries*, IED Discussion Paper, Boston University, Boston.

Boadway, R., M. Marchand, M. Vigneault (1998), *Decentralization of the Socialist State: Intergovernmental Finance in Transition Economies*, The World Bank, Washington D.C..

Cersosimo, D. e G. Wolleb (2006), *Economie dal Basso*, Donzelli, Roma.

Dahlby, B. (1996), Fiscal externalities and the design on intergovernmental grants, *International Tax and Public Finance*, 3, pp. 397-412.

Dixit, A. e J. Londregan (1998), Fiscal Federalism and Redistributive Politics, *Journal of Public Economics*, 68(2), May pp. 153-180.

---

[184] Best practice is invoked everywhere by the vast world of functionaries in international, European and national agencies. Often all it means is good intentions. To actually replicate best practice is another matter.

DPS (2005), *Documento Strategico Preliminare Nazionale: Continuità, discontinuità, priorità per la politica regionale 2007-2013*, Ministero dell'Economia e delle Finanze, Dipartimento per le Politiche di Sviluppo e Coesione, November.

Glaeser E. e A. Shleifer (2001), A Reason for Quantity Regulation, *American Economic Review*, 91,.n. 2, pp. 431-35.

Gillespie, S (2003), *Scaling up community driven development: a synthesis of experience*, World Bank Working Paper 31631.

Jütting, J., C. Kauffman, I. Mc Donnel, H. Osterrieder, N. Pinaud, L. Wagner (2004), *Decentralization and Poverty in Developing Countries, Exploring the Impact*, OECD Development Centre Working Paper n. 236.

Kreps, D (1990), Corporate Culture and Economic Theory, in J. Alt and K. Shepsle (ed), *Perspectives on Positive Political Economy*, Cambridge University Press, Cambridge.

Lippert, S. e G. Spagnolo (2004), *Network of Relations*, North American Winter Meeting of the Econometric Society, "Network: Theory and Applications", 10-12 June, 2004 Vaxholm, Working Paper.

Martinez-Vazquez, J. (2001), *The Design of Equalization Grants: Theory and Applications*, Georgia State university, Andrew Young School of Policy Studies, for World Bank, Washington, DC, USA

Majumdar, S.K. e A.A. Marcus (2001), Rules versus Discretion: the Productivity Consequences of Flexible Regulation, *Academy of Management Journal*, 44, pp. 170-179.

Milgrom P. e J. Roberts (1992), *Economics, Organization and Management,* Prentice-Hall International, New York.

Picci, L. (2004), Discussione, *Rivista di Politica Economica*, March-April, pp. 331-340.

Shah, A. (1998), *Balance, Accountability, and Responsiveness: Lessons about Decentralization*, Policy Research Working Paper Series, n. 2021, World Bank, Washington, DC, USA.

Shah, A. (ed.) (2003), *Handbook on Public Sector Performance Reviews, Volume V: Fiscal Federalism: Principles and Practices* (A. Shah, R. Boadway, I. Cromb)*,* World Bank web publication.

Sheuerman, W. (2004), Democratic experimentalism or capitalist synchronization? Critical reflections on directly-deliberative polyarchy, *The Canadian Journal of Law & Jurisprudence*, XVII, 1 January pp. 101-127.

Stato e Mercato. (2005), Politiche pubbliche, pratiche partecipative, democrazia deliberativa, governance territoriale, 1, pp. 3-130.

UNCDF (2002), *Policy Impact and Replication in Local Government and Microfinance*, ONU, New York.

Von Braun, J, e U. Grote (2002), Does Decentralization Serve the Poor? in IMF (ed.), *Fiscal Decentralization*, pp. 92-119, Routledge Economics, Washingtong, D.C.

Chapter Six - Conclusion

The number of academic publications documents produced by international and national organisms and programme frameworks devoted to local development is growing all the time.[185] And yet actual results are few and far between. There is little documentation on local development, and research is often based on incomplete data. There is, however, one thing most people involved in development agree on: whatever results are currently available, they are as yet uncertain. Only time will tell.

This is, in our view, true only in part. Results are disappointing because the theories to which reference is made - both explicitly and implicitly – in building expectations are often simplistic, if not downright wrong. The damaging effect of over-expectations is clearly illustrated by the gap between the set of convictions possessed by those involved in development and the beliefs held by commentators (especially if they are economists). There is widespread scepticism among economist commentators regarding development policies in general, and local policies in particular – unless these policies are highly regulated, or, even better, deregulated. They have no expectations and are quick to point out that if there are no results so far it is probably because no results are forthcoming. On the other side of the divide, those who are involved directly in development are convinced that policies – especially local policies - are the key to future development. Their expectations are high: significant, rapid results in generous proportion to the resources employed. The combination of these two contrasting, and, we believe incorrect, approaches inevitably generates a view that results are inadequate: either because the few results that are achieved are under-estimated, thus confirming a prejudiced view, or because the expectation of results was too high in the first place.

This assai came about from the idea that something needs to be done to avoid these two sides pitting themselves against each other. We felt that it would be a good idea if the expectations of those directly involved in development had more connection with the views of those undertaking research or social commentary. This book invites those

---

[185] For Italy alone, starting with a bibliography on local development that has been updated, more than 1,000 publications have come out between 1990 and 2004, produced by over 200 researchers in over 30 main groups based in universities in the faculties of Economics, Political Economics, Applied Economics, Business, Sociology, and, to a lesser extent, Economic History, Agrarian Economics, Law and Land Economics.

directly involved in development to understand that even though results are uncertain, delayed in time, and disproportionate to the effort made and the resources employed, at the same time local policies are both needed and potentially effective. It also invites those who study the subject and comment on it in the press to consider the fact that there are many other models other than the deterministic ones in which development is either easy or impossible, and in which development takes place only where material conditions allow it to, as if intentionality had nothing to do with it.

By now it will have become clear, as we announced in the Introduction, that the materials provided chapter by chapter tend to deconstruct rather than construct. Although there are many positive suggestions in the book, the arguments are more often against. To conclude, then, we would like to indicate some important subjects that need further investigation.

There are two preliminary questions. One is related to the local nature of development, and the other to the local nature of policies. The questions are distinct, in that answering one does not necessarily mean there is an answer to the other. The first question is: can we claim that economic development is local? The second question is: are local development policies necessary? These are separate questions because it is legitimate to consider local development policies necessary even if you believe economic development is not local[186], or to consider global policies necessary even if you are convinced development is local.[187] The concepts presented in this book are grounded in a positive answer to the first question (economic development is local) inasmuch as the potential for development is linked to immobile resources that have been underutilised, or neglected all together. If there is such a thing as immobile resources, then by definition development is local, as long as (and this is the second condition that supports our view) different resources generate different development mechanisms. Which brings us to the third condition: that the concept of 'local'

---

[186] It is possible to think of development as a process in which decisive elements work on a national scale, or even on a continental or global scale, and to believe that appropriate development policies for this process – including, for example, the accumulation of physical or human capital – should have a relevant local dimension because they depend on the efficiency of single organizations as well as on the general rules governing the various functions. An example of this could be intermediation in savings and knowledge for banks and schools.

[187] It is also possible to think development is local (i.e. that it responds to mechanisms that are specific to territories and therefore different from place to place, not secondary or less important) but at the same time believe that development policies should be national (European, global). An example of this could be national infrastructure that gives different local advantages in different places, but can only be organized at a central level.

corresponds to the concept of 'differentiated'. We also answer question two (whether policies are local) in the affirmative, for three reasons: I) the constraints that impede the productive use of potential resources are specific to the territory; II) the difficulties inherent in individual and collective action are specific to the territory; III) the capabilities and difficulties of institutions are specific to the territory. Development projects therefore need to be specific, and different in each territory. These assumptions all need further exploration, of course. Any further discussion should include an analysis of the fact that policies need to be local mainly because local communities have become more aware and more capable of mobilisation. It is therefore a democratic imperative (or a state of necessity) to design and implement development policies grounded in the territory, or at least involving those who live and work in the territory, over and above considerations of effectiveness or efficiency.

Further questions involve the nature of institutions. Should only rules be taken into consideration, or should institutional bodies also play a role in development? In this book we have adopted the latter view, in accordance with the hypothesis that public goods and services are an essential component of development, and that rules on their own are not enough. Our justification is that development always involves a transition towards a greater division of labour, in the widest sense, and that private agents, without the support of public goods and services, cannot sustain this transition. Another discussion point is the idea presented in the book that bottom up collective action in civil society without state intervention cannot deliver adequate solutions.

There is, too, the matter of institutional entrepreneurship. In this book we insist on the partial and inevitably indeterminate nature of choices underlying development policies, to the extent that we consider institutional entrepreneurship a decisive factor in their success. An objection to this approach could be: what is intended by the term 'undetermined'? Can choices really depend on chance? Our answer is that you can distinguish between something happening by chance and wanting something to happen, but this too is material for further discussion.

Negotiation is another matter that needs further consideration. Usually there are three different arguments in favour of negotiation and concertation in local development policy building. These are: I) the fact that interest groups are able to perceive the effects of single policies and are able thus to give or take away their consent according to their perception; II) the fact that useful knowledge for designing and implementing development policies is dispersed among local actors, which means

that concertation is needed to bring these actors' knowledge together and contribute to the decision making process; III) the fact that there are many different options for policies that are hard to prioritise, and that the problem is to align behaviour along one of these options. Analytically, a problem arises because it is not clear whether these reasons complement each other or are alternatives. There could be relevant consequences in terms of analysis and strategy for development. For example, if it was believed that there was such a thing as an optimal strategy that needs to be discovered and implemented, then perhaps the first two reasons (composition of interest groups, and reconstructing knowledge) should be considered complimentary and the third reason should be excluded. If this were the case, local capabilities and choices would become paramount. If, on the other hand, it were believed that there is no such thing as an optimal strategy but that there were a series of options impossible to choose from a priori, then things would be quite different. In this case, the institutional framework and arrangements would be considered more important. Chapters 4 (the difficulties of local institutions) and 5 (institutional arrangements) try to maintain a certain balance, but in the end we recognise that interventions from the centre are important. One could of course object with an approach at the other extreme.

There are many other issues that this assai has not tackled. The most obvious concerns the fact that when we talk about local development we do not only mean lagging areas taking off, but also re-launching areas that are already advanced so that growth can take place there too. It is possible that after an accurate analysis the role of intentional policies needs to be considered differently in these two different scenarios. In the case of development in advanced areas a important argument should concern innovation and the need to reconsider current paradigms critically. Another argument should regard the interaction between economic dynamics and international specialisation shifts.