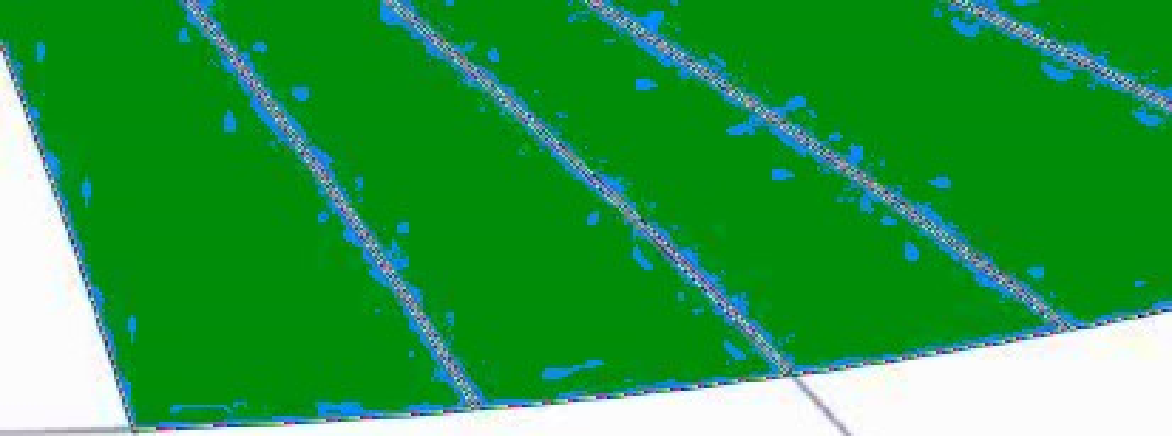

Gunther Maier
Sabine Sedlacek

Spillovers and In

Space, Environm and the Economy

WIRTSCHAFTS UNIVERSITÄT

SpringerWienNewYork

Interdisciplinary Studies in
Economics and Management

Vol. 4

Edited by the Jubiläumsstiftung
der Wirtschaftsuniversität Wien

Gunther Maier
Sabine Sedlacek (eds.)

Spillovers and Innovations

Space, Environment, and the Economy

SpringerWienNewYork

Univ.-Prof. Dr. Gunther Maier
Dr. Sabine Sedlacek
Institut für Wirtschaftsgeographie, Regionalentwicklung und Umweltwirtschaft
Wirtschaftsuniversität Wien, Vienna, Austria

Printed in Austria
SpringerWienNewYork is a part of Springer Science + Business Media
springeronline.com

Typesetting: Camera ready by authors
Printing: Novographic Druck G.m.b.H., 1230 Wien, Austria
Printed on acid-free and chlorine-free bleached paper
SPIN 10975992

With 18 Figures

Library of Congress Control Number 2004114866

ISSN 1615-7362
ISBN 3-211-20683-3 SpringerWienNewYork

Contents

Spillovers and Innovation, Environment and Space: An Introduction

Gunther Maier*, **Sabine Sedlacek**‡

1 Introduction

In recent years a lot of literature has dealt one way or the other with spillover effects in innovation and knowledge production. The introduction of the knowledge production function by Griliches (1979) provided the basis for a formal treatment of knowledge production in economics. It links R&D and human capital – the inputs – to some innovative activity – the output. As pointed out by Audretsch and Feldman (2004), the concept of the knowledge production function is strongly supported at an aggregate level of countries or industries, while analyses at a more disaggregate level show only a weak relationship. The firms with the largest inputs into knowledge production are not necessarily the ones with the highest level of innovation activity. This is interpreted as strong evidence for spillovers in knowledge production and innovation activities.

Since knowledge shows characteristics of a public good, a view going back to Arrow (1962), the hypothesis that knowledge spills over from one economic actor to another is widely accepted (Wheeler, Mitchelson, 1991). However, substantial disagreement exists about the mechanisms of these spillovers and whether they are geographically bounded (Audretsch, Feldman, 2004). "Furthermore, in those studies that deal with knowledge spillovers in a spatial context, the spatial element is seldom modelled in a satisfactory way and often suppressed" (Karlsson, Manduchi, 2001, p. 102).

This lack of precise understanding is in stark contrast to the many policies and policy-oriented concepts that attempt to utilize the positive impulses of knowledge spillovers. Motivated by the result of growth theory that in a developed economy the rate of economic growth is determined by the growth of technical knowledge, policy makers have looked for ways to promote technical progress. Romer's (1990) famous endogenous growth model introduces knowledge spillover effects through the argument that the increase of technical knowledge in a certain time period depends upon the stock of technical knowledge available at that time. Partly inspired by the success of Silicon Valley and other high-tech regions, concepts like innovative milieux, learning regions, etc. have been proposed and used for innovation-oriented regional development policies. Although they typically go well beyond innovation, the cluster

*Department of City and Regional Development, Vienna University of Economics and Business Administration, Vienna, Austria.

‡Department of Environmental Economics and Management, Vienna University of Economics and Business Administration, Vienna, Austria.

policies, which are currently used in many countries and regions (see e.g. Feser, 2004), also have a knowledge spillover argument embedded.

As a concequence, collaboration, network building, transfer of knowledge, etc. are key policy measures in most comprehensive policies. The EU framework programs are a good example. Sometimes networking seems to be a value per se, and its underlying reason has been forgotten.

While (regional) economists and policy makers currently concentrate on the positive aspects of knowledge spillovers, environmental economists are concerned with spillovers of a different type. Economic activities – production and consumption – tend to create unintended residuals that spill over into the public sphere. Environmental sciences have demonstrated (Bossel, 1998, Georgescu-Roegen, 1987) that in many cases the relationship between the creation of such a spillover and its effects is highly complex. The medium, to which they are emitted, transports the residuals over space and through time. Residuals from one or several sources accumulate over time, may reach threshold levels and interact with one another or the receiving medium in chemical or physical transformations.

These residuals have a potentially harmful effect on the eco-system, including human beings. Acid rain and the greenhouse effect are just the most drastic examples of processes where spillovers from the economic to the ecological sphere are the usual suspects (Costanza, 1997, Mabey, 1997). Their consequences – climatic change – may turn out so dramatic that the sheer existence of mankind is at stake.

Another, though related line of arguments in environmental economics is that economic activities as they are structured today, depend heavily upon non-renewable resources. Using these resources today makes it more difficult for future generations to follow this path of development. In this sense, today's activities generate negative spillovers for future generations. Both these arguments are at the heart of the sustainability discussion in environmental economics and policy (Daly, 1990, 1996, Hussen, 2000). To what extent can the current economic system be sustained in the long run and where does it have to be changed in order to become sustainable?

Regarded from a more general perspective, all these issues are related to questions of measuring spillovers, long term stability, the role of various policies, etc. With this volume we intend to contribute to the ongoing discussion. By bringing together arguments from different angles and contexts we hope to draw a more comprehensive picture than earlier, more specific publications did. The material put forward in the following chapters originates from two sources: First, from a workshop entitled "Spillovers and Innovation: Space, Environment, and the Economy", which was held in May 2002 honoring the 60th birthdays of our colleagues Edward M. Bergman and Uwe Schubert. In this workshop we brought together researchers from various parts of the world to discuss this topic, which was an essential one throughout Ed's and Uwe's academic careers. The second source of material is research work in this area done at the Vienna University of Economics and Business Administration partly within its pool of research areas, partly outside of it. We hope that this collection of material can achieve the attempted goal and advance the discussion in this area a step further. It is the purpose of this introductory chapter to set the stage and raise some of the issues that will be dealt with more thoroughly in the coming chapters.

2 Spillovers as Externalities

Spillovers, as we have discussed them above, are of course part of a more general phenomenon; that of externalities. They "arise whenever the value of an objective function, for example the profit of a firm or the happiness of an individual, depends upon the unintended or incidental by-products of some activity of others" (Lin, 1976, p. 1). The concept was introduced by Marshall (1890), but was given little attention until Pigou (1920), "where, developed and extended, it appears as one of the chief causes of divergence between 'private net product' and 'social net product'" (Mishan, 1971, p. 1).

Various classifications of externalities have been proposed. Scitovsky (1954) distinguishes between pecuniary externalities and technological externalities, the former operating via financial links between economic actors, the latter via technological relationships. In regional economics the differentiation first introduced by Ohlin (1933) into localization economies and urbanization economies is quite common. Localization economies operate among firms of the same sector, whereas urbanization economies link firms of different sectors and even different types of economic actors. Note that while Scitovsky's classification distinguishes externalities by the type of linkages that generate them, that of Ohlin puts the focus on the recipient of the external effect and its consequences. Localization economies lead to concentration or dispersion of one sector, urbanization economies concentrate or disperse economic activities in general.

Usually the terms 'positive' and 'negative' externalities are used to distinguish between those externalities that agglomerate and those that disperse economic activities. Positive externalities raise the objective of the recipient and are therefore attractive. Negative externalities, on the other hand, lower the objective and are therefore avoided. It is important to note that in this context positive and negative refers to the type of impact the externality has on the recipient's objective function and not to whether these externalities are 'good' or 'bad'. In the policy discussion, however, positive and negative externalities often have these normative component. One tries to generate positive externalities and avoid negative ones.

The distinction between positive and negative externalities also defines a line between regional economics and environmental economics, the two economic subdisciplines that mainly contribute to this volume. While the recent discussion in regional economics, as we have argued above, relies on positive externalities, environmental economics has always considered negative externalities more important. For environmental economists environmental problems result from a misplaced allocation of scarce natural resources (including allocation over time, i.e. intra- and intergenerational). In the 1970s and 1980s this had stimulated intensive policy discussions which in many European countries changed the political system. Green parties and many new NGOs entered the scene and their success forced traditional political parties to integrate ecological ideas into their programs.

The reason for this misallocation of scarce natural resources lies in the externality originating from their public good character. Since, in many cases, natural resources lack a well defined owner, their use cannot easily be restricted. This has led, on the

one hand, to overuse and, on the other, to recent attempts at establishing policies that assign property rights to environmental resources to prevent this overuse.

However, environmental economics is not exclusively concerned with negative externalities. As experience shows, environmental regulations had not only direct effects in form of higher production costs, but also indirect benefits from e.g. water supply of good quality and sewerage (Chapman, 1999). This implied a higher demand of specific environmental technologies which became important and provided new market niches for manufacturing and production sectors suffering economic crisis.

Our list of classifications of externalities is by no means exhaustive, many more have been proposed. In the following chapter Wang and Nijkamp (2004) discuss a taxonomy of externalities that explicitly takes into account the spatial component. In a recent paper Johansson (2004) proposes yet another classification. Starting from the – somewhat artificial – distinction of a firm's activities into supply and development activities, Johansson proposes to differentiate between efficiency externalities and development externalities. "An efficiency externality may be classified as a static problem and refers to how externalities may cause one industry or several industries to be more efficient in one location than another" (Johansson 2004, p. 3). Development externalities, on the other hand, are tied to the various innovation activities of firms and result from dynamic processes.

But, despite all these classifications, much of the discussion of externalities relies on examples. Since they are often defined as a residual category – all relationships between economic actors outside the price mechanism; all factors disturbing the perfectly competitive market – externalities remain difficult to define in a precise way. This becomes even more difficult when we consider the relation to other economic concepts: public goods, infrastructure, economies of scale, monopolistic markets are all closely related to externalities and it is often difficult to precisely draw the line between them. To give an example, positive localization economies in one region appear as economies of scale when we look at the sector's aggregate production function.

So, Bohm's (1987, p. 263) complaint is still valid today: "These cases are awkward to handle in traditional, well-structured economic analysis. So the main characteristic of these external economies, very much like those suggested by Marshall, is that we cannot yet say in any systematic way exactly what they represent."

3 Consequences of Externalities

As a consequence, much of the discussion of externalities has focussed on their consequences rather than the processes that generate them. A good example is the distinction between localization and urbaniziation economies common in regional science. (Positive) Localization economies, since operating among the firms of one sector, attract these firms to one another and therefore lead to their co-location in a certain region. Urbanization economies, on the other hand, attract different economic actors to one location, thus urbanizing this place. The symptom of a spatial concentration of one sector or of economic activities in general is seen as evidence that positive externalities are at work in that region. The type and nature of these externalities remains in the mist and is often only laid out in some speculative arguments with no real evidence.

In economics, for a long time the discussion of externalities was dominated by their allocative consequences. Since externalities – positive as well as negative – are unintended side-effects of one actor's decision on another actor's objective function, they are not taken into account in the first actor's decision making. As a consequence, this actor does not produce or consume the socially optimal quantity, but rather his or her private optimum which may be too high (in the case of negative externalities) or too low (positive externalities) as compared to the social optimum.

The implications are quite dramatic. The difference between private and social optimum implies that the economy will not converge toward a Pareto optimum, even if all the other assumptions of the competitive neoclassical economy hold. Externalities were therefore considered a nasty disease that keeps the market mechanism from doing its wonders. They belong to the category of market failures. Much of the earlier economics literature on externalities has therefore been concerned with the question of how to cure this disease (Mishan, 1971, Lin, 1976). The treatment of choice is to internalize the externality through taxes or subsidies, such that the private optimum coincides with the social one.

This approach has some obvious shortcomings. First, it assumes that the policy maker is neutral and not part of the system. However, it is often the public sector or policies that generate the largest external effects. The provision of infrastructure like roads, education, etc. often contains a substantial element of positive externality. Many technology policies attempt to stimulate just those components that generate spillover effects.

Second, as Mishan (1971, p. 8) notes, "the analysis of external effects has always been conducted within a partial equilibrium framework". Therefore, even if policy wants to internalize an external effect, the taxes and subsidies to internalize it are based on the assumption that the optimality conditions are met in the rest of the economy. This implies the assumption that no externality besides the one under investigation exists in the economy. In the light of our latter discussion this appears to be a rather heroic position. When this condition is not met, we know from the General Theory of the Second Best (Lipsey and Lancaster, 1957) that there is no certainty that the measures taken to internalize the one externality will move the economy as a whole closer to a Pareto optimum.

4 An Economy without Externalities – Is It Feasible?

In recent decades the view of externalities has changed substantially. Today, there are serious doubts that an economy without externalities – i.e. the neoclassical economy of perfect competition – can even exist in the long run.

These doubts originate from growth theory, a core area of economics. One of the problems with the standard growth model as it was developed by Solow (1956) is that technological progress is exogenous and thus remains unexplained. In the long run, however, the growth rate of an economy is determined solely by that technological progress. Neoclassical growth theory explains long-run growth by a factor that remains unexplained.

In the 1980s and 1990s numerous contributions were published that tried to in-

tegrate the production of technological progress into economic theory (e.g. Romer, 1987, 1990, Grossman and Helpman 1991a, b, Rebelo, 1991). They have become known as "new" or "endogenous growth theory". In addition to the theoretical problems with the standard model, this development was also motivated by the observation that in a modern economy a substantial and growing amount of resources is devoted to research and development (Grossman and Helpman, 1991b, OECD, 1986). Despite their different approaches, the endogenous growth models generate one common result, namely that the production of knowledge and technological innovations is incompatible with the neoclassical concept of perfect competition.

There are mainly two reasons for this. First, in an economy under perfect competition all income is divided between capital and labor. There are simply no resources left to be devoted to knowledge production. Second, there is no economic incentive for firms to invest in knowledge production. Due to the assumption of perfect information, all knowledge that is produced by the firm immediately spills over to its competitors who will enjoy the benefits without the costs. Therefore, no firm will be willing to invest into this public good. The obvious solution for this dilemma, the implementation of strict patent regulations, does not solve the conceptual problem, because it introduces technological monopolies and thus violates another condition of perfect competition.

The models of endogenous growth theory introduce different types of "market failures" to overcome these problems. Some consider knowledge production to be an external effect of capital investment (e.g. Romer, 1986, Rebelo, 1991), some consider monopolistic competition in a knowledge producing sector (e.g. Romer, 1990). In any case, externalities play a very different role in an economy with endogenous growth. Rather than a nasty disease that needs to be internalized, they are a necessary ingredient for sustained economic growth.

While the argument of endogenous growth theory raises doubts about the sustainability of a perfectly competitive economy in the long run, a spatial perspective raises similar doubts even in the short run. The assumptions of the neoclassical model of perfect competition, as formalized by Arrow and Debeu (1954) "are restrictive in themselves but in the context of a spatial economy become literally untenable" (Fujita and Thisse, 2002, p. 27). Building on Koopmans and Beckmann (1957), Starrett (1978) shows that an economy where resources are utilized for transporting goods from one location to another cannot be in equilibrium under perfect competition. The formal proof follows the intuitive argument that Mills put forward a few years earlier: "Consider a general equilibrium model in which an arbitrary number of goods is produced either as inputs or for final consumption. The only nonproduced goods are land and labor, each of which is assumed to be homogeneous. Assume that each production function has constant returns to scale and that all input and output markets are competitive. Utility functions have the usual properties and have as arguments amounts of inputs supplied and products consumed. Under these circumstances, consumers would spread themselves over the land at a uniform density to avoid bidding up the price of land above that of land available elsewhere. Adjacent to each consumer would be all the industries necessary – directly or indirectly – to satisfy the demands of that consumer. Constant returns assures us that production could take place at an arbitrarily

small scale without loss of efficiency. In this way, all transportation costs could be avoided without any need to agglomerate economic activity." (Mills, 1972, p. 113)

This completely contrasts the main characteristics of contemporary economies: high and increasing volumes of freight transport, regional specialization in production, coordination through just-in-time deliveries, concentration through mergers and takeovers, etc. The location pattern implied by this type of "backyard capitalism" is just the contrary of what Krugman (1991, p. 5) identifies as "the most striking feature of the geography of economic activity (...) production in many industries is remarkably concentrated in space". Already Marshall identified externalities "between firms of a certain region as one of the main reasons for the spatial concentration of economic activities" (Karlsson and Manduchi, 2001, p. 106). Starrett (1978) demonstrates that the obvious costs of this concentration in terms of land rents, transportation costs, and inflated costs of other resources need to be balanced by benefits resulting from the spatial agglomeration of these activities. The highly specialized and concentrated spatial structure of our economies is thus a direct indicator of the amount of externalities that are at work.

5 An Economy with Externalities – Is It Manageable?

When we look at the economy from this perspective, externalities appear to be one of its key components rather than a disturbing anomaly. In the last few years a growing body of literature deals with externalities of various types. The golden anniversary issue of Papers in Regional Science (Vol. 83, nr. 1, January 2004) summarizes a substantial part of this literature.

Fujita and Krugman (2004), sketching the development, goals, and achievements of the New Economic Geography, stress the need to develop a microeconomic foundation for the agglomeration forces of their models in order to avoid the sarcastic criticism that "economists believe that companies agglomerate because of agglomeration economies" (p. 141). Indivisibilities, production linkages and knowledge spillovers are prime candidates for future formal explanations of increasing returns. "Increasing returns, in turn, lead to the market structure characterized by imperfect competition" (Fujita and Krugman, 2004, p. 142).

Johansson and Quigley (2004) stress the connection between agglomeration and network linkages. "Indeed, networks among economic actors dispersed over space may act as a substitute for agglomerations of actors at a single point, providing some or all of the utility gains and productivity increases derived from agglomeration" (p. 166). Interestingly, they even view modern firms as a form of network: "When a firm is established – with Coasean motivation – the action is nothing but the formation of a network internal to the organization for interaction among a set of interdependent subunits" (Johansson and Quigley, 2004, p. 169).

Batabyal and Nijkamp (2004) link space to the environment and environmental externalities . "Space is normally the medium through which environmental externalities are distributed" (p. 292). They argue that the spatial dimension functions as "a vehicle through which social costs are often transferred from a polluter to a pollutee."

Infrastructure is drawn into the picture by McCann and Shefer (2004). "External-

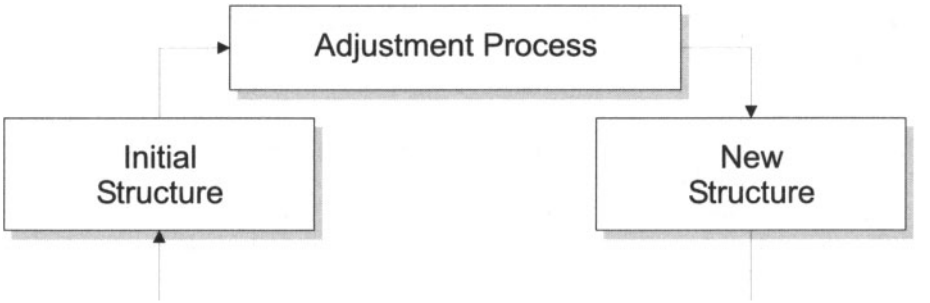

Figure 1: Feedback loop

ities are a common feature of infrastructure, and there are several aspects to this" (p. 179). "The crucial link here is between the local existence of the infrastructure and the generation of local positive externalities" (p. 179).

When we put all these arguments together, we see an abundance of both positive and negative externalities structuring the spatial economy. Heterogeneous economic agents will evaluate them and react to them differently. But heterogeneity may even be a source of externalities itself. "Diversity and variety of consumer goods or in producer inputs can yield external scale economies, even though all individual competitors and firms earn normal profits" (Johansson and Quigley, 2004, p. 170). Maier (1995) shows that random price fluctuations may cause suppliers to agglomerate in one location, even when homogeneous customers remain spread evenly.

However, it is not only the externalities that generate spatial structure, a spatially heterogeneous economy also generates externalities. Spatial proximity between firms linked in a production chain generates "input-cost externalities and delivery-cost externalities. Both cases emphasise the consequences of proximity for transaction costs" (Johansson and Quigley, 2004, p. 171).

In a more abstract view, the current structure of the economic system generates and shapes its externalities, which in turn generate the forces that advance the system over time. This link is sketched in figure 1. The key point is that in an economy with externalities the adjustment process depends upon the initial structure. Take the typical New Economic Geography models: the current distribution of economic activities endogenously determines the scale economies in the respective regions, which in turn determine investments within or capital/labor flow between the regions. A similar feedback loop between spatial structure and its dynamics is suggested by some of the innovation literature: "Agglomerations will have a more rapid development of technology, and hence faster productivity growth" (Johansson and Quigley, 2004, p. 174). Similarly, Romer (1990) argues: "The engineer working today is more productive because he or she can take advantage of all the additional knowledge accumulated as design problems were solved during the last 100 years" (p. S83f.).

The dynamic behavior of structures like the one sketched in figure 1 has been studied among others by W. Brian Arthur (1986, 1994). The consequences of introducing the above mentioned link between the structure and the adjustment process have

clearly been demonstrated by Arthur et al. (1987). They consider a repeated process with two possible outcomes at each step. When the probabilities for the two outcomes are independent in-between the repetitions, the strong law of large numbers guarantees that in the long run the observed frequencies of the outcomes converge toward the probabilities. However, when the process is changed such that the probabilities at each step equal the observed frequencies at that step, the long term outcome of the process is undefined. When this dependence between the current structure and the adjustment process – this externality – is introduced, the result is a so-called Polya-process. Polya (1931) proved that in the long run the relative frequencies resulting from this process tend toward a limit X with probability one, but "X is a random variable uniformly distributed between 0 and 1" (Arthur, 1994, p. 36). While in the first version of the process the long term outcome can be predicted with certainty at the beginning, in the second version, although we know that the process will converge, the outcome is completely unpredictable. There is an infinite number of possible outcomes – infinitely many multiple equilibria – and each one of them is equally likely.

To give some economic meaning to this abstract structure, think of it in terms of the location of shops in two shopping centers. The first version assumes that new shops will choose a center irrespective of the distribution of shops between the centers. The second version assumes that new shops are attracted by the center where shops have agglomeratted previously.

Arthur's model is clearly quite extreme. However, because of its simple structure and its remarkably different outcome it clearly demonstrates the long run effects of the links between structure and the adjustment process that we have identified with externalities. More realistic versions of the model have been applied by Arthur to competition between new technologies, industrial location, and transmission of information. Maier (2001) applies the same logic to a two region neoclassical growth model. In his model, capital is perfectly mobile and labor is immobile between the regions. Instead of being diffused to both regions equally over time, in each period of time increments of innovation are added at random to any one of the two regions. When the probability for the next increment of innovation depends upon the region's share of production, a region's share of production converges toward either 0 or 1 in the long run. From a certain point on, all economic growth is concentrated in the one region that has developed as the dominating one. The other region stagnates completely[1]. Interestingly, this tendency toward polarization does not exist at the very beginning of the process. Early on, production tends toward equal distribution between the regions. The tendency toward polarization only takes over at a later state in the process.

Irrespective of their structural details, such models of increasing return have some common characteristics. Arthur (1994, p. 112) lists the following:

- Multiple equilibria,
- Possible inefficiency,
- Lock-in,
- Path dependence.

[1] This happens despite the fact that because of the assumed immobility labor remains distributed equally between the regions, and despite the corresponding implications for the wage levels in the regions.

We have already discussed the possibility of inefficiency above. Multiple equilibria were mentioned at least implicitly. The Polya-model, for example, is a model with an infinite number of equilibria. The typical two-region core-periphery model of New Economic Geography, on the other hand, offers three possible equilibria: production split equally between the regions, and all production concentrated in one or the other region.

Lock-in is a by-product of multiple equilibria. When the system is at one of the possible equilibria – or converging toward one – this equilibrium may represent a *local* maximum of the respective objective function, but not the *global* one. Although another equilibrium may be clearly preferable, the stabilizing forces around the local maximum will prevent the system from skipping onto a path toward the global maximum. The system will be locked in at the local maximum. It would require a significant policy effort and reallocation of resources to overcome the barrier formed by the stabilizing forces.

Depending on the specific structure of the model, those barriers may change with the development process. They may become weaker over time or stronger, thus deepening the lock-in. When policy tries to overcome a lock-in situation, this has clear consequences for the timing of such a policy. In the first case, the earlier the policy is applied the better, in the sense that fewer resources are needed to overcome the lock-in, in the second case, waiting for the barrier to soften is clearly a better strategy.

Although, in a technical sense the term "lock-in" refers to the above described phenomenon, there is also another part of the story that may be called lock-in. As we know from New Economic Geography, an economy with externalities may converge toward a core-periphery structure. The division of labor between the regions in the system will generate an unbalanced distribution of economic activities between them. Each region will be "locked-in" to its position within the implied hierarchy of regions.

We have already seen path dependence at work in our discussion of Arthur's two model versions above – the "law of large numbers" version without externality and the Polya version with externality. When we disturb the process at an early stage, in the version without externalities its effect will wash out over time and in the long run the model will converge toward the same equilibrium as without it. When we add the externality, however, we add a self-reinforcing mechanism that sets the system off on another development path when disturbed and thus keeps the disturbance from being washed away. The Polya version of Arthur's model will reach another long term equilibrium when disturbed than without disturbance. The long term outcome depends on the path the system takes at early stages.

The Polya version of Arthur's model displays an extreme form of path-dependence, since at any time in the development process a disturbance will change the long term result. The difference is only that an earlier disturbance will have a larger impact than a later one. In more realistic models the long term consequence of a disturbance may differ dramatically depending on when it occurs. When the system is in a stable phase a disturbance may have no long term consequence at all. When the system is close to a bifurcation point where two or more development paths are possible, the consequence of a disturbance may be dramatic. It may set the system off on one path rather than another. The policy implications are obvious: depending on the state of the sys-

tem, whether it is on a stable path or close to a bifurcation point, a policy initiative attempting to move the system into a certain direction may either fail or be highly successful.

6 Externalities and Economics – A Paradigm Change?

At least since Marshall (1890) there have always been economists interested in positive or negative externalities. However, they have largely been marginalized within the profession as externalities were seen as market failure that needs to be corrected or – even better – avoided. The consequences that recent literature has demonstrated and that we have sketched above, also made economists shy away for a long time from accepting externalities as an integral part of the economic system. To Schumpeter the existence of more than one solution to the same problem seemed unscientific. Arthur (1994, p. 4) quotes him for writing in 1954 "Multiple equilibria are not necessarily useless, but from the standpoint of *any* exact science the existence of a uniquely determined equilibrium is, of course, of utmost importance, even if proof has to be purchased at the price of very restrictive assumptions". Even earlier, another great economist had seen the consequences of increasing returns and externalities, but shied away from them: "When John R. Hicks surveyed these possibilities in 1939 he drew back in alarm. 'The threatened wreckage' he wrote, 'is that of the greater part of economic theory' " (Arthur, 1994, p. 4).

While Schumpeter, Hicks and and others at their time felt that they had a choice to take externalities into account or not, the recent development has shown that there actually is no choice. As we have argued in this introduction, if economists want to understand the growth of an economy over time, they need to allow for externalities. These externalities lead to imperfect competition and tend to agglomerate production at certain locations in the economy. The resulting spatial structure leads to specialization, transportation and further externalities. This moves the spatial perspective closer to the core of economics. "As it turns out, if we can explain geographical concentration, then we can go a long way towards explaining important aspects of international trade and economic growth" (Hanson, 2000).

Accepting this, however, is equivalent to accepting that the economic system is nonlinear and cannot be modeled adequately by a linear model. "Nonlinearity admits the coexistence of multiple attractors, and hence of development paths, which means that a process, due to minor divergences in the process itself can take completely different courses" (Puu, 2000, p. 3). It also pushes the door wide open for all the complexities of non-linear dynamics: seemingly simple functional relationships that, depending on parameter values, converge to a unique steady state, fluctuate between two or more equilibria, or display full scale chaotic behavior[2]. Currie and Kubin (2003) show that the full menagerie of chaos theory – sensitive dependence on initial conditions, period doubling, chaotic regions with windows, and so on – can be derived from Krugman's core-periphery model. It only has to be formulated in discrete time.

[2]A famous example is the logistic function $y_{t+1} = ay_t(1 - y_t)$, which plays an important role in the modelling of biological and environmental processes. For different values of a it shows all those dynamic patterns (see e.g. Peitgen et al., 1998).

Puu (1997) discusses fractal networks and fractal land rent landscapes. In Puu (2000) he presents a more extensive account of nonlinear phenomena in economics. He analyzes monopoly, oligopoly, the business cycle, international trade, and development and finds in all of them traces of complex nonlinearity.

This not only makes economics more complex, it also raises at least two other issues:

1. With all these externalities, spillovers, scale-effects, and non-linearities, how can we assume that the various economic actors know all the relevant prices and parameters sufficiently well to make well informed decisions? Don't we have to drop this assumption and devote more thoughts to how information is collected, how perceptions are formed, and how they are influenced by such things as spatial structures?
2. Is it sufficient and justifyable to analyze the economy in isolation as it is standard practice in economics? "If we accept the separate study of subjects like 'economics' for example, then it means that we must believe that the flows of goods and factors, and the prices that are observed can form some kind of *closed* self-consistent system in themselves and explanations can be found from *within* this narrow structure." (Allen, Phang, 1993, p. 13). With non-linearities, even small changes in one of the related systems like environment, society, biology, etc., that spill over into the economic sphere can have dramatic consequences. In this book we point toward the environment and some aspects of spatial structure outside economics, but there are clearly more candidates.

All these arguments should not be misinterpreted as to completely discarding economics and its accumulated body of knowledge. Of course, prices still depend upon supply and demand, investment adds to the stock of capital, and long term growth requires innovation. However, when we accept externalities and all the other factors that we have discussed as integral parts of the economic system, the system is influenced by a larger set of factors and in a more complex way. The behavior of the system becomes less predictable since we cannot necessarily transfer the knowledge that we have accumulated under one set of circumstances to a new constellation. Policies that worked in one situation may fail miserably in another. Events and developments in areas outside the economic system may change its parameters and set it off on a new development path.

The more complex structure of such an open, non-linear system challenges the traditional mechanistic view of how it works. The economy is not the one big machine that dumps a certain amount of goods and services when one feeds it the necessary resources and pushes the right button. It is rather the arena for many economic actors who constantly learn about the constraints set by a macro structure, and use their knowledge and creativity to search for new opportunities within these constraints or create them by challenging the constraints. Through their interactive behavior the actors constantly modify and redefine the constraints at the macro structure thus developing the system further.

In this more evolutionary view of the economy which is partly inspired by Haken's concept of synergetics (see e.g. Haken, 1983, 1987, 1988, Haken and Mikhailov,

1993), competition, for example, is not so much the factor that enforces the one market price, but the mechanism that forces actors to look for new opportunities and niches, and thus to be able to depart from the one market price or to develop a new market of their own. Heterogeneity of actors, products, spatial structures, etc. is not a random fluctuation that can safely be pushed into an error term, but a constant source of challenge, inspiration, and innovation – in other words, an essential part of the system.

By opening the doors to externalities, economics has also opened a Pandora's box full of new questions and challenges, but also full of exciting new opportunities. As even this rough sketch – hopefully – has shown, many new questions await to be answered and some old questions need to be reconsidered. To answer them is much larger a task than can be performed in just one book. Many more books, articles, and conference discussions will be needed in this respect. Nevertheless, we hope that this book will make contributions that bring some of these questions a step closer to an answer.

7 Structure of the Book

With this brief introduction we hope to provide the context and framework for the more specific contributions in the remainder of the book. In the next seven chapters, an international group of scholars lays out their specific view of the relationship between spillovers and externalities on the one hand and environment and space on the other. The concluding chapter by Edward M. Bergman and Uwe Schubert points out what has been left open. They identify issues for future theoretical research as well as related policy issues.

The chapter by *Wang and Nijkamp* complements this introduction quite nicely. They focus specifically on the interaction between the economic and the ecological system, and try to take into account the spatial dimension of both these systems. Since externalities differ in their spatial reach and may be unidirectional or multidirectional, their approach naturally leads to a rich taxonomy of externalities. It becomes quite obvious from their presentation that the economic system and the environment are closely related and that these links can range, in terms of spatial scope, from quite localized to global.

Gumprecht, Gumprecht and Müller deal with the econometric implications of external effects. They use a seminal article by Coe and Helpman (1995) on international spillovers in total factor productivity and the corresponding dataset to discuss various aspects of econometric estimation in the context of spillovers. As it turns out, applying the most adequate model, a panel cointegration model with random coefficients and dynamic regressors, does not support the hypothesis of international spillover effects. This result contradicts the results of Coe and Helpman.

The following two chapters by *Goldstein and Renault* on the one hand and *Varga, Anselin and Acs* on the other, both deal with core aspects of the knowledge creation process. Both use regional data for the United States for their empirical investigation. Goldstein and Renault focus particularly on the role of universities in knowledge based regional development. They use the fact that research universities can be found only at very specific locations for a quasi-experimental research design. Pooled cross-

sectional and longitudinal information allows them to distinguish different mechanisms by which universities may influence a region's average earnings per worker. Goldstein and Renault find significant knowledge spillover effects, but only of moderate size. Varga, Anselin and Acs tackle the question of knowledge production in a more direct way by applying a knowledge production function. They measure the output of newly produced knowledge by the number of patent applications. Since they too use pooled data, they can identify regional and temporal differences in the knowledge production process. Since one of the key explanatory variables in the chapter by Varga, Anselin and Acs is university research, these two chapters provide complementary views of the same process.

In the next two chapters by *Townroe* and by *Gindl and Wukovitsch*, respectively, environmental externalities are at the focus of the attention. Both deal with environmental issues and with the problems policy has dealing with these issues. In the case of Townroe's chapter it is the question of how to take into account the environmental side effects of new roads in the respective appraisals and policy decisions. Townroe discusses the fundamental problems the policy maker faces in this context, the traditional cost benefit analysis approach which tries to internalize the externalities in an accounting framework, and contrasts it with a new approach that has been proposed and implemented in Great Britain. He shows convincingly that the most serious problems with both approaches are related to the adequate valuing of environmental effects, aggregating them over a certain population and evaluating their distributional implications. The chapter by Gindl and Wukovitsch deals with a different issue of environmental relevance, sustainable urban tourism. More specifically, they raise the question what forms of partnership – mainly between public and private sector – are most adequate in developing and managing sustainable urban tourism. Since they use a broad concept of sustainability which includes social and economic sustainability in addition to environmental sustainability, this chapter clearly points beyond the narrow scope of environmental issues and incorporates questions of public management.

Feser focuses on policy issues as well. As far as externalities are concerned, however, this paper returns to the positive spillover effects between localized economic actors. Feser's paper investigates regional policies in US states and Latin American countries that use – or at least claim to use – the business cluster concept. Reviewing policy documents and secondary literature, Feser categorizes the different approaches and in many cases finds a clear gap between the theoretical concept, the political rhetoric, and the actual policy. Not surprising, in the actual policies the guidlines of the theoretical concept get blurred with the specific needs of the existing industry.

As has been mentioned above, the book closes with a concluding chapter by *Bergman* and *Schubert*. They point out advances and gaps in the field of spatial externalities and identify – in their subjective view – promising directions for future research. In summarizing their summary they stress four general points that they see emerging: The central role of space in the recent literature; the need for a multidisciplinary approach to tackle multi-disciplinary problems; the need for higher education to integrate these questions in policy study programs; and the need for policy design to focus on the phases of a program, in order to take into account the dynamics of the underlying processes.

As this volume shows, much has been achieved in the investigation of spatial externalities, but much more needs to be done. The research has generated new questions and questioned some old answers. Since we are convinced that a good understanding of these mechanisms is essential for effective policy making, we hope that this volume will further stimulate discussion of these questions.

References

Allen, P.M. and Phang, H.K. (1993) Evolution, creativity and intelligence in complex systems. In: Haken, H. and Mikhailov, A. (eds.), Interdisciplinary approaches to nonlinear complex systems. Springer, Berlin: 12–31

Arthur, W.B. (1986) Industry location patterns and the importance of history. Center for Economic Policy Research Paper 84, Stanford University

Arthur, W.B. (1994) Increasing returns and path dependence in the economy. The University of Michigan Press, Ann Arbor

Arthur, W.B., Ermoliev, Y.M. and Kaniovski, Y.M. (1987) Path-dependent processes and the emergence of macrostructure. European Journal of Operational Research 30: 294–303

Arrow, K.J. (1962) The economic implications of learning by doing. Review of Economic Studies 29: 155–173

Arrow, K.J. and Debeu, G. (1954) Existence of an equilibrium for a competitive economy. Econometrica 22: 265–290

Audretsch, D.B. and Feldman, M.P. (2004) Knowledge spillovers and the geography of innovation. In: Henderson, J. V. and Thisse, J.-F. (eds.), Handbook of regional and urban economics. North Holland, Amsterdam: 2713–2739

Batabyal, A.A. and Nijkamp, P. (2004) The environment in regional science: An Eclectic Review. Papers in Regional Science 83: 291–316

Bohm, P. (1987) External economies. In: The new Palgrave dictionary of economics, vol. 2. Macmillan, London: 261–263

Bossel, H. (1998) Globale Wende: Wege zu einem gesellschaftlichen und ökologischen Strukturwandel. Droemer Knaur, Munich

Chapman, D. (1999) Environmental economics. Theory, application, and policy. Addison-Wesley Longman, Reading, Mass.

Coe, D.T. and Helpman, E. (1995) International R&D spillovers. European Economic Review 39: 859–887

Costanza, R. (1997) An introduction to ecological economics. St. Lucie Press, Boca Raton, Fla.

Currie, M. and Kubin, I. (2003) Chaos in the core-periphery model. The University of Manchester School of Economic Studies Discussion Paper Series 0307.

Daly, H.E. (1990) Towards some operational principles of sustainable development. Ecological Economics 2: 1–6

Daly, H.E. (1996) Beyond growth: The economics of sustainable development. Beacon Press, Boston

Feser, E. (2004) Industry cluster concepts in innovation policy: A comparison of U.S. and Latin American experience. In this volume

Fujita, M. and Krugman, P. (2004) The new economic geography: Past, present and the future. Papers in Regional Science 83: 139–164

Fujita, M. and Thisse, J.-F. (2002) Economics of agglomeration: Cities, industrial location, and regional growth. Cambridge University Press, Cambridge

Georgescu-Roegen, N. (1987) The entropy law and the economic process in retrospect. Schriftenreihe des IÖW 5/87

Griliches, Z. (1979) Issues in assessing the contribution of research and development to productivity growth. Bell Journal of Economics 10: 92–116

Grossman, G.M. and Helpman, E. (1991a) Quality ladders in the theory of growth. Review of Economic Studies 58: 43–61

Grossman, G.M. and Helpman, E. (1991b) Innovation and growth in the global economy. MIT Press, Cambridge, Mass.

Haken, H. (1983) Synergetics. An introduction, 3rd ed. Springer, Berlin

Haken, H. (1987) Advanced Synergetics, 2nd ed. Springer, Berlin

Haken, H. (1988) Information and self-organization. Springer, Berlin

Haken, H. and Mikhailov, A. (1993) Interdisciplinary approaches to nonlinear complex systems. Springer, Berlin

Hanson, G.H. (2000) Scale economies and the geographic concentration of industry. NBER Working Paper 8013, National Bureau of Economic Research

Hussen, A.M. (2000) Principles of environmental economics. Economics, ecology and public policy. Routledge, London & New York

Johansson, B. (2004) A menagerie of agglomeration and network externalities. CESIS Working Papers No 2, Centre of Excellence for Science and Innovation Studies, The Royal Institute of Technology, Stockholm

Johansson, B. and Quigley, J.M. (2004) Agglomeration and networks in spatial economics. Papers in Regional Science 83: 165–176

Karlsson, C. and Manduchi, A. (2001) Knowledge Spillovers in a Spatial Context – A Critical Review and Assessment. In: Fischer, M.M. and Fröhlich, J. (eds.), Knowledge, complexity and innovation systems. Springer, Berlin: 101–123

Koopmans, T.C. and Beckmann, M.J. (1957) Assignment problems and the location of economic activities. Econometrica 25: 1401–1414

Krugman, P. (1991) Geography and trade. Leuven University Press, Leuven

Lin, S.A. (1976) Theory and measurement of economic externalities. Academic Press, New York

Lipsey, R.G. and Lancaster, K. (1957) The general theory of second best. Review of Economic Studies 24: 11–32

Mabey, N. (1997) Argument in the greenhouse: the international economics of controlling global warming. Routledge, London

Maier, G. (1995) Spatial search, structure, complexity, and implications. Physica, Heidelberg

Maier, G. (2001) History, spatial structure, and regional Growth: Lessons for policy making. In: Johansson, B., Karlsson, Ch. and Stough, R.R. (eds.) Theories of endogenous regional growth: Lessons for regional policy. Springer, Berlin: 111–134

Marshall, A. (1890) Principles of economics. Macmillan, London

McCann, P. and Shefer, D. (2004) Location, agglomeration and infrastructure. Papers in Regional Science 83: 177–196

Mills, E.S. (1972) An aggregative model of resource allocation in a metropolitan area. In: Edel, M. and Rothenburg, J. (eds.) Readings in urban economics, Macmillan, New York: 112–123

Mishan, E.J. (1971) The postwar literature on externalities: An interpretive essay. Journal of Economic Literature 9: 1–28

OECD (1986) OECD Science and technology indicators No.2: R&D, invention and competitiveness. Organization for Economic Cooperation and Development, Paris, France

Ohlin, B. (1933) Interregional and international trade. Harvard University Press, Cambridge, Mass.

Peitgen, H.-O., Jürgens, H. and Saupe, D. (1998) Chaos, Bausteine der Ordnung. Rowohlt, Reinbek bei Hamburg

Pigou, A. C. (1920) The economics of welfare. Macmillan, London

Polya, G. (1931) Sur quelques points de la théorie de probabilités. Annales Institute H. Poincaré 1: 117–161

Puu, T. (1997) Mathematical location and land use theory. An introduction. Springer, Berlin

Puu, T. (2000) Attractors, bifurcations, and chaos. Springer, Berlin

Rebelo, S. (1991) Long run policy analysis and long run growth. Journal of Political Economy 99: 500–521

Romer, P.M. (1986) Increasing returns and long run growth. Journal of Political Economy 94: 1002–1037

Romer, P.M. (1987) Growth based on increasing returns due to specialization. American Economic Review 77: 56–62

Romer, P.M. (1990) Endogenous technological change. Journal of Political Economy 98: S71–S102

Scitovsky, T. (1954) Two concepts of external economies. Journal of Political Economy 62: 70–82

Solow, R.M. (1956) A contribution to the theory of economic growth. Quarterly Journal of Economics 70: 65–94

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

Wang, S. and Nijkamp, P. (2004) Spatial externalities in an open environmental-economic system. In this volume

Wheeler, J.O. and Mitchelson, R.L. (1991) The information empire. American Demographics 13: 40–42

Spatial Externalities in an Open Environmental-Economic System

Shunli Wang*, Peter Nijkamp*

1 Introduction

The United Nations Framework Convention on Climate Change and the associated Kyoto Protocol (KP) are examples of policy proposals which deal with transboundary environmental issues. Externalities, as was discussed in the introductory chapter of this book, are a nasty disease. For a better understanding of externalities in the case of transboundary environmental issues, a theoretical formulation for modeling the interactions between the economic actors and the environmental issues are needed. In the modeling of the fundamental relationships between the economic system and the ecological system, these systems are often treated as spaceless points (see e.g. Wang, Nijkamp and Verhoef, 2001). Important spatial characteristics – such as the distance between the individual consumers and producers, their living space, the location of the firms and households, the infrastructure, the geographical nature of the economic and the ecological systems (mountains, oceans, plains, forest, etc.) – are therefore ignored. However, in most real-world situations, spatial characteristics in fact are an important determinant of human behavior and ecological phenomena. Therefore, the inclusion of spatial characteristics in an analysis may lead to a better understanding of the real world (see Isard, 1956, Krugman, 1991a, Batten, 2001). Especially in the interaction between the economic and the ecological systems, spatial characteristics may cause additional complexities and complications for analyzing the impacts of environmental externalities (see Nijkamp, 1977, Siebert, 1985).

In the literature, we observe that most models that deal with the spatial characteristics reduce the complexities – which arise from adding spatial characteristics into the model – by ignoring some other aspects of real-world events. In addition to the simplified way that the ecological and the economic system are often represented in these models – i.e. as the interaction between both of these systems, and the interactions within each of these systems – spatial interaction models often simplify complexities concerning the issues of: (i) the determination of the borders of the regions; and (ii) the representation of the distance between two regions.

In this chapter, a taxonomy of models that have been developed for analyzing environmental-economic externalities in a spatial setting will be given. For this purpose, the interpretation of externalities in terms of framework of interaction between

*Department of Spatial Economics, Free University, Amsterdam, The Netherlands.

the economic and the ecological system, which was dealt in Wang, Nijkamp and Verhoef (2001), will be extended to include spatial characteristics. In this extended framework, the concept of spatial externality and forms of spatial interactions will be used.

This chapter is planned as follows. Section 2 formalizes the conceptual framework of interaction between the economic and the ecological system and extends this framework to include spatial characteristics. The issue of the determination of the borders of the regions is discussed in Section 3, by further elaborating on the formulation of the spatial landscape and on how this formulation gives rise to different kinds of spatial externalities, given the interaction between the spatial-economic and spatial-ecological aspects. The issue of the representation of the distance between regions will be discussed in Section 4, in order to categorize the set of models in terms of spatial interactions and spatial externalities. Section 5 summarizes the chapter.

2 Formalizing the Conceptual Framework

This section discusses the formalization of an extension of the conceptual framework of the interactions between the economic and the ecological system (as developed in Wang, Nijkamp and Verhoef, 2001) by the inclusion of spatial characteristics.

2.1 The spatial landscape

The spatial landscape ($\mathbf{L}$) is a continuous Euclidean space of three dimensions, i.e. $\mathbf{L} \in \Re^3$. Both the economic system and the ecological system operate in this spatial landscape. A two-dimensional representation of this landscape may be found in Figure 1. This landscape may be divided into a discrete number of areas $\mathbf{L}_m = (L_1, ..., L_M)$, which may, as pointed out by von Thünen, Lösch and Isard (see Isard, 1956), be envisioned as a set of points or as a grid system (see also Siebert, 1985). For notational convenience, this chapter will use the bold lettertype to denote sets or subsets, and the normal lettertype to denote specific elements in the set.

2.2 The ecological system

In the ecological system, there are:

(i) a set $\mathbf{Z}$ of inputs[1], e.g. water, air, metal, fish and other ecological amenities; this set will also denote the amounts of inputs (also known as the stock of the ecological goods). The distribution of the amount of these inputs over the spatial landscape $\mathbf{L}$ may be denoted by the matrix $\mathbf{Z}^m$, where the superscript m signifies the column of the grid of areas ($\mathbf{L}_m$);

(ii) a set of ecological regeneration processes $\mathbf{R}$, which transforms the set of inputs ($\mathbf{Z}$) into a set of output factors ($\mathbf{S}$). The distribution of the regeneration process over the spatial landscape is assumed to be dependent on the distribution of the presence of the amount of the inputs in the area;

[1] We use the superscript 0 for the inputs and superscript ' for the outputs to denote that the sets of the inputs ($\mathbf{Z}$) and the set of output factors ($\mathbf{S}$) may coincide, but not necessarily. This also applies for the notation in the economic system.

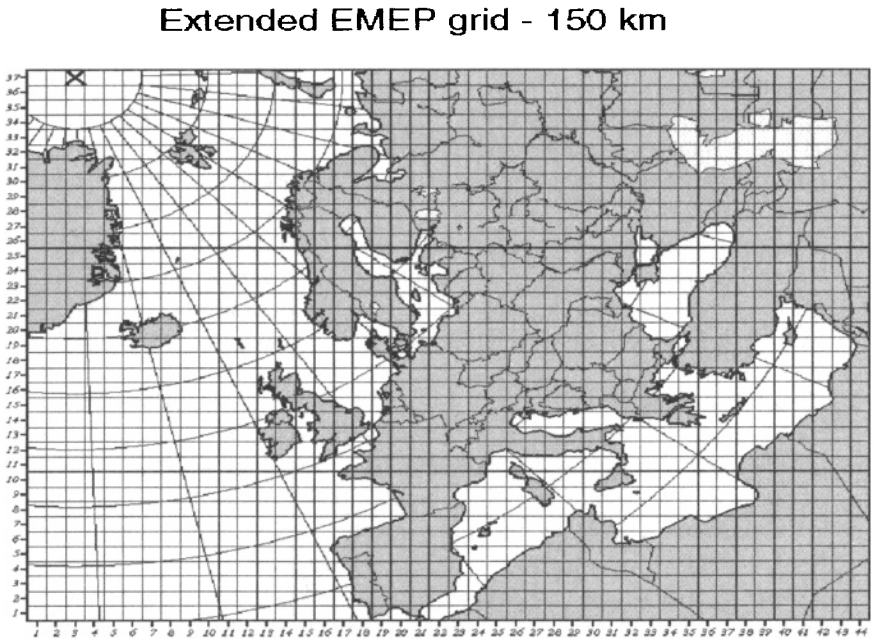

Figure 1: Example of a two dimensional landscape in a grid system
Source: EMEP grid-cell from the EMEP server (www.emep.int)

(iii) a set of outputs is defined as $\mathbf{S}$. The amount of outputs is the result of the regeneration processes and depends on the amount of the inputs distributed in the grid of space. Furthermore, the amount of output in a specific region L_m may depend not only on the amount of ecological goods in that area, but also on the amount in other areas. Thus, the relationship is as follows[2], $\mathbf{S}^m = R^m_{St}(Z^1, ..., Z^M)$. For example, in a river, we have the process related to the water that flows from area L_1 to area L_2, and the population of fishes that grows and regenerates.

(iv) the feedback in the ecological system implies that the set and the amount of outputs form the new inputs in the next phase. For an analysis in a temporal dimension, some relationships for the ecological feedback should be determined.

2.3 The economic system

The economic system operates in the same landscape $\mathbf{L}$, and consists of:

(i) a set $\mathbf{K}$ of elements of inputs. The distribution of the amount of these inputs over

[2]It should be noted that this notation only states that there is some kind of relationship. However, it does not assume an a priori functional relationship between the amount of input and the amount of output. The linearity of the system depends on the specification in a later phase.

the spatial landscape $\mathbf{L}$ may be denoted by the matrix $\mathbf{K}^m$, i.e. the amounts of each of the economic inputs $\mathbf{K}$ available in the grid of areas (L_m);

(ii) a set $\mathbf{J}$ of production processes, which describes the transformation of the amount of a set of inputs ($\mathbf{K}$) into the amounts of a set of outputs ($\mathbf{C}$). The distribution of the production process over the spatial landscape ($\mathbf{J}^m$) could either be homogenous or heterogeneous as well as being endogenously determined or exogenously given. It should be noted that for the production process, the inputs could be imported from other areas, i.e. $\mathbf{C}^m = \mathbf{J}^m(\mathbf{K}^1, ..., \mathbf{K}^M)$;

(iii) the set of outputs from the production process ($\mathbf{C}$) may be divided into intended – i.e. economic – outputs ($\mathbf{Q}$), and unintended – i.e. ecological – outputs ($\mathbf{E}$);

(iv) objective and feedback. The economic system functions in an environment partly created by goals determined by the economic subjects. In economic theories and in modeling terms, this is given by the preference or utility function of consumers $\mathbf{U}_I^m$, which stands for the matrix of the preferences ($\mathbf{U}$) of a set $\mathbf{I}$ of consumers in the grid of areas ($\mathbf{L}_m$). The utility function of consumers from a specific area L_m may depend on the amount of consumption of the outputs from all the areas and the presence of ecological goods which are not used as inputs in the economic system. However, the effective demand for the goods may depend on the economic inputs located in the area L_m ($\mathbf{K}^m$) and its production process ($\mathbf{J}^m$). Thus, we have preferences $\mathbf{U}_I^m(\mathbf{C}^m, \mathbf{S}^m|\mathbf{J}^m(\mathbf{K}), \mathbf{R}^m(\mathbf{Z}))$. In a temporal dimension, both the feedback in the economic system and the way the utility function evolves in time should be determined.

2.4 Economic-ecological interaction

Interaction, which is defined as a reciprocal action or influence, exists *within* the economic and the ecological system separately, as well as *between* the economic and the ecological system. The interaction between the economic and the ecological system may take place in the following ways (see also Wang, Nijkamp and Verhoef (2001)):

(i) a part of inputs in the economic system $\mathbf{K}_z \subseteq \mathbf{K}$ is retrieved from the elements of the ecological system, i.e. $\mathbf{K}_z \subseteq \mathbf{Z}$. Thus, $\mathbf{K}_z$ is the intersect of set $\mathbf{Z}$ and set $\mathbf{K}$ (or: $\mathbf{K}_z \subseteq (\mathbf{Z} \cup \mathbf{K})$);

(ii) a part of the outputs from the economic system, particularly the unintended output ($\mathbf{E}$), forms elements in the ecological system, i.e. $\mathbf{E} \subseteq S$; or,

(iii) a set of elements of the ecological system $\mathbf{S}_U \subseteq \mathbf{S}$ directly influences the utility $\mathbf{U_I}$ of the consumers (set $\mathbf{I}$) in the economic system, i.e. $\mathbf{U_I}(\mathbf{S}^m) \neq 0$.

Though the interaction between the economic and the ecological systems could be given in conceptual form, as has been formalized above (a summary is given in Table 1), it should be clear that the complexity is too great to permit the development of an applied model of the ecological and economic systems' interaction which takes into account the many characteristics of this interaction. As should be clear from this conceptual model, the number of interactions and their complexity increase according to (i) the number of inputs, i.e. the set $\mathbf{K}$ and the set $\mathbf{Z}$; (ii) the intersect between both these sets; (iii) the number of ecological and economic processes, i.e. the set $\mathbf{R}$ and

Description	Sets	Regional subset	Elements in the set
spatial landscape	$\mathbf{L}$	$\mathbf{L}^m$	$(L_1, ..., L_M, L_s)$
ecological system			
inputs	$\mathbf{Z}$	$\mathbf{Z}^m$	$(Z_1, ..., Z_Z)$
regenerative process	$\mathbf{R}$	$\mathbf{R}^m$	$(R_1, ..., R_R)$
outputs	$\mathbf{S}$	$\mathbf{S}^m$	$(S_1, ..., S_S)$
economic system			
inputs	$\mathbf{K}$	$\mathbf{K}^m$	$(K_1, ..., K_K)$
production process	$\mathbf{J}$	$\mathbf{J}^m$	$(J_1, ..., J_J)$
outputs	$\mathbf{C}$	$\mathbf{C}^m$	$(C_1, ..., C_C) =$ $(Q_1, ..., Q_Q, E_1, ..., E_E)$
intended outputs	$\mathbf{Q}$	$\mathbf{Q}^m$	$(Q_1, ..., Q_Q)$
unintended outputs	$\mathbf{E}$	$\mathbf{E}^m$	$(E_1, ..., E_E)$
consumers	$\mathbf{I}$	$\mathbf{I}^m$	$(I_1, ..., I_I)$
objective function	$\mathbf{U}$	$\mathbf{U}^m$	$(U_1, ..., U_I)$
E-E interaction			
inputs	$\mathbf{K_Z} \in (\mathbf{Z} \cup \mathbf{K})$	$\mathbf{K}_\mathbf{Z}^m$	$(K_1, ..., K_K) \cup (Z_1, ..., Z_Z)$
outputs	$\mathbf{E_S} \in (\mathbf{E} \cup \mathbf{S})$	$\mathbf{E}_\mathbf{S}^m$	$(E_1, ..., E_E) \cup (S_1, ..., S_S)$

Table 1: A summary of symbols and notations

J; and (iv) the number of consumers, i.e. the set (**I**). This complexity exists not only on the level of interaction between both systems, but also on the level of interaction within each separate system.

3 Spatial Landscape, Spatial Flow and Spatial Externality

3.1 Introduction

In this section, we will categorize various kinds of spatial externalities that arise from the conceptual framework of interaction between the economic and the ecological framework, with spatial characteristics. Roughly speaking, an externality exists if there is divergence between the marginal social costs and the marginal private costs for the economic subjects, i.e. if agents' actions that influence the utilities of other agents are not properly reflected in price signals. In the optimal case – i.e. no divergence exists between the marginal social costs and the marginal private costs for the economic subjects – climate change issues would not be so prominent, because, in the objectives of the economic subjects, the sustainability of the ecological system is already taken into account (assuming that marginal social costs would go to infinity if sustainability conditions were violated). However, due to the complexities of the ecological system and due to some characteristics of the ecological goods, i.e. the non-rivalry and the non-excludability properties of, for example, the air, these goods are not properly priced.

In essence, externalities apply to the level of economic decision units, which are the individual consumers and/or the individual producers. This implies that, strictly

speaking, every externality has some spatial characteristics if spatial interaction is taken into account. However, if the objects of study are not the individual decision-makers – i.e. individual producers and individual consumers – but are at a more aggregated level, for example, on the sectoral or regional level, the characterization of an externality within the aggregated units of decision makers may complicate the taxonomy of spatial externalities.

In accordance with the description of externalities, as discussed in Wang, Nijkamp and Verhoef (2001), a *spatial externality* implies that (i) some agents' action from a spatial area L_m influences the utility of other agents in other spatial area(s) $L_{s\neq m}$, without this effect being reflected in price signals; and (ii) the necessary conditions for a socially-optimal situation (i.e. the Pareto-optimality conditions) are violated (see, for example, also Papageorgiou, 1978a, b). Furthermore, we will speak of a *localized externality* if the effect takes place in the same spatial area, given the spatial aggregation.

3.2 Spatial landscape and regional borders: exogenous or endogenous?

For applied research on spatial externalities, usually, it is necessary to define a set of borders that identify the set of spatial units distinguished (here called 'regions', see also e.g. Isard, 1956, Beckmann, 1978, Fujita, 1999). The issue of the determination of the borders of regions is an important one, as many natural borders may, as a result of economic and political processes, cease to exist; or may become either more or less relevant for economic purposes over time. Therefore, an important question is whether the border of a region should – based on some criteria – a priori be determined, or whether this regional border should – as a function of economic processes – be endogenously determined.

In reality, the landscape $\mathbf{L}$ is a continuous space covered by heterogeneous characteristics, e.g. geographical differences, differences in the distribution of ecological resources, cultural, political and economic differences. The heterogeneous landscape may be decomposed into various areas that are assumed homogenous in some respects. Difficulties will arise when the landscape is heterogeneous in a number of characteristics, such that the borders of the areas in these categories differ from each other and could not a priori be determined. This is what Mennes et al. (1969) called *defining the space units*. In the case of an exogenous determination of the borders, we normally take the political units, e.g. states, regions, cities, villages, as the areas.

For some research questions – e.g. the optimal market area or location questions – the borders and the areas should be endogenously determined by the economic decision processes (von Thünen, Launhard, Weber). Where a firm – either agricultural, industrial or service – locates and how firms interact with each other will determine the economic border. As Lösch (1938, 1953) pointed out, even in a vast plain with an equal distribution of raw materials and a complete absence of any other inequalities, spatial differences would still arise in a certain order, and a specific form of areas and borders is the most efficient one (i.e. a hexagon).

Thus, the operational determination of the space units depends, as Nijkamp (1987) pointed out, on the research questions. When, as in our study, data availability dictates

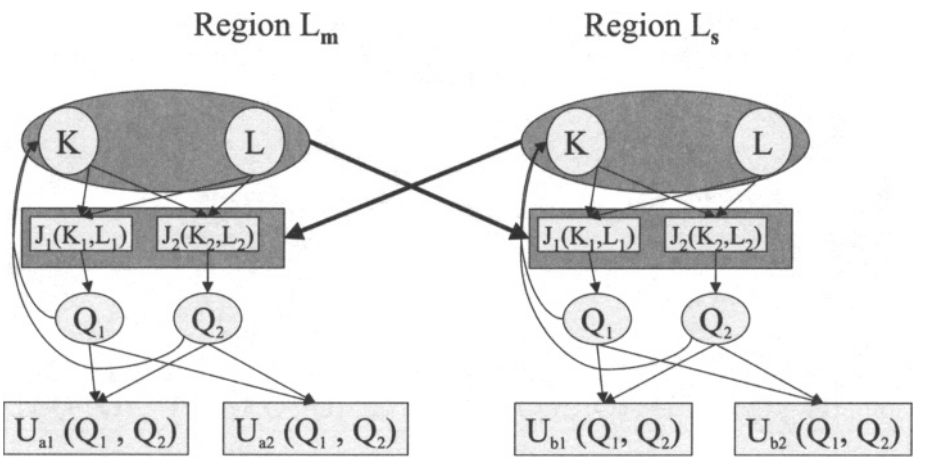

Figure 2: Spatial-economic interaction at input

the definition of space units, there is not much scope for the endogenization of borders. Nevertheless, one ought to be aware that issues of endogenous borders then cannot be captured in the modeling framework, which may imply limitations on the level of generality of the work carried out.

3.3 Spatial-economic interaction

Spatial interaction in the economic system, which is broadly defined in the literature as the flow of goods, people, or information between places that results from a decision process (Fotheringham and O'Kelly, 1989), plays an important role in analyzing spatial externalities.

In terms of the conceptual model, we may categorize spatial interaction in the economic system as follows:

(i) interaction at the level of inputs ($\mathbf{K}$), i.e. the input $\mathbf{K}^m$ from region L_m is demanded by the producers $\mathbf{J}^s$ at other locations $L_{s\neq m}$. This is addressed by the migration theories and international resource-use literature, as shown in Figure 2;

(ii) interaction at the location of the prodution process ($\mathbf{J}$), i.e. the location of producers $\mathbf{J}^m$ at the location L_m results from economic decision processes. This is shown in Figure 3 and is addressed by locational choice literature in the spatial economics and game-theoretic literature;

(iii) interaction at the level of outputs ($\mathbf{Q}$) or the objective (consumer's utility $\mathbf{U}$), i.e. the same amount of outputs of the products $\mathbf{Q}^m$ from region L_m is demanded by the consumers $\mathbf{I}^s$ from other regions $L_{s\neq m}$. This is shown in Figure 4 and is addressed by, for instance, the international trade theories;

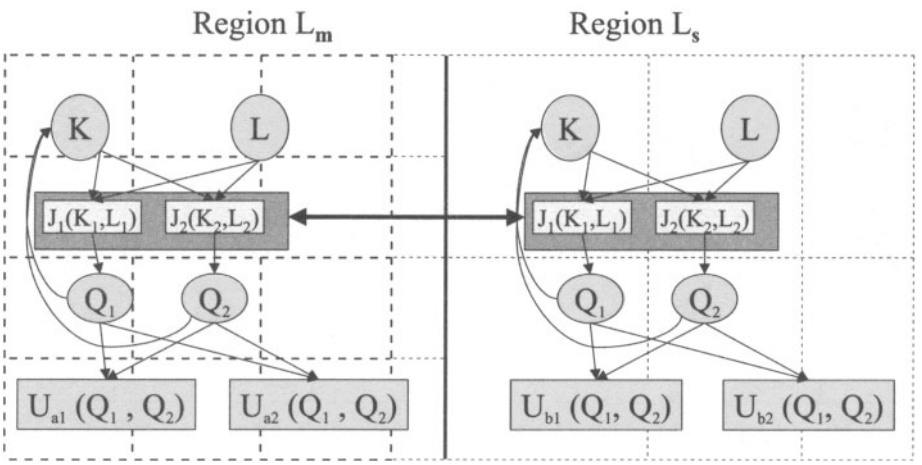

Figure 3: Spatial-economic interaction at the production level

(iv) a combined interaction of these factors (i-iii above). This is addressed by, for instance, the economic geography and spatial economics literature.

In the spatial economics literature, the ideal approach is, given an initial situation, to determine the spatial constellation of the economic system by considering interactions at the level of inputs, at the location of the production process, and the level of outputs in the context of general locational theory (Isard, 1956). From the traditional spatial-economic literature, which mainly concerns the locational behavior (e.g. von Thünen, Weber, Lösch, Isard, and, for an overview, see Beckmann and Thisse, 1986, Birkin and Wilson, 1986a, b, Isard, 1956, 1990a, Kilkenny and Thisse, 1999, Nijkamp, 1976, Sohns, 1978), it was also recognized that, after the firms have determined the location of the production process, the location itself would in its turn then affect the prices and flows of commodities which are important in the interaction of the inputs and the interaction of the outputs.

This influence is taken into account in the spatial general equilibrium model (Takayama and Labys, 1986, van den Bergh et al., 1996), which models complete and complex spatial interaction of the economic system in a spatial setting. However, analytical results sometimes may not easily be interpreted and empirically tested, e.g. because of a chaos-type of outcome, i.e. bifurcation of the results (see e.g. Nijkamp, 1987, Nijkamp and Reggiani, 1998).

3.4 The nature of spatial-ecological interaction

The ecological system relating to the aspect of climate change is a complex system and the diversity of interactions within this ecological system is also the reason for the complexity in the study of climate change: on the one hand, there is human action so that the dynamics in the ecological system could be characterized as interac-

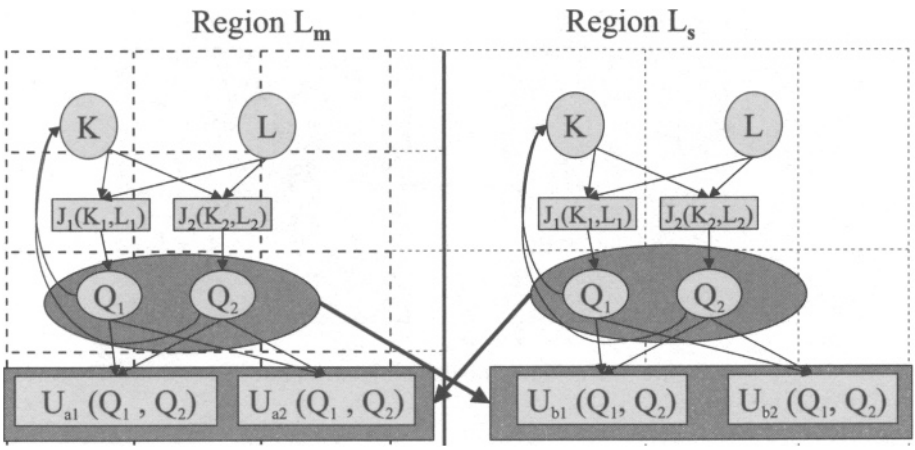

Figure 4: Spatial-economic interaction at output level

tion between the ecological and the economic system; on the other hand, there are some internal dynamics of the ecological system itself: the interaction within the subsystems of the ecological system has its own dynamics. For example, it covers not only the carbon cycle and external solar influences, but also the atmospheric process, the ocean, the terrestrial vegetation and the inland glaciers (IPCC, 1997). The study of interactions between both systems is a considerable study which is carried out by international programs (see e.g. Hibbard et al., 2001).

From a spatial point of view, we may, by generalizing the interacting processes of several subsystems within the ecological system, divide the spatial interaction of the ecological subprocesses in terms of:

(i) a localized ecological subprocess, i.e. there is no interaction between any part of the ecological subprocess in region L_m and regions $L_{s\neq m}$. This concerns immobile ecological systems, such as a mountain area, a forest, a lake, or an island;

(ii) a uni-directional ecological subprocess, i.e. the ecological subprocess in region L_m will have effects on the ecological process in other regions ($L_{s\neq m}$). This is depicted in Figure 5. An example of this is the waterstream in a river, which could be interpreted as flows in the perspective of spatial areas L_m and $L_{s\neq m}$; and

(iii) a multidirectional ecological subprocess, i.e. the ecological subprocess in regions L_m and $L_{s\neq m}$ are spatially interrelated, such that the ecological subprocess in both regions may be perceived as one common ecological subprocess. This is depicted in Figure 6.

From the perspective of economic subjects in an individual area, for example $\mathbf{I}^m$, both the unidirectional and multidirectional ecological subprocesses may cause spatial

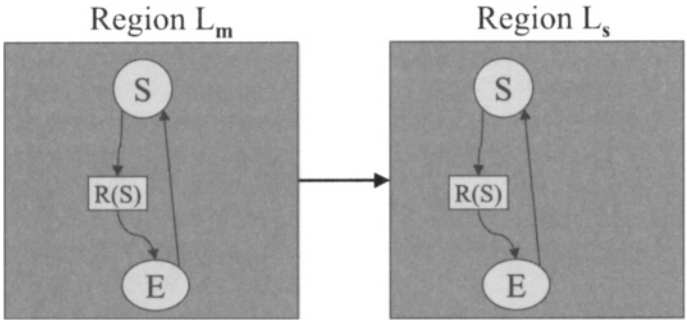

Figure 5: Uni-directional ecological subprocess

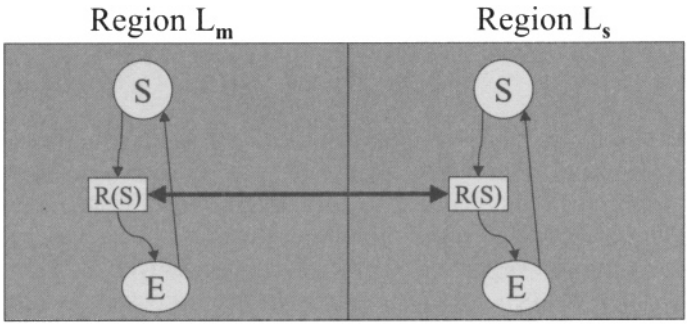

Figure 6: Spatially-interrelated ecological subprocess

externalities for economic subjects from other spatial areas, i.e. $\mathbf{I}^{s \neq m}$. This will be further discussed in the next subsection.

3.5 A taxonomy of spatial externalities

One of the main elements of the spatial aspects of the interaction between the ecological system and the economic system is that the mapping of borders in spatial area L_m differs for each system. In other words, an ecological process does not necessarily observe the borders between the regions or countries in the economic systems. Given that (i) the intersecting elements in the interaction between the ecological and the economic systems are known, and that (ii) the border of spatial areas L_m are determined, we have the following taxonomy of spatial externalities:

(i) spatial-localized effects

Spatial-localized effects are environmental externalities within region L_m itself and these externalities do not affect other regions. This form of externality arises as a result

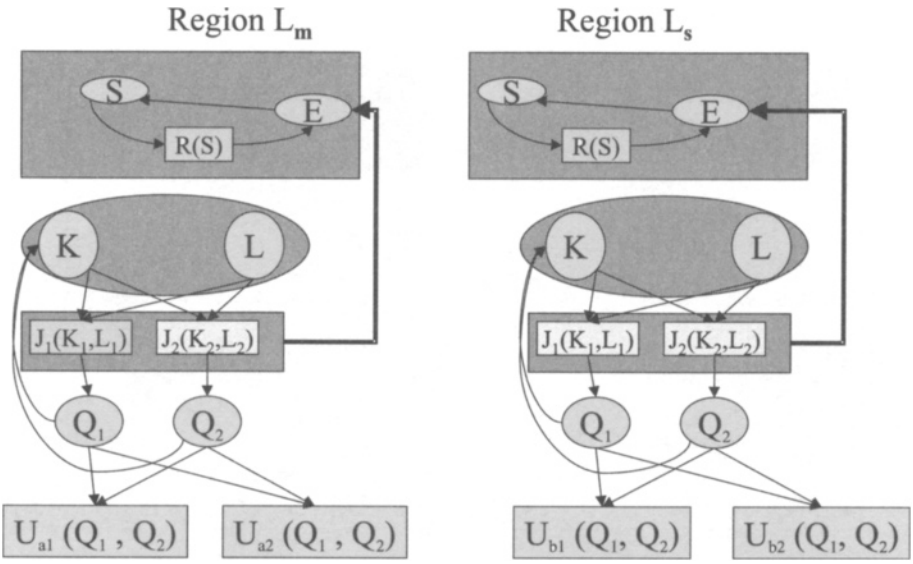

Figure 7: Localized externalities

of the aggregation of individual consumers and individual producers into a regional decision level. Figure 7 presents a graphical description of this form of externality.

(ii) spatial-economic environmental externalities

Spatial-economic environmental externalities are environmental externalities which are intensified by spatial-economic spillovers. This is known as 'point-externality' and is depicted in Figure 8. This may be subdivided into the following two categories:

(ii.a) localized environmental externalities in one (source) region.

In this category, environmental externalities occur only in the source region, i.e. localized ecological factors which are not properly priced in region L_m. For the source region, spatial-economic interaction may exacerbate environmental degradation, e.g. trade in garbage and waste disposal (see e.g. van Beukering, 2001). In this case, the source region (L_m) bears the impacts of spatial interaction because of localized environmental externalities.

For other region(s), i.e. $L_{s\neq m}$, spatial-economic interaction will cause a pecuniary environmental externality, as spatial-economic interaction with localized externality in region L_m may disturb the existing equilibrium in region(s) $L_{s\neq m}$. However, the market in region(s) $L_{s\neq m}$ would adapt and result in an optimal outcome in region(s) $L_{s\neq m}$ For example, in the case of tropical wood, a local environmental externality in region L_m results in a lower market price for this good, which, because of the spatial economic interaction, will in turn result in a higher demand for the import of tropical wood from region L_m in other region(s) $L_{s\neq m}$. As a consequence, region(s) $L_{s\neq m}$

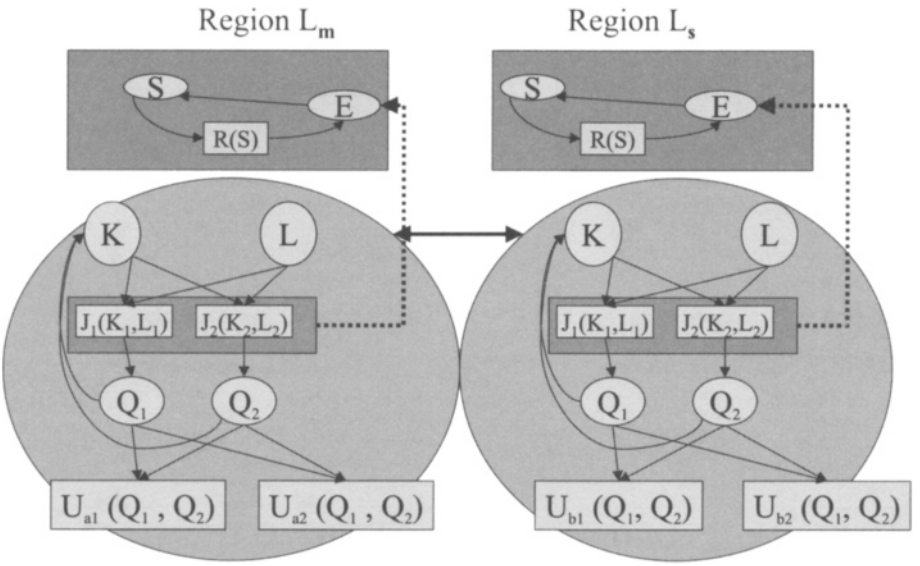

Figure 8: Spatial-economic environmental externalities

face a higher amount of waste caused by the higher demand for tropical wood. Then, because there is no externality in region(s) $L_{s\neq m}$, either a higher price for waste will cause a lower demand for tropical wood, or there will be trade-off between more waste and more demand for tropical wood.

(ii.b) localized environmental externalities in more regions

In this category, localized environmental externalities occur in both the source region (L_m) and in the other region(s) ($L_{s\neq m}$). In this case, the destination region(s) would also be faced with the impacts of the local environmental externality in the source region (L_m) – i.e. localized ecological factors which are not properly priced in region L_m – causing environmental degradation in region(s) $L_{s\neq m}$ through spatial-economic interaction.

The former example of tropical wood and waste illustrates this: the environmental externality in waste, for example, will not lead to a re-evaluation in the destination region(s). This environmental externality exacerbated by spatial-economic interaction is no longer a pecuniary externality, but a technological externality. Because of the higher amount of waste in region(s) $L_{s\neq m}$, the market in region(s) $L_{s\neq m}$ will not adapt and, hence, the result will not be an optimal outcome, i.e. the higher waste will cause environmental degradation. Thus, the environmental externality caused by spatial-economic interaction is a technological one and we will call this category of spatial externality 'spatial-economic environmental externality'.

One important issue here involves the causality of spatial-externality. As we may deduce from the illustrations above, spatial-economic interaction is not the cause of

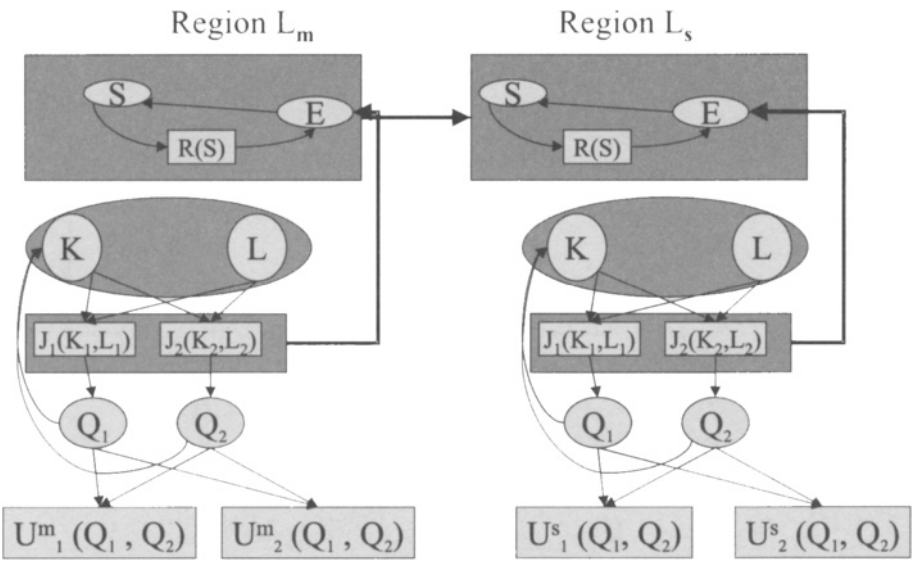

Figure 9: Transfrontier spatial-ecological environmental spillover

the externality, but it serves as an intervening variable for the spillover of the environmental externality. If the destination region has no localized externality, spatial-economic interaction will not cause technological spatial externality in the destination region(s), while if there is localized externality in both regions, there will be technological spatial-externality.

(iii) spatial-ecological environmental externalities

Spatial-ecological environmental externalities are environmental externalities intensified by spatial-ecological spillover. This category is the best-known form of spatial environmental externality. The term 'transboundary pollution' is mostly used in international trade literature, i.e. an internally inefficient price system in region L_m causes, because of the ecological system, environmental degradation in other regions. This category may be subdivided into the following subcategories (see e.g. Siebert, 1985):

(iii.a) unidirectional spatial-ecological environmental externalities

In this category, the ecological system is transfrontier, i.e. it transports pollutants from one region (L_m) to (an)other region(s) ($L_{s\neq m}$). This involves uni-directional ecological systems, e.g. the pollution from one region at the source of a river will affect a few other regions downstream. This is depicted in Figure 9.

(iii.b) multidirectional spatial-ecological environmental externalities

In this category, the ecological system of the regions is a spatially-interrelated system. Other terms are: global environmental system or common resource system. The environmental goods in this system have non-exclusive and non-rival characteristics. This

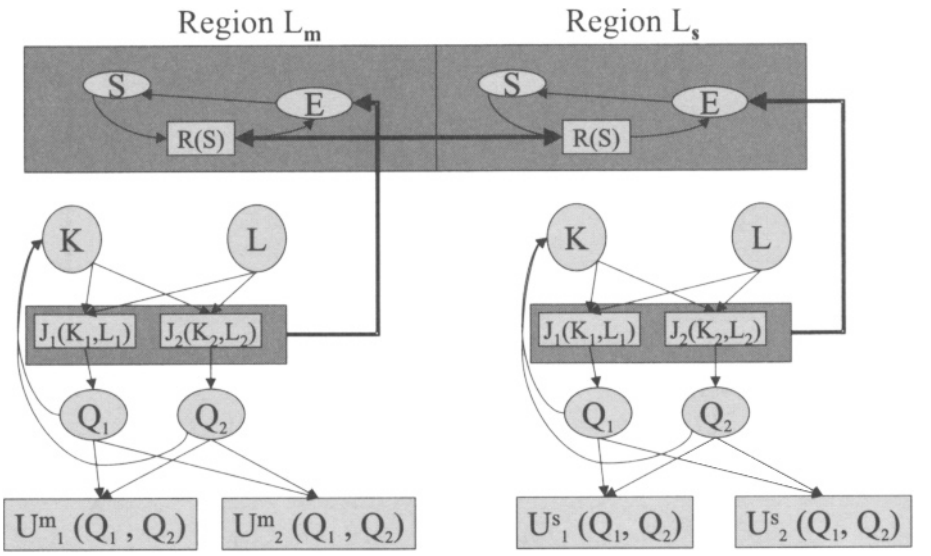

Figure 10: Multidirectional spatial-ecological environmental externalities

is called a global public good, e.g. clean air. This category of spatial-externality may be depicted as in Figure 10.

(iv) general spatial environmental externalities

General spatial environmental externalities are environmental externalities intensified by both spatial-economic interaction and spatial-ecological interaction. In this category, environmental externality in region L_m causes environmental degradation in other region(s) $L_{s\neq m}$, through both spatial-economic interaction (e.g. trade in energy) and spatial-ecological interaction (e.g. carbon emission).

The externalities may, as known from the environmental economic literature (Baumol and Oates, 1987, Opschoor and Vos, 1989, Opschoor et al., 1994), be internalized through (i) the equalization of the marginal private costs to that of the marginal social costs using a Pigouvian tax; or (ii) a property rights system, such that the equalization of both costs could be achieved by the parties involved themselves. Both instruments have their own shortcomings and restrictions (see e.g. Cropper and Oates, 1992). For the tax instrument, for example, a Pigouvian tax in the case of public goods is not easy to formulate and the willingness to pay and willingness to accept as compensation may differ for the same amount of pollution (see e.g. Bishop and Woodward, 1995, Bockstael and McConnell, 1993, Hanemann, 1991). Property rights also work badly for public goods because of their non-exclusivity and non-rivalry characteristics (see Baumol and Oates, 1987, Hanley et al., 1997, Samuelson, 1952). However, in the complete information case, both instruments result in the same equilibrium (see Baumol and Oates, 1987, Mas-Colell et al., 1995). Externalities characterized

by spatial-economic or spatial-ecological spillover are, however, even less easily internalized. Though the instruments are the same, spatial environmental externalities require cooperation between areas with different political orientations and objectives. In the next section, we categorize the literature that investigates these instruments and a priori presumes that such cooperation may arise.

4 Categorizing Spatial Externalities in Spatial Interaction Models

In this section, the literature on spatial externalities in spatial-economic interaction models will be categorized. Section 4.1 discusses the differences between a few approaches in terms of the way of linking the regions and the representation of the spatial characteristics in the models. Section 4.2 treats the spatial-interaction via the output. In this subsection, we consider the international trade approach (distance-less) as well as the spatial-price equilibrium approach (with transport costs). Spatial-interaction via the production process is treated in Section 4.3 by discussing locational choice and locational competition theories. However, for analyzing the internalization of the spatial externalities, the locational competition (or strategic interaction) models assume a distance-less space. Section 4.4 gives a review of interaction via the inputs; special attention is paid to the consumer-labor location and the literature on capital flight. Section 4.5 discusses a few multi-region models and the associated literature for internalizing spatial externalities.

4.1 Linking the regions: distance and transport costs

Before discussing the models of spatial-economic interaction, it is important to stress the treatment of space and distance in the literature. The decomposition of the continuous space of landscape $\mathbf{L}$ into homogenous units does not necessarily mean that there is no distance left between the areas L_m. In other words, (i) the connection between two areas may not only take place at the border, and (ii) the activities between both areas may be linked with each other through one or more routes from somewhere in the area L_m (Isard, 1990a).

This aspect is one of the major differences between the international trade literature and that of spatial economics (see Isard, 1956, Krugman, 1990, 1993, Fujita et al., 1999). In the traditional theories of international trade, where the spatial interaction, i.e. the flow of commodities between countries, is analyzed, the spatial landscape is typically decomposed, i.e. each nation is represented as a spot in the landscape $\mathbf{L}$ without measurable distance between the nations/regions. In such a distance-less representation of the economic system in a spatial setting, the inputs and the outputs from each area should necessarily be different, otherwise these goods will be perfect substitutes (see Isard, 1956, Krugman, 1990, 1993).

This spaceless treatment by international trade theory is criticized by Isard (1956), i.e. the distance between each nation/region is assumed to be equal and negligible. According to Isard (1956), distance affects the ultimate costs in the case of spatial-economic interactions. In this sense, the commodities produced in different regions would become imperfect substitutes: the characteristics of the goods are the same, but

not their location. Commodities from different space-areas could be substituted, and international trade may result in specialization. However, transport costs may prevent this from happening.

Armington (1969) tried to incorporate the imperfect substitute characteristics between the commodities from different countries, without explicitly taking the transport costs and distances between the countries into account in the model, by using substitution elasticities. A more explicit modeling of distance between two regions is represented by the gravity type of models, in particular, and the spatial interaction models, in general. These models assume that distance will directly affect the magnitude of interaction between the regions. The analogy is drawn from physics where masses attract each other (Isard, 1972). In the trade analyses in terms of gravity models, the trade flow between regions L_m and L_s is a positive function of the mass of the region, e.g. measured by the GDP, but is negatively correlated to the distance between both regions (see Fotheringham and O'Kelly, 1989, Isard, 1990b, Reggiani, 1990).

In the spatial economic literature, transport costs and distance are used in the models to reflect the spatial character. One often-used approach is the 'iceberg transportation costs'-approach (Samuelson, 1952). In this approach, transportation costs between the regions are represented as increasing functions of the distance. In the standard formulation, it is assumed that for every unit of distance, a given proportion of the (remaining) quantity of the good shipped 'evaporates' to reflect transportation costs. This is used, for example, by Krugman (1991a, b) in economic geography, where the agglomeration effect of regions is analyzed. Isard (1956), on the other hand, also introduced the term 'transport-unit', which directly includes transport as an input in the production process. Consequently, more transportation, keeping other inputs fixed, would result in a higher output.

4.2 Interaction through output

The international trade approach

The traditional international trade theories, such as the Ricardian comparative costs advantage and the Hecksher-Ohlin (HO-)model of trade, treat the output from the production process as mobile between the 'regions', while the production factors are assumed to be immobile between the regions. With respect to the initial location of the production process, the Ricardian approach assumes different (heterogeneous) production technologies between the regions, while the HO-model assumes identical (homogenous) production technologies across the regions. In both approaches, the endogenous character of locational behavior is implicitly assumed. In Ricardian models, differences in the production technologies lead, in the case of international trade, to specialization of that industry which has a comparative cost advantage. In other words, agglomeration in Ricardian models implicitly takes place in that industry which has a comparative cost advantage over another industry. In HO-models, relative factor endowments work, through price mechanisms in the output market, to produce an optimal allocation of factor endowments in the sectors. In other words, agglomeration depends on factor endowments and the relative preferences of the consumers for

the outputs (see Obstfeld and Rogoff, 1996).

The standard textbook examples are two-country models with two production factors (labor and capital). However, by including the environment as a production factor, Siebert (1981) extended the HO-model and argued that a country with relatively rich environmental endowments (e.g. natural resources) will have a comparative advantage in environmental intensive products and will thus produce and export these commodities. Van Beers and van den Bergh (1999) have discussed possible ways of extending the standard trade model for environmental research, including adding the role of utility and technological issues.

The impact of internalizing the spatial externality in the presence of trade

In the presence of spatial externalities – in the form of spatial-economic spillovers or spatial ecological spillovers – as a result of free trade, the generally accepted insight that international trade will have mutual benefits for all free-trade countries could be reversed, as this insight would implicitly assume that the environmental prices are incorporated in the market prices, i.e. that the externalities are internalized. The literature on this subject is enormous (see e.g. Rauscher, 1991, 1997, Siebert, 1981, van den Bergh and Nijkamp, 1995, Withagen, 1998, 1999). Therefore, our discussion only highlights some important conclusions on this subject.

The environmental effects of international trade could, according to the NAFTA Commission for Environmental Cooperation (1996), be boiled down into three effects: (i) product effects, (ii) scale effects, (iii) structural effects. The first of these effects could occur in the home market, if it gets environmentally superior goods that would not be supplied if there were no international trade (e.g. low-pollution engines), or – conversely – if it gets some environmentally harmful products, such as 'polystyrene' packaging for fast foods. The scale effect has to do with the expanded market when international trade occurs. The lower costs as a result of scale effects could lead to an ever-increasing use of natural resources, such that economic development becomes unsustainable. The structural effect concerns the patterns and processes of production that could be affected by the specialization process that will unavoidably arise after international trade: this could – like the product effect – be positive or negative.

Theoretically, in the case of spatial-economic environmental externalities – i.e. spatial economic spillover because of local environmental externalities – international trade would not necessarily be beneficial for countries that have not internalized the environmental externalities (Verbruggen, 1991, 1999). The 'first-best' assumptions that underlie the benchmark result of the positive effects of trade are violated, as an environmental externality exists. It turns out that, because of the existence of negative domestic environmental externality, the country would seemingly have comparative advantage in the environmental goods causing the externality. Consequently, Copeland (1994) analyzed the policies for a small, already polluted country in the case where the environment is a local public good and concludes that there would be environmental degradation.

In the absence of technological progress, the 'pollution-haven hypothesis' (Copeland and Taylor, 1999) may occur, as some countries would set too low a standard

for environmental policy in order to strengthen the competitiveness. However, it is also possible that the 'ecological dumping hypothesis' or the 'pollute-thy-neighbor strategy' (see e.g. Dean, 1992) is relevant, i.e. countries would dump the polluted industries in other countries. An example is the already-mentioned study on waste-disposal (van Beukering, 2001). In this example, it is shown that a developing country attracts polluting industries as it is 'competitive' in the pollution-intensive sector.

In the case of spatial-ecological spillover, singly or unilaterally internalizing the externalities is more difficult, as the loss of competitiveness (if we keep assuming that there will be no technological progress) is still present, while the gains from this policy would, to a certain extent, flow away as the result of spatial-ecological flow. Transboundary pollution may be analyzed in cases that where the pollution affects (i) the utility of the consumers (e.g. Rauscher, 1997); (ii) the ecological system as it is (e.g. Copeland and Taylor, 1995); and, (iii) the productivity of the producers (e.g. Benarroch and Thille, 2001).

The trade-based literature considers the effect of free trade on the quality of the environmental system[3]. On the one hand, because of the pollution-haven hypothesis or ecological dumping possibilities, there is the argument that trade would adversely affect the ecological system of some countries and allow rich countries to export their environmental problems to poor countries (see e.g. Baumol and Oates, 1987, Siebert, 1985). The 'trade-induced degradation hypothesis' suggests that international trade can play a key role in initiating a vicious cycle, in which trade-induced environmental degradation could lead to income losses (Copeland and Taylor, 1999). Moreover, these income losses can then lead to further degradation (Daly, 1995). On the other hand, Grossman and Krueger (1991) argue that growth may improve environmental quality, as trade promotes income growth and thus could lead to a cleaner environment. This is, in short, also the relationship that is studied by the empirically oriented studies on the environmental Kutznets curve (see de Bruyn, 1997, de Groot, 1999). Copeland and Taylor (1999) show that, in situations where the costs of pollution are small in the short run, but large in the long run the environmental degradation hypothesis would prevail.

However, technological progress would play an important role in overcoming environmental degradation in general (see e.g. van den Bergh and de Mooij, 1997). If there is environmental degradation as a result of spatial externalities, the Porter-hypothesis (Porter and van der Linde, 1995) argues that as a result of environmental regulations in the home country, firms could gain comparative advantages as they are forced to implement more advanced production technologies. Although technological progress is important in the endogenous growth literature, trade-based technological models are still rare[4].

[3]See Copeland and Taylor (1995) for further references on debates, empirical work, and theoretical issues on this subject.

[4]In a static model, it is straightforward to incorporate technological differences between the countries (see van Beers and van den Bergh, 1999).

Spatial price equilibrium

In addition to the international trade approach, the spatial price equilibrium models (see Samuelson, 1952, Takayama and Judge, 1964, Takayama and Labys, 1986) take transport costs into account. The spatial price equilibrium models are partial equilibrium models which analyze the equilibrium of prices in the interaction between the producers of goods, the consumers of goods and the shippers or traders of these goods in spatially-separated areas. The literature on environmental policy in a multi-regional context uses the spatial price equilibrium models to analyze the potential first-best or second-best policy rules (see Verhoef and Nijkamp, 2000, Verhoef and van den Bergh, 1995). In Verhoef and Nijkamp (2000), the spatial externality as a result of both economic and ecological spillover is discussed.

4.3 Interaction through decision process units: locational choice?

Generally speaking, the regional science-oriented literature explicitly analyses locational behavior from the viewpoint of spatial-economic interaction. The most intuitive way to analyze the interaction in a spatial context is to focus on firms that are settled or may settle in a certain area. The questions in this context are then: 'Where would a firm locate?' and 'Given the locations of some firms, how would these firms interact with each other?'.

In the most simplified locational analysis, the spatial diffusion of both the inputs and the demand for the output is assumed to be exogenously given. Then, the locational choice of an individual firm could be interpreted as the choice of a producer j for a location in the spatial area L_m which would maximize his objective, i.e. his profit.

Von Thünen analyzed the locational behavior and spatial distribution of agricultural firms in a homogenous area and came up with concentric rings of agricultural firms around a center. In his model – among other things – transport costs, distance and the input – viz. land – are immobile. These are important assumptions for a firm's location decision (see e.g. Beckmann and Thisse, 1986, Kilkenny and Thisse, 1999).

On the other hand, for the industrial firms, the location of the inputs need not necessarily be fixed. Weber analyzed the locational choice of such a firm, which is able to transport the input and the output. In this situation, locational choice is not bounded by the spatial distribution of the inputs; the transport costs and distances of both input and output also play an important role. In Weber, as the location of the inputs and the demand for the output are given, the problem boils down to a minimization of the transport costs (see Birkin and Wilson, 1986b, Isard, 1956).

Hotelling analyzed the locational behavior of more firms (see e.g. Birkin and Wilson, 1986a, Greenhut et al., 1987, Isard, 1956). A well-known example is the competition between the location of two ice-cream sellers on a beach. Under certain conditions, it seems that agglomeration may prevail, i.e. both sellers would locate near each other to serve the whole market, while, under other conditions, location at both ends of the market may be the equilibrium. The result also depends on the strategy that is chosen by the sellers.

The precursor of city models may also be found in Lösch (1953), who sought the optimal locational structure of firms, given some forms of strategic behavior of the firms, and given the initial situation of the spatial distribution of input factors as well as of the consumers. According to Lösch (1953), the optimal locational structure of a number of firms in a homogenous plain is a hexagonal structure. This structure may, given other settings of spatial distribution of inputs, be adjusted (Greenhut et al., 1987).

The result of a firm's locational decisions would affect the spatial economic behavior, such as the price of the input and output as a result of rent-seeking behavior. In this sense, Lösch's idea of optimal locational structure may be seen as a step prior to the analysis performed by the network models (van den Bergh and Nijkamp, 1995). It is, however, a challenge to integrate both approaches, as, in reality, a homogenous plain on a large scale is difficult to find, so that it is difficult to test whether the network structure is optimal or not.

Given the optimal locational structure, the network models represent the whole spatial constellation by a network of nodes, for which the interaction is accomplished by links. As van den Bergh and Nijkamp (1995) pointed out, in the modern network models, the nodes and links "*. . .may reflect the real transportation infrastructure consisting of a complex network with ports and cargo terminals*". In the network models, an important role is assigned to the transportation sector. In this sector, the shippers and carriers will determine the equilibrium outcome. The shipper is a decision-making entity that desires a particular commodity to be delivered to a particular destination and the carrier is a decision-making entity that actually executes the transport of freight.

Result of internalizing the locational spatial externalities: strategic interaction

In the literature, strategic interaction models that incorporate both distance and locational choice are still rare. Therefore, the studies in this subsection only discuss strategic interaction models on locational choice as a result of environmental policy or in the case of transboundary pollution.

'Strategic interaction' forms a part of the game theoretic literature[5], which analyses the strategies of the economic agents – which could be an individual, a firm, a government or some other groups – to which some economic conditions (rules) apply. As there are some imperfections in the market structure – e.g. market power, externality, increasing returns to scale, incomplete information – strategic considerations become important. In the studies of environmental policy, the object of study is mostly the strategic interaction between governments and/or between firms. In the first case, governments in both countries will choose their tax rates knowing that the

[5]The works in this subsection differ from the 'modern' trade theory, which applies game theory to imperfect competition within the international trade system. Elements of imperfect competition could be: monopolistic or oligopolistic competition (thus the firms are no longer price takers); increasing returns to scale; multinational enterprises, etc. A unifying work on game theory and perfect competition may be found in the model of Keyzer and van Wesenbeeck (1999) that incorporates the game-theoretic elements of the imperfect competition in a general equilibrium framework in order to have a better understanding of the welfare effects of various environmental policies.

choice of the domestic tax rate will affect the world price and therefore affect production both in the home country and in the foreign country. In the second case, firms will decide whether to locate in some places or not, given some environmental policy set by the government. Some examples of the literature are: Kennedy (1994), Barrett (1994), Copeland and Taylor (1995), Markusen (1997), Ulph and Ulph (1994).

Kennedy (1994) analyzed a symmetric two-country model with a homogenous polluting product and symmetric oligopolistic industries with sunk costs for new entrants in the industries. He came to the conclusion that imperfect competition will lead to inefficient distortions of pollution taxes in the case of free trade. He decomposed these distortions into a rent-capture effect and a pollution-shifting effect. The first effect lowers the equilibrium taxes as each country attempts to gain a competitive advantage over its trading partner, and the second effect raises the equilibrium taxes as each country attempts to transfer production and its associated pollution to the other.

Markusen (1997) also analyzed a symmetric two-country model (so that the specialization effect is neutralized) with imperfect competition, extending that of Kennedy (1994) with two goods: one competitive and one imperfectly competitive with increasing returns to scale and imperfect competition. Furthermore, his attention is focussed on the locational choice of the multinational firm that could decide to have a plant in one or both countries. The conclusion is that when these multinational firms are confronted with free trade (by lower transaction costs from the removal of trade barriers), there will be a regime shift away from multinationals toward national firms. Furthermore, the firms headquartered in the home country have an incentive to close plants in the foreign country.

Barrett (1994) analyzed government strategy in the case of environmental policy in terms of weak environmental standards for industries that compete for business in imperfectly competitive international markets. He concluded that when the domestic industry consists of one firm, the foreign industry is imperfectly competitive, and competition in international markets is of the Cournot-type, then the domestic government has an incentive to impose a weak environmental standard. 'Weak' in this sense means that the marginal damage from pollution exceeds the marginal cost of abatement. However, this conclusion is not robust, because when the international competition is of the Bertrand-type, then the conclusion would be just the opposite. In addition, other kinds of policy, e.g. reduction of subsidy, would then function better.

Also interesting are Ulph and Ulph (1994) and Ludema and Wooton (1997). The first combines both the firms' and the government's perspective in a three-stage game. It also incorporates the R&D investments by the firms and thus allows both the government and the producers to act strategically. Ludema and Wooton (1997) took another perspective by analyzing situations where the importing countries are faced with negative externalities and try to find international trade rules to prevent 'trade wars' relating to the setting of environmental standards.

Though these models analyze the locational behavior of the firms in the presence of environmental externalities and/or as a result of policies aimed at internalizing the environmental externalities (either local or spatial externalities in the form of spatial-ecological spillover), they also have the shortcoming of treating space in a distanceless manner.

4.4 Spatial interaction through inputs

Spatial interaction through inputs may be interpreted as international trade in resources. In general, this may be analyzed in terms of comparative cost theory, where demand for the inputs depends on the price and availability of the goods. Therefore, pure interaction on the level of inputs is not widely studied, though we may interpret modern trade theory which analyzes the trade in intermediates as models analyzing the interaction through inputs. Despite this, there are three important elements of inputs which deserve our attention.

Consumers and labor location

Consumers play a dual role in the spatial-economic interaction analysis. On the one hand, consumers provide labor and other inputs (i.e. endowments) for the firms, which may be viewed as inputs in the production process. On the other hand, consumers exercise demand for the output of the firms. Both roles could, of course, be separated. In the commuter-traffic analysis, consumers are assumed to have choices concerning spatial distribution for the demand role and the labor role. The labor market and the migration literature, on the other hand, analyse the mobility of this input-factor in spatial areas.

Though this literature is very broad and covers urban economic research, the main focus is not on the spatial effects of environmental externalities (Verhoef and Nijkamp, 2000). Therefore, we will not discuss this literature in detail. However, from the consumers' perspective, there is a vast body of literature on environmental externalities. The problem is that most of this literature is not directly related to spatial externalities per se.

Capital mobility and capital flight

Another important area is the literature on 'capital flight' or 'foreign direct investment'. In the presence of environmental externalities, these studies investigate the implications of mobility of the production factors, notably capital. Some survey papers on foreign direct investment are: Beghin et al. (1994), Beghin et al. (1996), Copeland and Taylor (1995), Dean (1992), and Jaffe et al. (1995). In this literature, the industrial-flight and the pollution-haven hypotheses are tested. In the case of relocation of industries, there is a fear that the relatively low environmental standards in developing countries compared with industrialized nations will lead 'dirty' industries to shift their operations to these LDCs, i.e. the industrial-flight hypothesis. In addition, LDCs may purposely undervalue the environment in order to attract new investment, i.e. the pollution-haven hypothesis (see Dean, 1992). Although theoretically very interesting, the general empirical conclusion is that the capital-flight effect is quite small (see e.g. Dean, 1992, Bouman, 1998).

Land use and the ecological footprint

Another important input is the land. Spatial interaction concerning this input is, except in terms of rent differences, difficult because of the immobility of land in a spatial context. Because of this characteristic of land, we may interpret the literature on land use in terms of location theory, i.e. the function of some piece of land in terms of the kinds of firms located on the land.

Exactly because of this characteristic of land, the 'ecological-footprint approach' (Wäckernagel, 1998, Wäckernagel and Rees, 1996) uses land to indicate the magnitude of ecological degradation in a spatial context, i.e. the use of other natural resources causing environmental externality as if it were related to land use. The ecological-footprint approach uses the idea that there is spatial interaction of inputs worldwide. Thus, the economic processes in each area affect scarce environmental resources anywhere in the world. This approach is, however, a normative one (see e.g. Nijkamp and Finco, 2001) as the ecological footprint is measured as the ratio of the use of the scarce environmental resources in a particular area in comparison with a normalized amount of land for the resource concerned. Thus, a country's land use according to the ecological footprint may be far larger than the actual area of that country. This happens, for example, if the country uses many natural resources. This approach is, however, criticized by van den Bergh and Verbruggen (1999) because the procedure for measuring the ecological footprint is aggregated and biased. Furthermore, it does not recognize the advantages of spatial specialization and concentration. Finally, the ecological-footprint approach uses a hypothetical optimal land use.

4.5 Spatial interaction through multiple levels: input-output, spatial interaction and general equilibrium modeling

Multi-regional input-output models

The input-output model was first proposed by Leontief and assumes a fixed technological coefficient (and fixed production process) for the transformation of the amount of inputs and intermediates into the amount of outputs. Isard (1956) operationalized the input-output analysis for spatial-economic research by using the interregional and regional input-output table. Since then, multi-regional input-output model analysis is one of the standard methods of empirical spatial economics (Bröcker, 1998). For this approach, the border should, as in all other multi-regional models, be exogenously determined. Furthermore, this approach requires stability in the relative supply prices of each output produced by several regions, as the substitution effects (for the inputs) are not incorporated (Isard, 1990b). This also implies that the 'economic' distance between two regions should not change, i.e. that the transport routes and costs would not change between the regions. Bröcker (1998) furthermore raised the following objections: (i) multi-regional input-output models do not sufficiently take account of income-expenditure interdependencies; and (ii) multi-regional input-output models are one-sidedly demand-driven, such that effects coming from the supply side can not be modeled appropriately.

The input-output analysis is applied in environmental studies by treating, for ex-

ample, carbon emission as an input. The interaction between the ecological and the economic system is realized by treating the natural resources as inputs provided by the environment (see e.g. Isard, 1972, Siebert, 1981, 1985).

Multi-regional general equilibrium models

The same interregional input-output tables may also be used by the general equilibrium models. The additional advantage of general equilibrium models compared with the input-output approach is that there is a theoretical foundation for the behavioral rules. In the general equilibrium models, the transformation of input into output is not only a result of a fixed technological coefficient such as in the input-output approach, but allows for substitution effects too. In this approach, competitive markets, utility-maximizing behavior of the consumers and profit-maximizing behavior of the producers are assumed to influence the relationship between the input and the output. As in input-output analysis, general equilibrium analysis typically assumes that the production function in the regions is given. Thus, the location of the production is predetermined, although the equilibrium production levels can change and can become equal to zero (Bröcker, 1995, Göttinger, 1998, Truong, 1999).

The internalization of the spatial environmental externalities in general equilibrium models may be realized in two ways: (i) via the price mechanism through taxes which will affect the demand for the taxed goods; or (ii) in the quantity space, which will give a shadow price comparable to the tax.

Spatial interaction models

Spatial-interaction models may also use data from input-output tables. However, the assumed causality and theoretical focus is rather different (Pooler, 1994a, b, Openshaw, 1998, Diplock, 1998). Though the spatial interaction models can also be related to some utility-based models, such as the general equilibrium models, with the additional characteristic that the distance is measured in terms of transportation costs (see e.g. Nijkamp and Reggiani, 1987, 1990, 1998, van Lierop and Nijkamp, 1978), we may see the difference more clearly in the gravity models, which, as Isard (1990a) already pointed out, may be perceived as a special case of spatial interaction models. In the gravity models, the focus is on the spatial flow of the commodities (input and or output) between two regions, which – in analogy with physics – is positively dependent on the mass, measured by indicators such as regional income, output or other macroeconomic quantity-variables of the regions, and negatively related to the distance between the regions. The transportation costs are thus implicitly incorporated.

Other multi-regional models

Multi-regional systems are a generic term for models that have an explicit regional element (Beckmann, 1978, Isard, 1956, Hafkamp, 1984, Hordijk and Nijkamp, 1980, Nijkamp, 1976, 1987). In a sense, the input-output and the multi-regional general equilibrium models are a subset of the multi-regional systems. In a survey of multi-regional economic models, Nijkamp et al. (1982) have already pointed out a wide

range of possibilities. Hafkamp (1984), for example, designed a triple-layer multi-region model for analyzing the impact of internalizing the environmental externalities in a multi-regional context. Currently, except for the models mentioned in the previous sections, multi-regional models which analyze the spatial ecological-economic interaction consider regions as open systems, both in terms of economic and environmental processes. This is done, among others, by Rembold (1975), van den Bergh and Nijkamp (1998), and Inoue (1998). Considering the contribution of each of these authors in turn, first, Rembold (1975) analyzed a regionalized multi-sector model, combining a trade, an environmental, a demand and a production model. Without endogenous knowledge accumulation, which is assumed to be the basis for technological progress, the raw material or pollution is the limiting factor for economic growth. In van den Bergh and Nijkamp (1998), growth could be engined by 'international' trade as well as technological progress (as in endogenous growth theory). However, in this approach the state of economic development is limited by the carrying capacity of the environmental system.

The sustainability issue is analyzed in van den Bergh and Nijkamp (1995) through the interlinkage of both the environmental and the economic system of two regions. In this two-region model, endogenous technological progress is engined in one of the regions and the process of knowledge diffusion occurs through mutual trade. On the basis of this approach, Inoue (1998) studied the sustainability of economic development when knowledge diffusion came through aid instead of trade and the emphasis is on the abatement technology.

5 Conclusion

By extending the conceptual framework, which was developed in Wang, Nijkamp and Verhoef (2001), to the spatial landscape, this chapter has categorized the existing literature on spatial interaction models that concern environmental externalities. It was shown that the literature may be subdivided into model categories focussing, on the location and transportation of the inputs, the outputs, and the location of the production possibilities, respectively.

In this categorization, it is evident that the models become more general the more factors they include. In this sense, a pure partial model would be one which, according to the framework developed in Section 2, considers, as given, all the inputs, outputs and production possibilities in all regions, except for one: for example, one of the input, output or production possibilities in one of the regions. A general model of spatial interaction would then assume that, given some initial situation, all the inputs, outputs and the production possibilities would be endogenously determined. In this sense, even the spatial general equilibrium models, which include the markets for all inputs and the outputs, are not truly general models of spatial interaction, as they presuppose a fixed location of the production possibilities across the regions. In other words, an operational 'generalized theory' of spatial interaction, as expressed by Isard (1956), is, given the complexities, not yet available in the literature.

From the point of view of reality, another important aspect is the time that is involved in the interactions between the economic and the ecological system. The inclu-

sion of time in the analysis may have quite an important influence on the interpretation and the method for analyzing the underlying system: when time plays a role, the feedback may become even more complex. Therefore, some additional assumptions about the feedback process (or the movement of the variables of interest) have to be made. In the optimal control literature – for example, Chiang (1992), and Kamien and Schwartz (1978) – this is typically expressed by the smooth growth or decline of the variables, such as capital accumulation, which presupposes rational or adaptive expectations of the economic subjects.

Despite the fact that the issues of dynamics and statics have been considered by many economists, there is a wide variety of definitions of dynamics and statics (for a comprehensive review of the definitions, see Machlup, 1963). To have a better understanding of the concepts of statics and dynamics and to clarify the differences between the various definitions, consider the case in which the system could move towards one or more states. Theoretically, each period (e.g. a generation, an economic cycle, or even a second) may be perceived as one state. Then, statics studies the system within one state, while comparative statics compares two or more states of the system, and dynamic analysis studies the transition of the system from one state to another state (or more states).

Schumpeter (1948) illustrated the relationship between dynamics and statics by the following two, different points of view. First, static theory involves a higher level of abstraction: while dynamic patterns ignore a good many things, the static patterns drop even more features of reality, for example technological progress, and statics is, therefore, still nearer to a pure logic of economic quantities than dynamics. Secondly, statics may be seen as a special case of a more general dynamic theory: as we may derive static patterns from dynamic ones by the simple process of equating the 'dynamizing factors' to zero. As the history of economic thought starts with static analysis, Schumpter stated that, when using dynamic analysis, under all conditions, it must be possible to restate the dynamic model as a static one (of course, with additional simplifications).

References

Armington, P. S. (1969) A theory of demand for products distinguished by place of production. International Monetary Fund Staff Papers 13: 159–176

Barrett, S. (1994) Strategic environmental policy and international trade. Journal of Public Economics 54: 325–338

Batten, D.F. (2001) Complex landscapes of Spatial interaction. The Annals of Regional Science 35: 81–111

Baumol, W. J. and Oates, W. E. (1987) The theory of environmental policy. Prentice-Hall, Englewood Cliffs, N.J.

Beckmann, M.J. (1978) Outline of a theory of regional structure. In: Funck, R. and Parr, J.B. (eds.), The Analysis of Regional Structure: essays in honour of August Lösch. London: Pion: 8–18

Beckmann, M. and Thisse, J.-F. (1986) The location of production activities. In: Nijkamp, P. (ed.), Handbook of regional and urban economics: regional economics. Elsevier, Amsterdam: 21–95

Beers, C. van and Bergh, J. C. J. M. van den (1999) An overview of methodological approaches in the analysis of trade and environment. Journal of World Trade: 143–167

Beghin, J., Dessus, S., Roland-Holst, D. and Mensbrugghe, D. van der (1996) General equilibrium modelling of trade and environment. OECD Technical Paper no. 116. OECD, Paris

Beghin, J., Roland-Holst, D. and van der Mensbrugghe, D. (1994) A survey of the trade and environment nexus: global dimensions. In: OECD, OECD Economic Studies 23. OECD, Paris: 167–192

Benarroch, M. and Thille, H. (2001) Transboundary pollution and the gains from trade. Journal of International Economics 29: 139–159

van den Bergh, J.C.J.M. and De Mooij, R.A. (1997) An assessment of the growth debate. TI 97-096/3, Tinbergen Institute, Amsterdam

van den Bergh, J.C.J.M. and Nijkamp, P. (1998) A multiregional perspective on growth and environment: the role of endogenous technology and trade. The Annals of Regional Science 32: 115–131

van den Bergh, J.C.J.M., Nijkamp, P. and Rietveld, P. (eds.) (1996) Recent advances in spatial equilibrium modeling: methodology and applications. Springer, Berlin

van den Bergh, J.C.J.M. and Verbruggen, H. (1999) Spatial sustainability, trade and indicators: an evaluation of the 'ecological footprint'. Ecological Economics 29: 61–72

van den Bergh, J.C.J.M., and Nijkamp, P. (1995) Models of Spatial Externalities and Networks, TI Discussion Papers 1995/139, Tinbergen Institute: Amsterdam

Beukering, P.J.H. van (2001) Recycling, international trade and the environment: an empirical analysis. Dissertation, Vrije Universiteit Amsterdam, Faculty of Economics and Business Administration

Birkin, M. and Wilson, A.G. (1986a) Industrial location models 1: a review and integrating framework. Environment and Planning A 18: 175–205

Birkin, M. and Wilson, A. G. (1986b) Industrial location models 2: Weber, Palander, Hotelling, and extensions within a new framework. Environment and Planning A 18: 293–306

Bishop, R.C. and Woodward, R.T. (1995) Valuation of environmental quality under certainty. In: Bromley, D.W. (ed.), Handbook of environmental economics. Blackwell, Oxford: 543–567

Bockstael, N.E. and McConnell, K.E. (1993) Public goods as characteristics of non-market commodities. Economic Journal 103: 1244–1257

Bouman, M. (1998) Environmental Costs and Capital Flight, Dissertation, Tinbergen Institute, Amsterdam

Bröcker, J. (1995) Chamberlinian Spatial Computable General Equilibrium Modelling: a theoretical framework. Economic Systems Research 7: 137–149

Bröcker, J. (1998) Operational spatial computable general equilibrium modeling. The Annals of Regional Science 32: 367–387

Bruyn, S.M. de (1997) Explaining the environmental Kuznets curve: Structural change and international agreements in reducing sulphur emissions. Environment and Development Economics 2: 485–503

Chiang, A.C. (1992) Elements of dynamic optimization. McGraw-Hill Inc., New York

Copeland, B.R. (1994) International trade and the environment: policy reform in a polluted small open economy. Journal of Environmental Economics and Management 26: 44–65

Copeland, B.R. and Taylor, M.S. (1995) Trade and the environment: a partial synthesis. American Journal of Agricultural Economics, 77: 765–771

Copeland, B.R. and Taylor, M.S. (1999) Trade, spatial separation, and the environment. Journal of International Economics 47: 137–168

Cropper, M.L. and Oates, W.E. (1992) Environmental economics: a survey. Journal of Economic Literature 30: 675–740

Daly, H.E. (1995) Against free trade: neoclassical and steady-state perspectives. Evolutionary Economics 5: 313–326

Dean, J.M. (1992) Trade and the environment: a survey of the literature. In: Low, P. (ed.), International trade and the environment, World Bank, Washington D.C.: 19–42

Diplock, G. (1998) Building new spatial interaction models by using genetic programming and a supercomputer. Environment and Planning A 30: 1893–1904

Fotheringham, A.S. and O'Kelly, M.E. (1989) Spatial interaction models: formulation and applications. Kluwer Academic Publishers, Dordrecht

Fujita, M. (1999) Location and spatial economy at half a century: revisting Professor Isard's dream on the general theory. Annals of Regional Science 33: 371–381

Fujita, M., Krugman, P.R., Venables, A. (1999) The spatial economy: cities, regions, and international trade, MIT Press, Cambridge, Mass.

Göttinger, H.W. (1998) Greenhouse gas economics and computable general equilibrium. Journal of Policy Modeling 20: 537–580

Greenhut, M.L., Norman, G. and Hung, C.-S. (1987) The economics of imperfect competition: a spatial approach. Cambridge University Press, Cambridge

Groot, H.L.F. de (1999) Structural change, economic growth and the enviromental Kuznets curve: a theoretical perspective. Mimeo, Amsterdam

Grossman, G.M. and Krueger, A.B. (1991) Environmental impact of a North American free trade agreement. NBER Working Paper 3914, Cambridge

Hafkamp, W.A. (1984) Economic-environmental modeling in a national-regional system: an operational approach with multi-layer projection. Elsevier, Amsterdam

Hanemann, W.M. (1991) Willingness to pay and willingness to accept: how much do they differ? The American Economic Review 81: 635–647

Hanley, N., Shogren, J.F. and White, B. (1997) Environmental economics in theory and practice. MacMillan Press Ltd., London

Hibbard, K., Steffen, W., Benedict, S., Busalachi, T., Canadell, P., Dickinson, R., Raupach, M., Smith, B., Tilbrook, B., Vellinga, P. and Young, O. (2001) The carbon challenge: an IGBP - IHDP - WCRP joint project. International Geosphere-Biosphere Programme (IGBP), International Human Dimensions Programme on Global Environmental Change (IHDP), and World Climate Research Programme (WCRP)

Hordijk, L. and Nijkamp, P. (1980) Integrated approaches to regional development models: a survey of some Western European models. Serie Research Memoranda, 1980-4, Economische Faculteit, Vrije Universiteit, Amsterdam

Inoue, T. (1998) Optimal environmental policies for sustainable growth: a two-region model. Journal of Regional Sciences 38: 599–620

IPCC (1997) An introduction to simple climate models used in the IPCC Second Assessment Report. IPCC Technical Paper 2

Isard, W. (1956) Location and space-economy: a general theory relating to industrial location, market areas, land use, trade and urban structure. The MIT Press, Cambridge, Mass.

Isard, W. (1972) Ecologic-economic analysis for regional development. The Free Press, New York

Isard, W. (1990a) Location analysis and general theory. MacMillan Academic and Professional Ltd., Hampshire

Isard, W. (1990b) Practical methods of regional science and empirical applications. MacMillan Academic and Professional Ltd., Hampshire

Jaffe, A.B., Peterson, St.R., Portney, P.R. and Stavins, R. (1995) Environmental regulation and the competitiveness of U.S. manufacturing. Journal of Economic Literature 33: 557–572

Kamien, M.I. and Schwartz, N.L. (1978) Optimal exhaustible resource depletion with endogenous technological change. Review of Economic Studies 45: 179–196

Kennedy, P. (1994) Equilibrium pollution taxes in open economies with imperfect competition. Journal of Environmental Economics and Management 27: 49–63

Keyzer, M. and van Wesenbeeck, L. (1999) Trade models of imperfect competition. In: Ranis, G. and Raut, L.K. (eds.), Trade, growth and development: Essays in honor of Professor T. N. Srinivasan. Contributions to economic analysis, vol. 242. Elsevier Science, Amsterdam: 89–124

Kilkenny, M. and Thisse, J.-F. (1999) Economics of location: a selective survey. Computers & Operations Research 26: 1369–1394

Krugman, P. (1990) Rethinking international trade. The MIT Press, Cambridge, Mass.

Krugman, P. (1991a) Geography and trade. Leuven University Press, Leuven

Krugman, P. (1991b) Increasing returns and economic geography. Journal of Political Economy 99: 483–499

Krugman, P. (1993) On the relationship between trade theory and location theory. Review of International Economics 1: 110–122

Lierop, W.F.J. van and Nijkamp, P. (1978) A utility framework for interaction models for spatial processes. Vrije Universiteit, Economische Faculteit 1978-12, Amsterdam

Lösch, A. (1938) The nature of economic regions. The Southern Economic Journal 5: 71–78

Lösch, A. (1953) The economics of location. Yale University Press, New Haven

Ludema, R.D. and Wooton, I. (1997) International trade rules and environmental cooperation under asymmetric information. International Economic Review 38: 605–625

Machlup, F. (1963) Statics and dynamics: kaleidoscopic words. In: Miller, M. (ed.), Essays on economic semantics. Prentice-Hall, Englewood Cliffs: 98–118

Markusen, J.R. (1997) Costly pollution abatement, competitiveness and plant location decisions. Resource and Energy Economics, 19: 299–320

Markusen, J.R., Rutherford, T. and Hunter, L. (1995) Trade liberalization in a multinational-dominated industry. Journal of International Economics 39: 95–111

Mas-Colell, A., Whinston, M. and Green, J. R. (1995) Microeconomic theory. Oxford University Press, New York

Mennes, L.B.M., Tinbergen, J. and Waardenburg, J.G. (1969) The element of space in development planning. North-Holland Publishing Company, Amsterdam

NAFTA Commission for Environmental Cooperation (1996) A survey of recent attempts to model the environmental effects of trade: an overview and selected sources. NAFTA Commission for Environmental Cooperation, Washington D.C.

Nijkamp, P. (1976) Harmonie en Konflikt in het Ruimtelijk-Economisch Denken. H. E. Stenfert Kroese B.V., Leiden

Nijkamp, P. (1977) Theory and applications of environmental economics. North-Holland Publishing Company, Amsterdam

Nijkamp, P. (1987) Regionale Economie: een erfenis met toekomst. Economisch-Statistische Berichten: 676–80

Nijkamp, P. and Finco, A. (2001) Urban fingerprints and spatial footprints in sustainable urban-rural development.

Nijkamp, P. and Reggiani, A. (1987) A utility framework for interaction models for spatial processes. Vrije Universiteit, Faculteit der Economische Wetenschappen en Econometrie, Amsterdam

Nijkamp, P. and Reggiani, A. (1990) Spatial interaction models as a unified framework for dynamic transportation models, Series Research Memorandum 1990-82, Vrije Universiteit, Faculteit der Economische Wetenschappen en Econometrie, Amsterdam

Nijkamp, P. and Reggiani, A. (1998) The economics of complex spatial systems. Elsevier, Amsterdam

Nijkamp, P., Rietveld, P. and Snickars, F. (1982) Multiregional economic models: an introduction to the survey. In: Issaev, B., Nijkamp, P., Rietveld, P. and Snickars, F. (eds.), Multiregional economic modeling: practice and prospect. North-Holland Publishing Company, Amsterdam: 1–10

Obstfeld, M. and Rogoff, K. (1996) Foundations of international macroeconomics. The MIT Press, Cambridge, Mass.

Openshaw, S. (1998) Neural networks, genetic, and fuzzy logic models of spatial interaction. Environment and Planning A 30: 1857–1872

Opschoor, J.B. and Vos, H.B. (1989) Economic instruments for environmental protection. OECD, Paris

Opschoor, J.B., Savornin Lohman, A.F. and Vos, H.B. (1994) Managing the environment: the role of economic instruments. OECD, Paris

Papageorgiou, G.J. (1978a) Spatial externalities I: theory. Annals of the Association of American Geographers 68: 465–476

Papageorgiou, G.J. (1978b) Spatial externalities II: applications. Annals of the Association of American Geographers 68: 477–492

Pooler, J. (1994a) A family of relaxed spatial interaction models. Professional Geographer 46: 210–17

Pooler, J. (1994b) An extended family of spatial interaction models. Progress in Human Geography 18: 17–39

Porter, M.E. and Van der Linde, C. (1995) Green and competitive: ending the stalemate. Harvard Business Review: 120–134

Rauscher, M. (1991) Foreign trade and the environment In: Siebert, H. (ed.), Environmental scarcity: the international dimension. J.C.B. Mohr, Tubingen

Rauscher, M. (1997) International trade, factor movement, and the environment, Clarendon Press, Oxford

Reggiani, A. (1990) Spatial interaction models: new directions. Vrije Universiteit Amsterdam

Rembold, G. (1975) Limitations of regional and sector specific economic growth by pollution restrictions and scarcity of raw materials: a regionalized multi-sector model. In: Beckmann, M. and Kuenzi, H.P. (eds.), Environment, regional science and interregional modeling. Springer, Berlin: 178–195

Samuelson, P.A. (1952) Spatial price equilibrium and linear programming. American Economic Review 42: 283–303

Schumpeter, J.A. (1948) History of economic analysis. Routledge, London

Siebert, H. (1981) Economics of the environment: theory and policy. Springer, Berlin

Siebert, H. (1985) Spatial aspect of environmental economics. In: Kneese, A. V. and Sweeney, J.L. (eds.), Handbook of natural resource and energy economics. North Holland, Amsterdam: 125–164

Sohns, R. (1978) Loesch and the theory of trade. In: Funck, R. and Parr, J.B. (eds.), The analysis of regional structure: essays in honour of August Loesch. Pion, London: 119–135

Takayama, T. and Judge, G.C. (1964) Equilibrium among spatially separated markets: reformulation. Econometrica, 32: 510–524

Takayama, T. and Labys, W.C. (1986) Spatial equilibrium analysis. In: Nijkamp, P. (ed.), Handbook of regional and urban economics. Elsevier, Amsterdam: 171–199

Truong, T.P. (1999) GTAP-E. Incorporating energy substitution into GTAP model. Purdue Center for Global Trade Analysis, Purdue University

Ulph, A. and Ulph, D. (1994) Trade, strategic innovation and strategic environmental policy – a general analysis. Discussion Paper no. 1063, CEPR.

Verbruggen, H. (1991) Political economy aspects of environmental policy instruments. In: Dietz, F., van der Ploeg, F. and van der Straaten, J. (eds.), Environmental policy and economy, North-Holland, Amsterdam: 141–149

Verbruggen, H. (1999) Environment, international trade and development. In: van den Bergh, J.C.J.M. (ed.), Handbook of environmental and resource economics. Edward Elgar, Cheltenham: 449–460

Verhoef, E.T. and van den Bergh, J.C.J.M. (1995) A spatial price equilibrium model for environmental policy analysis of mobile and immobile sources of pollution. In: van den Bergh, J.C.J.M., Nijkamp, P. and Rietveld, P. (eds.), Recent advances in spatial equilibrium modeling: methodology and applications. Springer, Berlin: 201–220

Verhoef, E. T. and Nijkamp, P. (2000) Spatial dimensions of enviromental policies for transboundary externalities: a spatial price equilibrium approach. Environment and Planning A 32: 2033–2055

Verhoef, E.T. and P. Nijkamp (2000) "Spatial dimensions of environmental policies for transboundary externalities

Wäckernagel, M. (1998) What we use and what we have: ecological footprint and ecological capacity. Centre for Sustainability Studies, 8, Universidad Anahuac de Xalapa, Xalapa

Wäckernagel, M. and Rees, W. (1996) Our ecological footprint: reducing human impact on the earth. New Society Publishers, Gabriola Island, BC

Wang, S., Nijkamp, P., Verhoef, E. (2001) Modelling externalities between ecological and economic systems. Tinbergen Institute Discussion Paper, 01-098/3

Withagen, C.A.A.M. (1998) Milieubeleid en Concurrentievermogen. Vrije Universiteit, Amsterdam

Withagen, C.A.A.M. (1999) De Porter Hypothese: een verkenning van de literatuur. In: De Porter-Hypothese Belicht, Raad voor het Milieu – en Natuuronderzoek, Rijswijk: 7–41

Some Current Issues in the Statistical Analysis of Spillovers

Daniela Gumprecht*, Nicole Gumprecht*, Werner G. Müller‡

1 Coe & Helpman's R&D Spillover Study

Most theories of growth explain economic growth in terms of the accumulation of capital and the growth of the labor force and exogenous technological progress captured by a time trend. In recent formulations these variables are quality adjusted (human capital, embodied technological progress). In contrast, the new growth theory (Romer, 1990, Grossman and Helpman, 1991) tries to explain the growth record in terms of endogenous R&D decisions. Productivity depends therefore on the amount of knowledge generated by innovation activities, and productivity increases depend on current R&D efforts which translate into increased technical knowledge. By building on these theories Coe and Helpman (1995) claimed that the productivity of an economy depends on its own R&D as well as the R&D spendings of its trade partners. A direct advantage is a more effective use of resources by the application of new technologies, materials, production processes and organisation methods. Indirect benefits come from the import of goods and services from trade partners.

In their meanwhile classical paper Coe and Helpman (1995) used a panel dataset to study the extent to which a country's productivity level depends on domestic and foreign stock of knowledge. They used the cumulative spendings for R&D of a country to measure the domestic stock of knowledge of this country. As a representative for the foreign stock of knowledge, Coe and Helpman used the import-weighted sums of cumulated R&D expenditures of the trade partners of the country. The importance of the R&D capital stock is measured by the elasticity of total factor productivity with respect to the R&D capital stock. A panel dataset with 22 countries (21 OECD countries plus Israel) during the period from 1971 to 1990 was used[1]. The variables total factor productivity (TFP), domestic R&D capital stock (DRD) and foreign R&D capital stock (FRD) are constructed as indices with basis 1985 (1985 = 1).

In their papers Coe and Helpman have used a variety of specifications to model the effects on TFP. To simplify the exposition we will here only regard one of those. Our conclusions, however, are not limited to this particular case but rather apply to

*Department of Statistics and Decision Support Systems, University of Vienna, Vienna, Austria.

‡Department of Statistics, Vienna University of Economics and Business Administration, Vienna, Austria.

[1] All data can be found on the homepage of Elhanan Helpman (Helpman, 2003), which is accessible via the internet address http://post.economics.harvard.edu/faculty/helpman/data.html

all of the suggested models (for a more complete analysis see D. Gumprecht, 2003). Our illustrative model contains three variables: total factor productivity (TFP) as the regressand, domestic R&D capital stock (DRD) and foreign R&D capital stock (FRD) as the regressors. The impact of domestic and foreign R&D expenditures is supposed to be the same for all countries. The equation – with regional index i and temporal index t – has the following form:

$$\ln F_{it} = \alpha_{it}^0 + \alpha_{it}^d \ln S_{it}^d + \alpha_{it}^f m_{i,t-1} \ln S_{it}^f,$$

where

F_{it} denotes total factor productivity (TFP),
S_{it}^d domestic R&D capital stock (DRD), and
S_{it}^f foreign R&D capital stock (FRD). FRD is defined as the import-share-weighted average of the domestic R&D capital stocks of trade partners.
α_{it}^0 stands for the intercepts, which are allowed to vary across countries for two reasons: first, there may exist country-specific effects on productivity that are not included in the variables of our model; and second, all variables are transformed into index numbers and TFP is measured in country-specific currency whereas DRD and FRD are measured in U.S. dollars.
α_{it}^d then denotes the regression coefficient, which corresponds to the elasticity of TFP with respect to DRD, and
α_{it}^f determines the elasticity of TFP with respect to FRD, which equals $\alpha_{it}^f m_{i,t-1}$. Finally
$m_{i,t-1}$ denotes the fraction of imports in GDP.

According to standard practice Coe and Helpman (1995) used for their estimations a panel data model with fixed effects, which is described in detail in the next subsection. They were especially focussed on the time dimension of the data and therefore used time series methods and analysis for their panel data model. As they were interested in identifying a long-run relationship between TFP and domestic and foreign R&D spendings, and as TFP, DRD and FRD showed a clear temporal trend, they estimated cointegrated equations.

"The basic idea of cointegration is that if there is a long-run relationship between two or more trended variables, a regression containing all the variables – the cointegration equation – will have a stationary error term, even if none of the variables taken alone is stationary. If the error term is not stationary, the estimated relationship may be spurious." (Coe and Helpman, 1995, p. 867–868 according to Granger and Newbold, 1974). Cointegrated equations have the important econometric property that OLS estimates are 'super consistent' (Stock, 1987). This means, when the number of observations increases, the OLS estimator of the cointegrating equation converges to the true parameter value much faster than in the case where the variables are stationary. The idea of cointegration comes from time-series analysis and it seems natural for Coe and Helpman to use this technique for their R&D spillovers problem. Because of the relatively small number of time-series observations for each country, Coe and Helpman estimated their equations from panel data and interpreted the results as pooled cointegration equations (Coe and Helpman, 1995, p. 868).

Conditions for the existence of cointegration are the following: first, the separate variables have to be nonstationary; and second, the error term of a linear combination of the variables has to be stationary. Nonstationarity of each single time-series was tested with the Dickey-Fuller, the augmented Dickey-Fuller (1979, 1981), the Levin and Lin (1992) and the Levin and Lin (1993) unit root tests. The Levin and Lin unit root tests on the pooled data confirm the nonstationarity of the variables. Nonstationarity of the error term was tested with Levin and Lin (1992), Levin and Lin (1993) unit root tests and a test from Engle and Granger (1987). These tests provided different results (Coe and Helpman, 1995, Table 3). Because of these mixed results and the fact that the econometrics of pooled cointegration were not fully worked out at that time, Coe and Helpman concentrated more on the theoretical model and on the a priori plausibility of the estimated parameters rather than on the tests for cointegration (Coe and Helpman, 1995, p. 870).

In what follows we present a recalculation of the OLS estimators for the model of Coe and Helpman (1995) with corrected degrees of freedom in the calculation of the t-values making use of the Helpman (2003) data. Kao et al. (1999) have re-estimated Coe and Helpman's equations (with corrected t-values, see a discussion of their approach later). However, they made a mistake when implementing the calculation in GAUSS (a commonly used statistically oriented matrix language package, see www.aptech.com). They used wrong degrees of freedom for the calculation of the t-values,

```
namely      vb1=inv(x1'*x1)*((u1'*u1)/(N*T-1)).
instead of  vb1=inv(x1'*x1)*((u1'*u1)/(N*T-N-k1),
```

The corrected estimation (with no substantial difference in significances) yields (t-values in parentheses)

$$\ln F_{it} = \hat{\alpha}^0_{it} + \underset{(12.8885)^{**}}{0.10511 \ln S^d_{it}} + \underset{(5.8011)^{**}}{0.2665\, m_{i,t-1} \ln S^f_{it}} + \text{rest},$$

with $i = 1, \ldots, N\ (= 21)$ and $t = 1, \ldots, T\ (= 20)$ and a coefficient of determination $R^2 = 0.5576$ and an Adjusted R^2 of 0.5331. Note that in panel models the definition of the coefficient of determination is not without ambiguity and we have calculated all R^2 throughout the paper as the squared correlations of $\hat{y}_{it}$ and y_{it}.

Coe and Helpman (1995) took these estimation results, with positive (and statistically significant) regression coefficients as a confirmation of their hypothesis that TFP of a country depends on both domestic R&D capital stock and foreign R&D capital stock.

A corresponding exploratory study seems to confirm these conclusions. The simple time-series scatterplots of TFP and DRD, and TFP and FRD are given in Figures 1 and 2, respectively. To simplify the plots we have only included the G7 countries[2], without restriction on generality. The plots show the time paths from lower left (1971) to upper right (1990), which all exhibit an upward slope as an indication of a positive relationship between these variables.

[2]U.S.A., Japan, Germany, France, Italy, U.K., Canada

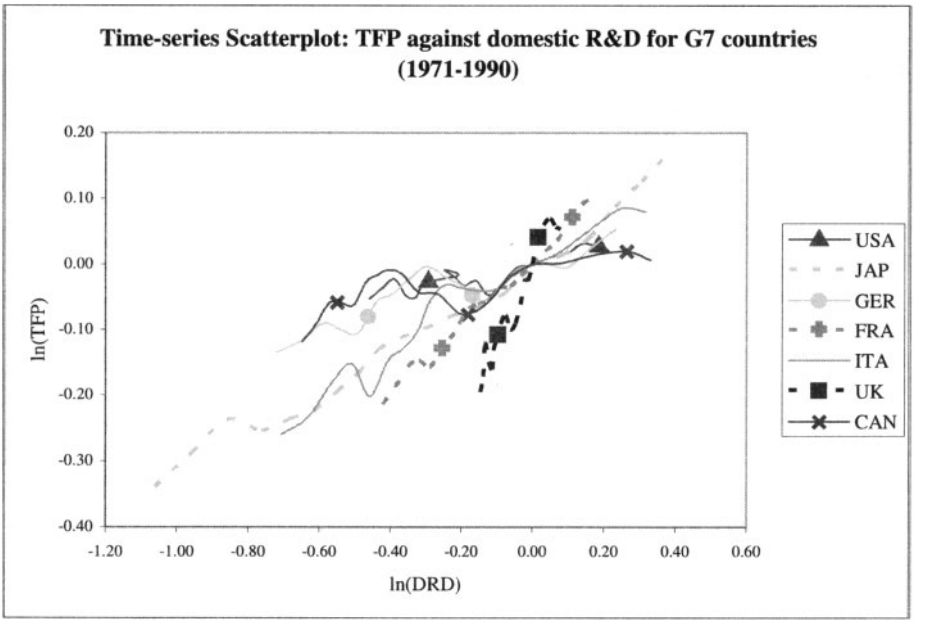

Figure 1: Time-series scatterplot TFP against DRD for G7 countries (1971–90).

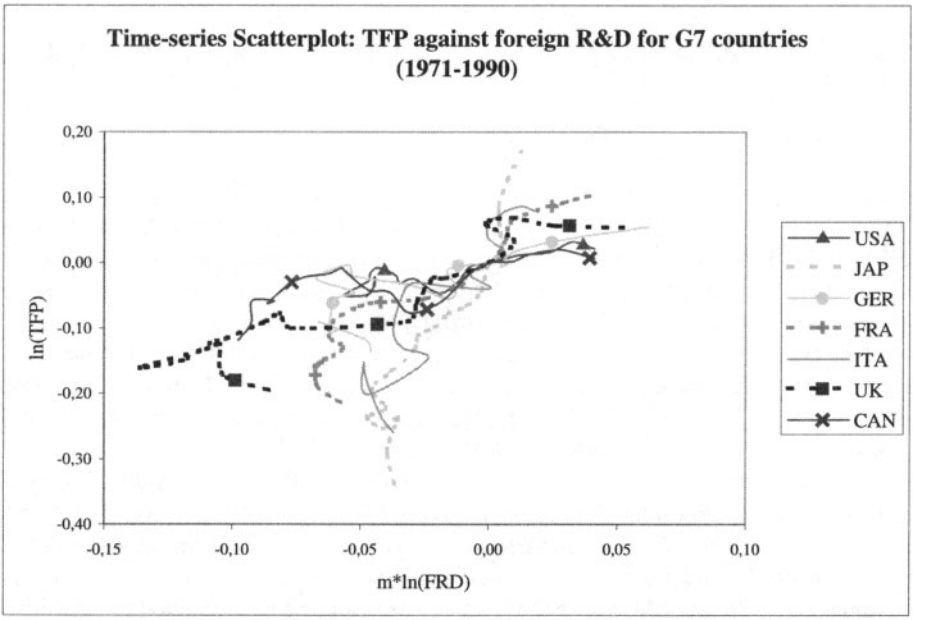

Figure 2: Time-series scatterplot TFP against FRD for G7 countries (1971–90).

2 Fixed Effects Panel Regression

In the following section we will review the estimation techniques employed in most of the spillover studies. We will hereby largely follow the exposition of standard econometrics textbooks such as e.g. Greene (2000). More detailed material on the various specifications used in panel data regressions can e.g. be found in the monographs by Hsiao (1986) and Baltagi (2001).

2.1 Simple OLS estimators

In fixed effect panel models differences between cross-section units (individuals, regions, etc.) are shown by differences in the constant terms. Each α_i is an unknown parameter and must be estimated. This approach is suitable for models where the differences between individuals can be interpreted as parametrical shifts of the regression function.

There are three different ways to specify the regression model.

1. **Original form**:

$$y_{it} = \alpha + \beta' \mathbf{x}_{it} + \epsilon_{it}$$

The total sums of squares and total cross products are given by

$$\mathbf{S}_{xx}^t = \sum_{i=1}^{N} \sum_{t=1}^{T_i} (\mathbf{x}_{it} - \bar{\bar{\mathbf{x}}})(\mathbf{x}_{it} - \bar{\bar{\mathbf{x}}})' \qquad \mathbf{S}_{xy}^t = \sum_{i=1}^{N} \sum_{t=1}^{T_i} (\mathbf{x}_{it} - \bar{\bar{\mathbf{x}}})(y_{it} - \bar{\bar{y}})$$

where:

$$\bar{\bar{\mathbf{x}}} = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T_i} \mathbf{x}_{it}}{\sum_{i=1}^{N} T_i} = \frac{\sum_{i=1}^{N} T_i \bar{\mathbf{x}}_{i.}}{\sum_{i=1}^{N} T_i} = \sum_{i=1}^{N} w_i \bar{\mathbf{x}}_{i.} \quad \text{and} \quad \bar{\bar{y}} = \sum_{i=1}^{N} w_i \bar{y}_{i.}$$

$$\text{with} \quad w_i = \frac{T_i}{\sum_{i=1}^{N} T_i}$$

and the LS-total estimator follows:

$$\mathbf{b}^t = (\mathbf{S}_{xx}^t)^{-1} \mathbf{S}_{xy}^t$$

2. **Departure from group-mean form**:

$$y_{it} - \bar{y}_{i.} = \beta' (\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.}) + \epsilon_{it} - \bar{\epsilon}_{i.}$$

Here the so-called sums of squares within and cross products within are given by:

$$\mathbf{S}_{xx}^{w} = \sum_{i=1}^{N}\sum_{t=1}^{T_i}(\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.})(\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.})' \qquad \mathbf{S}_{xy}^{w} = \sum_{i=1}^{N}\sum_{t=1}^{T_i}(\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.})(y_{it} - \bar{y}_{i.})$$

with the corresponding LS-within estimator:

$$\mathbf{b}^{w} = (\mathbf{S}_{xx}^{w})^{-1}\mathbf{S}_{xy}^{w}$$

3. **Group mean form**:

$$\bar{y}_{i.} = \alpha + \beta'\bar{\mathbf{x}}_{i.} + \bar{\epsilon}_{i.}$$

with only N observations because there are only N groups. The corresponding so-called sums of squares between and cross products between are given by

$$\mathbf{S}_{xx}^{b} = \sum_{i=1}^{N}\sum_{t=1}^{T_i} T_i(\bar{\mathbf{x}}_{i.} - \bar{\bar{\mathbf{x}}})(\bar{\mathbf{x}}_{i.} - \bar{\bar{\mathbf{x}}})' \qquad \mathbf{S}_{xy}^{b} = \sum_{i=1}^{N}\sum_{t=1}^{T_i} T_i(\bar{\mathbf{x}}_{i.} - \bar{\bar{\mathbf{x}}})(\bar{y}_{i.} - \bar{\bar{y}}),$$

respectively, and the LS-between estimator follows

$$\mathbf{b}^{b} = (\mathbf{S}_{xx}^{b})^{-1}\mathbf{S}_{xy}^{b}.$$

The group means are calculated in the following way:

$$\bar{y}_{i.} = \frac{1}{T_i}\sum_{t=1}^{T_i} y_{it}, \quad \bar{\mathbf{x}}_{i.} = \frac{1}{T_i}\sum_{t=1}^{T_i} \mathbf{x}_{it}, \quad \bar{\epsilon}_{i.} = \frac{1}{T_i}\sum_{t=1}^{T_i} \epsilon_{it}.$$

The terminology "within" and "between" stems from the fact that the estimators are determined by the variation within and between the specific groups as opposed to the "total" variation.

In a panel data model with fixed effects the within-estimator ($\mathbf{b}^{w}$) is the BLUE (best linear unbiased estimator). The proof follows directly from the Least Square Dummy Variable (LSDV) form of the fixed effects model (see e.g. Greene, 2000, chapter 14.3).

Suggestions for improvement of Coe and Helpman's estimation came – amongst others – from Kao et al. (1999). They criticized two things: First, Coe and Helpman presented their results without any t-values because the asymptotic distribution of the t-statistic for estimates in cointegrated panel data was not known at that time. Therefore no exact statements about the significance of the OLS estimators could be made. As Coe and Helpman's resulting estimates were both relatively small one cannot safely conclude that even one of the true coefficients was bigger than zero. Second, due to the

unit-root in the time dimension and in spite of the super consistence of the time-series estimator, the upward bias of the estimate can be quite substantial for small samples and there is no reason to assume that this bias becomes negligible by the inclusion of a cross section dimension in panel data. Kao et al. (1999) argue that it is quite possible that the estimators even change their sign when introducing a bias correction in the calculation.

For those reasons Kao et al. (1999) used different estimation methods for Coe and Helpman's International R&D Spillovers regression and compared the empirical consequences from the different estimation methods. They claim that the DOLS (dynamic OLS) estimation is the best solution for this problem because in the given setting the DOLS estimator exhibits no bias and is asymptotically normal.

2.2 Corrected OLS estimators

Kao et al. (1999) also used a panel data model with fixed effects for their estimations. The regression function has again the following specification:

$$y_{it} = \alpha_i + \mathbf{x}'_{it}\beta + \epsilon_{it}$$

where now

y_{it} again denotes the dependent variable,
β $M \times 1$ the vector of slope parameters,
α_i the region specific intercepts,
ϵ_{it} stands for a stationary error term, but now
$\mathbf{x}_{it}$ is regarded as an $M \times 1$ first order integrated process, with $\mathbf{x}_{it} = \mathbf{x}_{i,t-1} + \zeta_{it}$.

Under these assumptions the panel regression describes a system of cointegrated regressions, this means y_{it} is cointegrated with $\mathbf{x}_{it}$. Furthermore y_{it} and $\mathbf{x}_{it}$ are independent between different cross section units and $\mathbf{w}_{it} = (\epsilon_{it}, \zeta'_{it})'$ is a linear process that fulfils the assumptions of Kao and Chiang (1997). The asymptotic covariance matrix $\mathbf{\Omega}$ of $\mathbf{w}_{it}$ can be written in the following form:

$$\begin{aligned}\mathbf{\Omega} &= \sum_{j=-\infty}^{\infty} E(\mathbf{w}_{ij}\mathbf{w}'_{i0}) \\ &= \mathbf{\Sigma} + \mathbf{\Gamma} + \mathbf{\Gamma}' \\ &= \begin{bmatrix} \mathbf{\Omega}_{\epsilon} & \mathbf{\Omega}_{\epsilon\zeta} \\ \mathbf{\Omega}_{\zeta\epsilon} & \mathbf{\Omega}_{\zeta} \end{bmatrix},\end{aligned}$$

where

$$\mathbf{\Gamma} = \sum_{j=1}^{\infty} E(\mathbf{w}_{ij}\mathbf{w}'_{i0}) = \begin{bmatrix} \mathbf{\Gamma}_{\epsilon} & \mathbf{\Gamma}_{\epsilon\zeta} \\ \mathbf{\Gamma}_{\zeta\epsilon} & \mathbf{\Gamma}_{\zeta} \end{bmatrix}$$

and

$$\mathbf{\Sigma} = E(\mathbf{w}_{i0}\mathbf{w}'_{i0}) = \begin{bmatrix} \mathbf{\Sigma}_\epsilon & \mathbf{\Sigma}_{\epsilon\zeta} \\ \mathbf{\Sigma}_{\zeta\epsilon} & \mathbf{\Sigma}_\zeta \end{bmatrix}$$

are partitioned according to $\mathbf{w}_{it}$. The one-sided asymptotic covariance is defined as:

$$\Delta = \mathbf{\Sigma} + \mathbf{\Gamma}$$

$$= \sum_{j=0}^{\infty} E(\mathbf{w}_{ij}\mathbf{w}'_{i0})$$

with

$$\Delta = \begin{bmatrix} \Delta_\epsilon & \Delta_{\epsilon\zeta} \\ \Delta_{\zeta\epsilon} & \Delta_\zeta \end{bmatrix}.$$

With this "long run correction" the correct t-values can be calculated.

Kao and Chiang (1997) defined the limiting distribution of the OLS and a so-called DOLS (dynamic ordinary least squares) estimator of a cointegrated regression. They also showed that these limiting distributions are asymptotically normal and analysed the characteristics of these estimators in finite samples. They found that the OLS estimator has a non-neglible bias and that the DOLS estimator is therefore preferable for estimating cointegrated panel regressions. The OLS estimator is given by

$$\hat{\beta}_{OLS} = \left[\sum_{i=1}^{N}\sum_{t=1}^{T}(\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.})(\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.})'\right]^{-1} \left[\sum_{i=1}^{N}\sum_{t=1}^{T}(\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.})(y_{it} - \bar{y}_{i.})\right],$$

where $\bar{\mathbf{x}}_{i.}$ and $\bar{y}_{i.}$ are the respective group means (see Kao et al., 1999, p. 697). The asymptotic distribution of this estimator is, according to Kao and Chiang (1997),

$$\sqrt{NT}(\hat{\beta}_{OLS} - \beta) - \sqrt{N}\delta_{NT} \rightarrow N(\mathbf{0}, 6\mathbf{\Omega}_\zeta^{-1}\mathbf{\Omega}_{\epsilon\zeta}); \qquad \text{convergence in distribution,}$$

where

$$\mathbf{\Omega}_{\epsilon\zeta} = \mathbf{\Omega}_\epsilon - \mathbf{\Omega}_{\epsilon\zeta}\mathbf{\Omega}_\zeta^{-1}\mathbf{\Omega}_{\zeta\epsilon}$$

and

$$\delta_{NT} = \left[\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}(\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.})(\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.})'\right]^{-1} \times$$

$$\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{\Omega}_\zeta^{1/2}\left(\int_0^1 \tilde{\mathbf{W}}_i(\mathbf{r})d\mathbf{W}'_i(\mathbf{r})\right)\mathbf{\Omega}_\zeta^{-1/2}\mathbf{\Omega}_{\zeta\epsilon} + \Delta_{\zeta\epsilon}\right],$$

$\mathbf{W}_i(\mathbf{r})$ being a standard Brownian motion, and

$$\tilde{\mathbf{W}}_i(\mathbf{r}) = \mathbf{W}_i(\mathbf{r}) - \int_0^1 \mathbf{W}_i(\mathbf{r})d\mathbf{r}.$$

2.3 The DOLS estimator

This estimator, which was employed in Kao et al. (1999), can be obtained by running the regression:

$$y_{it} = \alpha_i + \mathbf{x}'_{it}\beta + \sum_{j=-q_1}^{q_2} \Delta\mathbf{x}'_{i,t+j}\mathbf{c}_{ij} + v_{it}, \qquad q_1, q_2 \in \{0, 1, 2, ...\}$$

The DOLS estimation as used by Kao et al. (1999) is also based on a fixed effect regression model:

$$y_{it} = \alpha_i + \mathbf{x}'_{it}\beta + \epsilon_{it} \qquad i = 1, ..., N, t = 1, ..., T.$$

We assume that $\{\mathbf{x}_{it}\}$ are $k \times 1$ integrated processes of order one for all i, where

$$\mathbf{x}_{it} = \mathbf{x}_{i,t-1} + \zeta_{it}$$

and

$$\Delta\mathbf{x}_{it} = \mathbf{x}_{it} - \mathbf{x}_{i,t-1},$$

where $\Delta\mathbf{x}_{it}$ denotes the difference of $\mathbf{x}_{it}$ to $\mathbf{x}_{i,t-1}$.

If we assume that the process $\{\epsilon_{it}\}$ can be projected on to $\{\zeta_{it}\}$, we get

$$\epsilon_{it} = \sum_{j=-\infty}^{\infty} \zeta'_{i,t+j}\mathbf{c}_{ij} + v_{it}$$

where

$$\sum_{j=-\infty}^{\infty} \|\mathbf{c}_{ij}\| < \infty,$$

$\{v_{it}\}$ is stationary with mean zero, and $\{v_{it}\}$ and $\{\zeta_{it}\}$ are uncorrelated, both contemporaneously and in all lags and leads (see Saikkonen, 1991, p. 11).

In practice, the lags and leads are restricted to a range from q_1 to q_2. Retaining the former assumption approximately, it follows that

$$\epsilon_{it} = \sum_{j=-q_1}^{q_2} \zeta'_{i,t+j}\mathbf{c}_{ij} + v_{it}.$$

This follows the assumption that $\{\mathbf{c}_{it}\}$ are absolutely summable, which means

$$\sum_{j=-\infty}^{\infty} \|\mathbf{c}_{ij}\| < \infty.$$

After substitution

$$\epsilon_{it} = \sum_{j=\infty}^{\infty} \zeta'_{i,t+j}\mathbf{c}_{ij} + v_{it}$$

and

$$\zeta_{i,t+j} = \Delta \mathbf{x}_{i,t+j} = \mathbf{x}_{i,t+j} - \mathbf{x}_{i,t+j-1}$$

into the initial model

$$y_{it} = \alpha_i + \mathbf{x}'_{it}\beta + \epsilon_{it} \qquad i = 1, ..., N, t = 1, ..., T$$

we yield the specification

$$y_{it} = \alpha_i + \mathbf{x}'_{it}\beta + \sum_{j=-q_1}^{q_2} \Delta \mathbf{x}'_{i,t+j}\mathbf{c}_{ij} + v_{it}.$$

This is the regression model for the DOLS estimation (Kao and Chiang 1997, p. 9). The asymptotic distribution of a corresponding (now unbiased) estimator $\hat{\beta}_D$ is given by

$$\sqrt{NT}(\hat{\beta}_D - \beta) \rightarrow N(\mathbf{0}, 6\boldsymbol{\Omega}_{\zeta}^{-1}\boldsymbol{\Omega}_{\epsilon\zeta}); \text{convergence in dist. as } N \rightarrow \infty \text{ and } T \rightarrow \infty.$$

For definition of $\Omega_{\epsilon\zeta}$ see the OLS estimator of β (section 2.2).

Kao et al. (1999) reported the following results for their DOLS estimation of Coe and Helpman's R&D Spillovers model:

$$\ln F_{it} = \hat{\alpha}_{it}^0 \quad + \underset{(5.9572)^{**}}{0.1237 \ln S_{it}^d} \quad + \underset{(0.6333)}{0.0682\, m_{i,t-1} \ln S_{it}^f} \quad + \text{ rest}$$

with an R^2 of 0.5016.

They concluded from this estimation that domestic R&D expenditures affect TFP of a country but foreign R&D expenditures do not have a significant effect on TFP of the country. Thus, they argue, Coe and Helpman's (1995) conclusions should be rejected. However, Kao et al. (1999) wanted to estimate a fixed effect regression model (a model with county-specific intercepts) but erroneously they implemented a common coefficient model (a model with a common intercept). Furthermore, they left out the lag zero, which is not backed up by the corresponding theory by Saikkonen (1991). Additionally, their R^2 is calculated by ESS (Explained Sums of Squares) divided by TSS (Total Sums of Squares) but the wrong numbers of degrees of freedom were used.

The result of the correct implementation of the fixed effect model with dynamic regressors is now the following:

$$\ln F_{it} = \hat{\alpha}_{it}^0 \quad + \underset{(18.3164)^{**}}{0.1284 \ln S_{it}^d} \quad + \underset{(3.8375)^{**}}{0.1321\, m_{i,t-1} \ln S_{it}^f} \quad + \text{ rest}$$

with an R^2 (according to our definition) of 0.8755 and an Adjusted R^2 of 0.8689.

The correct estimated coefficient for foreign R&D expenditures is again - as in the original paper - significant. Domestic- and foreign R&D expenditures still seem to affect TFP of a country, which supports Coe and Helpman's conclusions. The R^2, calculated as the square of the correlation between $\hat{y}_{it}$ and y_{it}, is much better than the R^2 of Kao et al. (1999), calculated as ESS divided by TSS.

Nevertheless, a considerable innovation of Chiang and Kao's (1999, 2002) implementation is the use of the so-called "long run correction" (see section 2.2) for the correct calculation of the t-values of the coefficients, a suggestion, which will be taken up in our final model.

3 An Alternative View

There are many debates in the panel data estimation literature whether regarding the region specific or other effects as random outcomes poses a valuable alternative to the fixed coefficient model. In the present context Müller and Nettekoven (1999) have suggested a so-called random coefficient model to analyse the R&D Spillovers model of Coe and Helpman (1995) and conclude that although the alternative specification is well compatible with the data, one astonishingly has to draw contradictory conlusions.

3.1 The Random Coefficient Model

Here, the parameters β_i are assumed to vary randomly around a common mean β. This model can be described in the form:

$$\mathbf{y}_i = \mathbf{X}_i\beta_i + \epsilon_i,$$

where

$$\beta_i = \beta + \mathbf{v}_i$$

with

$$E[\mathbf{v}_i] = 0, \qquad E[\mathbf{v}_i\mathbf{v}_i'] = \mathbf{\Gamma}.$$

Under the assumption that there is no autocorrelation and no correlation between the cross section units, β_i (that applies for a particular cross section unit) can be considered the result of a random process with mean β and covariance matrix $\mathbf{\Gamma}$.

If β_i is expressed by the relation $\beta_i = \beta + \mathbf{v}_i$ the following model results:

$$\mathbf{y}_i = \mathbf{X}_i\beta + (\epsilon_i + \mathbf{X}_i\mathbf{v}_i) = \mathbf{X}_i\beta + \mathbf{w}_i,$$

where

$$E[\mathbf{w}_i] = 0, \qquad E[\mathbf{w}_i\mathbf{w}_i'] = \sigma_i^2\mathbf{I} + \mathbf{X}_i\mathbf{\Gamma}\mathbf{X}_i' = \mathbf{\Pi}_i$$

The covariance matrix for all observations ($\mathbf{V}$) has the following form:

$$\mathbf{V} = \begin{bmatrix} \mathbf{\Pi}_1 & 0 & 0 & \dots & 0 \\ 0 & \mathbf{\Pi}_2 & 0 & \dots & 0 \\ & & \vdots & & \\ 0 & 0 & 0 & \dots & \mathbf{\Pi}_n \end{bmatrix}.$$

Now, the (best linear unbiased) GLS estimator can be expressed by a matrix weighted average of the OLS estimators:

$$\hat{\beta} = \sum_{i=1}^{N} \mathbf{W}_i\mathbf{b}_i$$

where $\mathbf{b}_i$ is the i-th OLS coefficient estimator and

$$\mathbf{W}_i = \left[\sum_{i=1}^{N}(\mathbf{\Gamma} + \mathbf{V}_j)^{-1}\right]^{-1}(\mathbf{\Gamma} + \mathbf{V}_i)^{-1}$$

where

$$\mathbf{V}_i = \sigma_i^2(\mathbf{X}_i'\mathbf{X}_i)^{-1}.$$

The estimator of β can also be expressed in the usual form of the GLS estimator:

$$\hat{\beta} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y}.$$

As $\mathbf{V}$ is a block diagonal matrix it follows that:

$$\hat{\beta} = \left[\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{\Pi}_i^{-1}\mathbf{X}_i\right]^{-1}\left[\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{\Pi}_i^{-1}\mathbf{y}_i\right]$$

where

$$\mathbf{\Pi}_i = \sigma_i^2\mathbf{I} + \mathbf{X}_i\mathbf{\Gamma}\mathbf{X}_i'$$

This representation of $\hat{\beta}$ follows the fact that $\hat{\beta}$ is a weighted average of the OLS estimators (for a detailed proof, see Greene, 2000, p. 610).

To estimate the unknown parameters in $\mathbf{\Gamma}$ and $\mathbf{V}_i$ Swamy (1971) suggested the following procedure. Let $\mathbf{b}_i$ be the group specific OLS coefficient vector and let $\hat{\mathbf{V}}_i$ be the sample covariance matrix,

$$\mathbf{s}_i^2(\mathbf{X}_i'\mathbf{X}_i)^{-1},$$

where

$$\mathbf{s}_i^2 = \frac{\mathbf{e}_i'\mathbf{e}_i}{T_i - K};$$

now

$$\bar{\mathbf{b}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{b}_i,$$

then

$$\hat{\mathbf{\Gamma}} = \frac{1}{N-1}\left(\sum_{i=1}^{N}\mathbf{b}_i\mathbf{b}_i' - N\bar{\mathbf{b}}\bar{\mathbf{b}}'\right) - \frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{V}}_i$$

If the second matrix in $\hat{\mathbf{\Gamma}}$ is quite big it is possible that $\hat{\mathbf{\Gamma}}$ is not positive definite anymore. In big samples the second matrix is negligibly small but in small samples

$\hat{\mathbf{\Gamma}}$ might become not positive definite. A simple and asymptotical valid solution for this problem is, just to drop the second matrix. For the calculations in this paper this asymptotical valid form of $\hat{\mathbf{\Gamma}}$ was used, i.e. the matrix $\hat{\mathbf{V}}_i$ was not included in the estimation of $\hat{\mathbf{\Gamma}}$.

Now predictors for the individual parameter vectors can be calculated. The best linear predictor for β_i is:

$$\hat{\beta}_i = [\mathbf{\Gamma}^{-1} + \mathbf{V}^{-1}]^{-1}[\mathbf{\Gamma}^{-1}\hat{\beta} + \mathbf{V}_i^{-1}\mathbf{b}_i] = \mathbf{A}_i\hat{\beta} + [\mathbf{I} - \mathbf{A}_i]\mathbf{b}_i$$

where

$$\mathbf{A}_i = (\mathbf{\Gamma}^{-1} + \mathbf{V}_i^{-1})^{-1}\mathbf{\Gamma}^{-1}$$

and

$$\mathbf{V}_i = \sigma_i^2(\mathbf{X}_i'\mathbf{X}_i)^{-1}.$$

This predictor is again a matrix weighted average. The weights are the inverse of the covariance matrix of $\hat{\beta}_i$ and $\mathbf{b}_i$. In practice the estimators $\hat{\mathbf{\Gamma}}$ and $\hat{\mathbf{V}}_i$ are used for $\mathbf{\Gamma}$ and $\mathbf{V}_i$.

The variance of the predictor $\hat{\beta}_i$ is given by

$$\mathrm{Var}[\hat{\beta}_i] = \begin{bmatrix} \mathbf{A_i} \\ \mathbf{I} - \mathbf{A}_i \end{bmatrix}' \begin{bmatrix} \sum_{i=1}^{n} \mathbf{W}_i(\mathbf{\Gamma} + \mathbf{V}_i)\mathbf{W}_i' & \mathbf{W}_i(\mathbf{\Gamma} + \mathbf{V}_i) \\ (\mathbf{\Gamma} + \mathbf{V}_i)\mathbf{W}_i' & (\mathbf{\Gamma} + \mathbf{V}_i) \end{bmatrix} \begin{bmatrix} \mathbf{A}_i \\ \mathbf{I} - \mathbf{A}_i. \end{bmatrix}$$

Parameters estimated according to this specification partly differ considerably from the OLS estimators of Coe and Helpman (1995) (as well as the DOLS estimators of Kao et al., 1999). Even to such an extent that the sign of single parameters may depend on the choice of the model – a model with fixed or a model with random coefficients (cf. Müller and Nettekoven, 1999). Other than the fixed effect model the random effect (coefficient) one assumes the existing array of countries represents a random draw from a (fictitious) population of similar economies.

A correct random coefficient estimation yields

$$\begin{array}{lcccc} \ln F_{it} = \hat{\alpha}_{it}^0 & + & 0.2475 \ln S_{it}^d & - & 0.0841\, m_{i,t-1} \ln S_{it}^f + \text{rest}\,, \\ & & (7.7578)^{**} & & (-0.5087) \end{array}$$

with an R^2 of 0.9122 and an Adjusted R^2 of 0.9074.

The estimates for the random coefficient model differ decisively from the fixed coefficient model and especially the estimator of the foreign R&D expenditures changed sign, although this is not statistically significant. Values for R^2 and Adjusted R^2 raised, both are now around 0.91, so the explanatory power of the model is quite good. Contrary to the other estimations so far (and Coe and Helpman's conclusions) this model indicates that the foreign spillover effect is not significant!

Note that although Müller and Nettekoven (1999) have already identified this effect, they report other estimates for the random coefficient model. This is due to erroneously relating foreign R&D expenditures of some countries to domestic R&D expenditures and TFP of other countries.

4 Time Series Added-Variable Plots

In this section we will demonstrate that by a proper exploratory analysis it could have been possible to detect the reported inconsistencies of the data with the posited models and the achieved results.

Although at first glance Figures 1 and 2 seem to confirm Coe and Helpman's conclusions, we propose in a second step to construct so-called Added-Variable-Plots of the data. Such plots are used to analyse the importance of additional variables for the explanation of the variation of a dependent variable and they are capable of discovering masking effects.

The relevant question that should be answered is whether foreign R&D expenditures can provide any explanation of the variation of TFP additional to the explanation provided by domestic R&D expenditures. Is it useful to add FRD into the model when it already includes DRD? To answer this question, additional to the time-series scatterplot of TFP against domestic R&D expenditures for G7 countries (see Figure 1) and the time-series scatterplot of TFP against foreign R&D expenditures for G7 countries (see Figure 2) a third time-series scatterplots is of relevance, the one of foreign R&D expenditures against domestic R&D expenditures (see Figure 3).

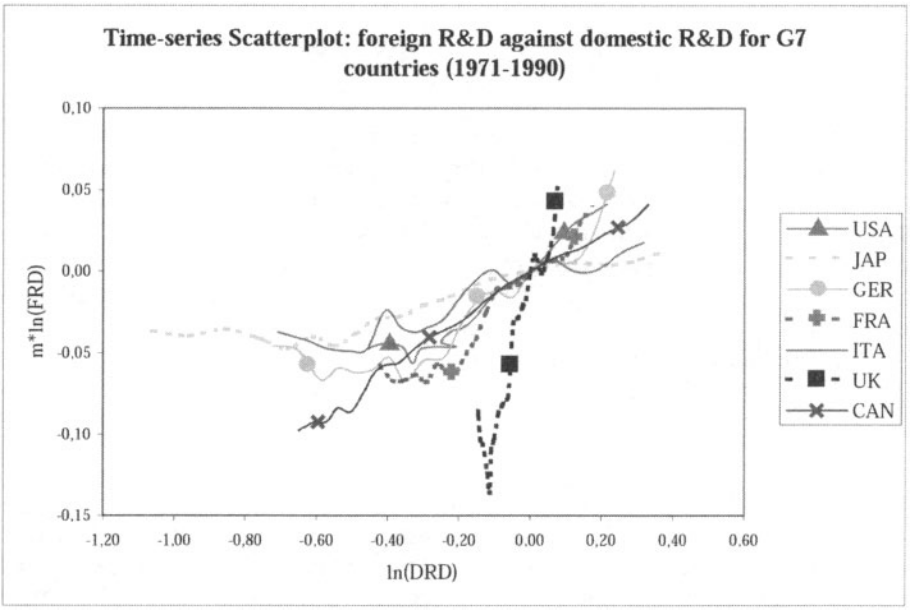

Figure 3: Time-series scatterplot FRD against DRD for G7 countries (1971-90).

The first two time-series scatterplots show positive correlation and this might lead to the (somewhat premature) conclusion that domestic as well as foreign R&D expenditures might be able to explain TFP. However, the time-series scatterplot of FRD against DRD also shows a positive correlation between those variables, which strongly

indicates that one of the regressors might carry mainly redundant information.

Added-Variable-Plots rather than plotting original variables like in usual scatterplots employ the partial effects of the considered regressors manifested by the residuals of corresponding OLS regressions (for a detailed description see Cook and Weisberg, 1994). Thus in our context two time-series Added-Variable-Plots will be useful: one shows the residuals from a simple OLS regression of the natural logarithm of TFP on the natural logarithms of the domestic R&D expenditures (y-axis) against the residuals of a simple OLS regression of the natural logarithms of the foreign R&D expenditures multiplied with the import-shares on the natural logarithms of the domestic R&D expenditures (x-axis), i.e. it will be displaying the partial effect of FRD on TFP (Figure 4). The other will be constructed vice versa for the partial effect of DRD on TFP (Figure 5). Both OLS regressions were calculated for each country separately.

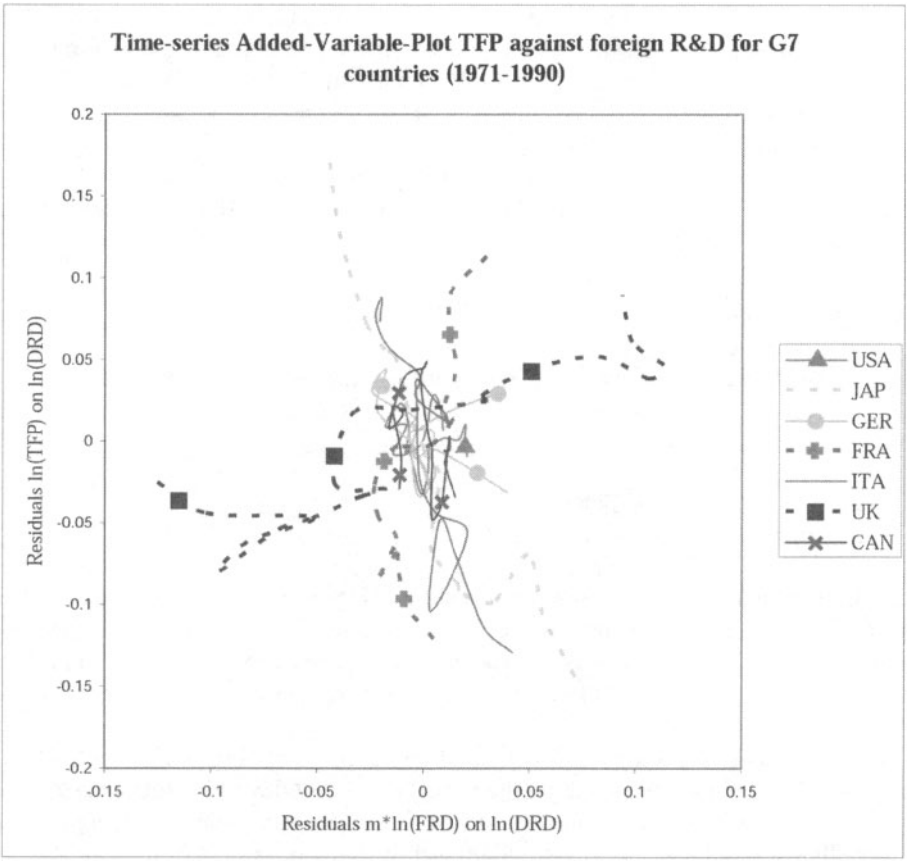

Figure 4: Time-series A-V-Plot TFP against FRD for G7 countries (1971–90).

For the interpretation of these time-series Added-Variable-Plots it is useful to have a look at the three extreme cases of Added-Variable Plots. If all points lie exactly on

a straight line with nonzero slope, this means all residuals from the A-V-P regression are zero, it is useful to add the second variable (e.g. in Figure 4: $m\ln(FRD)$) to the first variable (in Figure 4: $\ln(DRD)$) because this will give a perfect fit of the model. If all points lie on a horizontal line all variation of the dependent variable is explained by the first variable, so there is no need to insert the second variable into the model. If all points lie on a vertical line the second variable is a linear function of the first one and it is not useful to include the second variable into the model because it cannot explain the variation any further.

The A-V-P for FRD of the R&D Spillovers data for the G7 countries (Figure 4) shows that the lines for Japan, France and Italy are nearly vertical, for Japan and Italy they even show a negative trend, this means that FRD are nearly a linear function of DRD and therefore they are redundant for the model. The lines for the U.S. and Canada do not show any trend, they vary randomly and therefore FRD of the U.S. and Canada are also not able to provide additional explanation to the model with DRD only. The line for the U.K. shows a slight positive trend but it is still nearly horizontal and therefore can provide nearly no further explanation of the variation of TFP. In the A-V-P for DRD (Figure 5) on the other hand, with the exception of the U.K., all countries exhibit an upward slope, which confirms the relevance of this factor.

All in all, it is evident from this exploratory analysis that the use of A-V-P's in an early phase of their study would have prevented Coe and Helpman (1995) from premature conclusions.

5 A New Model

After a detailed examination of the model of Coe and Helpman (1995) and the comments of its various critics, the following changes and modifications for this model are suggested:

- use of a random coefficient model,
- use of DOLS regression.

The advantage of the random coefficient model over the fixed effect model is that there is no need for the assumption that there is no variation between the cross section units (countries). This parameter heterogeneity is regarded as a random variation. The fit of the model improves by allowing for the random variation of the single parameters β_i around β.

The use of dynamic regressors is based on the paper of Kao et al. (1999) where the DOLS estimator and its advantage over the simple OLS estimator is explained. Thus q_1 lags and q_2 leads of the first differences of the domestic and foreign R&D expenditures should be included as additional dynamic regressors in the model with domestic and foreign R&D expenditures. Kao et al. (1999) tested the assumption of cointegration of the estimated equations. They used the panel cointegration test of Kao (1999) and the test of Pedroni (1995). All test-statistics were significant and therefore the null-hypothesis of no cointegration was rejected (see Kao et al., 1999, Table 2). Edmond (2001) used cointegration tests of Pedroni (1997, 1998) and the augmented

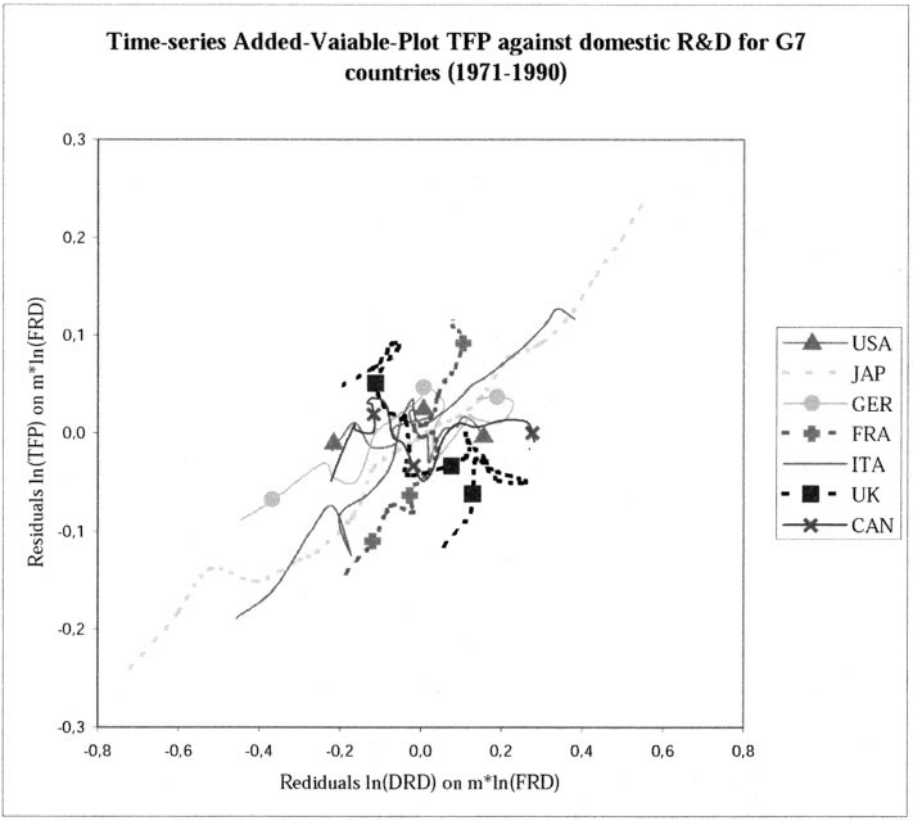

Figure 5: Time-series A-V-Plot TFP against DRD for G7 countries (1971–90).

Dickey-Fuller test to test the assumption of cointegration and came to the same result as Kao et al. (1999), see Table 2 of Edmond (2001). Because of these results the R&D Spillovers data can be regarded as cointegrated.

The suggestion for the analysis of the international R&D Spillovers is a random coefficient panel cointegration model with dynamic regressors. In this case Coe and Helpman's (1995) model has the following specification:

$$\mathbf{y}_i = \mathbf{X}_i \beta_i + \epsilon_i,$$

with

$\mathbf{y}_i$ being the regressand $T - (q_1 + q_2 + 1)$ vector; here the natural logarithm of TFP of country i: $\ln F_i$,

$\mathbf{X}_i$ denotes the $T - (q_1 + q_2 + 1) \times 1 + k + (q_1 + q_2 + 1)k$ regressor-matrix; here $\mathbf{X}_i = [\mathbf{1} : \ln \mathbf{S}_i^d : m \ln \mathbf{S}_i^f : \Delta \mathbf{X}_i]$, where $\ln \mathbf{S}_i^d$ and $\ln \mathbf{S}_i^f$ are the natural logarithms of the domestic- and foreign R&D expenditures and m are the import

shares. $\Delta \mathbf{X}_i$ is the group-specific matrix of the differences. For all elements of the original matrix $\mathbf{X}_i$ the values for the respective years get subtracted for all lags and leads (inclusive lag = 0). The dimension of matrix $\Delta \mathbf{X}_i$ is thus: $T - (q_1 + q_2 + 1) \times (q_1 + q_2 + 1)k$.

β_i is then the $1 + k + (q_1 + q_2 + 1)k$ vector of parameters. The first entry is the common intercept, the second and third entries are the parameters of the domestic and foreign R&D expenditures and the other entries are the parameters of the differences.

ϵ_i finally denotes the corresponding $T - (q_1 + q_2 + 1)$ vector of errors.

In contrast to Coe and Helpman's model, where all years T are included, q_1+q_2+1 years get lost in this model by forming the difference-matrix. Eventually, the DOLS random coefficient estimation yields

$$\ln F_{it} = \hat{\alpha}_{it}^0 + \underset{(8.0541)^{**}}{0.3062 \ln S_{it}^d} + \underset{(0.4051)}{0.0756\, m_{i,t-1} \ln S_{it}^f} + \text{rest}$$

with an R^2 of 0.9766 and an Adjusted R^2 of 0.9750. The computations were performed with a special GAUSS package, which is described in detail in N. Gumprecht (2003).

The coefficient estimate corresponding to $m \ln S_f$ is not significant. R^2 and Adjusted R^2 are even higher than in the random coefficient model without dynamic regressors.

The results of the panel cointegration model with random coefficient and dynamic regressors do not support Coe and Helpman's hypothesis that the TFP of a country depends on domestic and foreign R&D knowledge (measured by the R&D expenditures). The effect of the knowledge of the trade-partners of a country is marginal because it is not significant. It seems as if foreign R&D did rather not affect the TFP of a country.

6 Conclusions

Coe and Helpman's hypothesis that the TFP of a country depends on the domestic and foreign R&D knowledge can only be partly supported. Imported knowledge seems to have no effect on TFP of a country.

A summary of different articles about the relationship of imported knowledge and the TFP of a country is provided by Navaretti and Tarr (2000). They concluded that there is a strong evidence for the positive effect of imported technology on TFP of a county. The reason for this completely different conclusion might be the level of aggregation of the analysed data. Navaretti and Tarr (2000) used articles that cared especially about microeconomic relationships between trade and knowledge diffusion, whereas the articles discussed here care about a macroeconomic relationship between R&D and TFP. One referee noted the relation of the results to the theory of absorptive capacities (Cohen and Levinthal, 1989) – a respective interaction effect could be tested in a slightly alternative specification.

The preferred panel cointegration model with random coefficients and dynamic regressors confirms the positive effect of domestic R&D on TFP but it does not confirm the effect of foreign R&D on TFP.

Acknowledgements

We are most grateful to Werner Hölzl for a number of useful comments that led to an improvement in the paper. Many thanks go to Sonja Steffek and Gunther Maier for their technical help in the format conversion of the manuscript.

References

Baltagi, B. H. (2001) Econometric analysis of panel data. 2nd Edition. John Wiley & Sons, West Sussex

Chiang, Min-Hsien and Kao. Ch.(2002) Nonstationary panel time series using NPT 1.3 - A User Guide. National Cheng-Kung University and Syracuse University.

Chiang, Min-Hsien and Kao, Ch.(2002) "Chihwa Kao." <http://web. syr.edu/ cdkao>

Cohen, W. and Levinthal, D. (1989) Innovation and learning: The two faces of R&D. Economic Journal 99: 569–196

Coe, D. T. and Helpman E. (1995) International R&D spillovers. European Economic Review 39: 859–887

Cook, D. R. and Weisberg, St. (1994) An introduction to regression graphics. Wiley, New York

Dickey, D. and Fuller, W. (1979) Distribution of the estimators for autoregressive time series with a Unit Root. Journal of the American Statistical Association 74: 427–431

Dickey, D. and Fuller, W. (1981) Likelihood ratio statistics for autoregressive time series with unit root. Econometrica 49: 1057–1072

Edmond, Ch. (2001) Some panel cointegration models of international R&D spillovers. Journal of Macroeconomics 23: 241–260

Engle, R. F., and Granger, C. W. J. (1987) Co-integration and error correction: Representation, estimation and testing. Econometrica 55: 251–276

Granger, C. W. J. and Newbold, P. (1974) Spurious regressions in econometrics. Journal of Econometrics 2: 111–120

Greene, W. H. (2000) Econometric analysis. 4th Edition. Prentice Hall, New Jersey

Grossman, G. M. and Helpman, E. (1991) Innovation and growth in the global economy. MIT Press, Cambridge, Mass.

Gumprecht, D. (2003) Ein Panel Kointegrationsmodell für internationale Forschungs- und Entwicklungs Spillovers. Unpublished master thesis at the Department of Statistics and Decision Support Systems, University of Vienna.

Gumprecht, N. (2003) Regression mit zufälligen Koeffizienten: Software Implementierung. Unpublished master thesis at the Department of Statistics and Decision Support Systems, University of Vienna.

Helpman, E. (2003) Professor Elhanan Helpman's Data On The Web." <http://post.-economics.harvard.edu/faculty/helpman/data.html>

Hsiao, Ch. (1986) Analysis of panel data. 1st Edition. Cambridge University Press, New York

Kao, Ch. (1999) Spurious regression and residual-based tests for cointegration in panel data. Journal of Econometrics 90: 1–44

Kao, Ch. and Chiang, M.-H. (1997) On the estimation and inference of a cointegrated regression in panel data. Syracuse University

Kao, Ch., Chiang, M.-H. and Chen, B. (1999) International R&D spillovers: An application of estimation and inference in panel cointegration. Oxford Bulletin of Economics and Statistics 61: 693–711

Levin, A. and Lin, Ch.-F. (1992) Unit foot tests in panel data: asymptotic and finite-sample properties. Discussion paper 92-93, University of California, San Diego, CA

Levin, A. and Lin, Ch.-F. (1993) Unit root tests in panel data: New results. Discussion paper 93-56, University of California, San Diego, CA

Müller, W. G. and Nettekoven, M. (1999) A panel data analysis: research and development spillover. Economics Letters 64: 37–41

Navaretti, G. B. and Tarr, D. International knowledge flows and economic performance: A review of the evidence. The World Bank Economic Review 14: 1–15

Pedroni, P. (1995) Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Indiana University, Working Paper in Economics No. 95-013

Pedroni, P. (1997) Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis: new results. Unpublished manuscript

Pedroni, P. (1998) Approximate critical values for cointegration tests in heterogeneous panels with multiple regressors. Unpublished manuscript

Romer, P.M. (1990) Endogenous technical change. Journal of Political Economy 98: S71–S102

Saikkonen, P. (1991) Asymptotically efficient estimation of cointegration regressions. Econometric Theory 7: 1–21.

Stock, J.H. (1987) Asymptotic properties of least squares estimations of co-integrating vectors. Econometrica 55: 1035–1056

Swamy, P. (1971) Statistical inference in random coefficient regression models. Springer, New York

Estimating Universities' Contributions to Regional Economic Development: The Case of the U.S.

Harvey A. Goldstein*, Catherine S. Renault*

1 Introduction

Since the mid-1980s, academic interest in the relationship between knowledge production within a region and the region's economic growth and development performance and prospects has burgeoned. The reasons for this increased interest include dramatic changes in the global economy and conditions of regional competitiveness, the increased importance of knowledge inputs in the production of a wide range of goods and services, and more recently advances in "the new growth theory".

The relatively severe economic downturn of 1981–1982 hit the traditional manufacturing sectors particularly hard. In response, economic development officials at the state and regional levels began investing in a variety of programs and institutions aimed at strengthening their regions' knowledge infrastructure. Public universities were often the primary institutional recipient for many of the investments in knowledge infrastructure. Faced with budgetary squeezes, leaders at many public universities were only too eager to accept responsibility (and the funding that came with it) for adding economic development to their traditional missions of teaching, research, and public service. The passage of the Bayh–Dole Act of 1980 gave universities an even larger incentive to engage in entrepreneurial activities, by giving them intellectual property rights on patentable inventions originally stemming from federal government-sponsored research.

The purposes of these public investments in knowledge infrastructure have varied among states and regions, depending upon particular economic conditions and perceived needs. Generally, however, they have focused on attracting, nurturing, and retaining high tech industries and innovative firms, to provide existing residents with the range of skills and competencies they will need to be productive in the knowledge economy, and to help the region sustain its competitiveness into the future. Beliefs about the importance of such investments tended to be based neither upon sound empirical evaluations nor theoretical arguments. Instead they emanated from a few well-known and celebrated "success" cases such as Silicon Valley in California, Route 128 in Massachusetts, the Research Triangle in North Carolina: "We can be the region spawning the next Apple Computer." There was also a tendency for state and regional

*Department of City and Regional Planning, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

*Department of Public Policy, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

development officials to play "follow the leader," so as not to risk being the one left behind (Atkinson, 1989).

For any significant public investment, there is always a range of questions about the magnitude of impacts, their distribution, and the investment's effectiveness in achieving desired objectives. In the particular case of investments in the production of knowledge and knowledge infrastructure, the expected regional economic impacts revolve around the nature of spatial spillovers and other forms of externalities generated from knowledge producing organizations. If we focus on universities as the dominant type of knowledge producing organizations,[2] the following questions should be addressed:

1. To what extent do institutions of higher education, and research universities specifically, generate regional economic development outcomes that would otherwise not occur?
2. Which university-based activities, e.g. teaching, basic research, extension and public service, technology transfer, technology development, businesses spinning off from university research, etc, are most responsible for any net regional economic development impacts from the presence of universities?
3. Through what mechanisms, or channels, does knowledge production – broadly considered – within universities lead to economic development outcomes in the surrounding region? Is it, for example, through economic transactions between actors or units within the university and external organizations, through spillovers, or through milieu effects, which are particular kinds of localization economies?
4. What are the critical internal and external factors that condition the contribution of knowledge-producing organizations to regional economic development, and determine the share of total economic development impacts that are retained within the region, vis-à-vis the rest-of-the-world?

In this paper we focus on addressing the first, the second, and the fourth questions. Our empirical analysis employs a quasi-experimental design with a large sample of metropolitan regions in the U.S. The results provide us with some indirect insights on the third question, the modes of transmission of impacts from universities to the larger region.

The paper is organized into six sections. The next section lays out conceptually the different ways universities potentially may contribute to regional economic development. The third section provides a brief, critical review of the two primary methodological approaches that have been used to estimate university impacts on regional economic development, and introduces the advantages of using a quasi-experimental design. In the fourth section we describe the study population, the measures, the data, the hypotheses, and the models used in the present study, while the empirical results are presented in the fifth section. The sixth and last section is a conclusion that focuses

[2]Other knowledge producing organizations include federal laboratories and the full array of corporate and non-profit R&D labs. In terms of total U.S. R&D expenditures, however, universities comprised 14.0 percent in 1998 (National Science Board, 2000).

on the nature of university spillovers to their regional environment, given the empirical results.

2 Outputs and Potential Impacts of Research Universities

The literature on the economics of knowledge production and innovation, on knowledge spillovers, and on the changing role of institutions of higher education together suggest there is a wide range of ways that universities potentially can contribute directly and indirectly to regional economic development. The modern research university, at least in the United States and increasingly in Western Europe, can be considered as a multiproduct organization. Within legal, cultural, and political constraints, the research university can choose its optimal product mix in response to perceived changes in its markets and the availability and prices of its inputs (unlike corporate R&D organizations, universities do not consider changing locations as responses to changing market conditions).

Goldstein, Maier, and Luger (1995) synthesized a wide range of literature on institutions of higher education and identified and described the range of products, or outputs, from modern research universities. They then suggested how each of these types of outputs may potentially lead to specific economic development impacts. The outputs include: (i) knowledge creation, (ii) human capital creation, (iii) transfer of existing know-how, (iv) technological innovation, (v) capital investment, (vi) provision of regional leadership, (vii) coproduction of the knowledge infrastructure, and (viii) coproduction of a particular type of regional milieu. The potential impacts include: productivity gains, business innovation, new business start-ups, an increase in regional economic development capacity (for sustained, long-term development), regional creativity, and direct and indirect spending impacts. The hypothesized relationships between the university outputs and the hypothesized regional development impacts can be seen in Figure 1.

3 Available Methodological Approaches

The literature suggests there are two primary ways to investigate the link between knowledge production and regional economic development: the case study, and econometric models of knowledge production and spillovers.

3.1 The Case Study

We can trace case studies of the economic impact of specific universities in the published literature at least back to the early 1970s. These case studies attempt to estimate the contribution of a particular institution of higher education to the regional economy in one of two typical ways. In the first, a regional input-output model is used to estimate the direct and indirect spending effects of the organization on the region's output, earnings, and employment. In the second, a sample of other businesses or organizations in the region are asked in a questionnaire what the perceived importance of the

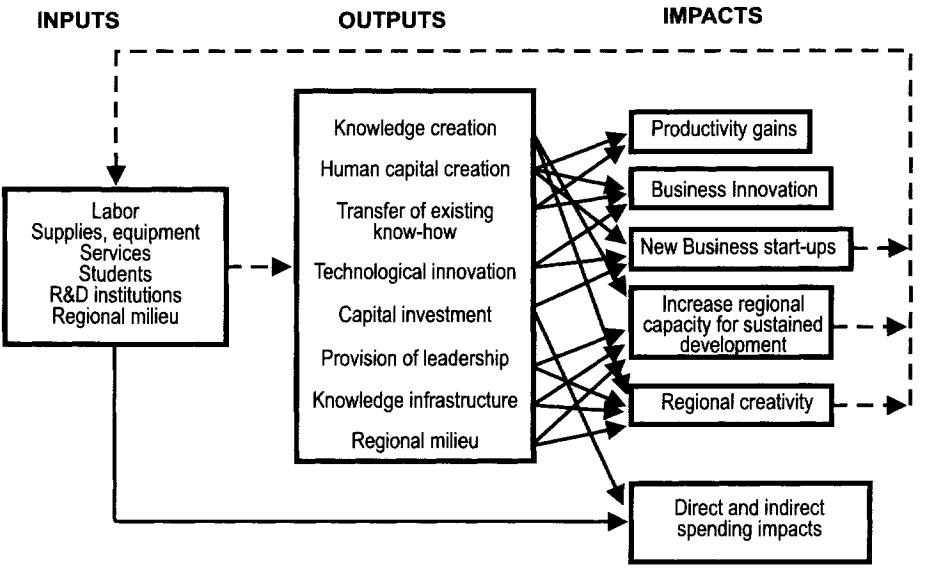

Figure 1: University outputs and expected economic impacts (Source: Goldstein, Maier and Luger, 1995).

university was to the location decision, productivity and competitiveness, innovation, and output levels of the respondent's organization.

A recent article by Thanki (1999) reviews much of the case study literature, particularly in the experience of the United Kingdom. The author finds that such studies todate have not been able to capture many of the potential ways that institutions of higher education can lead to regional economic development, either because of a too narrow understanding of how regional development occurs, or because of the limitations of the techniques used. For instance, regional input-output models and other multiplier techniques can only capture economic growth that stems from backward linkages induced by an institution of higher education's spending. The reliance on only such techniques fails to take into account the unique qualities of knowledge as a product of such organizations, essentially treating a university no differently from any other kind of organization that hires and pays labor and purchases supplies and equipment from both regional and outside sources. Felsenstein (1994a) reviews a large number of economic impact studies conducted in the U.S. using input-output and related multiplier techniques. For the most part, these studies suffer from the same set of limitations as discussed by Thanki (1999).

Within the case study framework, a more eclectic set of data collection and analytic techniques have been used to try to estimate a wider range of putative regional economic effects of universities, in addition to those measured by input-output models. Several researchers (Brett et al., 1991, Smilor et al., 1990, Ko, 1993, Steffenson et al., 2000) have focused on "counting" the number of spin-offs from university re-

search centers and academic departments, using newspapers and the regional business press, as well as personal interviews as sources. Luger and Goldstein (1991) and Goldstein and Luger (1992, 1993) have used targeted questions sent in a questionnaire to a stratified sample of high tech companies recently located in a region to estimate to what extent the region's research universities were an important factor in their regional location decision. By asking the counterfactual, "If the University of . . . were not in the region, how likely would you have located your business in this region: (1) Very likely, (2) Likely, (3) Perhaps, (4) Unlikely, (5) Very unlikely," and then applying a specific probability to each response, the researchers were able to estimate the total employment and output in the region's high tech sector that otherwise would not have been located there.

Goldstein and Luger (1992) also have estimated the human capital impacts on a region of a university by linking historical student registration and alumni records to estimate the percent of graduates, by degree and discipline, who have remained and been employed within the region over time. Felsenstein (1994b) estimates the amount of university-induced migration to a region (that would not have otherwise occurred) using Keynesian-type multipliers.

Saxenian's (1994) comparative case studies of Silicon Valley and Route 128 in Massachusetts, while not attempting to measure the impact of specific universities on regional development, tries to explain the difference in economic performance and success between the two areas by using ethnographic techniques focusing on the extent and qualities of inter-organizational linkages and collaborations within the respective regions.

In summary, the principal strengths of the case study approach are: (1) its flexibility in being able to collect primary data through surveys and interviews, and documentary evidence in the form of internal reports, in order to measure, or count, particular kinds of university outputs and some of the associated impacts; (2) to be able to generate information about the internal organization and culture of a particular university in order to be able to relate these internal attributes to effectiveness in contributing to regional economic development; (3) being able to generate information about the soft and "fuzzy" side of how universities may stimulate economic development: through the provision of leadership and the co-production of a creative milieu, concepts that are difficult to measure. The often intensive data gathering effort to estimate the spending impacts of the university on the regional economy with the use of input-output and related multiplier techniques is almost always conducted within a case study approach. The case study approach is most appropriate for counting or estimating incidences of technology transfer, technological innovation (e.g. in the form of patents issued and licenses sold), capital investment, and number of business spin-offs.

The principal weakness of the case study approach is solving the "attribution" problem, i.e. controlling for other putative factors when trying to make causal inferences between university activities and regional economic development outcomes. This is most problematic for estimating indirect impacts such as regional productivity gains, increases in innovative activity in the region, increased economic development capacity, and regional creativity. In principle one could conduct multiple cases studies using a standard set of data and techniques in order to mitigate the inability to statis-

tically control for other intervening factors, but this is likely to be impractical owing to the relatively large cost of conducting individual case studies. In several of the case studies discussed above, researchers have asked respondents about the "perceived" impacts, or "perceived importance" of the university, or have constructed counterfactuals, but the use of such data and techniques carries with them some obvious and significant validity threats.

3.2 Econometric Models of Knowledge Production and Spillovers

More recently, economists have tried to assess the economic impact of new innovation such as that produced at a university by using econometric models. Early models focused on production functions that contained an R&D variable. Griliches (1979) proposed a specific knowledge production function assumed to be Cobb–Douglas in form, relating Y, some output, to labor, capital and knowledge, measured by R&D expenditures. This work was based on the empirical work of Solow (1957), which showed the existence of a latent variable besides capital and labor in aggregate production functions, and Arrow (1962), who described the impact of learning (gaining knowledge) on the production function.

Jaffe (1989) modified Griliches' knowledge production function to the form that is generally used by researchers today: $LogK = \alpha_0 + \alpha_1 logRD + \alpha_2 logURD + \epsilon$ where K is some measure of innovation, *RD* is industry research and development expenditures and *URD* is university research and development expenditures. Jaffe found "geographically mediated spillovers from university research to commercial innovation." A significant effect of university research on corporate patents was found, particularly in drugs, medical technology, electronics, optics and nuclear technology. He states that university research appears to have an indirect effect on local innovation by inducing industrial research and development spending.

A number of researchers have modified and expanded this model in the last twelve years or so. Jaffe, Trajtenberg and Henderson (1993) followed up this work by looking at the geographic location of patent citations to understand the extent to which knowledge spillovers are geographically localized. They found that citations to domestic patents are more likely to be domestic, and more likely to come from the same state and MSA as the cited patents. Acs, Audretsch and Feldman (1994) used the Griliches–Jaffe model to understand the degree to which university and corporate R&D spills over to small firms as compared to large firms. They found that university spillover was more important to small firms than large firms. Varga (2000) uses the Griliches-Jaffe knowledge production function to try to model agglomeration. He develops a hierarchical linear model by modeling each α, to include the dependence of the knowledge transfer on the concentration of high technology production and business services.

Keilbach (2000) has taken a different path. Incorporating the work of Krugman (1991) and Fujita et al. (1999), he builds a dynamic model of regions using the endogenous growth model including human capital and knowledge spillovers, to be able to separate the effects of spillover and agglomeration. The model builds a class of discrete dynamic systems, an n-dimensional array of sites, each with local rules as well as

global rules. Assuming the initial endowments of capital and labor are identical in each region and that the distribution of knowledge is random, that capital moves globally, labor migrates locally, but human capital spills over, Keilbach observes agglomeration under assumptions of increasing returns to R&D only. The model demonstrates a self-reinforcing structure where labor agglomerates no matter what, but agglomerates quicker with increasing returns to R&D.

The various authors above, as well as others, have considerably moved forward our understanding of the nature of knowledge spillovers. These studies demonstrate that knowledge, as an output of an organization, has unique qualities that render knowledge producing organizations not just like any other large organization that purchases inputs locally and employs a large of number of people. Rather there are significant external benefits of knowledge production in the form of spatial spillovers that lead to higher levels of innovative activity among other firms within the region.

There are, however, some limitations in the use of this approach for estimating the contribution of universities to regional economic development. These include measurement and data issues, the limited scope of potential outputs and impacts stemming from universities, and thus the limited ability to separate different ways that universities contribute to regional economic development.

The primary challenge with using these types of models is how to measure the dependent variable, innovation. Jaffe (1989) quotes Griliches as saying, "The dream of getting hold of an output indicator of inventive activity is one of the strong motivating forces for economic research in this area." Since we do not know how to measure innovation directly, we use proxy indicators, each with their own inherent challenges.

Varga (2000) and Acs, Audretsch and Feldman (1994) use a one-time cross section data set from the Small Business Administration on new product introductions. They reason that these innovations had market value since companies introduced them and that this was a better measure than patents or other pre-market indicators. From a company point of view, product introductions are still indicators of "probable" market value. Until sales of a product are known, the market value is still a prediction. So, while it is true that product introductions, being further down the product development chain than patents, are probably better indicators of innovations with market value, they are still predictors, not measures. The other significant issue with this data set is that it is a one-time indicator, constructed in 1982, and now highly dated.

Jaffe (1989) and Jaffe, Trajtenberg and Henderson (1993) use the number of patents as their dependent variable. There are a lot of issues with this indicator, as Jaffe recognized at the time. For instance, there are factors outside knowledge generation that affect the number of patents that may arise from a given piece of research and development.

First, not all innovation is patentable. Other types of innovation not included in patents are codified knowledge embodied in copyrights, trade secrets and scientific papers as well as tacit know-how and shared expertise. Since the transmission of tacit information is an important explanation of why spillover has local effects (as compared to patents that can be licensed worldwide), our inability to measure this dimension is a significant handicap.

Second, especially within universities, changes in patenting policies and the uni-

versity's fiscal environment significantly affect the number of patents applied for and granted. Thursby and Thursby (2000), for instance, have shown that the recent growth in licensing activity at U.S. research universities has been based on input growth (i.e. more disclosures) of lowering quality and an increase in the propensity of administrators to patent and license faculty inventions as a revenue raising strategy.

In fact, the general hypothesis has been that patents are a proxy for inventive output and patent citations are a proxy for knowledge flows or spillover, the real source of innovation. Recent work (Hall, Jaffe and Trajtenberg, 2000) strongly suggests that patent citations are well correlated with the market value of knowledge and are the best indicator of innovation and knowledge spillover. To date, patent citations have not been used in the knowledge production function, although Jaffe and Trajtenberg (1996) do use them to find that diffusion of knowledge is geographically localized.[3]

The second major limitation of this approach for our purposes is that while it yields estimates on some of the outputs of universities, i.e. technological innovation, it does not take into account other outputs of universities discussed in section 2 above – human capital creation, building regional capacity, stimulating business start-ups, and the regional milieu effect – that potentially have regional economic development impacts.

4 A Quasi-Experimental Approach

Quasi-experimental designs are best described in comparison to true experimental designs. The hallmarks of true experimental designs are that the researcher can manipulate how and which cases receive a "treatment", such that the treatment can be randomly assigned to cases drawn from a sample of a given study population. Randomization means that probabilistically there should be no systematic difference on any rival causal factors between those cases receiving the treatment and those that do not. Thus any differences in the dependent variable between the treatment and non-treatment groups can normally be attributed to the treatment effect.

In field settings as opposed to laboratory settings, the ability of researchers to randomly select cases for treatment is rare. Instead they must accept the naturally occurring variation in a given treatment variable, in this case the regional distribution of universities by size and type. In other cases, it may be technically feasible for researchers to manipulate the selection of cases for treatment or nontreatment, but it is ethically prohibited. Quasi-experimental designs are a class of designs in which the researcher strategically manipulates the study population, the time period, and the sampling so as to construct control groups that enable one to control for some of the rival factors that cannot be explicitly included as variables in a statistical model. Cook and Campbell (1979) and Shadish, Cook and Campbell (2002) provide comprehensive and detailed descriptions of quasi-experimental designs.

The control group cases can not be assumed to be equivalent to the treatment cases in all respects except for the treatment variable itself. However, such selection differ-

[3] A third measure of innovation that has been used is R&D intensity. This is measured by the number or percentage of R&D employees. Keilbach, when testing his model on 327 Kreise (districts) of West Germany, used R&D employment as a proxy for R&D intensity.

ences can be taken into account by measuring and comparing gain scores – differences in the outcome variable(s) between a pretest and posttest – instead of just comparing posttest scores.

In the context of conducting research on the relationship between knowledge production in a region and regional economic development outcomes, the presence of research universities or other significant knowledge-producing organization in a region would represent the treatment variable. The dependent variable would be the gain, or difference, of some meaningful economic development outcome over a pertinent time period. The control group would consist of regions that did not have research universities (or other significant knowledge-producing organizations). Additional control groups may be selected so as to effectively control for other factors that might also explain variations in the dependent variable in addition to the presence or absence of a research university. Thus, we might select two treatment groups, one with small- to medium-sized regions with research universities and one with large-sized regions with research universities, and two control groups (small to medium without research universities, and large without research universities). We then compare the gains for all four groups, to take into account the effect of agglomeration economies and its interaction with the presence of universities.

It is important to distinguish quasi-experimental designs from more commonly used large sample cross-sectional statistical analyses. In the latter, the researcher would randomly select a large sample of regions from the full population available, enter measures on a number of variables that the literature has suggested may explain the dependent variable into a multiple regression model, and then estimate the model. The interpretation of the effect of research universities would be based upon the significance, sign, and magnitude of the coefficient estimate on the measure(s) selected for the presence or level of activity of a research university in the region. In effect, the researcher statistically controls for all other factors by entering these explicitly into the model. In contrast, quasi-experimental designs first attempt to control for other putative factors by the strategic design of treatment and control groups such that for some putative causal variables there is no variation allowed. For others there is an attempt to maximize variation. Only after these possibilities are practically exhausted does the researcher introduce statistical controls.

Although quasi-experimental designs have been used occasionally in published regional research (e.g. Isserman, 1987, Isserman and Merrifield, 1987), to the authors' knowledge, they have not been utilized in attempts to estimate the contribution of knowledge-producing organizations to regional economic development.

4.1 Study Objectives, Unit of Analysis, and Study Population

The objectives of this study are (1) to estimate the magnitude of the contribution of universities to changes in regional economic well-being, controlling for other factors; and (2) to try to separate the regional economic development impacts of different functions of universities.

The unit of analysis in our study is the metropolitan statistical area (MSA). All variables are measured for the MSA. The study population consists of all 312 MSAs

in the U.S. on the basis of their 1990 Census geographical definitions.[4] The temporal period is from 1969 to 1998.

4.2 Measures of Regional Economic Development

Our measure of regional economic development is average annual earnings per worker. This contrasts with many studies that use per capita income as a measure of regional economic well-being. The most important difference between the two measures is that average earnings per worker takes into account only earned income, while per capita income includes unearned income including dividends, rent, interest, and transfer payments. These unearned sources of income are particularly significant in regions with a large number of retirees. Also, average earnings per worker takes into account only the economically active persons in the denominator. Finally, average earnings per worker focuses on the quality of jobs in a region as the most important dimension of improvement in regional economic well-being.

To separate out changes in national macroeconomic conditions over the time period, including changes in the value of the dollar, we construct a normalized index for each MSA by dividing the MSA's average earnings per worker by the U.S. average earnings per worker for the same year, and then multiplying this ratio by 100. Thus, an index value of 110.0 for a particular MSA means that its average earnings per worker for that year is 10.0 percent higher than that for the nation.

The dependent variable is calculated as the difference in the index value for a given MSA between two years. Thus, if a particular MSA had an index value of 110.0 in 1969, and a value of 120.0 in 1999, then the value of the dependent variable would be 10.0. Of course it is possible for MSAs to have negative values on the dependent variable.

One benefit of using change in the index value is that it helps control for some local factors that are endemic to particular areas. For example, in many small MSAs with large universities, the average earnings per worker is distorted downward by a relatively large number of students employed in low-wage, part-time jobs. This distortion is minimized when the difference, rather than the level, of average earnings per worker is used.

4.3 Measures of University Presence

We measure the presence of universities in an MSA in three alternative ways, in order to test specific hypotheses described below. Our first measure is whether there is a top 50 research university located within the MSA at the beginning of the respective time period. A top 50 designation is based upon rankings of universities in terms of total research expenditures, compiled by the National Science Foundation on an annual basis[5].

[4]We did not include five MSAs in our study population because they were considered outliers because of their size or their physical isolation: Anchorage (Alaska), Chicago, Honolulu, Los Angeles, and New York.

[5]National Science Foundation, CASPAR database, 2001.

The second measure is the sum of the total research expenditures over all universities located within the MSA for a given annual period. The third measure is the sum of all degrees awarded in all institutions of higher education located within the MSA, for a given year. This third measure directly takes into account the magnitude of the human capital creation, or the teaching function of higher education institutions, separate from their research and economic development functions.

4.4 Control Variables

There are, of course, a number of other causal factors that may help to explain variation in the magnitude of change in a region's economic well being. First are agglomeration economies, measured by the total MSA employment at the beginning of the time period. We also use three MSA size categories: small, medium, and large employment levels to partially define control groups.[6]

Other factors include:

Region of U.S. The region of the U.S. (Northeast, Midwest, South, and West) is used to control for broad shifts in the regional distribution of population, employment, and capital investment from the Northeast and Midwest to the South and West over much of the time period. Region of U.S. is measured as dichotomous variables and enter into the statistical model as three dummy variables.

Industry research activity. The research activity of industry is measured as the number of industry R&D facilities located in the MSA. This variable controls for other R&D activity outside universities that may stimulate regional economic development.[7]

Industry structure. We include three dimensions of the industry structure of MSAs: percent earnings in manufacturing, percent earnings in business services, and change in percent earnings from manufacturing over the time period. These variables take into account that concentrations of manufacturing and business services may increase a region's capacity for regional economic development.

Entrepreneurial Activity. Lacking better measures for all regions of the U.S. over the time period of the study, we use as a proxy the percent earnings in the MSA from proprietorships (self-employed).

Base-year level of average earnings per worker. To control for a possible endowment effect (the rich get richer . . .), we include the average earnings per worker at the start of the time period.

[6]The size categories chosen for 1969 are: (i) under 75,000 employment, (ii) 75,000 to 499,999, and (iii) 750,000 and over. The size categories for 1986 are: (i) less than 250,000 employment, (ii) 250,000 to 999,999, and (iii) 1,000,000 and over. This yielded 160, 104, and 25 MSAs, respectively in the three size categories in 1969, and 192, 59, and 27 MSAs, respectively, in 1986.

[7]We relied on the listings in *Industrial Research Laboratories of the U.S.* (New York: R.R. Bowker & Co., 1970, 1985).

4.5 Hypotheses and Designs To Test Them

The first hypothesis is that research universities contribute significantly to regional economic development, controlling for other factors. This is tested in two ways. First, if true, then the gain in average earnings per worker should be higher for regions with top 50 research universities than those without. We thus compare the mean gain in average earnings per worker in regions with research universities against the same dependent variable for regions without top 50 research universities, controlling for size of MSA, region of the U.S., and industry structure, and apply a difference of means test for statistical significance. Second, in a multiple regression model, the coefficient on total university research expenditures in the region should be positive and significant, with other putative causal factors included as control variables.

The second hypothesis is that universities' technology development activity contributes significantly to regional economic development. Without a good direct measure of university technology development for all institutions of higher education within the MSA, we employ one of the strategies of quasi-experimental designs, i.e. manipulation of the time period. We split the full time period into two parts: 1969 to 1986 and 1986 to 1999. The rationale for this is that around the middle of the 1980s, many universities started to incorporate economic development missions, both in response to reductions in federal research support and the Bayh–Dole Act (Feller, 1990), as well as the initiatives of many state legislatures promoting universities as economic development actors in response to the 1981–82 recession and lagging competitiveness. Around the mid-1980s, universities began creating technology transfer offices to support patenting and licensing activities, building incubators and research parks and, occasionally, creating or investing in venture capital funds. Thus, if universities' involvement in technology development in various forms has been an important contributor to regional economic development, then the difference in the dependent variable (gain in average earnings per worker) between regions with research universities and those without should be larger in the 1986 to 1999 period compared to the 1969 to 1986 period.

The third hypothesis is that human capital creation (teaching) and the milieu functions are important contributors to regional economic development, controlling for the magnitude of research and technology development activity. Human capital creation and milieu, together, are measured by the number of degrees awarded annually by all institutions of higher education in the region. If this is true, then the coefficient in the multiple regression model will be positive and significant in both time periods.

The fourth hypothesis we investigate is whether agglomeration is more important than the presence of research universities as a contributor to regional economic development. If this is true, then the coefficient for size of MSA will be positive and significant, and stronger than coefficients on measures of university presence variables in the regression model, and more so in the 1986 to 1999 period than the earlier period.

The fifth hypothesis is a variant on the fourth: that there is interaction between presence of research universities and agglomeration. More specifically, our hunch is that they are substitutes. If this is true, then the effect of research universities will

be smaller the larger the MSA. This is tested by comparing the difference of means between MSAs with top 50 research universities against those without top 50 research universities, by size class. The differences should be largest for the smallest class of MSAs, and smallest for the largest class of MSAs. These differences should also be more prominent in the 1986 to 1999 period than the 1969 to 1986 period.

5 Empirical Results

As stated above, we conducted two different types of analysis of the data. The first is a set of difference-of-means tests between MSAs with and without top 50 research universities, overall, and classified by a number of control variables. These difference-of-means tests are done for each of the two separate time periods to be able to compare for temporal differences. The results are shown in Tables 1 and 2, respectively.[8]

There are no significant differences in change in average earnings per job between MSAs with top research universities and those without, for the 1969–1986 period. These results hold overall, and for each MSA size category, location, and type of industry structure. Two-way cross-tabs combining MSA size categories, location, and industry structure also showed that the presence of a research university was not associated with change in average earnings per job, except for the one case: large MSAs located in the South. Here, however, the small number of such MSAs without research universities outperformed those with top research universities.

The results for the 1986–1998 period tell a quite different story. In the entrepreneurial university period, the differences between MSAs with, and without, top research universities are significant and positive for the overall study population. When we disaggregate by type of MSA, we get the same results for small MSAs, those in the Northeast, South, and West, and with both high and low percentages of manufacturing. Within small MSAs, the positive presence of research universities is most prominent in the West.

The second type of analysis is based upon OLS multiple regression. Separate models are estimated for each of the two time segments. These results are shown in Tables 3 and 4.

During the 1969–1986 period, neither total university research expenditures nor the total number of degrees awarded are statistically significant in explaining variation in the MSAs' change in average earnings. The factors that are significant are: location in the Midwest (negatively related), location in the South (positively related), MSA size (positive), average earnings per job at the start of the period (negative), proportion of total earnings from the manufacturing sector (negative), proportion of self-employed earnings (negative), number of private R&D labs (positive), and the change in percent earnings from manufacturing over the time period (positive). The results suggest that general regional macroeconomic conditions, agglomeration economies, and aspects of industry structure, rather than the presence of universities, were the most important factors determining gain in regional economic well-being during this

[8] In a number of cases the n of one of the groups was small, a condition which often leads to unequal group variances. In these cases we used the approximate t statistic (SAS Institute, 1991).

Table 1: Change in Average Earnings/Job, 1969–86

Type of Area	MSAs with RU		MSAs without RU		Difference
	Mean	N	Mean	N	
All	−1.2	37	−0.67	252	−0.53
Small	−2.72	11	−1.21	149	−1.52
Medium	−1.8	13	−0.13	91	−1.67
Large	0.68	13	1.92	12	−1.24
Northeast	−0.72	6	−2.21	33	1.49
South	4.37	9	4.18	109	0.18
Midwest	−6.02	11	−6.21	62	0.19
West	−1.2	11	3.48	48	2.28
Low Mfg	0.05	17	0.83	105	−0.78
High Mfg	−2.26	20	−1.74	147	−0.52
Small, Low Mfg	−2.45	7	−0.72	68	−1.73
Medium, Low Mfg	−0.04	6	3.38	35	−3.45
Large, Low Mfg	4.55	4	8.77	2	−4.22
Small, High Mfg	−3.2	4	−1.62	81	−1.58
Medium, High Mfg	−3.32	7	−2.33	56	−0.99
Large, High Mfg	−1.03	9	0.55	10	−1.59
Small, Northeast	−7.9	1	−5.24	14	2.66
Small, South	3.88	3	3.95	66	−0.07
Small, Midwest	−5.26	4	−5.8	37	0.54
Small, West	−4.23	3	−4.78	32	0.55
Medium, Northeast		0			
Medium, South	4.95	4	3.99	40	0.96
Medium, Midwest	−6.26	5	−7.18	21	0.92
Medium, West	−2.99	2	−5.05	4	−1.94
Large, Northeast	0.72	3	1.33	5	−0.61
Large, South	3.92	2	11.92	3	−8.01*
Large, Midwest	−6.99	2	−5.05	4	−1.94
Large, West	2.86	4	1.75	2	1.1

period. Where knowledge producing activity mattered, it was outside universities.

The results are sharply different for the 1986–1998 period. Total university R&D activity is significant and positive, while total number of degrees awarded is significant and negative. Location still matters, but now the advantage is only for MSAs in the Northeast. MSA size is still significant and positive, and average earnings per job at the start of the time segment is still significant and negative. The effect of industrial structure attributes has changed. The relative size of the business services sector – a measure of the development of the business infrastructure – is now significant and positive, but the relative size of the manufacturing sector and the self-employed sectors are not significant factors. Neither is the number of private R&D labs significant in this later period. It should also be noted that the overall explanatory power of the model in the second time segment is lower than in the earlier one as indicated by the R-squared.

Table 2: Change in Average Earnings/Job, 1986–98

Type of Area	MSAs with RU		MSAs Without RU		Difference
	Mean	N	Mean	N	
All	3.37	36	−4.61	242	7.98**
Small	0.91	9	−5.63	185	6.54**
Medium	0.02	13	−2.77	46	2.8
Large	8.06	14	4.93	13	3.13
Northeast	7.38	8	−1.1	32	8.47*
South	1.88	10	−4.48	105	6.36**
Midwest	−3.08	9	−5.81	61	2.73
West	7.91	9	−5.82	44	13.72**
Low Mfg	−0.79	17	−4.49	116	5.28**
High Mfg	5.68	19	−4.72	126	10.4**
Small, Low Mfg	−3.03	6	−5.33	89	2.3
Medium, Low Mfg	0.89	5	−2.58	22	3.47
Large, Low Mfg	4.51	6	1.99	5	2.52
Small, High Mfg	8.79	3	−5.92	96	14.7**
Medium, High Mfg	−0.52	8	−2.96	24	2.44
Large, High Mfg	10.72	8	7.37	6	3.34
Small, Northeast	−3.37	1	−4.21	22	0.84
Small, South	−2.4	3	−5.71	81	3.31
Small, Midwest	−2.55	3	−5.7	47	3.16
Small, West	13.2	2	−6.26	35	19.46**
Medium, Northeast	5.81	3	2.83	7	3.98
Medium, South	2.29	4	−0.61	21	2.9
Medium, Midwest	−4.9	4	−7.82	11	2.92
Medium, West	−3.34	2	−6.94	7	3.59
Large, Northeast	11.23	4	12.26	3	−1.39
Large, South	5.62	3	1.67	3	3.96
Large, Midwest	−0.23	2	−0.08	3	0.15
Large, West	10.29	5	5.78	2	4.51

6 Conclusions and Suggestions for Further Study

Our first hypothesis, that research universities contribute significantly to regional economic development, controlling for other factors, was not supported throughout the full period of study. However, our second hypothesis, that it is the universities' economic development activities that matter most, is confirmed by the data. In the earlier 1969–1986 period, the presence or absence of a top 50 research university did not affect the gain in average earnings per worker, while it is a significant factor in the 1986–1998 period. These activities were generally absent in the pre-1986 time period and quite prevalent in the latter period.

Our third hypothesis is that the human capital creation and milieu functions of the university are important contributors to regional economic development. Since these

Table 3: Multivariate Regression Model Results: 1969–1986

Dependent Variable: Change in Average Earnings 1969–1986

Total DF	311	Root MSE	5.40
R-squared	0.554	Adj R-Sq	0.5361
F Value	30.96	Pr > F	<0.0001

Variable	Parameter Estimate	Standard Error	T value	Pr > t	Tolerance
Intercept	34.36	3.61	9.52	<0.0001	
UNRD72	−0.04	0.04	−1.02	0.31	0.23
Degree72	−0.19	0.19	−1.03	0.30	0.15
West	−2.34	1.28	−1.83	0.07	0.36
Midwest	−4.02	1.09	−3.69	0.00	0.42
South	2.27	1.09	2.08	0.04	0.33
Employ68	0.00	0.00	3.88	0.00	0.29
AVGEARN69	−0.24	0.03	−7.20	<0.0001	0.43
MFG_EARN69	−0.07	0.03	−2.13	0.03	0.45
BUS_EARN69	−0.30	0.22	−1.37	0.17	0.77
Prop_Earn69	−0.85	0.10	−8.30	<0.0001	0.75
LABS70	0.04	0.01	2.86	0.00	0.53
CHMFG69_86	0.34	0.07	4.53	<0.0001	0.53

Table 4: Multivariate Regression Model Results: 1986–1998

Dependent Variable: Change in Average Earnings 1986–1998

Total DF	311	Root MSE	5.90
R-squared	0.371	Adj R-Sq	0.345
F Value	14.68	Pr > F	<0.0001

Variable	Parameter Estimate	Standard Error	T value	Pr>t	Tolerance
Intercept	14.02	4.30	3.26	0.00	
UNRD86	0.02	0.01	2.35	0.02	0.32
Degree86	−0.39	0.16	−2.40	0.02	0.15
West	−4.20	1.26	−3.33	0.00	0.45
Midwest	−4.80	1.15	−4.18	<0.0001	0.46
South	−4.22	1.12	−3.77	0.00	0.37
Employ86	0.00	0.00	4.14	<0.0001	0.21
AVGEARN86	−0.20	0.04	−5.28	<0.0001	0.53
MFG_EARN86	0.03	0.04	0.82	0.41	0.58
BUS_EARN86	1.38	0.21	6.40	<0.0001	0.67
Prop_Earn86	−0.19	0.15	−1.26	0.21	0.67
LABS85	0.01	0.01	1.10	0.27	0.20
CHMFG86_98	0.10	0.06	1.83	0.07	0.79

factors are present throughout the two periods, and the university was not found to be a significant factor in the earlier period, it appears that these functions are not as important as we had thought originally. That MSAs with research universities outperform MSAs without top research universities, in the age of the knowledge-based economy and the "entrepreneurial university," controlling for other factors, is expected. That the presence of universities did not matter one way or the other in the earlier 1969–1986 period, supports the view that the teaching and milieu functions are not as important as the research and economic development functions of universities, since the former functions did not appreciably change over the full period, while research activity and economic development increased dramatically from the early to the later period. Indeed, the negative relationship between the number of degrees awarded by all institutions of higher education in the MSA and the dependent variable for the 1986–1998 period adds further support to this view. The negative relationship can be interpreted as a tendency toward an oversupply, or saturation, of highly educated workers in the average regional labor market. Since we do not have any indicator of the milieu function independent of teaching activity, unfortunately we cannot separate these factors in the interpretation of the results.

Our fourth and fifth hypotheses, the questions of whether the agglomeration economies are more important than research universities, and whether there research universities may serve as a substitute for agglomeration economies, have been debated in the literature for some time. Our results provide mixed evidence. On the one hand, in the 1986–1998 period, the only MSA size category for which the presence of research universities made a significant difference was small MSAs, suggesting a substitute effect. On the other hand, the coefficient estimate for MSA size in the regression model for both time periods indicates that agglomeration matters, independent of the presence and size of the university sector. It may be that both of these seemingly contradictory results are true; research universities in small areas can provide a number of external benefits that urban agglomerations generally provide.

Finally, although the presence of research universities and their scale of research activity are statistically significant factors in explaining gains in average earnings per job among MSAs in the later period, the strength of the causal relationship is quite modest. Controlling for other factors, it would have taken an increase of $10 million in research expenditures among universities in an "average" MSA to increase the index of average earnings per job by 0.2. To give these numbers some perspective, the "average" MSA had $30.7 million in R&D expenditures in 1986. If the universities in this hypothetical MSA had been able to increase their R&D expenditures by $10 million more (about a 33 percent increase), the MSA would have increased its index from 100.0 to only 100.2.

Overall, our results provide additional support to the view that universities' research and technology development activities do indeed generate significant knowledge spillovers that are captured within the regional environment, and result in enhanced regional economic development. Yet the magnitude of the contribution of those activities is small compared to other factors.

It would be interesting to conduct a similar study of the contributions of institutions of higher education to regional economic development in European countries.

A number of EU countries recently have had changes in higher education policy that would encourage universities to become more entrepreneurial and to become more engaged with the private economy. A study conducted in several years would allow enough time to observe if there have been changes in the magnitude and type of economic impacts of universities and other institutions of higher education on their respective regions. Such a comparative study would allow us to test the hypothesis that there is convergence, albeit with a time lag, between the role and impacts of universities in parts of Europe and in the U.S.

References

Acs, Z.J., Audretsch, D.B. and Feldman, M.P. (1994) R&D spillovers and recipient firm size. The Review of Economics and Statistics 76: 336–340

Arrow, K.J. (1962) The economic implications of learning by doing. Review of Economic Studies 80: 155–173

Atkinson, R. (1989) Some states take the lead: Explaining the formation of effective and ineffective state science and technology policies. Unpublished Ph.D. Dissertation, University of North Carolina at Chapel Hill

Audretsch, D.B. and Feldman, M.P. (1996) R&D spillovers and the geography of innovation and production. The American Economic Review 86: 630–640

Brett, A.M., Gibson, D.V. and Smilor, R.W. (eds.) (1991) University spin-off companies: economic development, faculty entrepreneurs and technology transfer. Rowman & Littlefield Publishers, Savage, Md.

Cook, T.D. and Campbell, D.T. (1979) Quasi-experimentation: design & analysis issues for field settings. Houghton Mifflin Company, Boston, Mass.

Feller, I. (1990) Universities as engines of R&D-based economic growth: They think they can. Research Policy 19: 335–348

Felsenstein, D. (1994a) The university in local economic development: Benefit or burden? Unpublished paper, Department of Geography and Institute of Urban and Regional Studies, Hebrew University, Mt. Scopus, Jerusalem, Israel

Felsenstein, D. (1994b) Dealing with 'induced migration' in iniversity impact studies. Unpublished paper, Department of Geography and Institute of Urban and Regional Studies, Hebrew University, Mt. Scopus, Jerusalem, Israel

Fujita, M., Krugman, P.R., Venables, A. (1999) The spatial economy: cities, regions, and international trade. MIT Press, Cambridge, Mass.

Goldstein, H.A. and Luger, M.I. (1992) Impact Carolina: The University of North Carolina at Chapel Hill and the state's economy. Final report prepared for the University of North Carolina at Chapel Hill Bicentennial Observance

Goldstein, H.A. and Luger, M.I. (1993) The economic impact of North Carolina State University's College of Engineering on the state. Final report prepared for the College of Engineering, North Carolina State University

Goldstein, H.A., Maier, G. and Luger, M.I. (1995). The University as an instrument for economic and business development: U.S. and European Comparisons. In: Dill, D. and Sporn, B. (eds.), Emerging patterns of social demand and university reform: Through a glass darkly. Pergamon, Oxford, England

Griliches, Z. (1979) Issues in assessing the contribution of research and development to productivity growth. Bell Journal of Economics 10: 92–116

Hall, B.H., Jaffe, A. and Trajtenberg, M. (2000) Market value and patent citations: A first look. A working paper of the National Bureau of Economic Research #7741

Isserman, A. (1987) An automated quasi-experimental control group method and its application to a tourism economic development project. Paper presented at the North American Regional Science Association Meetings, Baltimore, MD, November 6, 1987

Isserman, A. and Merrifield, J. (1987) Quasi-experimental control group methods for regional analysis: An application to an energy boomtown and growth pole theory. Economic Geography 63: 3–19

Jaffe, A.B. (1989) Real Effects of Academic Research. American Economic Review 79: 957–970

Jaffe, A.B. and Trajtenberg, M. (1996) Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. Proceedings of the National Academy of Science, USA, 93: 12671–12677

Jaffe, A.B., Trajtenberg, M. and Henderson, R. (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108: 577–598

Keilbach, M. (2000) Spatial knowledge spillovers and the dynamics of agglomeration of regional growth. Physica-Verlag, Heidelberg/New York

Ko, S. (1993) The incidence of high technology start-ups and spin-offs in a technology-oriented branch plants complex: The case of the research triangle region, North Carolina. Unpublished Ph.D. dissertation, University of North Carolina at Chapel Hill

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

Luger, M.I. and Goldstein, H.A. (1991) Technology in the garden: Research parks and regional economic development. The University of North Carolina Press, Chapel Hill, N.C.

National Science Board (2000) Science & Engineering Indicators – 2000. National Science Foundation, Arlington, VA (NSB-00-1)

SAS Institute (1991) SAS/STAT User's Guide, Release 6.03 edition. SAS Institute Inc., Cary, NC

Saxenian, A. (1994) Regional advantage: culture and competition in Silicon Valley and Route 128. Harvard University Press, Cambridge, Mass.

Shadish, W.R., Cook, T.D., Campbell, D.T. (2002) Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin, Boston, Mass.

Smilor, R.W., Gibson, D.V. and Dietrich, G.B. (1990) University spin-out companies: Technology start-ups from UT-Austin. Journal of Business Venturing 5: 63–76

Solow, R.M. (1957) Technical change and the aggregate production function. Review of Economics and Statistics 39: 312–320

Steffensen, M., Rogers, E.M. and Speakman, K. (2000) Spin-offs from research centers at a research university. Journal of Business Venturing 15: 93–111

Thanki, R. (1999) How do we know the value of higher education to regional development? Regional Policy 33: 84–89

Thursby, J.G. and Thursby, M.C. (2000) Who is selling the ivory tower? Sources of growth in university licensing. Working paper of National Bureau of Economic Research # 7718

Varga, A. (2000) Local academic knowledge transfers and the concentration of economic activity. Journal of Regional Science 40: 289–309

Appendix 1: Average Earnings by Job by Type of Region

Area	Change 69–86	Change 86–98	Change 69–98
I, Raleigh-Durham, NC	12.7	4.9	17.6
II, Richmond, VA	7.5	1.1	8.6
III, Southeast United States	6.1	0.1	6.2
IV, Small-Medium MSAs with Research Universities			
Ann Arbor, MI	−9.8	−4.6	−14.4
Boulder, CO	−6.1	19.6	13.5
Bryan-College Station, TX	0.1	−6.6	−6.5
Champaign-Urbana, IL	−6.8	−5.4	−12.2
Gainesville, FL	5.2	0.2	5.4
Madison, WI	−12.3	1.8	−10.5
Raleigh-Durham, NC	12.7	4.9	17.6
State College, PA	−7.9	−3.4	−11.3
Tallahassee, FL	3.2	4.8	8.0
Tucson, AZ	−6.3	−5.0	−11.3
Average	−2.8	1.1	−1.7
V, Large MSAs with Research Universities			
Atlanta, GA	6.7	9.7	16.4
Austin, TX	10.9	13.3	24.2
Baltimore, MD	1.2	0.5	1.7
Boston, MA	5.7	12.9	18.6
Columbus, OH	−7.5	−0.7	−8.2
Minneapolis-St. Paul, MN	−3.0	2.3	−0.7
Pittsburgh, PA	−5.0	−0.6	−5.6
San Diego, CA	3.6	−2.8	0.8
Seattle, WA	−6.2	15.2	9.0
St. Louis, MO	−1.9	−2.7	−4.6
Average	0.6	4.7	5.3
VI, Small-Medium MSAs without Research Universities			
Albuquerque, NM	−1.9	−5.1	−7.0
Burlington, VT	−3.4	0.9	−2.5
Charlottesville, VA	4.5	4.8	9.3
Columbia, MO	−1.1	−0.4	−1.5
Columbia, SC	5.9	−2.1	3.8
Eugene, OR	−13.3	−2.3	−15.6
Knoxville, TN	4.8	−3.4	1.4
Lexington, KY	4.3	1.1	5.4
Lincoln, NE	−4.3	−0.9	−5.2
Richmond, VA	7.5	1.1	8.6
Tuscaloosa, AL	15.7	−11.9	3.8
Average	1.7	−1.7	0.0
VII, Large MSAs without Research Universities			
Charlotte, NC	8.8	8.1	16.9
Dallas, TX	11.9	9.2	21.1
Denver, CO	5.5	6.8	12.3
Detroit, MI	−5.8	−2.1	−7.9
Indianapolis, IN	−5.4	1.1	−4.3
Kansas City, MO	−0.3	−1.8	−2.1
Milwaukee, WI	−8.8	0.5	−8.3
Portland, OR	−7.1	4.8	−2.3
Sacramento, CA	−4.5	−1.7	−6.2
Tampa-St. Petersburg, FL	0.2	3.8	4.0
Average	−0.5	2.9	2.4

Regional Innovation in the US over Space and Time [1]

Attila Varga*, Luc Anselin‡, Zoltan J. Acs§

1 Introduction

Knowledge plays a central role in economic development as recently emphasized especially in endogenous growth models (e.g. Romer, 1986, 1990, Aghion and Howitt, 1999). Therefore, explaining the process of knowledge production is crucial to understand modern economic growth. Innovation activities have a predominant tendency to cluster spatially as demonstrated by recent empirical studies (e.g. for the US in Varga, 1998 and for the European Union in Caniels, 2000). Sensitivity of the transmission of tacit knowledge to distance provides a principal reason for the development of regional innovation clusters since the transfer of non-codified knowledge elements frequently requires close personal interactions (Polanyi, 1966, Dosi, 1988). Thus, relative spatial position of the actors in knowledge creation is a potentially significant factor of innovation. Endogenous growth theories provide models to study the role of knowledge in macroeconomic growth but leave out the regional dimension despite the substantial evidence provided in the recent empirical economics literature that a significant fraction of knowledge spillovers tends to be localized (Acs and Varga, 2002).

Four approaches have been developed in the recent empirical economics literature to estimate the role of localized knowledge flows in the process of innovation: surveys of industrial researchers (Mansfield, 1995), the study of the spatial patterns of patent citations (Jaffe, Trajtenberg and Henderson, 1993), regional innovation surveys (Cooke, 2000, Koschatzky and Sternberg, 2000) and econometric analyses within the knowledge production function framework. This framework has been widely applied in empirical studies of regional innovation in the US (e.g. Jaffe, 1989, Acs, Audretsch and Feldman, 1991, Acs, Anselin and Varga, 2002, Anselin, Varga and Acs, 1997, and Varga, 2000), in Italy (Audretsch and Vivarelly, 1994, Capello, 2001), in France (Autant-Bernard, 1999), in Germany (Fritsch, 2002 and in Austria (Fischer and Varga, 2003).

Building on a recently developed cross-sectional time-series data set of US inno-

[1] Research assistance in data collection by Oleg Smirnov (University of Illinois at Urbana Champaign), Dapeng Chen, Nicolay Nedev, Baishali Majumdar and Vladimir Starkov (West Virginia University) is highly appreciated.

*Center for Research in Economic Policy (CREP) and Department of Economics, University of Pécs, Pécs, Hungary

‡Regional Economics Applications Laboratory (REAL) and Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA.

§Merrick School of Business, University of Baltimore and U.S. Bureau of the Census, USA.

vation, private and university research and high technology employment, we continue our previous work in this paper. We provide a first-cut analysis of the data to shed some new light on the spatial and temporal aspects of US innovation. The novelty of this data set is that it opens up the possibilities to incorporate the time dimension into knowledge production function analysis at an appropriate level of spatial aggregation (i.e. US metropolitan areas) that has not been possible in empirical research yet. The following section introduces the methodology and the applied data, while the third and fourth sections highlight some important space-time aspects of US innovation. A summary concludes the paper.

2 Methodology and Data

The knowledge production function (KPF) framework was initiated by the work of Griliches (Griliches, 1979, 1986) and was first implemented in the spatial context in Jaffe (1989). Since then it has become a major methodological approach to understand the geography of innovation. A critique against knowledge production function studies (i.e. that the model does not allow for an explicit modeling of the way knowledge spillovers occur and as such it is difficult to separate spillovers from the correlation of variables at the geographical level as expressed e.g. in Feldman, 2000) is certainly valid to some extent. However, an important advantage of the knowledge production function analysis is that it can provide an account of innovation-related interactions on the basis of large number of geographical areas with the fraction of the costs of a similarly designed survey-based research given that KPF studies rely on secondary data sources. On the other hand, since the applied data do not refer to actual interactions, much care should be taken on econometric specification.

Formally, the knowledge production function is expressed as:

$$\log(K) = \alpha + \beta \log(R) + \gamma \log(U) + \delta \log(Z) + \epsilon \tag{1}$$

where K is a proxy for knowledge (either patents or innovation counts), R is industry R&D and U is university research, with ϵ as a stochastic error term. Z typically includes a measure of the concentration of a given activity (a proxy for innovation networks of manufacturing firms). The analysis is usually carried out for aggregate cross-sectional units (e.g. states, MSAs), possibly for several points in time and/or disaggregated by sector. Positive and significant coefficients for β, γ and δ indicate positive effects of different regional knowledge sources on industrial innovation.

We aggregated the data to the "high technology" sector, that is a set of industries where the intensity of knowledge inputs to production exceeds the industrial average. Table 1 provides more information on the set of specific industries included. Our panel data set comprises variables observed for three years (1985, 1988 and 1991) and aggregated to the level of US metropolitan statistical areas (MSAs). K is measured by patent applications (US Patent Office, 1998), R is professional employment in high technology industrial laboratories compiled from three editions of the Directory of American Research and Technology (1986, 1989, 1992), U is university research expenditures obtained from CASPAR data files (National Science Foundation, 1997) and

Z is high technology employment (Bureau of the Census, 1999).

As in our previous studies we apply the methodology of spatial econometrics in studying the geography of innovation. Spatial econometrics (Anselin, 1988, 2001, Anselin and Florax, 1995) turns out to be a very powerful analytical tool in empirically modeling localized knowledge spillovers when cross sectional data are applied. Spatial econometrics supplies both the appropriate statistics to test for potential misspecifications as well as different modeling approaches of spatial dependence with a high intuitive value in actually measuring inter-regional knowledge spillovers. SpaceStat, the software for spatial data analysis developed by Luc Anselin is used for spatial regressions throughout this paper.

Table 1: High technology industries

SIC (1972)	PTO
Drugs	
283 Drugs and medicines	14
Chemicals	
281 Industrial inorganic chemistry	6
282 Plastic materials and synthetic resins	8
286 Industrial organic chemistry	7
289 Miscellaneous chemical products	13
Information Technology	
357 Office computing and accounting machines	27
361, 3825 Electrical transmission and distribution equipment	35
365 Radio and television receiving equipment except communication types	42
366, 367 Electronic components and accessories and communications equipment	43
High Technology Machinery and Equipment	
351 Engines and turbines	23
353 Construction and related machinery	25
356 General industrial machinery and equipment	30
362 Electrical industrial apparatus	36
363 Household appliances	38
364 Electrical lighting and wiring equipment	39
369 Miscellaneous electrical machinery, equipment and supplies	40
Defense and Aerospace	
372 Aircraft and parts	54
376 Guided missiles and space vehicles and parts	47
Professional and Scientific Instruments	
38 Professional and scientific instruments	55

Notes: The list of industries is based on Acs (1996). Concordance between SIC codes and PTO sequence numbers is provided by the US Patent and Trademark Office

Table 2: Comparative statics. OLS knowledge production estimates with contemporaneous and lagged dependent variables

Variable	PATHT85 X85	PATHT88 X88	PATHT88 X85	PATHT91 X91	PATHT91 X88
Constant	-4.826	-3.676	-3.822	-4.284	-3.719
	(0.488)	(0.440)	(0.452	(0.482)	(0.475)
Log(RD)	0.166	0.224	0.218	0.163	0.189
	(0.043)	(0.039)	(0.040)	(0.039)	(0.041)
Log(URD)	0.086	0.067	0.071	0.093	0.090
	(0.026)	(0.024)	(0.024)	(0.027)	(0.026)
Log (EMPHT)	0.697	0.599	0.615	0.679	0.618
	(0.066)	(0.059)	(0.062)	(0.064)	(0.063)
CON50	0.244	0.260	0.236	0.328	0.268
	(0.127)	(0.121)	(0.118)	(0.128)	(0.130)
SOUTH and	0.254	0.002	-0.002	0.149	0.010
WEST	(0.125)	(0.118)	(0.116)	(0.127)	(0.127)
R^2-adj	0.80	0.81	0.82	0.79	0.79
Number of obs.	143	143	143	143	143

Notes: All dependent variables are in logarithm. Estimated standard errors are in parentheses; X denotes the dependent variables; RD is professional employment at industrial research and development laboratories; UR is university research expenditures; EMPHT is high technology employment; CON50 is a dummy variable: it takes 1 if at least one MSA is located within a 50 mile distance band and 0 otherwise; SOUTH and WEST is a dummy variable: it takes 1 if the MSA is situated in the South or West and 0 otherwise.

3 Space–Time Patterns of U.S. Innovation – Some Methodological Issues

Two important methodological issues are considered in this section. First, an examination of the extent to which parameters of lagged independent variables in the knowledge production function are stable over time with different time lags applied and, second, an exploration with respect to the stability of estimated parameters over spatial units.

The issue of the stability of estimated parameters for different time lags applied between the dependent variable and the explanatory variables is important in evaluating regression results when single cross sections are used and data constraints do not allow to apply time lags between innovation inputs and outputs (as, for example, in Anselin, Varga and Acs, 1997). In principle, time lags of 2–3 years are recommended (see Edwards and Gordon, 1984) when patent data are used in order to account for the time difference between the actual development of an invention and the approval of its patent.

In Table 2 the knowledge production function of equation (1) is extended with two additional dummy variables. CON50 accounts for potential effects of agglomeration on the intensity of localized knowledge spillovers (in case of a single metropolitan

Table 3: Pooled OLS estimates of the knowledge production function with regional dummies Variable Log(PATHT)

Variable	PATHT	PATHT	PATHT	PATHT
Constant	-4.079	-4.020	-4.069	-4.060
	(0.284)	(0.267)	(0.263)	(0.266)
Log(RD)	0.197	0.195	0.184	0.198
	(0.023)	(0.023)	(0.023)	(0.023)
Log(URD)	0.084	0.084	0.089	0.082
	(0.015)	(0.015)	(0.015)	(0.015)
Log (EMPHT)	0.635	0.628	0.646	0.632
	(0.038)	(0.037)	(0.037)	(0.037)
CON50	0.313	0.267	0.290	0.334
	(0.072)	(0.076)	(0.071)	(0.073)
Mid-West	0.019			
	(0.078)			
North East		0.150		
		(0.082)		
South			-0.264	
			(0.078)	
West				0.139
				(0.091)
R^2-adj	0.79	0.79	0.80	0.79
Number of obs.	429	429	429	429

Notes: All dependent variables are in logarithm. Estimated standard errors are in parentheses; for variable definition see notes to tables 2; PATHT is patent application counts in high technology; Mid-West, North East, South and West are dummy variables taking 1 if the MSA is situated in a given region and 0 otherwise.

area this variable takes the value of 0 and it is 1 if the MSA is part of a larger cluster of cities). The SOUTH and WEST dummy is included to test for potential differences between patterns of localized knowledge production in the US industrial heartland (the North East and the Mid-West regions) and the recently emerging "new economy" in the South and the West[2] of the country (Suarez-Villa, 2000). The connectivity dummy stays consistently significant, whereas the regional dummy remains insignificant.

A three-year time lag is applied between the date of patent approval and invention in the third and fifth columns. A comparison of the results with a time lag applied (third and fifth columns) to those without time lags (second and fourth columns) shows no significant differences between sizes, signs and significances of parameter estimates as well as regression fits. It is also shown in the table that the relative importance

[2]The North-East consists of Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, Pennsylvania, New Jersey, Delaware, Maryland, Washington DC, Virginia and West Virginia. The Midwest states are Minnesota, Michigan, Wisconsin, Iowa, Missouri, Illinois, Indiana, Ohio, Kentucky, North Dakota, South Dakota, Nebraska and Kansas. The South consists of Oklahoma, Texas, Arkansas, Louisiana, Mississippi, Tennessee, Alabama, Georgia, Florida, North Carolina and South Carolina. States in the West are Washington, Montana, Arizona, New Mexico, Wyoming, Idaho, Oregon, California, Nevada, Utah and Colorado.

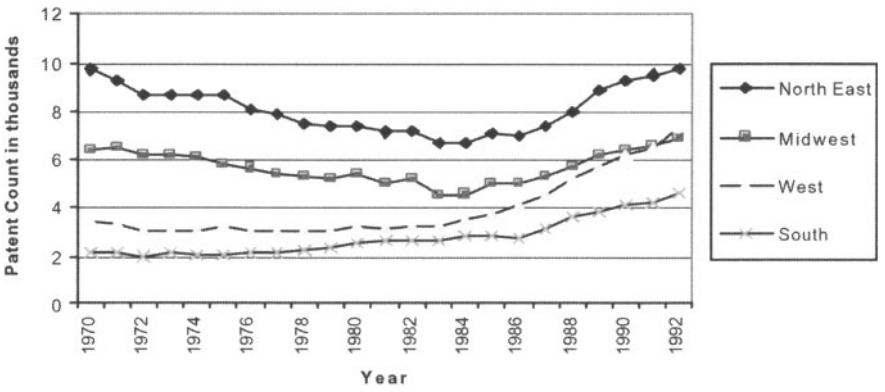

Figure 1: Regional trends in knowledge production in high technology (Source: Varga, 1999)

of different local sources of innovation remains the same no matter whether lagged or contemporaneous explanatory variables are used (i.e. interfirm knowledge flows dominate over research spillovers among local R&D laboratories and both are more important than knowledge transfers from regional universities).

The second research question relates to parameter stability over space. Compared to the South and West dummy a finer distinction among US regions is applied in Table 3 with the four regional dummies. In order to increase the level of information extracted from the data we run pooled time series cross-sectional regressions with 429 observations. Parameter values for local knowledge inputs as well as the connectivity dummy do not differ meaningfully, however, there are important differences as to the effect of regional dummies. Whereas no significant differences are reported for Mid-West, North East and the West, the significant (and negative) dummy for the US South suggests that local innovation systems in the newly emerging Southern high technology centers might differ in structure from the rest of the country. The following section focuses on this problem in more details.

4 Changing Geography of U.S. Innovation: Is There Any Role of Localized Knowledge Spillovers?

Perhaps one of the most fascinating issues in economic development is the recent emergence of high technology centers in the traditionally non-manufacturing sectors dominated US West and South, most notably in California, Texas, Arizona, Utah and Florida. Understanding the extent to which the impressive growth of these US regions is a result of consciously designed regional economic development policies (that can be learned and might be replicated in other parts of the World) may have relevance for currently lagging regions not only in the US, but in Europe as well. In Suarez-Villa

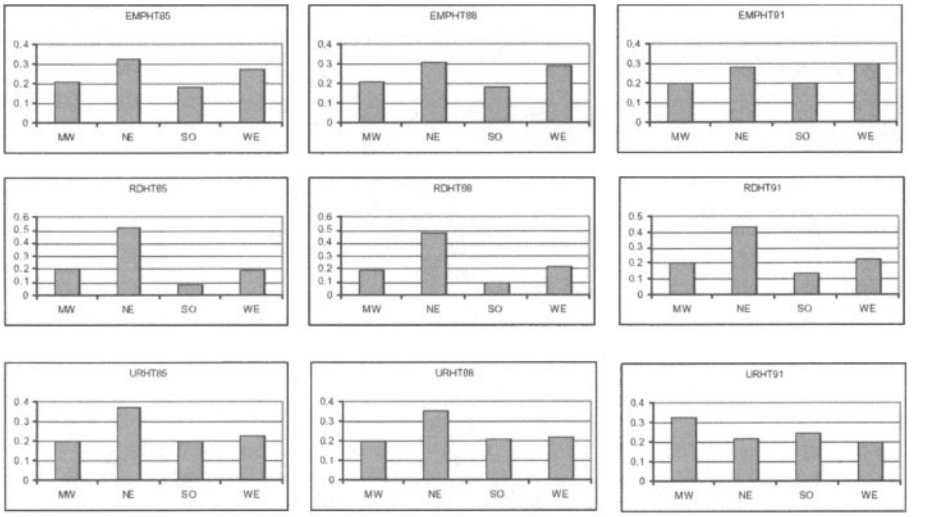

Figure 2: Geographical distribution of regional inputs to knowledge production between 1985 and 1991

(2000) the hypothesis that this growth is induced by previous investments in education and infrastructure is tested. In this section the focus is more on an exploration as to the potential differences in the relative importance of different regional factors of knowledge production.

Figure 1 shows regional trends in high technology knowledge production (measured by patent application counts) between 1970 and 1992. Whereas patenting activity followed a decreasing trend in the traditional manufacturing areas of the US (i.e. the North East and the Mid-West regions) until the early eighties, innovation activities of states in the South continuously increased, while in the West it stagnated during this period. However, after 1983 the differences among regional growth trends are dramatic and perhaps surprising. Although the North East maintained its traditional leading position in innovation during the whole time period, it seems that this position was increasingly challenged by the West, especially after 1989 when the rate of growth started to diminish in the North-East. Figure 1 shows that the North-East and the Midwest regions, which have been traditionally considered as leading manufacturing centers of the US, increasingly started losing their dominance in high technology innovation after 1983.

Differences in regional growth rates of patenting activity over the period of 1983–1992 also support this observation. While the North-East and the Midwest increased patenting by 45 and 53 percents, respectively, for the same time period growth rates of the West and the South were 128 and 79 percents. Moreover, while the North-East and the Midwest lost their share in total patents by 14 and 9 percents, the West and the South produced a substantial increase, 35 and 6 percents, respectively (Varga, 1999). This changing pattern might be induced by changes in the spatial distribution

Table 4: Maximum Likelihood Spatial SUR Regression Results for Log(Patents) at the level of US MSAs

Variable	National	North-East	Midwest	South	West
Constant	-4.783	-3.694	-3.534	-7.013	-4.955
	(0.449)	(1.066)	(0.822)	(0.949)	(0.526)
Log(RD85)	0.111	0.246	0.160	-0.001	0.123
	(0.035)	(0.064)	(0.071)	(0.060)	(0.054)
Log(URD85)	0.097	0.007	0.091	0.187	0.156
	(0.024)	(0.044)	(0.043)	(0.053)	(0.041)
Log(EMPHT85)	0.711	0.624	0.570	0.906	0.688
	(0.059)	(0.130)	(0.120)	(0.105)	(0.090)
CON50	0.288	0.208	0.620	0.456	-0.304
	(0.125)	(0.241)	(0.246)	(0.280)	(0.223)
SOUTH and WEST	0.279				
	(0.123)				
Constant	-3.484	-2.639	-3.806	-3.934	-3.720
	(0.398)	(0.777)	(0.818)	(0.838)	(0.621)
Log(RD88)	0.194	0.255	0.141	0.188	0.161
	(0.030)	(0.046)	(0.075)	(0.050)	(0.054)
Log(URD88)	0.073	-0.003	0.069	0.135	0.061
	(0.022)	(0.036)	(0.044)	(0.047)	(0.039)
Log(EMPHT88)	0.588	0.565	0.632	0.570	0.666
	(0.051)	(0.092)	(0.123)	(0.085)	(0.096)
CON50	0.293	0.043	0.773	0.222	-0.079
	(0.118)	(0.197)	(0.250)	(0.272)	(0.248)
SOUTH and WEST	0.013				
	(0.116)				
Constant	-4.059	-3.384	-4.011	-5.239	-3.315
	(0.441)	(0.870)	(0.827)	(1.038)	(0.770)
Log(RD91)	0.141	0.171	0.112	0.125	0.195
	(0.031)	(0.047)	(0.073)	(0.049)	(0.065)
Log(URD91)	0.098	0.027	0.102	0.196	0.170
	(0.024)	(0.433)	(0.045)	(0.057)	(0.071)
Log(EMPHT91)	0.662	0.689	0.664	0.668	0.512
	(0.056)	(0.106)	(0.122)	(0.104)	(0.118)
CON50	0.353	0.098	0.627	0.588	-0.122
	(0.125)	(0.206)	(0.258)	(0.287)	(0.306)
SOUTH and WEST	0.153				
	0.124)				
R^2-adj	0.63	0.58	0.62	0.58	0.87
Number of observations	429	117	126	111	75
	Tests on spatial dependence				
D50 LM (error)	2.512	2.184	0.481	6.531*	0.834
D50LM (lag)	2.802	1.575	1.208	2.951	2.370
	Wald tests on parameter stability				
Log(RD)	8.875**	7.110**	0.349	10.682***	1.004
Log(URD)	2.975	2.026	1.025	2.907	14.612***
Log(EMPHT)	7.852**	4.037	0.577	15.220***	2.670
CON50	0.619	1.367	0.925	5.047*	2.100
SOUTH and WEST	9.246***				

Notes: Estimated standard errors are in parentheses; for variable definition see notes to tables 2 and 3; D50 is distance-based contiguity matrix for 50 miles; * denotes significance at least at 0.10; ** denotes significance at least at 0.05; *** denotes significance at least at 0.01.

of regional sources of innovation. However, a closer inspection of Figure 2 does not support this hypothesis. With the exception of the difference in the spatial patterns of university research between the last two time periods, no meaningful changes can be observed.

An alternative explanation is that there might be meaningful differences as to the "efficiency" with which the different local innovation systems combine their local knowledge resources (e.g. differences in local cultures with respect to the propensity of the actors to interact with each other as exemplified in Saxenian, 1994 for Silicon Valley and Route 128, or differences in the effectiveness in regional economic development policies). Comparison of sizes, signs and significances of parameter estimates over space and time might suggest some clues in this respect.

Table 4 lists spatial Maximum Likelihood Seemingly Unrelated Regression (SUR)

results for the four large US regions and the nation for 1985, 1988 and 1991. This regression technique opens the possibility to compare estimated parameters over space as well as to test the stability of the coefficients. Perhaps the most striking difference relates to the university research parameter. This parameter is consistently non-significant in the North-East, which is perhaps a surprising result. This finding certainly needs a closer examination in the future, however heavy restructuring of the local economies of some North-Eastern metropolitan areas (such as Boston and New York as shown in Acs, 1996) characterized by major losses in high technology jobs during this time period could be behind this observation. On the other hand, parameter estimates of university research in the South are consistently higher than anywhere in the rest of the regions, which might suggest a more intensive local role of universities in economic development in the South than anywhere else in the country. This observation would certainly need further investigations, however it is definitely an interesting finding.

Regarding the rest of the parameters of local innovation inputs no comparable differences can be found across large regions. A further interesting result is the non-significant connectivity dummy for all the regions but the Mid-West. For this region CON50 stays consistently significant, indicating differences in local innovation systems between large agglomerations and smaller metropolitan areas. With the exception of the university research parameter, all the rest of the parameters of local innovation inputs are unstable in the South (as shown by the significant values of the Wald tests in Table 4). This might be taken as an additional support to the important role of local innovation inputs in the restructuring of metropolitan areas in the US South.

5 Summary

Local dimensions of knowledge production are gaining increasing attention in both theoretical and empirical research in economics. However, our understanding is still constrained by the availability of appropriate data on knowledge production-related activities. In this paper we presented results of a first-cut analysis based on a recently developed space-time data set of US innovation activities. The most important findings can be summarized as follows.

- No significant differences were observed between the regression results with lagged and contemporaneous explanatory variables, suggesting that within a relatively short period of time (e.g. in about three years) no meaningful changes occur in the performances of local innovation systems. This result has an important technical consequence: at least at the level of spatial aggregates the use of contemporaneous dependent and independent variables is acceptable in knowledge production function studies.
- Differences in the trends of knowledge production across large US regions do not seem to be the result of a changing spatial distribution of local innovation inputs.
- Differences are found regarding the importance of universities as local sources of new technological knowledge. Perhaps the most surprising result is the con-

sistently insignificant university effect in the North East.

- Compared to the rest of the country, the recently emerging US South seems to follow different patterns in combining local innovation inputs especially with respect to the role of local universities in supporting production of new technological knowledge. However, instability of most of the parameters indicates that the metropolitan areas in the region are in a reconstruction process of their innovation systems.

References

Acs, Z. (1996) American high technology centers. In: De La Mothe J. and Paquet G. (eds.), Evolutionary economics and the new international political economy. Pinter, London: 183–219

Acs, Z. and Varga, A. (2002) Geography, endogenous growth and innovation. International Regional Science Review 25: 132–148

Acs, Z., Anselin, L. and Varga, A. (2002) Patents and innovation counts as measures of regional production of new knowledge. Research Policy 31: 1069–1085

Acs, Z., Audretsch, D. and Feldman, M. (1991) Real effects of academic research: comment. American Economic Review 81: 363–367

Aghion, P. and Howitt, P. (1998) Endogenous growth theory. MIT Press, Cambridge, Mass.

Anselin, L. (1988) Spatial econometrics: methods and models. Kluwer Academic Publishers, Boston

Anselin, L. (2001) Spatial econometrics. In: B. Baltagi (ed.), A companion to theoretical econometrics. Basil Blackwell, Oxford: 310–330

Anselin, L. and Florax, R. (1995) (eds.) New directions in spatial econometrics. Springer, Berlin

Anselin, L., Varga, A. and Acs Z. (1997) Local georaphic spillovers between university research and high technology innovations. Journal of Urban Economics 42: 422–448

Audretsch, D. and Vivarelli, M. (1994) Small firms and R&D spillovers: evidence from Italy. Discussion Paper 953, Centre for Economic Policy Research

Autant-Bernard, C. (2001) Science and knowledge flows: evidence from the French case. Research Policy 30: 1069–1078

Bureau of the Census (1999) County business patterns. Data obtained from ICPSR online data services

Caniels, M. (2000) Knowledge spillovers and economic growth. Edward Elgar, Cheltenham, Glos.

Capello, R. (2001) Spatial and sectoral characteristics of relational capital in innovation activity. Paper presented at the 41th Congress of the European Regional Science Association meetings, Zagreb August 29–September 1, 2001.

Cooke, P. (2000) Business processes in regional innovation systems in the European Union. In: Acs, Z. (ed.), Regional innovation, knowledge and global change. Pinter, London: 53–71

Directory of American Research and Technology for 1986, 1989, 1992. R.R. Bowker New York.

Dosi, G. (1988) Sources, procedures and microeconomic effects of innovation. Journal of Economic Literature 26: 1120–1126

Edwards, K. and Gordon, T. (1984) Final report: Characterization of innovations introduced on the U.S. market in 1982. Prepared for the U.S. Small Business Administration. The Futures Group

Feldman, M. (2000) Location and innovation: the new economic geography of innovation, spillovers, and agglomeration. In: Clark G., Feldman M. and Gertler M. (eds.), The Oxford handbook of economic geography. Oxford University Press, Oxford: 373–394

Fischer, M. and Varga, A. (2003) Spatial knowledge spillovers and university research: evidence from Austria. Annals of Regional Science 37: 303–322

Fritsch, M. (2002) Measuring the quality of regional innovation systems - A knowledge production function approach. International Regional Science Review 25: 86–101

Griliches, Z. (1979) Issues in assessing the contribution of research and development to productivity growth. Bell Journal of Economics 10: 92–116

Griliches, Z. (1986) Productivity, R&D, and basic research at the firm level in the 1970s. American Economic Review 76: 141–154

Jaffe, A. (1989) Real effects of academic research. American Economic Review 79: 957–970

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108: 577–598

Koschatzky, K. and Sternberg, R. (2000) R&D cooperation in innovation systems some lessons from the European Regional Innovation Survey (ERIS). European Planning Studies 8: 487–501

Mansfield, E. (1995) Academic research underlying industrial innovations: sources, characteristics and financing. The Review of Economics and Statistics 77: 55–65

National Science Foundation (1997) Academic science and engineering: R&D expenditures. Data obtained from CASPAR data files

Polanyi, M. (1966) The tacit dimension. Doubleday Anchor, New York

Romer, P. (1986) Increasing returns and long-run growth. Journal of Political Economy 94: 1002–1037

Romer, P. (1990) Endogenous technological change. Journal of Political Economy 98: S71–S102

Saxenian, A. (1994) Regional advantage: culture and competition in Silicon Valley and Route 128. Harvard University Press, Cambridge, MA

Suarez-Villa, L. (2000) Invention and the rise of technocapitalism. Rowman & Littlefield Publishers, Inc., New York

US Patent Office (1998) Inventor and Patsic data files for years 1962–1996

Varga, A. (1998) University research and regional innovation: a spatial econometric analysis of academic technology transfers. Kluwer Academic Publishers, Boston

Varga, A. (1999) Time-space patterns of US innovation: stability or change? A detailed analysis based on patent data. In: Fischer M., Suarez-Villa L. and Steiner M. (eds.), Innovation, networks and localities. Springer, Berlin: 215–234

Varga, A. (2000) Local academic knowledge spillovers and the concentration of economic activity. Journal of Regional Science 40: 289–309

Weighing Environmental Factors in the Appraisal of Major Highway Investments

Peter M. Townroe*

1 Introduction

The environmental impact of building even a short length of a major new highway is always significant. Or at least, so it seems to an increasingly voluble slice of public opinion, whether represented by a national pressure group or by a newly formed local coalition of NIMBYs[1]. For some opponents of a scheme complete abandonment of the proposal is the only option. For others, re-routing of the road or some form of monetary compensation or environmentally friendly offset investment will provide sufficient satisfaction. For all, there is a clear demand that their objections be heard, in public and in a democratic forum. Road builders, transport investment analysts, and their political decision-making masters, must be able to present the strongest possible case of benefits from the investment to offset against the claimed environmental damage.

All governments in what used to be called the 'Western' nations, and many among the less economically advanced nations, have wrestled with this issue: of making a case for road investments. A case is needed for public acceptance. And within available road building budgets a procedure has been needed to rank alternative potential schemes in their relative desirability in terms of overall prospective public benefit. Using a fairly standardised Cost-Effectiveness Analysis, promoted internationally by such bodies as the OECD and the World Bank, the monetary calculus of direct transport costs and benefits has provided a simple and apparently unambiguous framework for this purpose. However, once a political agreement has been reached that the environmental impacts of alternative schemes may significantly bias the rankings, or even push the available funds towards other forms of transport investment, then transport economists turn to their environmental economist colleagues for assistance.

This paper uses a discussion of recent developments in the approach taken by the British government towards these issues to highlight the analytical dilemmas that sit in behind the political furore that now so often seems to accompany the possible externality effects of major highway investments. In 1998 the new-at-that-time Blair government launched what it called *The New Approach to Appraisal* under the general brave heading of *A New Deal for Trunk Roads in England* (although it was also

* Visiting Professor, Sheffield Hallam University, Sheffield, Great Britain.

[1] "Not In My Back Yard".

meant to apply to the rest of the United Kingdom as well). In fact the *Approach* was not so new, and the *New Deal* has not been noticeable in its impact; but the supporting developments in analytical approaches and in the presentation of the appraisals of alternative schemes have served to further clarify what are the seemingly inevitable local complexities of each scheme. These complexities include the environmental impacts. The paper considers the options currently available in the weighing and evaluating of these impacts.

The next section of the paper counter-poses, in summary only, the case for further major highway investments in the United Kingdom with the arguments against, both in general and in relation to specific localities. Then, when once a scheme has been accepted in principle, it faces further design problems. Any builder of trunk roads in those regions of Europe with high levels of residential density faces a routing problem, normally around rather than through urban areas. In addition, most parts of most of those regions are characterised by valued environmental qualities, natural and man-made.

Section 3 of the paper outlines, again briefly, the evolution of the standard Cost-Benefit Analysis (COBA) as used by the British government to appraise road schemes for more than three decades, against increasing criticism that it was too simple and too uni-dimensional. This leads to Section 4 which outlines the *New Approach to Appraisal* and its application to a tranche of British motorway and trunk road improvement schemes.

A review of the outstanding issues in both evaluating environmental impacts and in Cost Benefit Analysis for road schemes is the basis of the discussion in Section 5. This goes through the familiar territories of approaches taken to gain monetary values, and approaches to reflect those impacts which do not have monetary values. The final section of the paper is then a brief discussion of the role of analysts and of decision makers in a pluralistic democratic environment in reaching major decisions on transport investments.

2 Requirements and Restraints

In a population of 57 million Great Britain has approximately 24 million cars and 5 million goods vehicles and buses. Although the number of cars per head of population is lower than in France, Italy or Germany, the annual mileage per vehicle is greater. The intensity of use on major trunk roads is very high. Travel by car accounts for 85% of total passenger mileage, and road freight takes 65% of the tonne kilometre total in all forms of land transport. In the particular case of Britain the intensity of road use has been exacerbated by the urban form of housing developments throughout the twentieth century. This has been dominated by the suburban detached or semi-detached house with a garden, with a consequential relatively low urban density and therefore difficult economics for public transport and a premium on the convenience of the household motorcar (or cars). The related forms of retailing and wholesaling have encouraged logistics systems dominated by road rather than rail, further encouragement coming from the relatively short inter-urban distances between the major concentrations of population. This contributes to the intensity of use of both local roads and the inter-

urban routes.

At the same time, by the standards of other major EU nations, the proportion of the UK public expenditure invested in all forms of transport has been low throughout the past fifty years. This may be seen in the miles of motorway standard highway per 1,000 sq km (under half that of Germany for example, and less than one quarter of The Netherlands), or in the present poor state of the rail infrastructure. The motorways and major trunk roads account for 4% of the road mileage but take 36% of the traffic. Arguably more effort has been put into road traffic management than into building new roads.

The result of these trends has been growing road congestion, as found elsewhere in the EU. This congestion is systematic on many key routes and is allowed for by many drivers in planning their journey times. It is also unsystematic, due to accidents, road works or one-off events held at particular locations, and therefore more difficult to allow for, leading to disruption of schedules and both drivers and passengers counting the value of their time spent unexpectedly sitting in a vehicle. With rising real incomes, this travel time has a rising premium. A special issue also arises in those settlements without a bypass that are reliant on a local road form that has a long history but which now carries an important through route. Congestion here also has been rising, with a consequential adverse environmental impact for local residents.

Pressure to build more roads in the United Kingdom is therefore a political reality, from industry and commerce, from frustrated longer distance commuters, from leisure time visitors to extended family and friends, and from the residents of the non-by-passed settlements. The economic development councils and the industrial and urban regeneration agencies are also part of this chourus[2]. But ranged against this loose coalition is a politically voluble phalanx of opposition.

If the United Kingdom "lost" Scotland north of the Clyde-Forth Valley and most of Central Wales, it would be one of the most densely populated nations in Europe. Already the North West region of England has a residential density greater than The Netherlands, the nation often held up as the most densely populated member state in the EU. The South East region outside of London and the West Midlands region are not far behind[3]. Many parts of these and the other regions have stretches of very beautiful countryside, and love of the countryside rather than the town is a deeply embedded trait of the British (and especially the English) character. There is an almost instinctive reaction to resist any form of development in open countryside. New roads are typically portrayed as "scars on the landscape". Bodies such as the Campaign for the Protection of Rural England (the CPRE), the National Trust[4], and the Royal Society for the Protection of Birds (the RSPB)[5] have very large memberships and considerable financial resources to fight road proposals thought to be inappropriate.

Blanket opposition to road proposals also comes from smaller but very voluble and

[2]Although the apparent certainty of the contribution of major road investments to local economic regeneration is not well supported by economic research.

[3]In 1998 The Netherlands 464 persons per sq km, North West England 486, South East region 419, and West Midlands region 410.

[4]The National Trust has 2.7 million members, and exercises stewardship over 200 properties and 345,000 hectares.

[5]The RSPB has over one million members and owns 168 bird reserves, covering 115,000 hectares.

pro-active 'green' groups. They are against any further road building on grounds that range from fears of global warming to localised carcinogenic effects of vehicle emissions, as well as loss of trees, landscape, animal habitats etc. Local NIMBY protesters, worried about impacts on property values, on the growth of subsequent 'urban' facilities at junctions, and the significance of noise etc., frequently find themselves as unlikely allies of the fiercer green protesters.

Blanket protest to a road scheme requires a political response. In contrast perhaps, local protests about one route as opposed to another or about the design details of a scheme will require an analytical basis for the eventual political decision, a basis that can portray the cost factors and the various impact factors of the options[6]. An analytical basis is also required by the road builder in choosing routes around or through areas of land or through buildings that are protected, by law or by convention.

In the UK some 560,000 buildings are "listed". They are therefore protected by law for their architectural and/or historical significance. There are also 31,000 protected Scheduled Monuments. And large areas of countryside have varying degrees of statutory protection against development: 6,755 Sites of Special Scientific Interest, 392 National Nature Reserves, 759 Local Nature Reserves, 11 National Parks and 40 Areas of Outstanding Natural Beauty (AONB)[7]; plus 44 Ramsar Wetlands, 19 Environmentally Sensitive Areas, 17 Forest Parks, 200 Country Parks and 13 Biosphere Reserves; and 765 Special Areas of Conservation covering 33,000 hetares. The total is over 60,000 sq km. In addition, non-statutory but land use planning protection is given to National Trust land, reserves of the RSPB, the Woodland Trust and the local (county based) naturalist trusts, to Heritage Coasts, and to land owned by the Wildfowl and Wetlands Trust and the Field Studies Council. In many parts of the country there are also significant revealed and recorded archeological sites. These date from the Iron and Bronze Ages, the Roman occupation, the Saxon and Viking centuries, the Medieval Era after 1066, through to the more recent Industrial Age. These sites are lost reluctantly[8]. In addition, there are 550,000 hectares of 'Common Land', and 'Green Belts' around the major British cities. These and other local land use designations exercise further constraints. There are over 10,000 Conservation Areas for example. These are localities within towns and cities of special architectural or historic interest designated by local governments. All of these elements present a challenge to a cost-benefit or multi-criteria analyst advising on alternative routes.

3 The COBA Approach

The Department of Transport of the British government has used a form of cost-benefit analysis to appraise major highway schemes since the mid- nineteen sixties. The successive editions of what became a computerised package were known as 'COBA'. This takes a "time-saving plus" approach. It comes out of a "predict-and-provide" tradition

[6]In the UK the Highways Agency, the body responsible for building and maintaining major highways, normally puts forward three routes for public consultation and review at a public inquiry.

[7]Between them National Parks and AONBs cover 23% of the land area of England.

[8]Road builders are also required to allow 'rapid' archeological recording of remains discovered in earth moving operations.

of investment decision making for roads, albeit always against severe public sector budget constraints. The costs of construction and of maintenance and servicing a new road are balanced in a benefit-cost ratio against the time saving benefits to traffic flows forecast over a 30-year period against the use that would be made of that traffic on the existing network. Changes in vehicle operating costs, as a user cost, and accident savings are also allowed for, and, as the computer models have improved, full network effects and allowance for traffic generation have been incorporated. Discounting to net present values was (and is) achieved with a standard public sector test discount rate.

Criticisms of COBA have come at three levels: the sweeping dismissal that it was too reductionist to put everything (time, human life) into monetary terms; or the economic differences of opinion that the key parameters were inappropriate (the choice of discount rate, the length of the life span, the posited vehicle mix, the seasonality assumptions, the path of future land-use planning strategies in the local area etc); or the systems criticism that the technology of the modelling for the required forecasts and traffic assignments was too crude and uncertain. Demand was to be met by supply, without allowance for traffic restraint policies or possible changes in attitude to or the cost of motoring. The analysis, critics claimed, too often seemed to be based on a statistical base that was already out of date by the time the decision came to be taken, let alone when construction would begin or the road would open. And there was a fear that it seemed to take decisions away from the political arena into the black box of the specialist analyst. This was perhaps most forcefully expressed in relation to schemes where environmental damage loomed large.

COBA has an inter-urban focus. From the late seventies it has been complemented by URECA, a procedure of intra-urban schemes. COBA does not value environmental impacts. Originally these were left to a generalised statement for discussion at the public inquiry, and then since the mid- eighties put into the form of an Environmental Impact Appraisal and Statement. COBA also does not place a value on the quality of life benefit of a by-pass road to the residents of the settlement that is by-passed (unless the residents use the new road and get counted into the time savings).

Early protests at the narrowness of the COBA assessments, the seemingly too simple engineering focused reduction to a cost-benefit ratio, led the Department of Transport to commission a review of its procedures. This was published as the *Leitch Report* (Leitch, 1977). This report argued for a more balanced appraisal process and for greater openness in the assumptions and in the uncertainties. The government's response was guided by the 1979 report of its independent Standing Advisory Committee on Trunk Road Assessment (SACTRA). The forecasting methodology improved, especially in the assignment modelling and in model validation checks which included sensitivity tests. SACTRA has produced further reports, *Urban Road Appraisal* in 1986, *Assessing the Environmental Impact of Road Schemes* in 1992, *Trunk Roads and the Generation of Traffic* in 1994, *Transport Investment, Transport Intensity and Economic Growth* in 1996, in response to public criticisms of the Department[9].

Politically, by the mid-nineties it was recognised by analysts and politicians alike that the appraisal of highway schemes in the United Kingdom not only had to draw

[9]The current *COBA Manual* was published by the Department of Transport in 1996.

in consideration of environmental impacts more centrally, but that appraisal also had to link trends and investments in private road transport to trends and investments in public road transport and the use of highways by cyclists and pedestrians, as well as to trends and investments in rail. By the time the Blair government came to power in 1997 the stage was set for the claims of a new approach to highway appraisal, one that was both socially and environmentally aware and multi-modal.

4 The New Approach to Appraisal

One year after it was elected the Blair government was ready to allow the ministers and the supporting civil servants in the Department of Transport to publicly admit the following:

1. That plans for each individual major highway scheme must be able to demonstrate that it is set into an overall national transport strategy, in order to defend the long held position that the public inquiry for a scheme is there to consider the routing of the new road, not the requirement.
2. That the appraisal of each such scheme must incorporate consideration of the impact of the investment on other modes of transport, on both the local and the wider environment, and on regeneration and community severance.
3. That the COBA methodology needs further refining, and needs to be drawn into consistency with the cost benefit analyses used to appraise investment using public funds in other modes of transport.
4. That greater attempts to find monetary values for environmental impacts using a Contingent Valuation technique are worth pursuing (see Bateman, et al., (2002), the guide sponsored by the Department of Transport).
5. That non-CBA methodologies (such as Multi-Criteria Analysis) may be useful in contributing to decision-making (see DETR, 2000a).

Items 1 and 2 above have been taken the furthest. Following a 'Roads Review', re-appraising nearly 100 schemes in-the-pipeline in England[10], in July 1998 the government published *A New Deal for Trunk Roads in England* (DETR, 1998a) with two supporting documents: *Understanding the New Approach to Appraisal* (DETR, 1998b) and *Guidance on the New Approach to Appraisal* (DETR, 1998c). The new approach was applied to each of the schemes. A ten-year 2010 *Transport Plan* followed (DETR, 2000b). This Plan projects an expenditure of £21.3bn (at 2000 prices) on 30 trunk road by-passes, on relief at 80 major bottlenecks, and on widening 5% of the key network; as well as even greater expenditure on both road and rail public transport investments. Further expenditures, mostly on motorway widening schemes, were announced in 2003.

Items 3, 4 and 5, after four years, are still within a research agenda. And thinking aloud about congestion charging and trunk road pricing is expressed with great political caution and sensitivity to voter reaction. A congestion charge cordon (of £5

[10]Scotland, Wales and Northern Ireland take decisions on road investments through their own agencies and representative bodies.

per vehicle) has been introduced in London under a new freedom granted to towns and cities to introduce traffic restraint policies under their 'Local Transport Plans'. This has suprised many commentators in the success of the technology used and in the degree of public acceptance. And the first major toll motorway in Britain, the privately financed Birmingham Northern Relief Road, is due to open in 2004. Experiments have been conducted for some time into the feasibility of directly tolling existing motorways with number plate reading equipment[11]. More recently a longer range programme of research has been launched into operating a road pricing system on all roads using a vehicle-mounted GPS aerial. It has been publicly recognised by the government that it is politically unacceptable to further raise the tax on petrol as a traffic restraint policy.

The *New Approach to Appraisal* is more presentationally new than it is analytically new, but the development in presentation has pushed forward the analytic debate. The *New Deal* document sets out five over-arching objectives for transport:

- to protect and enhance the built and natural *environment*;
- to improve *safety* for all travellers;
- to contribute to an efficient *economy*, and to support economic growth in appropriate locations;
- to promote *accessibility* to everyday facilities for all, especially those without a car; and
- to promote the *integration* of all forms of transport and land use planning, leading to a better, more efficient transport system.

These five objectives are expressed as criteria to be applied to each major highway scheme, with sub-criteria:

- Environmental impact
  - Noise
  - Local Air Quality
  - Landscape
  - Biodiversity
  - Heritage
  - Water
- Safety
- Economy
  - Journey Times and Vehicle Operating Costs
  - Journey Time Reliability
  - Scheme Costs
  - Regeneration
- Accessibility
  - Access to Public Transport
  - Community Severance
  - Pedestrians and Others
- Integration

[11] As used in London.

The impacts each of these in any given scheme are portrayed on an *Appraisal Summary Table (AST)*, qualitatively, quantitatively and as a summary assessment. Where quantitative data is not available a seven point scale is used: large, moderate or small negative, neutral, and slight, moderate or large positive. Occasionally 'very large negative' is used where the environmental impacts are deemed to be exceptionally severe.

Portraying the environmental impacts in a consistent manner between schemes is clearly a difficult issue. The environmental sub-criteria broadly reflect the structure of the Environmental Impact Assessment used for trunk road schemes in the past fifteen years, as required under EC Directive 85/337. Taking these in turn:

Noise: data on the number of properties that would experience significant increases or decreases in noise levels (±3dB(A)) in the design year if the road was built.

Local Air Quality: also using the number of properties affected, positively and negatively, weighted by distance from the road, using NO2 and PM10 at the National Air Quality Strategy standards. An estimate is also given for the net change in the level of Carbon Dioxide as a reflection of the impact on global emissions.

For landscape, biodiversity and heritage the concept of Environmental Capital has been used, applying the seven point scale, with features appraised against indicators of scale, importance, rarity, substitutability and impact; and, in certain cases, possible mitigation.

Landscape: the impact of the scheme is scaled against the national classification of landscape in terms of the following features: pattern, tranquillity, cultural features, and land cover by different uses.

Biodiversity: an assessment in terms of a nature conservation evaluation of species and habitats, or of natural features affected and the ecological impact of the proposed scheme.

Heritage: the impact on the built historic environment of the scheme uses a standard national classification in terms of: the physical form of the site, the survival of original fabric, the condition of the site, the complexity of the elements, the contextual setting of the site within its immediate surroundings.

Water: the scaling is applied to the impact on water quality and on land drainage and flood defence on the basis of a risk-based approach to the sensitivity of the receiving environment, allowing for mitigation.

Further details of the approach taken to these environmental criteria and to the other indicators included in the ASTs are set out in the *Guidance* document (DETR, 1998c).

The AST is clearly a considerable aid to the decision maker choosing between schemes from a limited budget; and it allows members of the general public, both supporters and opponents of a scheme, to compare their scheme with others and to judge specific areas of strength or weakness in making their case to the media, to their elected representatives, or to the public inquiry. Its design has been a response to environmental protests as much as a desire to widen the transport implications away from just road traffic flow. It could be said to reflect an objective of reducing use of the motor car; while analytically it may be said to resemble the first steps of a Multi-Criteria Analysis. However it does have its limitations, as discussed in the next section.

5 Outstanding Issues

The AST is primarily a political tool. It is recognised that it would ideally be desirable to have a weighting between the criteria used. However, that runs into the issue of whether the weights should be those of the (ill-informed?) public, or of the technical (blinkered?) specialists, or of the (sagacious?) politicians. And weights would yield implicit valuations for each criteria. The judgements reached on major highway schemes are essentially for the long term and they are non-reversible. It is the interests of society, today and tomorrow, that are to be served. The same weighting dilemma applies within the environmental criteria to its sub-criteria.

The danger without weights is that a single sub-criteria could be deemed to be all-important in turning down a scheme, without acceptance of the principle of trade-offs. For some members of the public (and one suspects for some politicians) the sanctity of a Site of Special Scientific Interest or of land held by the National Trust is absolute, with the implication of an infinite value. The same attitude is sometimes put forward towards the habitats of rare species, or to what is claimed to be a 'unique' landscape. A recent partial response to this stance is reflected in a declared willingness to place an upgraded stretch of the A303 trunk road into a very expensive tunnel as it passes close to the historic site of Stonehenge in Wiltshire.

The AST retains avoidance of any approved attempt to place monetary values on environmental impacts of new roads in Britain. This is in spite of the fact that two of the sub-criteria apply property-based information, and differences in property values have a wide acceptance as proxies for environmental values. Other proxy approaches, such as the Clawson travel budget approach are less applicable. Confidence in Contingent Valuation as a consistent technique is increasing, as lessons are learnt in the formulation of the Willingness-to-Pay and the Willingness-to-Accept questions to sample populations; but small variations in survey responses can amount to very large financial numbers over the lifetime of a road scheme, even with discounting (See Bateman et al., 2002).

Inevitably there is concern about the appropriateness of the definitions used in the environmental sub-criteria. Two examples illustrate the point. Should the particulate measure of local air quality be based on PM4 rather than PM10? There is dispute over which is the more carcinogenic. And why 3dB(A) for the noise contour, and with no recognition of possible variations in the nature and duration of the noise? There is a difference between output of noise and the behavioural response to it. Also, ideally, a full picture would require an assessment of the environmental changes expected on other roads or through changes in the use of other modes of transport as a result of building the given scheme. It is important that the seven point scale is verbal and not numerical, without implied multiples in the scale.

Among the non-environmental criteria in the AST there is also a definitional issue for 'Reliability', measured by an indicator of 'road stress'. This reflects changes in the relationship between the traffic flows and the capacity of a road, either through junctions or on links. And while this criteria has a quantitative base, the criteria of 'Regeneration' is strictly qualitative. Awkwardly this criteria is currently set to reflect two rather different aspects of regeneration: whether the new road may be deemed

to be potentially beneficial for designated regeneration areas, or whether there are significant developments within or adjacent to the regeneration area which are likely to be dependent upon the road investment going ahead. Further, there is a sense of double counting here, with the regeneration advantages already reflected in the forecasts of use and time savings of the new road.

Issues also remain within the COBA procedure. The concerns noted in paragraph 16 above will always be there, but some of those concerns and some new ones open up when consideration is given to the desirability of consistency in CBA for road investments with CBA applied to other public expenditures in the transport arena. The assessment of four areas of such expenditure in addition to trunk road investments have recently been reconsidered on behalf of the British government: route subsidies paid to loss making bus[12] and train operators[13], investment subsidies paid to rail infrastructure, freight and passenger providers (now in the private sector in the UK), investment subsidies given to light rail and guided bus urban passenger transport investments[14], and in relation to support given to harbour authorities for port developments (DTLR, 2002).

There are also a number of issues which arise in striving to have the methodologies used for different types of public sector transport appraisals being mutually consistent. Section 56 agreements for example require a 'restricted CBA' in order to demonstrate that the level of investment grant (to urban light rail or guided busways) be no greater than the portion of the benefits of the scheme to arise for *non-users* (eg road users on the less congested roads which result from the improvement in public transport). Assessment is also needed for the subsidies that are paid to support private sector train operating companies on the railway system for what are deemed to be non-commercial passenger services (see Mills and Howe, 2000). The issues now have a new relevance in relation to the current round of "Multi-Modal Studies", which are considering proposals for packages of projects with a full or partial public sector financial contribution (road and rail infrastructure improvements, traffic management, public transport service upgrades) along key congested inter-urban transport corridors and at key nodes.

A general point in an AST type of presentation of an appraisal, where there are items that are not measured as monetary costs and benefits, is the need to be clear as to whether these are *additional* costs and benefits, or whether they are *redescriptions* of costs and benefits included elsewhere (many of the regeneration impacts might be an example).

More fundamentally is the question as to whether the CBA that is to be used is drawn up as a *calculus of willingness-to-pay* or as a *calculus of social costs and benefits*. These should be equivalent in total but the latter is based on real resource costs or benefits, while the former is based on a summing of the net welfare changes for each individual that is brought about by the project being considered. In this summing there might well be items that are benefits to one person while being a cost to another.

[12]Through the Transport Grants for 'Local Transport Plans', paid through County Councils.

[13]Through OFRAF, the Office of Passenger Rail Franchising. The Strategic Rail Authority, concerned with supporting investments in new rail infrastructure, also undertakes CBA.

[14]Through Section 56 Agreements with sponsors of the investment.

These are transfer payments or pecuniary externalities, commonly found with environmental impacts. The principle advantage of the WTP approach is that it allows the preparation of a balance sheet with the net social benefit of a project disaggregated into impacts upon different social groups. Financial and non- financial impacts can be distinguished (important where private firms are involved). The difference in the two concepts lies in presentation.

The two concepts present a unit of account issue: a resource view would measure at *factor cost*, net of indirect taxes, while a WTP approach would be at *market prices*. Therefore, for consistency an indirect tax correction factor should be applied. This is a particular issue in assessing vehicle operating costs in road schemes, where different rates of VAT and petrol or diesel duty apply to different categories of vehicle. COBA uses the factor cost unit of account. In contrast the OFRAF methodology does not take account of differences in indirect taxes. Sugden argues that it is appropriate that the public decision maker should be aware of the impact of the proposed project on indirect tax revenues, just as he should be aware of the differences between user benefits and non-user benefits (Sugden, 2002, Sections 5 and 6).

Further issues in transport CBA include the distinction between *behavioural* and *equity* values of non-working time, and between *perceived* and *unperceived* private costs. It is the behavioural values and the perceived costs that are used for forecasting changes in travel patterns. The COBA treatment of accident risks as unperceived costs of travel, similar to the non- fuel costs of car trips, means that accident rates play no part in forecasting travel patterns or modal choice. All accidents are treated as negative externalities of travelling, whereas, it may be argued, they are an element in WTP.

Such considerations lead to a wider issue in the presentation of highway and other transport appraisals. This is the general desirability of disaggregation, in three directions, each of which has relevance to the consideration of environmental impacts. The *first* is disaggregation by the *recipient* of the benefit or cost, the distributional impact of the project. The final incidence of a project may be very different from the initial incidence. The *second* disaggregation is by *source* of the benefit or cost, by the activity of the impact, important in distinguishing between user and non-user benefits. The *third* disaggregation is by the *nature* of the benefit or cost, the proportions of benefit for example that come as time savings, price reductions, frequency improvements etc. These disaggregations can be applied conceptually to the environmental impacts, even if the absence of monetary valuation keeps the appraisal of these impacts outside the CBA. It is an open question as to how worthwhile this is for any but the largest, most costly and contentious schemes.

6 The Democratic Process

> *"The decision on any scheme will always be an exercise in political judgement in the end, but the quality of that decision is critically dependent upon the quantity, quality, and accuracy of the material on which it is based".*

The last sentence of the 1992 SACTRA Report perhaps is based on optimism for

the contribution of the analyst in a democratic environment to come up with a 'best' answer in relation to a single highway project or in relation to ranking a portfolio of highway and non-highway transport projects.

While in the interests of transparency it is good that the AST used by the British government is not too complicated or sophisticated, many might argue that the AST 'single page' summary is an inadequate basis for making choices committing millions of pounds of public expenditure. This is a question for the detail of the decision making process. One criticism is that the AST does not reflect any parameters of risk and uncertainty for the scheme as a whole: there are no alternative scenarios or sensitivity analyses supporting the view taken on each criteria, especially the COBA estimates of PVB and PVC.

The AST reflects considerable uncertainty in the social choice process as to how to disentangle and value the desirability to commerce and industry of having access to an uncongested trunk road system in the interests of economic efficiency, while rising real incomes increase the attractiveness of the use of that same system by private motorists. A system of differential road pricing is probably the only answer to that dilemma. Acceptance of this solution is slowly gaining ground in British public opinion[15].

The AST approach also gives a prominence to environmental impacts that many would argue is undue in relation to what road users would actually be willing to pay for. This reflects a present political reality in a pluralistic democracy, where elected representatives shy away from leadership and the unpopular longer term view in the face of a critical media and voluble pressure groups. Arguably it is exacerbated by a current regime of inadequate compensation paid to households and environmental bodies and low levels of expenditure on mitigation for the adverse effects of new highway schemes.

References

Bateman, I. J., Carson, R. T., Day, B., Hanemann, M., Haley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Özdemiroglu, E., Pearce, D. W., Sugden, R., Swanson, J. (2002) Economic valuation with stated preference techniques: A manual. Edward Elgar, Cheltenham

DETR (1998a) A new deal for trunk roads in England. Department of the Environment, Transport and the Regions, London

DETR (1998b) A new deal for trunk roads in England: Understanding the new approach to appraisal. Department of the Environment, Transport and the Regions, London

[15]It is a sign of greater political willingness to confront the issue of direct charging for the use of major roads that the British govenment issued a Discussion Paper in 2003: Managing Our Roads (Department of Transport, 2003) and has invested in a major feasibility study.

DETR (1998c) A new deal for trunk roads in England: Guidance on the new approach to appraisal. Department of the Environment, Transport and the Regions, London

DETR (2000a) Multi-criteria analysis: A manual. Department of the Environment, Transport and the Regions, London

DETR (2000b) Transport Plan 2010. Department of the Environment, Transport and the Regions, London

DoT (1996) COBA 9 Manual. Department of Transport and HMSO, London

DoT (2003). Managing our roads. Discussion paper, Department of Transportation, London http://www.dft.gov.uk/stellent/groups/dft_about/documents/pdf/dft_about_pdf_022865.pdf

DTLR (2002) A project appraisal framework for ports: A consultation document. Department for Transport, Local Government and the Regions, London

Leitch, G. (1977) Report of the advisory committe on trunk road assessment. HMSO, London

Mills, G., Howe, M. (2000) Appraisal of non-commercial passenger rail services in Britain. Journal of Transport Economics and Policy 43(1): 113-130

SACTRA (1986) Urban road appraisal. Standing Advisory Committee on Trunk Road Assessment, Department of Transport, London

SATRA (1992) Assessing the environmental impact of road schemes. Standing Advisory Committee on Trunk Road Assessment, Department of Transport, London

SACTRA (1994) Trunk roads and the generation of traffic. Standing Advisory Committee on Trunk Road Assessment, Department of Transport, London

SACTRA (1996) Transport investment, transport intensity and economic growth. Standing Advisory Committee on Trunk Road Assessment, Department of Transport, London

Sugden, R. (2002) Developing a consistent cost-benefit framework for multi-modal transport appraisal investment,. Department for Transport, Local Government and the Regions, London

Partnerships Contributing to Sustainable Urban Tourism

Michaela Gindl, Florian Wukovitsch*

1 Introduction

This paper is based on the Deliverable no. 4 (Gindl et al., 2002) of the SUT-Governance project[1] funded by the European Commission under the Fifth Framework Programme. The project was conducted between May 2000 and June 2003. Further information can be found on the project homepage at http://sut.itas.fzk.de/.

In the project it is strongly argued for partnerships as innovative form and instrument of local governance to deal with externalities and sustainability in the urban field. Therefore the project presents an effort to work with public-private partnerships and urban governments in Europe to develop, validate, and deploy a 'general framework for urban sustainable tourism partnerships' that is applicable in a variety of urban municipal and development contexts. The overall goal of the project is to elaborate and promote innovative forms and instruments of local governance to improve urban tourism development involving the principles of sustainability and participatory decision-making.

A wealth of literature (Paskaleva-Shapira, 2000) focuses on partnership cooperation. In the last 20 years 'public-private partnership' has become a catchword and was presented as a remedy against sub-efficient policy coordination (Lowndes, Skelcher, 1998). In this line, the main goal of this paper is to enhance the understanding of the complexity characterising the forms of cooperations in the area of sustainable urban tourism and to contribute to the classification and definition of, as will be shown, this still vaguely defined concept. The core question of the proposed paper refers to the key factors determining forms of partnership collaboration in the field of sustainable urban tourism.

In the first section of the paper we trace back the advent of the discussion on governance and new forms of collaboration between the public and the private sector and embed the discourse into a wider politico-economic development context. By present-

*Department of Environmental Economics and Management, Vienna University of Economics and Business Administration, Vienna, Austria.

[1] 'Sustainable Urban Tourism. Involving Local Agents and Partnerships for New Forms of Governance' (SUT-Governance). Research project of Key Action 4: 'City of Tomorrow and Cultural Heritage' of the 'Energy, Environment and Sustainable Development' Program within the 'Fifth Framework Program' of the European Union. The project extends from May 2000 to June 2003. Further information can be found on the homepage of the project: http://sut.itas.fzk.de/.

ing a detailed literature review on current research in urban studies we illuminate the academic contribution to the debate on governance. In the following section different types of partnership collaboration in the field of sustainable urban tourism, based on empirical evidence, are developed. The paper is concluded by some final remarks on the opportunities and drawbacks of the governance-boom and its consequences.

2 Literature Review

2.1 Economic Restructuring and Urban Politics

For more than twenty years there has been a big academic discussion on the process of global economic restructuring, i.e. a shift of sectoral contributions to macro-economic output and an ongoing international division of labor in front of new technological and financial regimes. On the part of political sciences, economic geography and new urban sociology many scholars have engaged in analyzing the impact of the transformation of the international economic regime on urban, national and international politics. Among the most inspiring terms are the 'hollowing out of the nation state' with its counterpart of 'glocalisation', indicating the lost power of the nation state as opposed to increasing regulation effort on supra-/international and local level to find 'post-national' solutions (e.g. Lipietz 1992, Swyngedouw 1992). Even more absorbed in academic and political discussions became processes of cultural, economic and political 'globalization' in a increasingly 'networked society' (e.g. Altvater, Mahnkopf, 1999, Amin, 1992, Castells, 1998). Jessop (2002) speaks of the transition from the 'Schumpeterian workfare postnational regime' to the 'Keynesian welfare national state', indicating that while Keynes was the leading economist in Post-World War II politics in Europe, Schumpeter is the leading economist of today.

In the field of urban studies a lot of research has been done on the analysis of the impacts of globalization on social dynamics, economic restructuring and new formulation of politics in cities. The core hypothesis is that in the global economic regime cities take a crucial role as centers of innovation and nodal points of global power but are also the foci of new global challenges in the form of increasing social polarization and environmental problems. From a macro perspective, new urban sociologists coined the field of 'global city research' (e.g. Sassen, 1994, 2001) where also issues of economic competition between cities ('cities in competition') can be found on the agenda (e.g. Brotchie, 1995). On a micro level, much effort has been channeled to 'post-Fordist city politics' (Mayer, 1995), important questions being how agglomeration effects and economic prosperity could be promoted by local politics while, at the same time, further progress in terms of social equity and the internalization of environmental externalities could be achieved. Therefore, research has been particularly focused on institutional efficacy, power-struggles, democracy and the new share of responsibilities between public and private actors to account for externalities and public goods (Borja, Castells, 1997, Eischenschitz, Gough, 1998, Ekins, Newby, 1998).

2.2 Urban Management and the Discourse on Governance

The new catchword to deal with the rupture of government's sovereignty and transformation of policy making and policy implementation on local, national, supranational and international scale has been referred to as 'governance'. Originally deriving from the field of development and foreign-aid politics, after the breakdown of actually existing socialism the term 'good governance' rapidly turned into a key concept in several policy fields of interand supranational institutions (Fürst, 2001, Raffer, Singer, 2001). Also in urban development an explicit break with the predominance of physical planning could be observed, the idea being that pre-set objectives of the functionalist approach to urban planning have lacked adequate problem-solutions in times of growing unemployment, urban segregation and environmental pollution. Enhanced awareness of those urban problems as well as the putative pressure to compete with locations around the world for the 'global dollar' called for new institutional settings beyond strict state control on the one hand and pure market solutions on the other. New forms of flexible regulation should help to manage externalities and public goods in a more efficient way (Andersen, van Kempen, 2003, Mayer, 1995).

As we have already indicated above, the growth of governance into a major issue of analysis in social and political science in recent years produced an overwhelming stock of literature and somehow a 'Babylonian confusion of tongues', complicating an exact definition of the concept. 'Good governance' as advocated by the Worldbank aims at the reduction of corruption and empowerment of local communities in development countries and, thus, targets at government policies, while 'corporate governance' denotes the other extreme of governance as an exclusive business concept, dealing with the steering of large corporations. Because of this high degree of elusiveness, Theys (2000) admits for the political arena of environmental issues that the debate on '*governance is generally locked into two contradictory discourses. For some, good governance is the only solution to current environmental problems and its vocation is to replace traditional public policies that are considered inadequate. For others, in contrast, governance is the problem [. . .] as it does no more than reinforce collective powerlessness in front of challenges which are increasingly ungovernable*' (Theys, 2000, p.4).

But beside asking for the its political relevance it seems most crucial to challenge the concept from a comparative perspective on current theories in urban and regional studies. Fürst (2001) elaborates on the differences between the concepts of 'Regional Governance', the French 'Milieu-approach' and 'Regulation-school', and the US-founded 'regime-concept'. As all those theories deal with the topic of regional co-operations to foster economic development, they self-evidently overlap in one way or another, but also have their relative theoretical merits. From this point of view, while e.g. the French 'Regulation school' suits best for a structural analysis of capitalist development (see the references to authors like Lipietz and Jessop above), the concept of governance has its strength in highlighting interaction between actors in the arena of regional and urban politics.

In contrast to this, Schneidewind (1997) analyses the contribution of the concept of 'Public Private Partnerships' in the field innovative environmental governance.

Schneidewind starts with the reflection that 'Public Private Partnerships' once connoted a form of collaboration between public and private partners in the field of huge infrastructure projects for which in times of deregulation and zero-deficit policies the public sector was financially too weak to provide resources exclusively on its own. He claims, however, that this approach has to be contrasted with the concept of environmental governance, where the term always stood for a much wider concept, including round tables and mediation processes with governments, companies and NGOs as well national and international agreements. Therefore, he suggests that neither a universal model of 'partnership' nor a clear demarcation line between governance and 'partnership' can be identified.

2.3 Empirical Studies on Urban Governance and Open Questions

For all the problems with finding a clear definition for the concept of governance, empirical data to confirm or modify existing findings and to understand in which concrete forms governance is or has been organized seems essential. Hitherto presented considerations gave insight into the larger context of partnership formation but lacked the provision of empirical evidence of co-operation models in different settings of urban politics and culture. Therefore, before we can discuss our findings about governance models for innovations in urban tourism, the questions arises whether there can be one universally valid and applicable model of sectoral urban governance in diverse and place-specific settings. To exemplify problems with the dissemination of governance models we refer to a comparative analysis of urban regeneration policies in the USA and the UK. Because this study deals with the transferability of a US partnerships model into an relatively similar cultural setting, it clearly shows the potentials and drawbacks of dissemination approaches.

Davies (2002) discusses the applicability of the US type 'regime governance', also referred to as 'governance without government' or 'governance by network', in UK urban regeneration policies. Although during the 1980s there was a strong trend of directly importing US regeneration policies to the UK, his perspective on the transferability of the concept to UK cases is rather sceptical. According to Davies' analysis, UK type regeneration partnerships 'are a distinctive mode of governance which fit neither the old model of governance by government, nor the new model of governance by network' (Davies, 2002, p. 302). While in the ideal-typical case of US type 'regime governance', voluntary networks between local authorities and business elites aim at achieving otherwise unattainable goals under a high degree of autonomy and thus form a highly hegemonic project influencing a whole borough, town or city, co-operation in the UK is as well characterized by hierarchical relationships between local actors and/or between extra-local and local actors and externally (i.e. not locally) determined objectives. Moreover, in many cases interaction between the business and public sector remains to be primarily short-term, instrumental and determined by law and to have no influence on the mechanisms of local policy making at all (Davies, 2002, p. 306). And finally, contrasting to what the model of 'regime governance' would suggest, instead of increasing autonomy for local institutions, recent transitions of urban regeneration policies in the UK even resulted in growing political centralization.

But if empirical models of governance resemble a large variety of concrete forms of public-private collaboration while at the same time narrow ideal-typical concepts in the literature lack empirical relevance, the only thing one can do is synthesizing the disperse observations and create a highly generalized model of governance, covering all aspects of individual observations. A prominent attempt to clarify the meaning of the concept in this way is made by Jan Kooiman (2000). He defines governance as '*all those interactive arrangements in which public as well as private actors participate aimed at solving societal problems, or creating societal opportunities, attending to the institutions within which these governance activities take place, and the stimulation of normative debates on the principles underlying all governance activities*' (Kooiman 2000).

Although those generalized definitions are most useful for the academic debate, it still (or even more) seems relevant to ask for the differences (see also Davies, 2002, Fürst, 2001). We strongly belief (and research results of our project justify our assumption) that there is a variety of patterns of public-private collaboration, depending on culture, the stage of capitalist development and last but not least the sector tackled by intervention. For this reason, in the second part of the paper we shed light on aspects of governance in the field of innovative problem-solutions for urban tourism.

3 Partnerships for Sustainable Urban Tourism – The Research Process and Results

Tourism figures among the industries with major growth in Europe and, as Law (1993, p.1) argued, '*large cities are arguably the most important type of tourism destinations*'. Paskaleva puts the ensuing challenges as follows: '*Steadily increasing investments in urban regeneration, heritage conservation and improving the quality of urban life to adapt the city to the needs of visitors of attractive facilities, comfortable transport, diverse events, and capitalisation of historical sites, among other activities, create new prospects for the industry and the urban communities.*' (Paskaleva-Shapira, 2000) This development involves complex decision making problems for the key stakeholders, among them city officials, planners, (tourism) industry and the public. Involving sustainability considerations and long-term community advancement poses serious challenges for policy makers and tourism developers for tourist functions are very rarely produced for, or consumed by, tourists but a whole range of users (Shaw/Williams, 1994, p. 201).

This research assumes that multi-stakeholder cooperations can be useful means in dealing with the issues raised above. Yet, as aforementioned, a weak point in recent theories is the provision of empirical evidence of models of co-operation in different settings of urban politics and culture. Moreover a cogent basis is lacking in existing theory to treat questions of stakeholder participation in a partnership framework as an operational mechanism in the pursuit of sustainable urban tourism. Therefore an inductive, exploratory approach was chosen to stepwisely narrow the research process towards understanding the complexity of partnerships for sustainable urban tourism and thus to contribute to the understanding of new forms of governance in the field of innovative problem-solutions for urban tourism.

This chapter illustrates the above literature review with results of empirical research on interactive arrangements in which public as well as private actors participate aimed at solving societal problems, or creating societal opportunities, i.e. partnerships for sustainable urban tourism (SUT-partnerships).

3.1 Preliminary Hypotheses

One of the initial hypotheses of this paper maintains that governance models such as of public-private collaboration strongly depend on the national and sectoral context in which they are established. It is assumed that basic characteristics of public-private partnerships for sustainable urban tourism differ in the four study countries of the SUT-Governance research project (Austria, Bulgaria, Greece, and Germany) and that similarities are only to be found at an abstract level of generalisation.

Beforehand, it should be mentioned that *partnership*, in the context of the present research, is defined as a process of sustained collaboration, in which distinct organisations come together to define, to resource and to achieve a shared vision. Talking about *SUT-partnerships* in detail, they are understood as characterised by

- favourable framework conditions, involving
- a viable partnership process and
- a successfully implemented activity, resulting in
- diverse sustainable development benefits.

The interest of this paper is in analyzing actual characteristics of SUT-partnerships, more precisely in understanding the nature of partnerships in order to provide an avenue for generalisations. This entails that the analysis of the process of collaborating as a partnership becomes decisively important. Yet, differences in collaborative capability between organisations can be a crucial barrier in establishing a partnership. Moreover, lacking legitimisation of partners within their organisations of origin can seriously constrain the collaboration in terms of maintaining the partnership process, involving arrangements and procedures durable over time. Here it is hypothesised that those risks can be overcome if the approach used for building up a partnership is opportune in developing a strategic framework jointly between the partnership actors and/or organisations.

3.2 Searching for SUT-Partnership Cases to Study

During autumn of 2001, each of the four national research teams of the SUT-Governance consortium (Austria, Bulgaria, Germany and Greece) conducted and analysed two detailed partnership case studies representing successful examples of multi-stakeholder cooperations in sustainable urban tourism (SUT-partnerships). The partnership cases (eight in total) were selected from the cities of Graz (Austria), Veliko Turnovo (Bulgaria), Heidelberg (Germany) and Thessaloniki (Greece).[2]

[2] Austrian partnership cases (from the City of Graz): 'Strategy Forum Tourism', 'OeKOPROFIT for Tourist Companies'. / Bulgarian partnership cases (from the City of Veliko Turnovo): 'Beautiful Veliko

At the beginning, the process of searching for study cases focused on partnerships between public and private actors (SUT-PPPs), the initial hypothesis being that those partnerships are mostly initiated by local/urban administrations seeking collaborative opportunities with other stakeholders, the private sector in particular, to promote urban tourism. Yet, finding eight (comparable) international individual cases of collaborative pursuits towards sustainable tourism in urban environments in four European countries posed some unexpected problems:

- Potential study cases represented a wide spectrum of how the sectors are combined into a partnership arrangement ranging from forms where the public sector dominates the collaboration to cooperations between public institutions or between private organisations only.
- Moreover, the multifaceted nature of 'partnership-content' in regard to sustainable urban tourism turned out to be largely heterogeneous (e.g. tourism related traffic management; environmental management; tourist information; tourism network development; product promotion; activity and service development; preservation of historical and cultural heritage; enhancement of residents' life quality; controlling urban development; improvement of urban space; representing tourists' and tourism industries' interests towards the local governments).
- Partnership actors were motivated to enter a collaboration for sustainable tourism for a variety of reasons (e.g. enhance tourism development to overcome existing economic problems); partnerships were also formed in reaction to specific pressures and demands of local development (e.g. the need of developing effective tourism practices, preservation of the cultural and historic heritage, fundraising for new activities, etc.).

While the initial focus of the SUT-Governance project was to develop, analyze and validate a generally valid model of public-private-partnerships for sustainable urban tourism (SUT-PPPs), the research consortium was soon confronted with the problem that narrowly defined PPPs were not represented in all case study locations. Therefore it was decided to widen the definition of partnership: For we have learned that SUT-partnership arrangements can range from (as initially searched but hard to find) ideal-typical PPP-forms to cooperations between public institutions or between private organisations only, the actual analysis focussed on multi-stakeholder partnerships in the field of sustainable tourism. Reflecting the constitutive importance of the partners' motivations instead of the sectoral affiliation the initial hypothesis was accordingly modified.

For all the heterogeneities, the domain, i.e. the object of study, for which the case studies were aimed to derive general results, needed to be articulated in order to be able to derive results that transcend the particularities of each case, to be comparable and to provide an avenue for generalisations beyond the immediate (Gomm, et

Turnovo', 'Council of Tourism'. / German partnership cases (from the City of Heidelberg): 'Healthy Food in Heidelberg's Restaurants', 'Heidelberg City Card'. / Greek partnership cases (from the City of Thessaloniki): 'Pilot Project for the Renewal and Development of the Historical and Commercial Centre of Thessaloniki', 'Inter-Municipal Co-operation: Linking Places of Natural Beauty'. For detailed partnership description, visit http://sut.itas.fzk.de/.

al., 2000, Hamel, 1993, p. 44) The common interest of the research was to elaborate and promote innovative forms and instruments of local governance to improve urban tourism development involving the principles of sustainability and participatory decision-making. For this purpose, the object of the best practice study cases discussed here is the partnership activity and the process of cooperation with its impacts on urban sustainability.

In order to be able to select the ideal cases to grasp this object of study, choice criteria had to be defined: the cases are

1. successful multi-stakeholder cooperations,
2. dealing with tourism, resulting in
3. positive impacts on urban sustainability, and last but not least showing
4. high readiness to co-operate with the research teams.

The selected cases satisfy those requirements particularly well in practice, as they are multi-stakeholder cooperation in the field of urban tourism with identifiable outcomes for sustainable urban development (the selected cases had to, whether intentionally or not, contribute to at least one dimension of sustainability[3] and to have no negative effects on the other two dimensions). Concerning their targets the eight partnership initiatives purpose to enhance sustainable management of tourism, urban renewal, development of suburban areas for recreation, and social progress and value of the urban communities.

3.3 Typology of SUT-Partnerships

Innovative problem-solution for urban tourism affairs is the key impetus for SUT-partnership establishment. More precisely, a 'local shortcoming' in tourism-development (e.g. stagnant tourism development, weaknesses in marketing, lacking attractiveness of the destination, etc.) is identified by touristic and tourism-related actors and in addition some of those driving actors are (explicitly or implicitly) aware of the obligation to pay regard to the principles of sustainable development. This is the common element of all SUT-partnerships investigated; beyond that a large variety of forms and contents was observed.

It was the explicit aim of the research to develop a common model of SUT-partnerships and their success and by doing this to enrich the scope of the typology of partnership collaboration. Although the case studies made general features of multi-stakeholder partnerships apparent, it turned out to be equally important, as Davies suggests, for comparative local studies to place sufficient emphasis on difference. "The fashion for highlighting processes of convergence, which the governance thesis implicitly encourages, could obscure important processes of divergence" (Davies 2002, p. 318).

Different problems evoke – depending on the local shortcomings and the state structure – different forms and types of partnership cooperation: In Bulgaria, for

[3]Economic, Environmental and Social Dimension as specified in the 'Brundtland-Report' (1987) and the 'Agenda 21', the concluding document of the 'Earth Summit' of the United Nations in Rio de Janeiro, 1992.

instance, the investigated partnerships would not have been established without external (international) financial support. In Greece, experience with public-private co-operation and residents' involvement in local decision-making has been limited and only recently becoming of increasing interest. Partnerships among public actors are, however, quite common. Moreover, the domination of the public sector and the wide scope of governmental intervention generally hinders public-private co-operations. In these conditions public-public partnerships are results of European policies, providing opportunities for additional financial support. Only in Austria and Germany or at least in the municipalities of Graz and Heidelberg, citizen involvement and public-private collaboration has been working well for years. In these communities, the public actors have realised that the efficiency of certain public initiatives would be increased, if public and private actors worked on a shared agenda.

Summarising, the modes of partnership formation vary significantly among the four countries analysed in this project. Many possible and reasonable ways of partnership categorisation were discussed during the case-analysis, reaching from a classification derived from country characteristics, sectoral particularities, the partnership roles or partnership content. Instead of developing a generally valid model of public-private-partnerships for sustainable urban tourism (SUT-PPPs), as initially intended, we have learned that the form of collaboration depends on the respective national and urban contexts. The formation of a local partnership initiative has to be traced back to various origins: In many cases, the public sector is the coordinator, sometimes the initiator, but not necessarily always the innovator who offers the decisive stimulus. In many cases external agents with scientific and/or development capacities provide the critical kick-off potentials. Local administrations usually provide the organisational frameworks for the partnerships, while other actors design and implement the activities.

Despite the divergences it holds true for the entire variety of partnership models identified that the specific form of partnership is defined in the early stage of first partner-contact and is based on the motivations for partnership formation and the supporting conditions. Considering the above perspectives, the initially suggested classification according to the involved sectors and their roles was revised in favour of a partnership typology based on the actors' motivations and objectives for partnership formation resulting in a jointly developed strategic partnership framework.

Development Partnerships

One pattern of the analysed partnership cases was characterised by the public sector stimulating and supporting (mainly financially) the implementation of co-operative initiatives for sustainable development of urban tourism. The general idea in this model is to create an (economic) win-win situation for the community and the participating private actors (mainly companies, enterprises). The public sector either aims the solution of long-term community problems (like unemployment) or to find new organisational approaches to tackling typical urban externalities, like environmental or townscape improvement. These are long-term development goals; but the duration of the partnership or, at least, the timeframe of public sector participation has a date

of expiry from the start. In most cases, once the private partners 'adopted the desired behaviour' or the investment programmes are completed, the public sector withdraws from the partnership or becomes solely a subsidiser of the private initiatives. More precisely, the relations between public and the private actors can be characterized as Mentor/Financier/Principal' as opposed to 'Learner/ Beneficiary/Target-Group'. Generally speaking, the public sector behaves as the principal player attempting to motivate the private sector to improve performance and adopt innovative practices by involving new know-how, establishing networks with other businesses or increasing the number of jobs via subsidized investments. That is to say, the public partner invests or co-finances the build-up of stocks, ranging from utilisation of know-how to improving infrastructures and buildings or recreational areas. Since the public sector typically lacks crucial know-how and skills for these tasks, in most cases in-between mediators are involved to facilitate partnership implementation. Summarising, development partnerships are based on the public sector's aim to stimulate and support the implementation of co-operative initiatives. The general idea is to create an (economic) win-win situation for the community and the participating private actors (like environmental or townscape improvement), but the duration of the partnership or, at least, the timeframe of public sector participation is limited right from the start. Once the private partners 'adopted the desired behaviour', the public sector withdraws completely or to the residual role of a financial contributor.

Marketing Partnerships

Like in the pattern described above, the organisations forming the partnerships remain distinct in this model, especially with regard to strategy making, but service delivery is combined and carried out by a jointly owned partnership agency. Hotel and restaurant owners, public events organisers, and local tourism development authorities join efforts to improve service delivery to tourists in their communities. Compared to the former category, the main differences here refer to the level of division of actors' roles, in this case, the relationships between the public and the private partners are more equal (although the public partners usually bear a larger share of the costs) and the partnership activity is a continuous undertaking. The latter can be regarded as inherent to the nature of this partnership type's objectives, (i.e. tourism marketing in the study cases) requiring a long-term co-operation using a common cooperative framework. In contrast to 'development partnerships', 'marketing partnerships' usually adapt the content of the co-operation and, if one objective is completed, a new one is set up. Therefore, an on-going process of collaboration is necessary, enabling swift adjustments to changing environments. While 'development partnerships' achieve sustainability targets by sustained stock enhancement, the 'marketing partnerships' achieve sustainability goals by sustaining the partnership process itself and are aimed at long-term community benefits. Summarising, marketing partnerships are founded for combined service delivery which is carried out by a jointly owned partnership agency. The partnership activity is a continuous undertaking as the nature of the set objectives requires a long-term co-operation using a common cooperative framework. The content of the co-operation is frequently adapted and, if one objective is completed, a new one is

set up. Therefore, an on-going process of collaboration is necessary, enabling swift adjustments to changing environments.

4 Partnerships as New Form of Governance? – A Critical Conclusion

The core question of the proposed paper refers to the key factors determining forms of partnership collaboration in the field of sustainable urban tourism. The classification drafted above according to the 'reasons and objectives for partnership establishment' seems most appropriate to provide a better understanding of diverse forms of partnership cooperations in the area of sustainable urban tourism. The typology tries to shed light on the rationale for partnership establishment and could serve as basic pattern for the understanding of different partnerships' particularities.

The conclusions of this paper focus on opportunities and drawbacks of partnership cooperations as new and innovative form of local governance to deal with externalities in the field of urban sustainability. Generally speaking, the case studies decisively support the assumption that multi-stakeholder partnerships can be effective means for pursuing sustainability targets in urban tourism development. However, to make a pointed remark, basic impetus to form a partnership (independent of the motivation and objectives determining the type of cooperation) is that the actors are willing to get involved in a partnership to maximise their benefit individually as well as collectively. Normally, participation in such networks is based on mutual interest, exchange of resources, and commitment, although the relations between the participants do not have to be balanced. (Andersen/van Kempen 2003, p.80) Or as Jessop puts it: Partnerships as new form of local governance can be exploited as a "flanking, compensatory mechanism for the inadequacies of market mechanism" (Jessop, 2002, pp. 454f). This involves a range of merits and advantages of partnerships for public policy, community gains and the enhancement of collaborative practices themselves (e.g. enhancement of resource availability, increasing effectiveness and efficiency of individual organizations, integrating sectors, substituting a lack of formal institutional structures, confidence and trust among partners and other stakeholders).

Despite their merits, partnerships also present multiple problems and, as Andersen and van Kempen argue, clear disadvantages, mostly in terms of lacking democratic legitimacy (Andersen, van Kempen, 2003, p. 81):

- Firstly, many partnerships are not accessible to everybody or even completely closed. Only those who can add resources, including political power and/or legitimacy, will normally be let in.
- Secondly, in terms of internal risks, goals may conflict between partners in partnerships. Such contradictions can cause severe difficulties for the partnerships themselves and even more for the project in their hands.
- Thirdly, partnerships as a new form of local governance are only suitable for specific projects or policy fields, not for a holistic view of policy as partnership frameworks easily effect that people focus on their own interests; effects on other areas and stakeholder groups might easily be ignored. This aspect is of particular importance for urban sustainability for it has to be tackled in a

holistic, multidimensional way.

- Finally, it might be difficult to find a good balance between (new) partnerships (generally aimed at specific tasks and/or areas) and existing governmental bodies like local governments. Even if there is agreement on the existence of a partnership, contradictions and conflicts about responsibility, carrying out the tasks, evaluation etc. might still emerge.

As we mentioned at the beginning of this paper, 'public-private partnership' is frequently presented as a remedy against sub-efficient policy co-ordination and a new way of handling externalities and public goods on the urban level. To satisfy these expectations involvement of actors from multiple levels and sectors is required. Generally speaking, partnership co-operations can boost urban development but they rarely substitute missing structural prerequisites preventing social polarization and environmental decay.

Concluding, having all those drawbacks and open questions in mind, the following venues of further research in the field seem important:

1. Theoretical replication: Testing the SUT-partnership typology in different national and/or organisational settings would endorse the classification's validity.
2. Bridging levels of analysis: The evaluation of SUT-partnerships' development and success should be linked in more detail to the theoretical debate on the governance-boom, its opportunities and drawbacks, and its consequences.
3. Increasing the sample of SUT-Partnership study cases: Larger samples of partnerships could greatly facilitate generalisations and improve the knowledge of key factors determining SUT-partnership models.

References

Altvater, E. and Mahnkopf, B. (1999) Grenzen der Globalisierung. Ökonomie, Ökologie und Politik in der Weltgesellschaft. Westfälisches Dampfboot, Münster

Amin, A. (ed.) (1994) Post-Fordism. A reader, Blackwell, Oxford

Andersen, H. T. and van Kempen, R. (2003). New trends in urban policies in Europe: evidence from the Netherlands and Denmark. Cities 20, No. 2: 77–86.

Berg, van den, L., van der Borg, J. and van der Meer, J. (1995) Urban tourism: performance and strategies in eight European cities. Avebury, Aldershot, U.K. and Ashgate, Brookfield, Vt.

Borja, J. and Castells, M. (1997) Local and global. Management of cities in the information age. Earthscan, London

Brotchie, J. (ed.) (1995) Cities in competition. Productive and sustainable cities for the 21st century. Longman, Melbourne

Camagni, R., Capello, R. and Nijkamp, P. (1998) Towards sustainable city policy: An economy-environment technology nexus. Ecological Economics 24: 103–118

Castells, M. (1998) The information age. The rise of the network society. Blackwell, Cambridge

Davies, J. S. (2002) The Governance of urban regeneration: A critique of the 'Governing without Government' thesis. Public Administration 80, No. 2: 301–322

Dunford, M. and Kafkalas, G. (eds.) (1992) Cities and regions in the new Europe: the global-local interplay and spatial development strategies. Belhaven, London

Eisenschitz, A. and Gough, J. (1998) Theorising the state in local economic governance. Regional Studies 32.8: 759–768

Ekins, P. and Newby, L. (1998) Sustainable wealth creation at the local level in an age of globalization. Regional Studies 32.9: 863–871

European Commission (1999) Sustainable urban development in the European Union: A framework for action. Brussels

Fürst, D. (2001) Regional Governance zwischen Wohlfahrtsstaat und neo-liberaler Marktwirtschaft. (http://www.laum.uni-hannover.de/ilr/publ/fuerst/governan.pdf)

Gindl, M., Paskaleva-Shapira, K., Stuppäck, S., Schubert, U. and Wukovitsch, F. (2001) Pilot partnerships for sustainable tourism. Cross-country: Synthesis and indicative factors of success. (SUT-Governance Deliverable 4a)

Gindl, M., Paskaleva-Shapira, K., Stuppäck, S., Schubert, U. and Wukovitsch, F. (2002) 'Best Practice' partnerships for sustainable urban tourism. International Cross-Case Synthesis and Success Factors (SUT-Governance Deliverable no. 4) (http://sut.itas.fzk.de/)

Glaser, B. (1992) Basics of grounded theory analysis. Emerging vs forcing. Sociology Press, Mill Valley

Glaser, B. and Strauss, A. L. (1967) The discovery of grounded theory: Strategies for qualitative research. Aldine, Chicago

Goel, R., Brown, M. and Berry, L. (1991) Guidelines for successful transferring government-sponsored innovations. Research Policy 20, No. 2: 121–143

Gomez, P. Y., Korine, H., Masclef, O. (2002) Generating cooperative behaviour between the unacquainted: A case study of the Renault/Nissan Formation process. Paper presented at the 2nd annual conference of the European Academy of Management, Stockholm (http://www.aegionline.it/euram/paper/Gomez%20et%20alii.pdf)

Gomm, R., Hammersley, M. and Foster, P. (eds.) (2000) Case study methods. Sage Publications, Thousand Oaks, Ca.

Hamel, J. (1993) Case study methods. Sage Publications, Thousand Oaks, Ca.

Hart, M. (1999) Guide to sustainable community indicators, 2nd ed. Hart environmental data, North Andover, Mass.

ICLEI (1994) Charter of european cities and towns towards sustainability (The Aalborg Charter), (http://www.iclei.org/iclei/la21.htm)

Jessop, B. (2002) Liberalism, neoliberalism and urban governance: a state-theoretical perspective. Antipode 34, No. 3: 452–472

Kafkalas, G., Yiannakou, A. and Tasopoulou, A. (2002) Successful partnerships for sustainable urban tourism: unified framework model. SUT-Governance Deliverable no. 7 (http://sut.itas.fzk.de)

Kafkalas, G., Paskaleva-Shapira, K., Demetropulou, L. and Voultsaki, M. (eds.). Country framework assessment report. SUT-Governance Deliverable no. 2) (http://-sut.itas.fzk.de)

Kanter, R.M. (1994) Collaborative advantage: the art of alliances. Harvard Business Review 72, No. 4: 96–108

Kooiman, J. (2000) Governance. A social-political perspective. Paper presented at the conference Democratic and Participatory Governance: From Citizens to Holders?' EUI Florence, 14-16 September 2000

Law, C.M. (1993) Urban tourism: attracting visitors to large cities. Mansell, London

Le Galès, P. (1999) Is Political economy still relevant to the study of the culturalization of cities? European Urban and Regional Studies 6 (4): 293–302

Lipietz, A. (1992) Towards a new economic order – post-fordism, ecology and democracy. Polity Press, Cambridge

Love, L. L. and Riley, R. W. (2000) The state of qualitative tourism research. Annals of Tourism Research 27, 1: 164–187

Lowndes, V. and Skelcher, C. (1998) The dynamics of multi-organizational partnerships: an analysis of changing modes of governance. Public Administration 76: 313–333

Luke, T. (1994) Placing power/siting space: The politics of global and local in the New World Order. Environment and Planning D: Society and Space 12: 613–628

Mackintosh, M. (1992) Partnership: issues of policy and negotiation. Local Economy 7, 3: 210–225

Masberg, B. and Morales, N. (1999) A case analysis of strategies in ecotourism development. Aquatic Ecosystem Health and Management 2, 3: 289–300

Mayer, M. (1995) Post-Fordist city politics. In: Amin, A. (ed.), Post-Fordism – A reader. Blackwell, Oxford: 316–387

O'Riordan, T. (2001) Globalism, localism and identity. Fresh perspectives on the transition to sustainability. Earthscan Publication, London

Paskaleva-Shapira, K. (2000) Innovative partnerships for effective governance of sustainable urban tourism: Framework approach. SUT-Governance Deliverable no. 1. (http://sut.itas.fzk.de)

Paskaleva-Shapira, K. (2001) Innovative partnerships for sustainable urban tourism: Framework Approach and the European experience. SUT-Governance Conference presentation (http://sut.itas.fzk.de)

Paskaleva-Shapira, K. (2001) Promoting partnerships for effective governance of sustainable urban tourism: The case of Germany. SUT-Governance Conference presentation. (http://sut.itas.fzk.de/)

Paskaleva-Shapira, K. and Kaleynska, T. (2001) Bulgaria's tourism industry: a thriving and sustainable future? SUT-Governance journal publication (http://sut.itas.fzk.-de/)

Pearce, D. (1996) Tourism development. 3rd Edition. Longman, Harlow, Essex.

Priestley, G.K., Edwards, J.A. and Coccossis H. (eds.) (1996) Sustainable tourism? European experience. CAB International, Wallingford

Raffer, K. and Singer H.W. (2001) The economic north-south divide. Six decades of unequal development. Edward Elgar, Cheltenham

Rhodes, R.A.W. (1998) Transforming British government: the ECSR's Whitehall Programme, in: Paper to the Workshop 'Les hauts fonctionnaires et la politique'. IEP de Paris, 1998

Rhodes, R.A.W. (1996) The new governance: Governing without government. Political Studies 44: 652–667

Sassen, S. (1994) Cities in a world economy. Pine Forge Press, Thousand Oaks, Ca.

Sassen, S. (2001) The global city – New York, London, Tokio, second edition. Princeton University Press, Princeton, N.J.

Schneider, V. (2002) Regulatory governance and the modern organizational state: the place of regulation in contemporary state theory. Paper presented at the workshop 'The Politics of Regulation', Universitat Pompeu Fabra Barcelona, 2002

Schneidewind, U. (1997) Public private partnership. In: Steger, U., (ed.) Handbuch des integrierten Umweltmanagements, Oldenbourg, München, Wien

Shaw, G. and Williams, A. (1994) Critical issues in tourism: A geographical perspective. Blackwell, Oxford

Stake, R.E. (1995) The art of case study research. Sage Publications, Thousand Oaks, Ca.

Steger, U. (ed.) (1997) Handbuch des integrierten Umweltmanagements, Oldenbourg, Munich, Vienna

Swyngedouw, E.A. (1992) The mammon quest. 'Glocalisation', interspatial competition and the monetary order: the construction of new scales. In: Dunford, M. and Kafkalas, G. (ed.) Cities and regions in the new Europe: the global-local interplay and spatial development strategies. Belhaven, London: 39–67

Theys, J. (2000) Environmental governance. From innovation to powerlessness. EUI Florence, 2000.

Willis, K.G., Turner, R.K. and Bateman, I.J. (eds.) (2001) Urban planning and management. Edward Elgar Publishing, Northampton

Industry Cluster Concepts in Innovation Policy: A Comparison of U.S. and Latin American Experience

Edward Feser*

1 Introduction

The increasing knowledge-intensity of production and the progressive elimination of barriers to trade have led many to conclude that a strong base of science, technology, and innovation is essential for sustained economic prosperity (Mytelka and Farinelli, 2000). Advanced industrialized countries are responding to increasingly open markets by seeking competitive advantage in general knowledge infrastructure: universities and colleges, public and private laboratories, educated workers, advanced physical infrastructure, and comparatively stable social, political, and market institutions. Interest in innovation is also heightened by fears of an emerging "two-tiered economy," that two sectors will come to dominate long-term employment growth prospects in industrialized countries: high skilled technology-intensive activities that are dependent on advanced knowledge infrastructure and low-skilled basic consumer services that serve immediate local market needs (Mowery, 2001). In developing or transition countries, fears of falling further behind the highly industrialized world as well as optimism borne of widely publicized examples of high technology success provide the principal motivation to designing ways to boost innovation and technology-related activity.

In this context, of growing interest are the phenomena of high technology industry clusters and their potential value as an innovation policy focus. Mainstream economic theory argues that technology-related activity often agglomerates in specific regions because knowledge spillovers are localized (Glaeser, 2000). Knowledge spillovers – the primary engine in the most recent theories of long-run economic growth – are the ability of economic agents to utilize a new technology or innovation without fully compensating its original source or owner (Grossman and Helpman, 1991). Innovations initially occur in companies, universities, and laboratories located in specific places. The subsequent spread (or diffusion) of such innovations, as well as the spillovers they generate, may occur more readily among economic actors located in close proximity, either because the innovation is tacit in nature or because its successful utilization requires an element of hands-on learning-by-doing. Increasing returns to innovation, coupled with a localized diffusion effect, imply that technology-oriented activity and R&D are likely to concentrate geographically. Technology businesses lo-

*Department of Urban and Regional Planning, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA.

cate near other high tech companies and R&D performers in order to share in the spillovers, further enhancing the attractiveness of the growing cluster for still more high tech enterprises. The cluster may then expand through a process of cumulative advance.

The emergence of new growth theory more or less coincided with Michael Porter's (1990) research on clustering and national competitiveness as well as an exploding literature on industrial districts. Early on, there was comparatively little cross-fertilization of ideas from these perspectives. However, they all emphasized the tendency toward localization of economic activity and the critical role of knowledge spillovers (albeit described differently by each perspective). The concurrent development of the literatures, all offering varying perspectives on a similar story, contributed strongly to the rise of industry clusters as a concept in development policy debates.

Such debates have been bolstered by stylistic qualitative analyses in highly industrialized economies that suggest that a combination of geographically co-located private sector producers of R&D, related manufacturing and services industries, linked or related suppliers and producer services providers, leading research universities and teaching institutions, and government sponsored labs and technology programs can combine to create powerful spatial clusters of technology-related activity that continue to expand through initial market leadership (often called "first-mover effects") and economies of scale (Saxenian, 1994, Porter, 1990, 1998, 2000, den Hertog et al., 2001a, b). Well-known examples in the United States are California's Silicon Valley and Boston's Route 128 (in information technology and biotechnology), greater Seattle (in software and aircraft), and North Carolina's Research Triangle region (in electronics, pharmaceuticals, and biotechnology). Such clusters have contributed to substantial increases in the local economic prosperity while also supplying the innovations that drive national and, in some cases, global economic growth. Such clusters are not restricted to the U.S. or other advanced industrialized countries, although they tend to be smaller and have much less depth in less developed countries (e.g. see den Hertog et al., 2001b, Melo, 2001a, Chairatana and Vorrakitpokartorn, 2001, Voyer, 1997b). Recent studies of Latin America have identified innovation clusters of differing varieties and size in Argentina, Brazil, Costa Rica, Cuba, Peru, and Mexico (Quandt, 1997, Voyer, 1997a, Altenburg and Meyer-Stamer, 1999, Bortagaray and Tiffin, 2000).

An important issue is what clusters imply for the design and implementation of innovation policy, particularly in newly industrializing countries and lagging regions in developed ones where technology-intensive activity and basic knowledge infrastructure are limited. Innovation policy constitutes strategies designed to build basic and applied research capabilities, raise the rate of advanced technology adoption and product innovation among home country firms, and generally increase the complement of higher wage knowledge- and technology-intensive industries in a country or region.[1]

[1] Temple (1998) identifies five determinants of technological change that may be the focus of innovation policy: the generation of new knowledge; the translation of new or existing knowledge into products and processes; the diffusion of innovation; the exchange of knowledge-intensive goods and services; and the absorption of knowledge or learning. All of the processes are subject to market failure. Therefore, the more knowledge- intensive an economy becomes, the more important institutional (i.e. policy) mechanisms for

The strategies might include, among others, the provision of R&D subsidies and incentives, the development of university research competencies, the improvement of basic education, the supply of training, the promotion of business development services, the encouragement of firm networks, the provision of industrial extension, the facilitation of technology transfer, and the targeting of public sector procurement (Leyden and Link, 1992, Malecki, 1997, Gambardella and Malerba, 1999, Tidd and Brocklehurst, 1999, Conceição et al., 1998, Geroski, 1990).

Many of those same interventions have been described, at one time or another, as industry cluster policies (Jacobs and de Jong, 1992, Jacobs and de Man, 1996, Rosenfeld, 1997, Enright, 2001, Rosenfeld, 2001). What unique insights, then, does the cluster concept bring to the innovation policy debate? Is an industry cluster policy merely the application of a conventional development initiative, such as an R&D incentive or procurement strategy, to a geographically concentrated group of firms? Does the process of clustering, as opposed to the phenomena of clusters, imply a specific and unique kind of policy intervention? Have governments formed an alternative model of intervention that utilizes findings from research on clusters but does not force them to pick favorite sectors, research concentrations, or regions? Those fundamental questions are raised from the explosion of literature on clustering and closely related sister concepts and perspectives such as learning regions, innovation systems, networks, districts, and innovative milieux.

This paper does not attempt a general discussion of the wide range of definitions, views, and theories of industry clusters. Such generalized reviews are already numerous.[2] Instead, it focuses on the empirical question of how cluster concepts are being utilized in economic development policy making, especially related but not limited to innovation, at least as could be determined with a review of secondary sources, government documents, and expert opinion. The focus is on Latin America and the United States, with one important aim being to consider how different institutional frameworks and stages of development link to differences in the way cluster concepts are applied in the policy arena. While the Latin American and U.S. cases are examples of the developing and developed economy contexts, respectively, I make no claim that they are representative of those contexts.

2 Industry Clusters and Innovation Policy Making

Examining how governments around the world are actually invoking cluster concepts in economic development planning and policy making, especially with regard to innovation, is no easy task. The active or at least nominal use of cluster ideas in policy making at all levels – local, regional, national, and international – continues to grow. Clusters have been debated at the national and regional levels in the U.S., Canada, Australia, New Zealand and most of the European Union since the mid 1990s (Roe-

resolving failures will be.

[2]The most relevant theoretical literatures include endogenous growth theory, new industrial districts, technology districts and technopoles, innovative milieux, industrial location and agglomeration economies, strategic management and industrial organization, and innovation systems. See McCann (1995), Feser (1998a, b), Bergman and Feser (1999), and Moulaert and Sekia (1999), and Gordon and McCann (2000).

landt and den Hertog, 1999, den Hertog, Bergman et al., 2001b). Enright (2001) claims that cluster initiatives have been pursued in ten countries in Latin America, as well as in Malaysia, Singapore, Morocco, South Africa and Senegal.

International organizations have been particularly important players in the evaluation, dissemination, and utilization of cluster ideas. The OECD, the World Bank, UNIDO and UNCTAD have all been engaged in considering, developing, funding, and/or evaluating cluster or networking policies of one form or another. For example, the World Bank organized a workshop of cluster practitioners in Chihuahua, Mexico, in November 1997, which helped inspire the founding of The Competitiveness Institute, a non-profit international association of practitioners that aims to disseminate best practices via a website, newsletter and annual conference. The OECD has considered clusters as part of its National Innovation Systems (NIS) project since 1996, an effort that has resulted in several international workshops and two edited volumes of best practice (Roelandt and den Hertog, 1999, den Hertog, Bergman et al., 2001b). UNIDO considers clusters part of its small and medium-sized enterprises (SME) programs (Ceglie and Dini, 1999, Fisher and Reuber, 2000, Russo, et al., 2000, UNIDO, 2001, Nadvi, 1995). UNCTAD has also focused on clustering and networking as modes of competition for small firms (UNCTAD, 1998).

A simple Internet search on the phrase "industry clusters" turned up over 10,000 references at the time of this writing, a huge figure considering that common estimates of the share of web content that current search engines are capable of tapping is about 15 percent (Reich, 2002). Yet at least some efforts are being made to understand broader trends in how the concept is being interpreted and applied. Claas van der Linde and Michael Porter have assembled a collection of over 350 studies that examine some 700 clusters in roughly fifty countries. The majority of those studies have been conducted or commissioned by public agencies interested in applying clusters to policy. Classifying studies of clusters is, however, very different from documenting the utilization of clusters in policy making, since industry clusters have been the subject of far more study than practical action. More apropos in the present context is a recent study by Sølvell et al. (2003).

A major problem with efforts to describe "cluster policy" is that many types of development interventions are targeted to specific sectors, regions, or both, and thus could be loosely interpreted as cluster-oriented strategies. For example, the establishment of research parks and technopoles could be considered cluster strategies, even though many such efforts around the world preceded the modern cluster literature. In the United States, North Carolina's development of Research Triangle Park in the 1950s, which subsequently became the anchor of substantial information technology and bioscience clusters, is sometimes viewed (and cited favorably) as a cluster strategy (Rosenfeld, 2001), even though it was initially designed as an industry recruitment tool. Melo (2001b), referencing Quandt (1999), describes Brazil's establishment of thirteen innovation centers in 1982, as well as a subsequent science park program in 1984, as among the earliest cluster strategies in Latin America. Business incubators, industrial parks, targeted recruitment, enterprise zones, foreign trade zones, and a large variety of other common economic development interventions could similarly be assessed as cluster policy if they aim to foster growth in specific industries or regions.

As a way of limiting the scope of the analysis, this section focuses strictly on U.S. and Latin American trends in the explicit use of cluster ideas. Only efforts that directly reference the concept of clusters, even if they do so in only a nominal way, are therefore considered. The assessment is based on existing literature, Internet searches, and personal communication with experts and policy makers. While it is invariably non-exhaustive, hopefully it is representative. The aim is to gain an empirical sense of how public officials are drawing on the notion of clusters. Given the highly malleable nature of the concept, it is to be expected that the utilization of clusters in policy making tends to reflect industry characteristics and mix, views of economic development, prevailing institutional frameworks, and political and economic constraints in the jurisdiction at hand. More specifically, it appears the most common use of cluster ideas is as a way to organize and undertake strategic planning exercises that yield a flexible set of policy options, a clear target group of beneficiaries, and a logical set of private sector partners to planned interventions.

2.1 Clusters and Innovation Policy in Latin America

In Latin America, views of innovation are influenced by a general debate about industrial policy. In a recent survey of economic policies in the region, Melo (2001a) documents two phases in the reforms that have followed the import substitution era. In the first phase – from roughly the late 1980s to the mid 1990s – Latin American countries sought to implement basic structural reforms related to export trade, privatization, domestic market liberalization, and regulation. At the same time, they curtailed explicit (sector targeted) industrial policies. The logic was that government intervention in liberal market economies is necessarily very modest and that industrial policy is generally prone to distortion and corruption. Yet Melo finds that by the mid-1990s many Latin American states had already begun to abandon that strictly hands-off philosophy in favor of explicit public sector strategies aimed at enhancing the competitiveness of particular sectors, value chains, and firms. This second phase, which is ongoing and still without definitive results, reflects a view of government intervention that is more nuanced, particularly as it pertains to technology. Latin American countries are recognizing that global competitiveness ultimately implies continuous learning and innovation, processes on which the public sector might exert considerable positive influence through its role as catalyst, source of demand, and supporter of research, basic education, and training. It is in this context that there is growing interest in Latin America in the phenomena of industry clusters and their potential value as an innovation policy focus.

Early interest in clusters and clustering focused overwhelmingly on advanced industrialized countries. That is no longer the case. Clusters, districts, and networks are now being systematically studied all over the developing world. Latin America, in particular, has been the subject of considerable research, with Brazil and Mexico receiving most of the attention. Among the clusters (or districts) studied in the former are leather shoe producers in the Sinos Valley (Nadvi, 1995, Schmitz, 1995a, b, 1999), various high technology sectors in Campinas (Quandt, 1997), the wood furniture industry of Ceará (Tendler and Amorim, 1996), and the textiles and clothing,

metal engineering and electromechanical, and ceramic tiles clusters in Santa Catarina (Meyer-Stamer, 1998). Rabellotti (1999) analyzes the footwear sector in Guadalajara, Mexico, and Visser (1999) describes the results of a case-control study of clustered and dispersed garment producers in Lima, Peru. None of the aforementioned studies are focused specifically on innovation or technology. In contrast, Bortagaray and Tiffin (2000) attempt a systematic identification of innovation clusters across the region, concluding that while firms in clusters seem to grow faster and generate more profits than those outside of clusters, no Latin American innovation clusters can be reasonably described as mature in the sense of a Silicon Valley. The authors do identify a number of significant developing and potential innovation clusters, including two in Argentina, twenty-seven in Brazil, two in Costa Rica (both in San Jose), one in Cuba (biotechnology), six in Mexico, and one in Uruguay (wine). Unsurprisingly, most of the clusters are in heavily urbanized areas. A lack of investment capital, weak inter-firm and inter-institutional networking, and an absence of adequate business development services are cited as the primary impediments to the further development of innovation clusters in the region.

A review of government documents and web sites gives a sense of how cluster ideas are either informing or being incorporated into economic policy in the region, both within and outside the area of innovation (see Table 1). Immediately noticeable is the breadth of interventions that Latin American governments themselves describe as cluster policy. They range from marketing and business networking to targeted export assistance, infrastructure development, and training. Overall, public agencies in the region tend to be invoking or actively applying cluster concepts mainly in three broad policy areas: export promotion and attraction of inward investment, value chain integration, and networking/SME policy. Those emphases reflect views of what is appropriate given the current industrial structure and stage of development in much of Latin America, including the continued dependence on the location of the manufacturing concerns of large multinational companies, a desire to diversify existing industry by filling out supply chains, and a predominance of uncompetitive small and medium-sized producers (particularly in peripheral and lagging areas). Notable is the lack of many innovation programs based on cluster concepts. That does not mean that Latin American governments are not targeting S&T investments to specific sectors, research competencies, and/or regions, but rather that there is only modest evidence that cluster ideas are being used explicitly to guide such initiatives. Current cluster interventions in the region seem to be focused on traditional sectors for the most part.

So exactly what value-added are clusters bringing to economic policy making in the Latin American region, even if their role in innovation policy has been limited? The answer appears to have less to do with the identification of specific interventions than with the defense of general approaches and setting of strategic priorities. First, as mentioned above, many Latin American governments are attempting to identify the right balance between the implementation of free market structural policy and activist (often local and regional) strategies designed to promote the competitiveness of strategic sectors and potential strengths in science and innovation. The region's move to open its markets to international competition while dismantling the protection of inefficient domestic industries has not yielded the gains initially anticipated (Melo,

Table 1: Latin America: examples of cluster policy initiatives by country

Country	Source	Policy Type	Example
MEXICO	Online information and document (SE, matrix by State)	Promotion of bus. associations	Creation of the Mexican Council of Cotton Production (Durango).
		Export promotion	Establishment of the Export Development Center (Chiapas).
		Infrastructure development	Construction of a fire station (Hidalgo).
		Direct investment	Development of a silver processing/refinement plant (Guerrero).
		Training	Training workshop for lime producers (Colima).
		Marketing	Trade shows, promotion, marketing, etc.
	Altenburg and Meyer-Stamer (1999)	Networking	Empresas Integradoras Program: private corporations owned by groups of small firms that provide demand aggregation, purchasing, and marketing services. Firms are lured to participate with tax deductions, loans, and training programs.
	Melo (2001a), Altenburg and Meyer-Stamer (1999)	Production chain integration	Eight broad sectors targeted: high-tech industries, automobiles, light manufacturing, petrochemicals, mining, agribusiness, forestry, and public sector. Aims to coordinate private-sector efforts with the following goals: formation of clusters; rationalization of internal corridors; identify supplier opportunities and customer technical requirements (through compilation of information sources such as directories and organization of supplier and export promotion conventions; subcontracting exchange schemes or SES).
PARAGUAY	Online document (STP)	Export promotion	Establishment of international production standards (e.g. animal feed processing (supporting grains cluster in Itapua y Alto Parana).
		Promotion of bus. association	Establishment of food production committee (grains cluster in Itapua y Alto Parana).
		Investment catalyst	Promotion of public-private partnerships for vegetable production.
		Support for suppliers/producers	Support existing agriculture cooperatives through credit (Caaguazu Dept.).
		Marketing	Carry on studies to identify potential export products (vegetable cluster).
		Infrastructure development	Development of highways to transport exported products (oranges in Itapu).
		Applied research, extension	Enhance productivity of sector by introducing new varieties of cotton.
		Training	Training for metal mechanics sector; promotion of tech. specialization, etc.
		Other strategies suggested	Reduce production costs; set up a revitalization program for the restructuring of the sector; establishment of production of a type of wood
	Farinelli and Kluzer (1998)	SME modernization, promotion, networking	EC supported establishment (1993) of small firm consortia for small firms in the furniture, textiles/clothing and leather goods industries; implemented by Ministry of Industry and Trade (Paraguayan Support Centre for Enterprises).

Table 1: Latin America: examples of cluster policy initiatives by country – cont.

Country	Source	Policy Type	Example
HONDURAS	Online information (UNIDO), Altenburg and Meyer-Stamer (1999)	Promotion of business associations	Establishment of networks of firms with common needs, like joint establishment of shops to sell finished products, sharing large orders of products, etc.
		Support for suppliers/producers	A UNIDO program to provide technical support to help SME obtain credit. Establishment of a cluster promotion center (CERTEC).
		Production chain integration	Promote vertical networks between small and large firms.
		Training	Training for metal workers (Tegucigalpa); training for cluster/network "brokers" to diffuse networks.
NICARAGUA	Online information (UNIDO)	Promotion of business associations	Establishment of networks of firms with common needs, like the handicraft hammock production sector.
		Export promotion	Technical assistance to standardize production and pricing systems among a network of firms in the same industry/sector.
		Training	Training for workers as well as cluster/network "brokers" to diffuse networks.
		Regulation assistance	Technical assistance in legal issues to formalize economic activity.
		Brokering, catalyst	Coordination between UNIDO and the National Institute for SMEs.
		Production chain integration	Promote firms' integration along production chains.
BRAZIL	Online information (SEBRAE) Melo (2001a)	Training	Via Brazilian Support Service of Micro and Small Enterprises (SEBRAE). Training for cluster promoters.
		SME support	Technological Support Program for Micro- and Small-size Enterprises (PATME): financing of product and process improvements and equipment upgrades; provision of training; assistance with quality control. SEBRAEtec grants for the purchase of business development services from consultants, universities, and technical institutes.
COSTA RICA	Document (CLACDS-INCAE)	Training	Courses offered by the National Institute for Learning according to the needs of firms.

Table 1: Latin America: examples of cluster policy initiatives by country – cont.

Country	Source	Policy Type	Example
URUGUAY	Document (EU); Farinelli and Kluzer (1998)	Production support and export promotion of SMEs; modernization	EC supported promotion of SME integration in order to share large contracts, as well as manage export activities, etc. (furniture industry); implemented by Comisec
		Promotion of business association, networking	Establish a common strategic vision of the industrial restructuring needed in the textile and clothing industries was agreed among the main trade associations.
COLOMBIA	Online information (MINDESA); Melo (2001a)	Production chain integration, expansion	Targets existing and potential chains. Existing target chains include exporters facing stiff external competition (textiles/apparel, leather goods/footwear, automotive cars/parts, sugar cane/products, oleaginous seeds/oil/soap, aquaculture, tuna, toiletries and cleaning products, and potatoes) as well as chains with significant internal trade and linkages (petrochemicals, plastics, steel, electronics, among others). Potential include IT, biotech, chemistry, and communications, among others. Provision of specialized and general support infrastructure; preparation of "sector competitiveness agreements"; education and training, fostering dialogue between private and public sectors.
VENEZUELA	Online information (MPD)	Production chain integration	
CHILE	Online information (CORFO)	Recruitment, inward investment	Marketing, diffusion of information, financial support for investment studies, etc. (by region)
		Extension	Training for firms on internet usage (on-line).
		Location incentives	Land incentives, wage subsidies; credit support; bonuses for project investment to banks, co-financing of risk studies, etc. (Arica)
		Strategic planning	Comprehensive development plan for the Lota region; includes a wide variety of policies.
	Altenburg and Meyer-Stamer (1999)	Networking	Proyectos de Fomento (PROFOs): Contracts between a small group of small firms' public or private agency network brokers that provide resources for market research, industry studies, and participation in trade fairs.

2001a). The result is a search for interventions that will address the shortfall while also meeting the approval of multinational lending institutions, key trading partners such as the United States and Canada, and international investors. Industry clusters are widely viewed by both public and corporate officials in the developed world as a key feature of international competitiveness, i.e. cluster promotion efforts have attained a level of legitimacy as market-friendly industrial policy that other (differently labeled but sometimes quite similar) perspectives have not. Thus while clusters may hold out the promise of a substantive route toward a more activist competitiveness strategy that does not threaten the region's continued shift toward free markets, it is also significant that they are viewed favorably from a symbolic perspective.

Second, the cluster concept is persuading some Latin American governments to place more emphasis on the diagnosis of problems and prescription of interventions for existing industries, and to avoid focusing exclusively on the attraction of inward investment. Knorringa and Meyer-Stamer (1998) note that industrial diversification continues to dominate economic development thinking in many developed countries, often to the detriment of existing businesses. They argue that "... it is unrealistic to expect local and regional policy-makers to embark enthusiastically on a cluster strengthening policy" (p. 18). They believe that governments are more inclined to try to attract major foreign assemblers in new sectors, even if the probability of success is low, in order to avoid locking into a narrow set of specializations. The result is neglect of the concerns of local businesses and the potential to expand the existing industrial base. Balanced attention to the needs of existing industry is especially valuable even aside from the growth prospects of that industry because it often exposes policy reforms and legitimate investments in infrastructure, education, and other basic factors that could improve the general business climate. The evidence suggests that industry cluster concepts are providing a useful framework for Latin American governments to think about how to address weaknesses and threats to the competitiveness of existing industry and to encourage corporate interests to participate – and even drive – the process. This utilization of clusters as a strategic planning and organizing device in Latin America parallels the experience in many developed countries.

2.2 Clusters and Economic Policy in the United States

As in Latin America, the utilization of cluster concepts in economic policy making in the United States reflects local economic conditions as well as views of appropriate industrial policy. In the U.S., since there is no explicit domestic economic development strategy at the federal level, industry cluster strategies have chiefly been a concern of states, regions, and metropolitan areas. Four different trends can be detected in U.S. cluster practice, some of which are represented in the selected illustrative examples in Table 2.

First, economic development at the state level in the U.S. remains dominated by business recruitment strategies coupled with the provision of location incentives in the form of direct grants, tax credits, and loans. Many states have therefore used industry clusters primarily as a means of promotion and marketing, often of highly desired technology-oriented sectors such as information technology, electronics and biotech-

Table 2: United States: selected examples of cluster policy initiatives by state

State	Background	Intervention	Source
Arizona	Development of Arizona Strategic Plan for Economic Development in 1992, resulting in adoption of state cluster strategy. Renamed Governor's Strategic Partnership for Economic Development (GSPED); administered by Arizona Department of Commerce. Public-private partnerships represent each cluster.	Analysis. Strategic planning. Targeted marketing for recruitment, tax policy (incentives).	Waits (1992), Morfessis (1994), Vieh (2002)
Connecticut	In 1998, established Connecticut's Industry Cluster Initiative under the Department of Economic and Community Development. Also established Governor's Council on Economic Competitiveness and Technology to monitor cluster progress.	Strategic planning. Creation of lead organizations. Biotech facilities fund. Workforce training in metalworking.	CDECD (2001)
Iowa	Legislation in December 2000 creates the New Economy Council to develop strategic planning and mobilize public and private resources in three clusters: life sciences, advanced manufacturing, and information solutions.	Analysis. Strategic Planning. Marketing. Workforce development (planned). Telecommunications Infrastructure (planned). Establishment of Plant Sciences center at U. of Iowa.	Iowa Governor's Office
Kentucky	Office of the New Economy established five research area clusters as priorities for development.	Strategic planning. Coordinated effort to obtain federal research dollars. Grants and loans for high tech industries.	KIC (2002)
Massachusetts	Massachusetts Technology Collaborative, an independent organization (organized in present form in 1994), coordinates technology policy for the state. Industry cluster concepts used in ongoing economic analysis and strategic planning.	Strategic planning. Economic analysis (tracking of tech sector growth).	MTC (2001)
Mississippi	In 2000, private sector commissioned study of communications and information technology cluster by Michael Porter. Subsequently, state funded follow-up studies of automotive and plastics/polymers industries. Effort managed by Mississippi Development Authority.	Analysis. Strategic planning. Establishment of cluster organizations. Workforce development planning (in progress).	RTS (2001)

nology, but also of advanced manufacturing sectors that promise substantial wage increases. For example, in the U.S. south, Mississippi, Alabama, and South Carolina have invoked clusters as a rubric for identifying and recruiting vehicle industry suppliers. In many states, the term cluster is synonymous with "industry" and economic development practice is little different in any substantive way.

Second, as is the case in some Latin American countries, clusters are commonly used as an organizational and analytical device for implementing a model of collaborative strategic planning and public-private engagement. Arizona's cluster initiatives are the earliest example of this trend (Ffowcs-Williams, 2000). In the early 1990s, Arizona used basic descriptive techniques to identify nine clusters around which it set up advisory groups, working groups, and town meetings to develop growth strategies (Rosenfeld, 2001). Private sector "buy-in" is a major feature of the state's approach, in contrast to the usual top-down implementation model characteristic of most states' development efforts. At the same time, Arizona has tended to apply a standard set of policy interventions to the clusters, some of which lack a strong central logic. An example is the state's "senior living" cluster. The value for the state seems to be the way cluster concepts are used to motivate the coordinated effort of multiple public agencies and private sector stakeholders and not as a means to design unique policy interventions. The utilization of cluster concepts in California and North Carolina provide similar examples (Feser and Luger, 2003), while a recent survey of California economic development practitioners by that state's Trade and Commerce Agency found that the cluster concept is being used mainly as part of a broader effort toward comprehensive economic development planning, interagency collaboration, and public-private partnership building: "a systems change is underway in how people conceive of and perceive economic development. To stay competitive in this 'global' information economy, better economic information is needed. The fast pace of change and global competition make timely, accurate information critical. The industry cluster analytical process, regional outlook and regional collaboration are tools assisting in this knowledge gain process" (Kawahara, et al., 2000, p. 8).

Much of the power of clusters as a strategic planning device derives from the traction the concept has in the corporate sector. Thus economic policy makers are able to gain more legitimacy with business leaders when using the language and logic of clusters than with more conventional sector-based approaches and esoteric development theories. This legitimizing function of the cluster concept compares with the Latin American case where the concern is with convincing lenders and multinational financial organizations of the appropriateness of certain industrial policies that once might have been viewed as protectionist but now are keyed toward enhancing local competitiveness. Either way, governments are using cluster ideas extremely effectively to bridge the difficult divide between public and corporate imperatives.

Third, the most recent trend in the United States is the utilization of clusters for the implementation of workforce development strategies, an approach almost entirely absent in Latin America. Again, the chief motivator is not extant theories of clustering, but rather pressing public policy issues coupled with the general flexibility of the cluster concept. Welfare reform, the Workforce Investment Act of 1998 (WIA), and the recent recession (resulting in considerable worker displacements and associated

re-training needs) have forced state and local agencies to seek ways to better target training, both geographically and by sector. WIA requires states to streamline workforce development programs by better coordinating the delivery of different kinds of services (e.g. job search and training). In most cases, cluster analysis serves as an analytical tool for detecting the occupational and training requirements of projected growth industries (based on a labor pooling argument), though it may also provide a general framework for strategic planning as noted above. The application of clusters to workforce development issues also reflects an increase in the use of cluster concepts by non-traditional economic development organizations, such as universities and community colleges.

Finally, many states and larger regions are using applied cluster analysis to identify localized concentrations of technology-related industry and research activity, so-called innovation or technology clusters. Such efforts usually motivate the design of innovation policy, although examples of sizable investments in detected clusters are few and specific interventions are largely conventional. One of the reasons for that is that in many states, high tech activity remains modest (at least compared to major concentrations such as Silicon Valley, Austin, and Boston). Therefore, the findings of cluster analyses are often too ambiguous to justify ambitious cluster building efforts. Moreover, the competing interests of various sectors and constituencies in the U.S. (as in most other countries) almost always mean that development resources must be spread relatively thinly across sectors and regions. The result is that clusters again become more of a strategic planning device, helping to reveal strengths and weaknesses facing local businesses and potential interventions that could improve the general business climate, than a rigid guide or model of development.

It is important to realize that the U.S. case is not reflective of the entire advanced industrialized world. Indeed, this should be obvious since by now it should be clear that the institutional and policy context in which cluster initiatives are pursued is central to their design and implementation (Sølvell et al., 2003). In Europe, for example, the experience with clusters reflects the much stronger historical role of national governments in development policy than in the U.S., continuing realignment of national policies in the face of European integration, and the heavy influence of research on famous small firm clusters/districts in Europe itself. More centralized development policy – at least in smaller countries such as the Netherlands, Austria, Denmark, and Finland – has produced cluster initiatives that have been more sustained and of somewhat greater sophistication than in the U.S. Integration is forcing European governments to reorient conventional macro policy. Increasingly, the jurisdictional expansion of the EU and the influence of broader global economic forces are harmonizing the general factors that influence a nation's relative business climate (what are often called "framework conditions," such as inflation, regulation, and product standards). Both national and local/regional governments are therefore focusing on local factors that remain under their control, including research competencies and institutions, educational institutions, financing institutions (e.g. venture capital organizations), and general infrastructure (Dalsgaard, 2001). In some countries, clusters and cluster analysis (or "cluster mapping") has become a means of achieving that policy reorientation. The institutional landscape in which economic development is pursued in Europe remains

complex despite integration. Viesti (2002) provides a discussion of the complexity of coordinating local, regional, national and international (e.g. EU) development policies, particularly those aimed at promoting local externalities. Cluster concepts and related theories are seen as one useful source of guidance.

Unsurprisingly, the literature on industrial districts and flexible specialization has been more influential in Europe than elsewhere. In the 1990s, several European countries undertook substantial experiments in the use of business networking schemes as a mechanism for encouraging collaborative competition and learning economies among small firms (Helmsing, 2001, UNCTAD, 2002). That experience has subsequently influenced the programs of multinational organizations such as UNIDO and the World Bank, which now are active in many LDCs, including Latin America. Indeed, most networking schemes in Latin America were initially pushed by international agencies and not national or regional governments. The findings of subsequent evaluations of business networking initiatives in Europe have been disappointing, with the chief problem being that few firms opt to remain in formalized networks after initial public sector incentives are exhausted (Hallberg, 1999, Lagendijk, 1999, 2000).

3 Discussion

So what can be said in the way of general trends, as well as similarities and differences between the U.S. and Latin American cases? First, a scan of initiatives in both Latin America and the U.S. finds no dominant type of policy intervention that is being used to establish or expand technology-based industry clusters or substantially influence innovation policy, aside from targeting perceived technology strengths or potentials. From the perspective of many public officials, what appears to make a policy a "cluster policy" is not the economic behavior the initiative is trying to influence but rather the target of the intervention as a loosely identified set of related companies and institutions. From this perspective, deregulation and workforce training may be just as much "cluster policies" as establishing business networks or other schemes to boost interfirm cooperation. In Latin America, traditional sectors are easily the most common target of interest, while both high tech and traditional industries have received attention in the United States. There is also some bias toward focusing on SMEs in Latin America, and similarly in the U.S. at the sub-state level.

Second, public officials are using the cluster concept liberally to identify and motivate the participation of key "partners" in the policy process and to legitimate general public sector intervention in the development arena. In the U.S., a focus on clusters is being used to secure corporate support and assistance with policy design (and thus to facilitate a general move toward policy making via public-private partnerships). The modern notion of clusters has its genesis in strategic management theory (e.g. Porter, 1990), a body of concepts that many business people find much more understandable and compelling than academic theories of the firm or the development process. In Latin America, industrial policy as cluster policy finds sanction with key trading partners and lending agencies concerned with promoting a shift toward free markets. Given a world in which industrial policy carries the taint of the protectionist strategies of the past, it appears to be easier to make the case that cluster policy is about

competitiveness, even if the specific interventions retain a certain protectionist flavor.

Third, applied cluster analysis – the detection of the presence of clusters and/or the strengths, weaknesses, and opportunities facing clustered enterprises – probably accounts for most of the current policy effort associated with cluster concepts. In most instances, governments in the U.S. and Latin America are not following up cluster analyses with major cluster building or expansion initiatives reflected in distinct policy changes. Rather, they are using the analyses to identify various problems facing current local or future businesses that could be addressed by interventions of relatively limited scope. The findings of cluster analyses are also occasionally being used to motivate support for general shifts in strategy, such as improvements in education or the provision of advanced infrastructure, that are increasingly viewed as key preconditions for the competitive success of industry in general (not just clusters). At the same time, in other cases the pursuit of clusters may be distracting policy attention from more basic needs. The latter is a particular concern in Latin America, where technology-oriented concepts like clusters can prove much more glamorous to pursue than very necessary basic infrastructure anti-poverty programs (Melo, 2001b).

Finally, despite a major policy implication of Porter's concept of clusters that higher rates of innovation and growth can be achieved by actively nurturing localized concentrations of linked businesses in selected promising industries rather than seeking a more diversified sectoral and spatial distribution of economic activity, it is surprisingly hard to find examples of governments in either Latin America or the U.S. (whether state, regional or municipal) making substantial investments in specific clusters to the exclusion of other local businesses and industries. It is the tendency of economic activity in general – and innovative and knowledge- intensive activity in particular – to concentrate functionally and geographically that suggests to so many that an effective S&T strategy might be to target specific groups of related high tech sectors in specific regions for development attention. The goal is to replicate elements of successful innovation clusters from around the world. It is as a result of that interest that various typologies of clusters and associated guides for how to expand them have been developed. The implication is a model of policy design, implementation, and evaluation that looks like the following: 1) identify or "map" groups of sectors that qualify, by some definition, as clusters; 2) assess strengths and weaknesses (or impediments to growth) in said clusters; 3) prescribe and implement policies to rectify weaknesses, maximize strengths, and spur growth; and 4) evaluate policies for overall impact on cluster expansion and performance. Usually left unsaid is that some sectors lose while others win, but the implication is clear.

Porter's descriptive theory of the determinants of competitiveness came to be interpreted as a narrow model of how to build localized clusters in specific regions. In fact, a careful reading shows that Porter set up a number of intriguing hypotheses that stand apart from the question of geography: namely, the links between sectoral economic growth, on the one hand, and sophisticated home demand, rivalrous yet cooperative competition, and the presence of related and supporting industries, on the other. Porter suggested that many of the industries characterized by such features tend to be localized in specific regions. He did not offer a systematic explanation of causes of localization, grounded in any theory of industry location or externalities, but essen-

tially an empirical observation of a tendency toward spatial co-location of competitive firms. This point is important because merely the observation was sufficient to set in motion a conviction among many analysts that building regional clusters – as opposed to raising productivity, boosting innovation, redressing market failures, or other more conventional objectives – is an appropriate goal of development policy. Indeed, in the cluster building view, innovation, productivity, and growth are an assumed indirect outcome of the expansion of the cluster.

Whether targeted development of identified clusters at the expense of a largely sector- neutral approach is a good idea is an empirical question that has received comparatively little attention in the cluster literature to date. It is also a strategy that has distinct distributional consequences that have to be evaluated as much on ethical as efficiency grounds. But, in any case, a review of the Latin American and U.S. cases suggests that few governments are actively buying into the specialization strategy, at least at present. This may be function of limited resources, lingering concerns about the risk of over-specialization, or, most likely, political realities that lead to the diffusion of development resources even where targeting makes sense. Or, perhaps policy makers have learned that the language and theory of cluster building is more compelling than the actual practice.

Acknowledgements

Special thanks to Marcela Gonzales Rivas and Henry Renski for assistance with assembling information on cluster-related development policies in Latin America and the United States.

References

Altenburg, T. and Meyer-Stamer, J. (1999) How to promote clusters: Policy experiences from Latin America. World Development 27: 1693–1713

Bergman, E.M. and Feser, E.J. (1999) Industrial and regional clusters: Concepts and comparative applications. Web Book of Regional Science. RRI-West Virginia University, Morgantown (http://www.rri.wvu.edu/WebBook/Bergman-Feser/contents.htm)

Bortagaray, I. and Tiffin, S. (2000) Innovation clusters in Latin America. 4th International Conference on Technology Policy and Innovation, Curitiba, Brazil (August 28–31)

CDECD (2001) Industry cluster progress report. Connecticut Department of Economic and Community Development, Hartford, CT (http://www.ct.gov/ecd/LIB/-ecd/20/14/cluster%20progress%20report%20feb%202001.pdf)

Ceglie, G. and Dini, M. (1999) SME cluster and network development in developing countries: The experience of UNIDO. Working paper, PSD Technical Working Papers Series, United Nations Industrial Development Organization, Vienna

Chairatana, P. and Vorrakitpokartorn (2001) Cluster and regional innovation system of Chiang Mai/Lampoon Twin City. 5th International Conference on Technology, Policy and Innovation, The Hague (June 26–29)

Conceição, P., Heitor, M.V., Gibson, D.V. and Shariq, S.S. (1998) The emerging importance of knowledge for development: Implications for technology policy and innovation. Technological Forecasting and Social Change 58: 181–202

Dalsgaard, M.H. (2001) Danish cluster policy: Improving specific framework conditions. In: den Hertog, P., Bergman, E.M. and Charles, D. (eds.), Innovative clusters: drivers of national innovation systems. OECD, Paris: 347–360

Enright, M.J. (2001) Regional clusters: What we know and what we should know. Kiel Institute International Workshop on Innovation Clusters and Interregional Competition, Kiel (November 12–13)

Farinelli, F. and Kluzer, S. (1998) Industrial and technical assistance in Latin America: Lessons from the field. The IPTS Report 21: 1–6

Feser, E.J. and Luger, M. (2003) Cluster analysis as a mode of inquiry: Its use in science and technology policymaking in North Carolina. European Planning Studies 11: 11–24

Feser, E.J. (1998a) Enterprises, externalities and economic development. Journal of Planning Literature 12: 283–302

Feser, E.J. (1998b) Old and new theories of industry clusters. In: M. Steiner (ed.), Clusters and regional specialisation. Pion, London: 18–40

Ffowcs-Williams, I. (2000) Policy for inter-firm networking and clustering: A practitioner's perspective. OECD/Italian Ministry of Industry Bologna Conference for Ministers, Bologna, Italy (June 13–15)

Fisher, E. and Reuber, R. (2000) Industrial clusters and SME promotion in developing countries. Commonwealth Trade and Enterprise Paper, Commonwealth Secretariat, London

Gambardella, A. and Malerba, F. (eds.) (1999) The organization of economic innovation in Europe. Cambridge University Press, Cambridge

Geroski, P.A. (1990) Procurement policy as a tool of industrial policy. International Review of Applied Economics 4: 182–198

Glaeser, E.L. (2000) The new economics of urban and regional growth. In: Clark, G.L., Feldman, M.P. and Gertler, M.S. (eds.) The Oxford handbook of economic geography. Oxford University Press, Oxford: 83–98

Gordon, I.R. and McCann, P. (2000) Industrial clusters: Complexes, agglomeration and/or social networks? Urban Studies 37: 513–532

Grossman, G.M. and Helpman, E. (1991) Innovation and growth in the global economy. MIT Press, Cambridge, Mass.

Hallberg, K. (1999) Small and medium scale enterprises: A framework for intervention. Working paper, Small Enterprise Unit, Private Sector Development Department, The World Bank, Washington, DC

Helmsing, A.H.J.B. (2001) Externalities, learning and governance: New perspectives on local economic development. Development and Change 32: 277–308

den Hertog, P., Bergman, E.M. and Charles, D. (2001a) Creating and sustaining innovative clusters: Towards a synthesis. In: den Hertog, P., Bergman, E.M. and Charles, D. (eds.), Innovative clusters: Drivers of national innovation systems. OECD, Paris: 405–419

den Hertog, P., Bergman, E.M. and Charles, D. (eds.) (2001b) Innovative clusters: Drivers of national innovation systems. OECD, Paris

Jacobs, D. and de Jong, M.W. (1992) Industrial clusters and the competitiveness of the Netherlands: Empirical results and conceptual issues. De Economist 140: 233–252

Jacobs, D. and de Man, A.-P. (1996) Clusters, industrial policy and firm strategy: A menu approach. Technology Analysis and Strategic Management 8: 425–437.

Kawahara, E., Kelly, T., Worley, S.C., Pochy, G. and Shane, B. (2000) Collaborating to succeed in the New Economy: Findings of the regional economic development survey. California Trade and Commerce Agency, Sacramento, CA (http://commerce.ca.gov/ttca/pdfs/detail/ersi/regionalecondevsurvey.pdf)

KIC (2002) New economy strategic plan. Kentucky Office of the New Economy, Lexington, KY (http://www.one-ky.com/plan.html)

Knorringa, P., and Meyer-Stamer, J. (1998) New dimensions in local enterprise cooperation and development: From clusters to industrial districts. Institute of Social Studies, ATAS Bulletin XI, The Hague

Lagendijk, A. (1999) Good practices in SME cluster initatives: Lessons from the 'core' regions and beyond. Centre for Urban and Regional Development Studies, University of Newcastle Upon Tyne, Newcastle, UK (http://http://curdsweb1.ncl.ac.uk/files/3137adapt_su.pdf)

Lagendijk, A. (2000) Learning in non-core regions: Towards 'intelligent clusters'; addressing business and regional needs. In: Bakkers, S., Boekema, F., Morgan, K. and Rutten, R. (eds.), Learning regions, theory, policy and practice. Edward Elgar, London: 165–191

Leyden, D.P. and Link, A.N. (1992) Government's role in innovation. Kluwer, Dordrecht

Malecki, E.J. (1997) Technology and economic development. Longman, Harlow, Essex

McCann, P. (1995) Rethinking the economics of location and agglomeration. Urban Studies 32: 563–577

Melo, A. (2001a) Industrial policy in Latin America and the Caribbean at the turn of the century. Working paper, Research Department, Inter-American Development Bank, Washington, DC

Melo, A. (2001b) The innovation systems of Latin America and the Caribbean. Working paper, Research Department, Inter-American Development Bank, Washington, DC

Meyer-Stamer, J. (1998) Path dependence in regional development: Persistence and change in three industrial clusters in Santa Catarina, Brazil. World Development 26: 1495–1511

Morfessis, I.T. (1994) A cluster-analytic approach to identifying and developing state target industries: The case of Arizona. Economic Development Review 14 (2): 33–37

Moulaert, F. and Sekia, F. (1999) Innovative region, social region? An alternative view of regional innovation. European Meeting on Applied Evolutionary Economics, Grenoble, France (June 7–9)

Mowery, D.C. (2001) Technological innovation in a multipolar system: Analysis and implications for U.S. policy. Technological Forecasting and Social Change 67: 143–157

MTC (2001) Index of the Massachusetts Innovation Economy, 2001. Massachusetts Technology Collaborative, Westborough, MA (http://www.mtpc.org/2001index/-about.htm)

Mytelka, L. and Farinelli, F. (2000) Local clusters, innovation systems and sustained competitiveness. Local Productive Clusters and Innovation Systems in Brazil: New Industrial and Technological Policies for their Development, Rio de Janiero (September 4–6) (http://www.utoronto.ca/isrn/documents/intechppr2005.pdf)

Nadvi, K. (1995) Industrial clusters and networks: Case studies of SME growth and innovation. Working paper, Small and Medium Industries Branch, United Nations Industrial Development Organization, Vienna.

Porter, M.E. (1990) The competitive advantage of nations. Free Press, New York

Porter, M.E. (1998) Clusters and the new economics of competition. Harvard Business Review: 77–90

Porter, M.E. (2000) Location, competition, and economic development: Local clusters in a global economy. Economic Development Quarterly 14: 15–34

Quandt, C. (1997) The emerging high-technology cluster of Campinas, Brazil. Technopolis 97, Ottawa, Canada (September 9–12)

Quandt, C. (1999) The concept of virtual technopoles and the feasibility of incubating technology-intensive clusters in Latin America and the Caribbean. International Development Research Centrepages, Ottawa, Canada.

Rabellotti, R. (1999) Recovery of a Mexican cluster: Evaluation bonanza or collective efficiency? World Development 27: 1571–1585

Reich, R.B. (2002) The future of success: Working and living in the New Economy. Vintage Books, New York

Roelandt, T.J.A. and den Hertog, P. (eds.) (1999) Boosting innovation: the cluster approach. Organisation for Economic Co-operation and Development, Paris

Rosenfeld, S.A. (1997) Bringing business clusters into the mainstream of economic development. European Planning Studies 5: 3–23

Rosenfeld, S.A. (2001) Backing into clusters: Retrofitting public policies. Symposium on Strategic Responses to Integration Pressures: Lessons from Around the World, JFK School of Government, Harvard University (March 29–30)

RTS (2001) Skills and workforce development for Mississippi's CIT cluster. Regional Technology Strategies, Inc., Chapel Hill, NC (http://www.decd.state.ms.us/pdf/-cit/CIT_report_10_01.pdf)

Russo, F., Clara, M. and Gulati, M. (2000) Cluster development and promotion of business development services (BDS): UNIDO's experience in India. Working paper, PSD Technical Working Papers Series, United Nations Industrial Development Organization, Vienna

Saxenian, A. (1994) Regional advantage: culture and competition in Silicon Valley and Route 128. Harvard University Press, Cambridge, Mass.

Schmitz, H. (1995a) Collective efficiency: Growth path for small scale industry. Journal of Development Studies 31: 529–566

Schmitz, H. (1995b) Small shoemakers and fordist giants: Tale of a supercluster. World Development 23: 9–28

Schmitz, H. (1999) Global competition and local cooperation: Success and failure in the Sinos Valley, Brazil. World Development 27: 1627–1650

Sølvell, O., Lindqvist, G. and Ketels, C. (2003) The cluster initiative greenbook. Ivory Tower AB, Stockholm

Temple, P. (1998) Clusters and competitiveness: A policy perspective. In: Swann, G. M. P., Prevezer, M. and Stout, D. (eds.), The dynamics of industrial clustering. Oxford University Press, Oxford: 257–297

Tendler, J. and Amorim, M.A. (1996) Small firms and their helpers: Lessons on demand. World Development 24: 407–426

Tidd, J. and Brocklehurst, M. (1999) Routes to technological learning and development: An assessment of Malaysia's innovation policy and performance. Technological Forecasting and Social Change 62: 239–257

UNCTAD (1998) Promoting and sustaining SMEs clusters and networks for development. Expert Meeting on Clustering and Networking for SME Development, Geneva (September 2-4) (http://www.unctad.org/en/docs/c3em5d2.pdf)

UNCTAD (2002) Partnerships and networking in science and technology for development. United Nations Conference on Trade and Development, Technology for Development Series, Geneva. (http://www.kiet.re.kr/files/econo/20020325-partner.pdf)

UNIDO (2001) Development of clusters and networks of SMEs: The UNIDO Programme. United Nations Industrial Development Organization, Vienna

Vieh, J. (2002) The role of the public sector in Arizona's economy. Arizona Department of Commerce, Phoenix, AZ

Viesti, G. (2002) Economic policies and local development: Some reflections. European Planning Studies 10: 467–481

Visser, E.-J. (1999) A comparison of clustered and dispersed firms in the small-scale clothing industry of Lima. World Development 27: 1553-1570

Voyer, R. (1997a) Emerging high-technology industrial clusters in Brazil, India, Malaysia and South Africa. International Development Research Centre, Ottawa, Canada

Voyer, R. (1997b) Knowledge-based industrial clustering: International comparisons. Working paper, Nordicity Group Ltd., Montreal, Canada

Waits, M.J. (1992) Arizona: Preparing for global competition through "Industrial Clusters". Spectrum 65: 34–37

Spillovers and Innovation, Environment and Space: Policy Uncertainties and Research Opportunities

Edward M. Bergman, Uwe Schubert

1 Introduction and Overview

The previous chapters demonstrate convincingly that the concept of externalities remains a fruitful source of research and policy studies worldwide and it shows few signs of being exhausted. As the introductory chapter makes clear, spatial and other externalities once constituted a series of esoteric assumptions that helped account for deviations from perfectly competitive markets. However, the concept has since expanded to account for a wide and growing set of issues whereby the unintended actions of certain agents affect other independent agents, negatively or positively. This expansion had two additional effects. First, other disciplines and theories have adopted the concept of externalities in various forms to reformulate unintended interdependencies in realms parallel with economics, often enriching the theoretical insights of several such fields simultaneously. Second, the expansion of issues, disciplines and analysts introduced these concepts to more areas of policy and decision-making, particularly as previously unrecognized positive externalities brought to light wholly new areas of beneficial side-effects (Putnam, 1993).

Other book chapters illustrate a selected subset of topics, mainly a subset with which the authors of this chapter have been concerned. We will not attempt here to systematically summarize or classify the contents of these chapters; rather, we wish to focus selectively on aspects of externalities as reflected in the chapter title. Our aim is to highlight certain areas we feel merit further or more intense attention on the part of scholars and policy analysts. We begin first with the general field of environmental externalities, which is one where a considerable body of policy and theory evolved more or less together, mainly from a concern to mitigate or eliminate negative externalities. However, as we point out, environmental policy is increasingly based on incentives that reflect a sophisticated understanding of applicable positive externalities as well. Positive externality concepts have multiplied exponentially in theories, literature, research and increasingly policy-fields affected heavily by new and endogenous growth theories, particularly the effects of spatial distributions of growth factors. The following highly selective comments will be limited to these topics.

Externalities linked with environmental matters are usually thought of as typical examples of negative externalities constituting a barrier to development. There are usually two ways these are dealt with in the framework of economic thinking, i.e. the various policy options to internalize these negative consequences of economic

activities (to be discussed below) and the positive preventive strategies available in principle to overcome these obstacles. The issue to be tackled in the second case starts from the simple thesis that natural resources are necessary production factors, the productivity of which can be enhanced by R&D and innovation in companies. It is particularly the notion of "sustainable development", postulating in various more or less rigorous forms that a stationary state of the stocks of natural resources is to be maintained for the benefit of future generations, that is central to the discussion. If this policy aim is to become compatible with economic development, the stocks of "human capital" need to increase in the future. The technical progress necessary to promote "eco-efficiency", from invention to diffusion follows an analogous logic to the general challenge posed by making any scarce production factor more productive. The major difficulty arising in this context is the different objective function steering the process. While the aim to increase productivity in a Schumpeterian world serves as a vehicle for pro-active companies to get ahead of the competitors, this is usually not seen as a successful strategy with respect to environment-oriented innovation, at least without the proper internalization of negative environmental externalities. Full internalization is difficult to achieve, as was already pointed out in the introductory chapter to this volume. The first best solution, as suggested by Pigou (1932), faces all the theoretical drawbacks alluded to in the introductory chapter of this volume, but also hinges on the ability of researchers to establish the full social cost of emissions caused by polluters. The instruments of environmental policy typically considered generally represent "second-" or "lesser-best" solutions.

An enormous change of view concerning what externalities represent in regional economies has taken place, which continues to redefine research and policy agenda alike. What were once seen as market imperfections or failures ripe for remedy are now considered evidence of how firms, individuals and governments logically seek advantage in a globalizing economy. The new growth theory assumes monopolistic rather than purecompetition, which helps better explain why private and, indirectly, public productive assets accumulate persistently in specific cities and regions; it also helps clarify firms' indifference to, or avoidance of investments in, peripheral areas that lack basic pre-conditions, despite heavy policy interventions. It is no coincidence that business strategists such as Michael Porter, whose rivalry concepts reflect a qualified form of monopolistic competition, are now taken more seriously when considering advantage- seeking behavior of firms that prefer regions and clusters capable of reinforcing or sustaining privileged market positions. At the same time, our altered view of how growth occurs reflects both endogenous innovation and the emerging knowledge economy that have become leading forces in restructuring industries and regions open to global frameworks of trade, capital and labor mobility, transport and communications.

Many different strands of research are rapidly converging from various disciplines that attempt to understand the still opaquely-perceived growth dynamics now underway; these appear, in turn, to have propagated wholly new varieties of development policy that address the partially-understood forces that governments at all levels hope somehow to shape. The present collection of essays provides good insight into several important policy objects and the underlying development forces policies hope to

influence. The second part of this chapter will focus on selected aspects of *industrial clusters, universities and related knowledge spillovers*, in which key policy uncertainties beckon understanding and research opportunities abound to study embedded mechanisms of endogenous innovation that lead to development.

2 Negative Externalities and Space

Negative externalities are potentially created by any transformation process such as the economic activities of production and consumption. In the transformation of inputs into outputs, residuals are created that have a potentially harmful effect on the eco-system, including human beings. Space enters this process naturally as the residuals produced are fed back into the eco-system by being emitted into nature's receiving media. These constitute natural resources such as air, water and soil which serve multiple uses. Once having been deposited in these media, the residuals diffuse over the space the resources cover. During this diffusion process chemical and physical transformations tend to occur and the emissions from various sources accumulate in the receiving medium and are partly or totally absorbed in various sinks. The consequential residuals' concentration, if total natural regeneration is not possible, produces potentially negative impacts upon the eco-system including the anthropogenic economic sub-system, thus causing negative externalities. As pointed out in the contribution by Wang and Nijkamp in this volume, this diffusion process can take various forms and thus can create different basic starting points for policy design. The best known examples are the one-way and the reciprocal trans-frontier pollution cases; the former are typical of river systems where the water flow makes polluters easily identifiable, while the latter, often associated with air-pollution, converts a spatial unit polluter into a victim simultaneously in many cases.

In policy design, space is often neglected as a consequence of assumptions made in the theory of environmental economics. Within a policy area, it is frequently (often tacitly) postulated that the ambient concentration of residuals is uniform, thus constituting the case generally investigated, i.e. the existence of a "public evil", which is an equal load for everybody, but may affect people and nature differently. In principle, permanent and temporary residents of such an area cannot escape the negative impacts, a fact believed to constitute an important incentive for common action. In many cases, however, this assumption does not hold, as the residuals' concentration actually varies over space, thus providing incentives for individual action to improve the personal environmental quality by relocating. Especially in urban regions, this motive for migration of households can be an essential driving force behind processes of urban sprawl. Although this fact has been known for some time, the strength of this phenomenon, which obviously varies between regions and countries (a well documented fact, e.g. Berry and Horton, 1974 use "isopleths" analysis to show the varying distribution of pollutants over urban space for some U.S. and European cities; Stanners and Bourdeau, 1995), still needs to be explored further. A particularly interesting feature of scientific and political relevance is the nature and elasticity of the trade-offs in relation to other factors of location choice (see e.g. Schubert, 1979). A multi-disciplinary research approach is also warranted in this case.

A topic less often explored is the concept of "social space" and its bearing on the development of environmental policy approaches. One of the most interesting propositions in this respect was made by Coase (1960), in which the attribution of property rights to a party in a negative externality conflict involving a limited number of persons provides the basis for direct negotiations and the option of designing a contract in which the conflict is settled and internalization is achieved. The start of such contract negotiations depends critically on the social relations between the persons involved, as the "social distance" between them has a bearing on the trust the future partners have to develop to design and comply with the contract. A similar argument can be made with respect to other co-operative activities requiring mutual trust, such as public-private partnerships (see the contribution by Gindl and Wukovitsch in this volume). The "social distance" is related to the friction encountered in communication, the determinants of which still merit attention by researchers.

3 Space and Environmental Policy

The problems of internalization and the various policy approaches to achieve it have been studied quite well in the literature (an excellent review is presented in this volume by Wang and Nijkamp). An important policy problem still remains somewhat elusive in this context, i.e. the delimitation of the relevant policy areas within which institutions are to be set up: spatially sensitive policies remain limited to the administrative competence defined within institutional boundaries (see e.g., Hoel and Shapiro, 2004). Obviously the reach of residuals' diffusion fields is neither confined to administrative boundaries, nor does it remain constant, partly due to short term variations in the physical movements of the receiving media (e.g. winds, water currents), but also because of changing overall volumes of emissions. These, in turn, depend on the economic growth process in the area and abatement measures taken by companies and households. The problem is further aggravated by the fact that the basic spatial units aggregated to form the "pollution management districts" often do not have the appropriate size or physical features. Relevant areas are sometimes excluded because they are within the jurisdiction of another country. Clearly these critical issues pose challenges to new forms of governance and they could greatly benefit from (necessarily multi-disciplinary) research.

4 Environment-oriented Policy and Innovation

As was pointed out above, the necessary long-run condition for sustainable development is the increase of the productivity of "natural capital". Product and process innovation with an environmental focus is the vehicle towards achieving this goal in companies (e.g., Fischer et al., 2004). The question immediately arises whether there is a difference between innovation in general and environment driven innovation activities.

Cleff and Rennings (2000) maintain that the main difference is the importance of the regulatory set-up. The general innovation literature emphasizes technology push factors and market pull factors as the main driving motives. The regulatory set-up is

not only defined as a command and control mechanism, but also as general interaction between governmental and non-governmental actors. This implies a new management oriented approach where "public administration is reorganized and fully integrated within the whole process" (Schrama and Sedlacek, 2003, p. 228). Additionally, policy co-ordination or integration, including environmental objectives in other non- environment related policy fields, need to be considered (see Gouldson and Murphy, 1998, Schrama and Sedlacek, 2003, p. 235ff.). Regulatory incentives, hence, have a particular role in stimulating environment-oriented innovations (e.g. Blazjeczak et al., 1999, Klemmer et al., 1999). The analysis of the impacts of policy measures, especially regulation, has confirmed the important role played by the political realm in innovation processes (Marin and Mayntz, 1991).

The role of policy, hence, has been clearly established by recent research. A question remains, however, about the effectiveness and efficiency of policy instruments. This query has various aspects. First, different relevant policy fields have to be distinguished, such as environmental policy, technology policy, particularly those programs with an environmental focus, and various sectoral and spatial policy areas (e.g. economic, transportation, energy etc. policy, urban and regional planning). Secondly, significant differences between instruments available within a policy area are claimed to exist. Economists are generally convinced that price-related environmental policy (e.g. green taxes, tradable emission permits) tends to be most efficient in the long-run (e.g. Pezzey, 1992, 2003), while administrators and political scientists tend to favor stringent command and control approaches. The empirical evidence available to reject any of these hypotheses is still weak (an empirically-based analysis is presented in Gale et al., 1995), in part simply due to the fact, that price related policies have not been applied widely yet, and where they do exist they have usually been introduced rather recently, making good analysis difficult.

An assessment of the effectiveness of technology policy faces great uncertainty and demanding efficiency assessments remain well out of reach. Specific technology policy initiatives with an explicit environmental focus are rare. If they do exist, the time span in which they can be studied is generally still very short (see e.g. Schrama and Sedlacek, 2003, Ulph, 1997). An important dimension to consider in such studies is the goals of the policy. The process of technical progress is characterized by two principal elements: the creation of new technologies and their diffusion in the economy (see e.g., Isik, 2004). The incentives provided in pertinent policy programs can have quite different effects on these elements. They could even be contradictory.

The instruments available to promote environment-oriented innovation in general and specific elements in particular would certainly merit more attention by researchers, as considerable funds (mostly public) are invested and the risk of betting on the wrong horses could be considerably decreased.

It must also be said that the complex interactions among other policy fields mentioned and the resulting effects on innovation and the environment are even less known. Analyses to detect and trace these effects through the economy hinge critically on the development of a more general "system's model".

The influence of *market forces*, however, must not be neglected in the context of environment-oriented innovation. These incentives play a role in different types of

markets. There is an influence of "green consumer demand", in the form of products desired that are environmentally superior to others, or the production process is critically viewed concerning possible negative effects on the environment.

Regulation often plays a critical role in the creation of new markets, which would hardly exist without it. The "technology developing" organizations (public, private, non-profit, etc.) rely on policy incentives that stimulate effective demand from "technology using" companies (the technology developers can be integrated in the same organization, e.g. as "R&D departments, etc.) for their products and services. Environment-related technology markets are particularly sensitive to changes in policy.

The relative strength of the various market and non-market forces that exert an influence on environment-oriented innovation is still being intensively debated. It seems to vary by country and historical development phase (Sedlacek and Schubert, 2004). More empirical evidence based research could make a considerable contribution to better policy making.

Positive externalities produced in innovation networks

One of the important questions arising in the context of the production of positive environment related externalities (in the case discussed here, innovation) is the form of organization which is best suited for this purpose. Discussions on this query have increasingly focused on the usefulness of networks. Characteristic contributions to this debate are e.g.:

> *"Innovation is increasingly recognized as requiring the convergence of many sources of knowledge and skill, usually linked in the form of a network."* (Pyka et al., 2002, p. 169, similar arguments can be found in Porter, 2000, Rosenkranz, 1996). The *"networks themselves emerge as a new form of organization within the knowledge production"* (Küppers and Pyka, 2002, p. 6).

The network structure and the type of partners involved seem to be the most essential element that distinguishes between innovation in general and innovation with an environmental focus. Various groups of actors are involved: those directly active in the innovation process (i.e. technology developer, technology user, supplier firms, etc.) and those supporting the network (i.e. policy makers, financial services, consultants, etc.). Both groups form a *"self maintaining social structure"* (Küppers and Pyka, 2002, p.7) - the innovation network. As pointed out above, the creation of such environment-oriented innovation networks is mainly influenced by external stimuli, provided by regulation and the instruments applied in the various policy fields.

The question of organization of innovation activities has two elements to consider, i.e. the structure of the network and its change during the different phases of an innovation project. The various functions necessary (such as knowledge creation, financing, etc.) to make the network operational and the partners who represent it determine the *structure*. One of the crucial elements is the flexibility and readiness to work towards a common goal in an atmosphere of mutual trust. The spirit of co-operation is particularly difficult to maintain for the representatives of governmental agencies, as they are

usually trained and socialized in their job to represent the law and thus are often not prepared for planning and development tasks. This fact certainly presents a challenge for the development of appropriate training programs.

Innovation projects take time. The work takes place in *phases*, which constitute necessary elements (idea - information - decision - implementation - monitoring). These stages do not necessarily have to follow each other in a simple, linear sequence (see e.g. Kline and Rosenberg, 1986). Defining innovation as a dynamic process, it can be described as a trial and error sequence in feedback-loops. Many open questions remain about the dynamics of innovation network activities. An essential feature is the readiness of partners to learn from failures, but which are the factors that determine successful learning behavior needs to be investigated further. The absence of hierarchical relations between the partners and the degree of formality in co-operation are seen to play a role, but in some cases the absence of formal contracts (particularly "division of the cake" issues in success case) have been blamed for failure during final project phases.

5 Clusters as Externality Arenas

Regions and agglomerated economies are the classic "externality arenas", where the accumulated events of long path-dependent histories produce location-specific assets that now yield economic advantage. A widely discussed contemporaneous arena within which positive externalities produce strong development potentials is an industrial or regional cluster. Clusters appear to be face-valid expressions of developmental externalities to even the casual observer. However, the literature on clusters and their variants has one of the lowest value-added-to-effort ratios of any subject on our general topic, which is another way of saying that genuine contributions per publication have been disappointing. Some of this can be blamed simply on its sheer popularity among many audiences, which stimulates the release and distribution of much derivative material by markedly different agents, often for promotional, marketing or policy advocacy purposes, rather than advancing the stock of original research and general understanding.

Feser's chapter describes this popular phenomenon and documents its size and growth, which has shown little subsequent sign of abating. He then examines the penetration of cluster ideas into the policy portfolios of Latin American governments and compares it with U.S. cluster policy and practice. The Latin American experience is instructive because it simultaneously reflects the popularity of clusters and their suitability in changing systems of political economy that have suffered with the rapid advance of globalization and attempts to liberalize economic institutions and practices. Other studies of cluster development in transitional and developing economies cited by Feser reveal a genuine hunger for locally-oriented development policies that also fit within global systems of production and trade. This raises the question of how suitable industrial clusters are to a variety of political economic settings, or perhaps whether certain aspects of clusters fit some political economies better than others.

This question results logically from the diverse origin of industrial cluster ideas in advanced economies of Europe and North America. In both settings, a large but

partially- connected stock of conceptualizations in economic geography, business and regional science has been drawn upon to identify and provide evidence for the following range of beneficial clustering effects:

1. home-market region and potential scale-economies permit establishment and growth of traded sectors (general agglomeration/scale externalities),
2. input cost reductions and specialized outsourcing possibilities permit unique locally-realized operational efficiencies and expansion of traded products or services (pecuniary spillovers), and
3. locally captured knowledge and innovation gains generated by interaction with competitors, suppliers, customers or organizations help drive dynamic improvements in production and product development (technological spillovers).

European and U.S. analysts differ somewhat concerning how these effects arise and are sustained: U.S. analysts tend to favor historical chance, technological disruption, external impacts, and first-mover advantage in markets while European views emphasize the creation of favorable conditions for firms and clusters through government or 3rd party policies, although all factors can be documented in the case experience of both continents. The unanswered question is how and which of these benefits can be expected to arise in the transition and developing (or even peripheral EU/U.S. regions) economies that lack key features of U.S. or EU environments.

The first effect results from historical patterns of settlement, urbanization and concentrations of demand that favor the emergence of final markets. As Schmitz (1995, 1999) and others have noted, clusters operate frequently and most effectively in urbanized portions of developing countries where production scale arises logically from urbanized demand patterns. Scale efficiencies must be "created" in rural and peripheral areas of developing countries where raw material factors are often located and many people still live at low densities. Transitional economies of former USSR have the opposite problem: existing production sites were initially selected and output scaled to targets based on some collective need, which was generally measured only in quantity (not quality). Inherited output capacities for low quality goods far exceed evolving demand within newly reorganized nation states that often remain incompletely developed in terms of product and factor markets. It is therefore important to recognize that cluster policies could differ radically in various countries and settings.

Indeed, Krugman (1991) hypothesized that a general reordering of site specialization and sectoral concentration will eventually result from EU integration of its then 12-member states, thereby implying a longer-term shift in cluster development *within EU countries*. Similar hypotheses might apply to the cities and regions in NAFTA countries, although changes of lesser degree would probably result from expanding trade and capital mobility, rather than labor mobility. Additionally, there are substantial variations in cultural practices, social capital, and business systems among regions in the U.S. or the EU that deserve far more attention concerning their potential impact on cluster development and success. Taken together, one might readily conclude that further study of the core processes of city and regional development could yield potentially useful insights into the formation of home markets, scale economies and export markets.

The second cluster effect could be considered a Marshallian consequence of the first: commonly shared factor inputs such as specialized and cost-efficient labor, unique inputs or capital goods, and tailored public infrastructure and services are eventually offered at "pecuniary" discounts to local cluster firms that specialize in scale- efficient export production. These are often characterized as localization economies that arise naturally within trajectories of capitalist development, the externalities of which accrue to a particular cluster's subset of firms and industries. National and regional differences do affect the degree to which these effects can be secured through various policies to benefit local clusters. Since the pecuniary factors mentioned here can be strengthened through a mix of policies concerning local business practice and public expenditure, regime differences may have potentially profound effects upon cluster viability. As only one example, the creation, attraction and retention of skilled labor implies locally relevant policies that affect residential environment and quality of life, mobility of innovatively creative workers, well-springs of entrepreneurship, etc. These are hotly contested issues in the U.S., mainly at journalistic levels, and are receiving attention recently in Europe as well. Since little serious scholarship has been devoted to the subject, there is not much convincing evidence concerning the relative importance of amenity-milieu factors identified by Goldstein and Renault in their chapter. Stronger efforts to understand the policies that affect successful stimulation of new firm startups or the attraction and accommodation of inward investment in supporting sectors would also prove valuable. Further, the willingness and ability of relevant governments to expend wisely on key public infrastructure and services is poorly understood at best.

The presence of pecuniary effects in specific clusters may also produce extra-cluster incentives, whereby otherwise unrelated sectors and firms are able to take selective advantage of pecuniary advantages borne of the original localized cluster. Jane Jacobs and others who stress the importance of urbanization economies point to this possibility, particularly in larger urban regions with robust local market demand, although the relationships *between* localization and urbanization externalities has not been examined closely to date. These widely recognized factors have more often been studied in isolation, with surprisingly little research dedicated to investigating *groups* of pecuniary factors, even though all are considered important. The far more common practice is that policy analysts and even some researchers assume these factors are necessarily present in regions that support above-average concentrations of output, as revealed through simple cluster-mapping efforts (usually location quotients or similar concentration indices).

Pecuniary advantages of the type described here are generally absent in developing countries, since the institutional framework conditions or pre-requisites must be established before firms acquire typical capitalist incentives or policies to promote cluster-supporting activities. Clusters in emerging economies of Asia and some Latin American economies appear to have made rapid improvement in these conditions, but this is a largely untested observation as well. The case could be made that research lags in these places because suitable measures of pecuniary factors are undeveloped, or that unambiguous measures of clusters or their viability are similarly absent, but this is another way of stating that necessary research concerning basic measurement

and conceptualization lags well behind the premature application of assumption-based policies.

The third cluster effect refers to the uncompensated flows of knowledge and innovative practices that leak from some firms to benefit others within a cluster. These inter-firm leaks or spillovers are said to be among the most significant sources of competitive advantage enjoyed by cluster firms, who expect to lose and gain spillovers in a positive-sum game. The reasons such spillovers do not devolve to zero- or negative-sum games is based on the observation that only firms with active programs of internal R&D are in a position to understand and apply innovative knowledge that might spill their way and that temporary product market niches thereby earned provide sufficient incentive to pursue innovations. The ability to conduct internal R&D usually depends upon minimal levels of profitability and the value of tax- deductible R&D expenditures, both of which are supported by cluster effects 1 and 2 as outlined above. Therefore, clusters could be said to provide hospitable conditions and incentives to innovate, the benefits of which accrue to it and neighboring cluster firms. Combined with the insights of endogenous innovation as a key component of the new growth theory, innovation is now seen far more as the natural outcome of monopolistically-competitive capitalist systems and the driving force behind growth. From this perspective, it is a short step to the basic principles behind what are known as innovation systems, where national versions (NIS) were first elaborated in a series of key studies, which were then followed by regional (RIS), generalized to spatial innovation systems (SIS), and eventually to learning regions[1]. In two large OECD studies of innovation systems, clusters were equated with "reduced-form innovations systems" in which key elements function identically in both concepts. The convergence of innovation systems and clusters in a policy repertoire has attracted more attention from all OECD member countries except the U.S., primarily because much more active national government involvement in both is envisioned, although activities at the state level in the U.S. are significant and growing.

6 Universities as Externality Agents

It is by now an article of faith that the "knowledge economy" could not have arisen in the U.S. as it did without the research function of universities - particularly elite public and private research universities - becoming heavily endogenized as market-responsive knowledge-producing institutions. A common denominator was the growing dependence on new knowledge and in particular knowledge generated by local university R&D, often functioning within self-organized regional innovation systems. Research universities are generally perceived as location factors of growing importance to corporate investment decisions (Dorfman, 1983; Andersson, 1985, Andersson, Anderstig, and Harsman, 1990, Hall, 1987), either as sources of public good "spillovers" or perhaps, as Breschi and Lisson (2001) have insisted, as increasingly

[1]Concepts pursued here represent a logical regional subset of the full set of processes and institutions seen as key elements in the innovation systems literature (Braczyk, et. al., 1998, Cooke, 1998, de la Mothe and Paquet, 1998, Cooke, et. al., 2000, Bergman and Feser, 2001, den Hertog, Bergman and Charles, 2001), which assumes the presence and importance of university spillovers.

organized suppliers of knowledge using a myriad of possible commercial transactions (Sampat, 2003).

The warp-speed of this process as the 20th century came to an end all but dispensed with quaint ideas of aloof and insular universities as "ivory towers" focused purely on the pursuit and profession of ultimate truths, nearly always holding indifferent to casual interest in (or mild disgust with) their commercial surroundings. Economic "endogenization" of universities has also gotten underway in Europe, as the Bologna Process continues to harmonize higher education practices throughout the EU. Generally missing from this process, however, are the institutional incentives and competitive pressures to expand research and innovation- creating functions that now account for the majority of budgetary support in the best U.S. research universities. This is also reflected when one reviews the widely-scattered case studies of clusters on the two continents: in comparison with the U.S., the role of universities appears markedly less important in EU cluster dynamics. The generally lower levels of EU research (% GDP) that prevail preponderantly in commercially-untainted national research laboratories, academies of science or similar state-sponsored innovation institutions seldom drive clusters.

The chapter by Goldstein and Renault summarize the impacts research universities could have on their regional host economies: *productivity gains, greater business innovation, new sources of business startups, regional creativity, and an overall increase in capacity for sustained regional development.* To investigate the general effects of universities - particularly their changed role during the last third of the 20th century - on U.S. metropolitan regional economies, they conduct a quasi-experimental test of regional changes in labor productivity: "Our measure of regional economic development is average annual earnings per worker. [A]verage earnings per worker focuses on the quality of jobs in a region as the most important dimension of improvement in regional economic well being" (p. 80). Their results demonstrate convincingly that the presence of a research university is significantly related to greater increases in local earnings per capita (or labor productivity) *after* universities became much more active research agents capable of exploiting knowledge gains, although it is not possible to know from the evidence available *which university-related mechanism is responsible.* While human capital creation is perhaps the most likely mechanism, the productivity impacts estimated here could have been realized through several possible mechanisms, as the authors' questions make clear:

> "*Through what mechanisms, or channels, does knowledge production – broadly considered – within universities lead to economic development outcomes in the surrounding region*? Is it, for example, through economic transactions between actors or units within the university and external organizations, through spillovers, or through milieu effects, which are particular kinds of localization economies?" (italics added, p. 72)

Universities appear to be important sources of knowledge to firms in host U.S. regions, a general proposition which can be tested empirically with only slight risk of mis-specifying the probable causal direction: knowledge is far more likely to flow *from universities to firms* that can exploit such advantages commercially. However, univer-

sities are not the only nodes from which knowledge and related factors might spill into a regional economy, which then complicates our understanding and testing of knowledge spillover mechanisms. As the chapter by Varga, Anselin, and Acs demonstrates, knowledge-generation nodes include other research organizations that go well beyond universities alone. The authors deploy a knowledge-production function approach to modeling regional high-tech patent applications that might arise from university *research* expenditures, industrial *research* and development (proxied by high-tech research laboratory employees), and total high-tech employees (proxy for networks of innovative firms). Even though research from both universities and corporate research laboratories is likely to spillover to firms, the authors are careful to acknowledge that such a

> "... model does not allow for explicit modeling of *the way knowledge spillovers occur* and as such it is difficult to separate spillovers from the correlation of variables at the geographical level.." (italics added, p. 94)

In other words, while regional units of analysis do permit one to infer generally that spillovers often flow from known knowledge generators to knowledge commercializers, variations in regional patent applications could just as easily reflect correlated variations in the "internalized intellectual property" of regional universities (Bayh-Dole Act beneficiaries), corporate research laboratories, or high-tech firms, all of which seek and avidly hold patents with *absolutely no spillovers having occurred.* Moreover, use of regions, clusters or other spatially defined units as "externality arenas" do not permit one to distinguish the extent or effects of spillovers when the flow and mechanism of knowledge transfer within these arenas remain unanswered questions. Clearly, we need better concepts of spillover mechanisms within regions to avoid attributing internal innovation to external spillovers. This inherent limitation poses the familiar problem of how spillovers are best conceptualized as theoretical constructs or mechanisms and as empirical objects suitable for measurement and testing.

7 Spillovers as Externality Mechanisms

General externality concepts in the form of positive spillovers that propel commercial innovation have become the object of intense analysis, although clear spillover definitions remain elusive and resist specification. Consequently, spillover research papers and articles repeatedly mention the need to specify better the actual knowledge spillover mechanisms. Audretsch and Feldman (2004) argue that better conceptual understanding of basic geographic location and agglomeration externalities

> "... was a significant step in generating innovative activity, [however] it provided little insight as to *how and why knowledge spills over*, nor did it illuminate the mechanisms that serve as conduits for the transmission of knowledge." However, recent research "... *literature on knowledge spillovers and the geography of innovation has begun to consider the mechanisms by which knowledge spills over and is put into economic use and*

> *the degree to which these processes are geographically localized.*" (underlining added) p. 2 and p. 19.

Their review establishes from the outset how geography and spatial units of analysis were first introduced in knowledge production function studies, the results of which improved steadily with numerous refinements, particularly later introduction of spatial lags and citation trails. Spatial arenas properly "internalize" core externalities and permit one to test theoretically expected evidence of spillovers that cannot convincingly be established by relying upon firms or industries as units of analysis. Evidence grew stronger and more convincing as the spatial units declined in size, from nations to states to regions. It is this evolving pattern of studies that led to concepts of "localized knowledge spillovers", referred to frequently as LKS.

Efforts to specify spillover mechanisms associated with an LKS view might be expected to evaluate the character or nature of interactions among aggregated agents that lead to spillovers. Seminal studies and literatures surrounding so-called "MAR", Jacobs and Porter effects fall clearly in this camp. Progress helped distinguish between pecuniary effects and other market mechanism effects, the spillover potential from specialized vs. diverse industrial structures, and the nature of competition for local pecuniary and technological advantages or in monopolistically competitive export markets. Subsequent literatures that stress the formation and value of social capital, affinity networks, and venture-capital density also fit well within this de facto framework, although there has been no systematic effort to elaborate this framework. Were one to do so, it might be worth inventorying and classifying quite systematically all spillover-enabling features of: *market mechanisms* (e.g., pecuniary advantages, rivalry, competitiveness, output and factor markets, etc.), *sectoral structures* (specialized vs. diversified, SME, FDI and entrepreneurial start-up ratios, cluster components), *labor pools* (skills, occupations, cohort detail, mobility), *residential quality of life* (amenities, opportunities, risks, well-being), *social fabric* (social capital, affinity networks), *public-good externalities* (political homogeneity, Thiebout effects, club goods) and so forth.

Even a fully-elaborated LKS arena framework is unlikely to permit one to infer precise mechanisms of *what* spills over, between *whom* and to *which* proximate effect; moreover, such attempted inferences might indeed lead to confusion. The LKS arena framework permits stronger specification of the overall conditions that enable or stimulate spillover potentials within externality arenas; however, improved specification of the spillover mechanisms per se between the agents are best acquired from other perspectives. These confusions are apparent from their working list of "mechanisms for spillover transmission," which indiscriminately mix agent-level concepts with arena-level concepts of spillovers (Audretsch and Feldman, 2004). Their list contains familiar items mentioned by others in the literature, most of which mention vaguely-described spillovers; these fall considerably short of understandable and testable mechanisms, despite serious intentions to do so. Perhaps some of the difficulty results from attempts to infer internal mechanisms at work *within* the geographic units of analysis that first enabled scholars to detect and estimate the net impact of such

spillovers.[2] As Caniels and Romijn (2003) put it,

> "... so far, all the existing approaches in the debate have adopted a regional (i.e., meso) level of analysis, without looking closely at the behaviour of individual actors (notably firms) that make up the region. Yet firms are the key actors in innovation and learning processes. It follows that a good grip on the micro-economic processes that underpin innovativeness is essential for gaining a better understanding of the driving forces of regional dynamism" (p. 6).

Accordingly, Caniels and Romijn (2003) attempt to open the LKS black box of firm spillovers. They do this by introducing a firm's view of micro-motives, a view that draws heavily upon evolutionary economics to demonstrate the importance of continual learning processes to innovative processes. In brief, they apply two widely accepted learning mechanisms ("trial and error" and "organizational search") to a collapsed set of Marshallian agglomeration advantages ("specialized labor pool and supplier base" and "technological spillovers" to identify six potential spillover mechanisms. Another somewhat more successful attempt to disentangle and classify these dimensions starts from conventional meso-level concepts of regional economics, gradually adding firm behaviour extracted from the evolutionary economics paradigm (Johansson, 2004). Johansson essentially triangulates in conventional economic terms key externality features (sources, consequences, and nature of meso-externalities) and spillover types (intra-market/pecuniary, quasi-market/club, extra-market/technological) to create a geometry of the evolutionary firm's vertical and horizontal relations in which spillovers might arise. A resultant matrix yields the economic nature and essential features of specific mechanisms classified by the firms' horizontal and vertical geometry. The result is a nearly exhaustive taxonomy of spillover mechanisms, with accompanying suggestions of how some might be operationalized. Further investigations and expansions along these lines hold promise for disentangling and better specifying the potential spillovers that might arise among firms and with their economic environment.

However, both approaches omit consideration of other economic agents previously implicated in spillovers, except very indirectly. Omitted are universities, independent or corporate research laboratories, public bodies, and other interested knowledge-generating or transferring organizations discussed earlier, which are typically among the key elements discussed in the innovation systems literature. What are their micro-motives and are they changing? Even answers to these questions would not account for the motivations that animate "human agents" who are known to be very heavily involved in spillovers.

[2]The difficulty is similar to that of understanding what causes regional employment to rise or decline when one measures the change only in net terms, rather than decomposing the total employment changes by sectoral and occupational components, including entry and exit to the region of both employers and employees.

8 Knowledge Workers as Spillover Agents

Organized legal agents (i.e., firms, research universities or laboratories, etc.) purposefully develop intellectual properties, pursue core R&D activities and actively manage knowledge as part of their ongoing activities. Knowledge and information is transferred between organized agents when actions are taken by some human inside the organization, although *knowledge workers* as spillover agents generally receive much less attention than the stocks and flows of knowledge they embody, discover, synthesize or manage. Their specific activities may include participation as staff on joint research projects, working with external consultants (firms, universities, etc) hired to provide specific knowledge and technology inputs, attendance at conferences and symposia to extract new information, systematic reconnaissance of libraries, journals, data archives, etc., and the casual industrial espionage absorbed effortlessly in the tacit- Marshallian "atmosphere" of social contact circles and among value-chain partners.

Tacit-knowledge transfers and acquisitions remain extremely distance-sensitive, which effectively reduces the potential contact field of knowledge workers to a local region capable of being traversed by auto or public transport, although occasional reassignment of a firm's knowledge workers to other locales has also been proposed (Rallet and Torre, 1998, pp. 44–45). Intelligence generated or absorbed by these and similar means accrue directly to – and is embodied in – knowledge agents. Only indirectly and perhaps later is such knowledge transferred to parent organizations, which attempt to stimulate its collection and internal exploitation by improving the internal incentives system and by adjustments in corporate culture.

Much intelligence remains embodied tacitly (Karlsson and Zhang, 2001) in knowledge agents and is potentially mobile, as originally envisioned by Marshall. Once (potentially) mobile, agents are no longer simply knowledge agents employed in some organization; they are the "knowledge spillover agents" (KSAs) who increasingly appropriate and profit from their embodied knowledge through mobility. Audretsch and Keilbach (2003, p. 5) express the incentive as follows:

> "When the lens is shifted away from the firm to the individual as the relevant unit of observation, ... the question becomes '*How can economic agents with a given endowment of new knowledge best appropriate the returns from that knowledge?*'"

They go on to argue that the most rational appropriation - at least in the U.S. - is the entrepreneurial route, i.e. spinoffs and startups originated by university scientists. A seminal series of research studies has focused on precisely this issue by examining new firm startups in advanced fields of science (e.g., biotechnology) as a function of highly qualified co- located scientists in the same fields, where various combinations of patents or articles in journals of record (including citations) are used to determine qualifications of what are termed "star scientists" (Zucker, Darby and Brewer, 1998, Zucker and Darby, 2001, Zucker, Darby and Armstrong, 1998, Zucker, Darby and Torero, 2000, Zucker, Darby and Torero, 2002). In distinguishing the unique value of scientists, the authors argue such scientists posses *more* than routine human capital,

which is:

> "... seen as earning a normal return on the cost of investment, both direct costs and foregone earnings. We believe that some innovations, particularly "invention as a method of inventing" [Griliches, 1988] may be better characterized as creating rivalrous human capitalintellectual human capitalcharacterized by natural excludability as opposed to a set of instructions for combining inputs and outputs which can be protected only by intellectual property rights" (p. 291).

Clearly, those who posses such talents and abilities are able to appropriate and capitalize on them to benefit themselves and often their surrounding localities: "The primary pattern in the development of the industry involved one or more scientist-entrepreneurs who remained on the faculty while establishing a business on the side." (p. 291). Even if renowned scientists leave their original location, others may find continuing local advantages embedded in the form of "enduring social capital" (Agrawal, Cockburn, McHale, 2003).

Property rights can also be appropriated in other ways without leaving one's post. Torero (2000) applied similar methods to locate the officers, executives, scientific advisers or advisory boards, etc. filled by "star scientists" in established semi-conductor firms reporting initial public (equity) offerings or other activities requiring public disclosure. To these corporate involvements must be added spatially unrestricted sale or licensing of patents, equity participation or board membership in acquiring firms, etc. Appropriation of embodied knowledge by scientists may, however, be somewhat less likely in Europe, since the U.S. model of intellectual property (Bayh-Dole, see Sampat, 2003) is seen to promote more spinoffs from U.S. universities: "A major obstacle to better application of research results is the way intellectual property issues are handled in Europe" (Commission of European Communities, 2003, p. 15 and 2004). There are other possibilities for KSA to exploit knowledge endowments as well, including better employment conditions at a new organization or within one's original organization, which is usually possible only with alternate employment prospects that a KSA might reasonably be expected to consider.

Moreover, mobility need not be considered as transitive networking among future employers, as mobility could indeed re-circulate KSAs among a sub-set of likely institutions. Indeed, Rallet and Torre (1998) argue that infrequent mobility of corporate researchers to other sites is often sufficient to transfer key tacit knowledge inputs at critical R&D stages. Impediments to elective mobility among European knowledge workers are far less frequently encountered in the U.S. (Drenth, 2003). Strong cultural, familial and linguistic preferences are of course important factors that bind people to organizations and locations everywhere, but so too are the very high professional and financial risks one takens if: when the most skilled knowledge workers become frozen solid in present positions due to:

1. mobility penalties implied by moving between increasingly unstable European health and pension systems (TIAA-CREF provided uniformly portable pension

systems for academic and research personnel in U.S. since the early 20th century),
2. academic credentials not fairly assessed by distant employers (university degrees and accreditation presently being harmonised under Bologna process),
3. lateral mobility/experience remains unrewarded by subsequent employers (i.e., loyalty valued).

The role of knowledge spillover agents in the knowledge transfer process has not thus far received the attention it deserves by European or U.S. policy-makers. Mobility of knowledge spillover agents includes the reciprocal movement of scientists between universities, laboratories and industry[3], as well as the initial mobility of newly trained students (e.g., ERASMUS and SOCRATES programs, plus new EC student mobility programs). In this respect, the role of continuing, life-long education and the potential contribution of underrepresented social groups among such agents also deserves greater recognition.

9 Summary and Outlook

The volume presented covers a wide variety of aspects of the theoretical and policy consequences of external effects. Advances and gaps in the research in this field were pointed out and discussed and the subjective views of this chapter's authors about future research directions were identified. Given the level of detail and the wide spectrum of the issues presented, which general points emerge in the opinion of the authors regarding the future of this field?

- *Space* plays an essential role in the theory of externalities. Space can be defined in physical and in social terms, distance between different "nodes" constitutes an essential element of analysis. Modern network theory profits from this consideration, while the "governance" and local knowledge spillover issues now given major attention would benefit from better spatial conceptualizations. Policy design needs to take space into account if they are to become more effective.
- Policy studies of programs to promote positive externalities and prevent or reduce negative external effects by their very nature touch upon various academic fields. The organization of research, hence, has much to gain from integrating *multi-disciplinary concepts* in its approach. The openness of the scientists representing various fields to multiple concepts and methodologies of research team members is an essential prerequisite for further progress.
- Given the importance of externalities in theory and practice, and the inseparable link to public policy, higher education has responded to this fact only very modestly. Policy studies programs in existence tend to emphasize a portfolio of narrow policy realms, which themselves are seldom multi-disciplinary, nor do such programs incorporate externality concepts to their full potential.

[3]Many European countries have programs designed to stimulate the KSA mobility among sectors and regions or to repatriate former "brain drain" KSAs.

- Policy design must focus on the *phases* of a program, from the creation of the program objectives and ideas to implementation. The dynamics implied across program phases still have received very little attention in the application of externality concepts. Thus only minor guidance is provided by theory for policy formulation. Policy studies tend to neglect the significance of the phases, analyses pertain principally to single elements, and the results often remain unintegrated.

It is the firm belief of the authors of this contribution, that the topic areas discussed in this volume have a promising future and offer researchers a broad spectrum of theoretically fascinating themes to address. Increasing demand from the public to increase the effectiveness and efficiency of public policy should attract the attention of scholars, policy analysts and managers as well.

References

Agrawal, A. K., Cockburn, I. M. and McHale, J. (2003) Gone but not forgotten. Labor flows, knowledge spillovers, and enduring social capital, NBER Working Paper 9950

Andersson, A. E. (1985) Creativity and regional development. Papers of the Regional Science Association 56: 5–20

Andersson A., Anderstig, C. and Harsman, B. (1990) Knowledge and communications infrastructure and regional economic change. Regional Science and Urban Economics 20: 359–376

Audretsch, D. B. and Feldman, M. P. (2004) Knowledge spillovers and the geography of innovation. In: Henderson, J. V. and Thisse, J.-F. (eds.), Handbook of urban and regional economics, vol. 4. North Holland, Amsterdam: 2713–2739

Audretsch, D. B. and Keilbach, M. (2003) Entrepreneurship capital and economic performance. Industrial Organization. Discussion Paper series No. 3678

Berry, B. J. L. and Horton, F. E. (1974) Urban environmental management. Prentice-Hall, Englewood Cliffs, N.J.

Bergman, E. and Feser, E. (2001) Innovation system effects on technological adoption in a regional value chain. European Planning Studies 9: 629–648

Blazejczak, J., Edler, D., Hemmelskamp, J. and Jänicke, M. (1999) Umweltpolitik und Innovation: Politikmuster und Innovationswirkungen im internationalen Vergleich. Zeitschrift für Umweltpolitik 1/99: 1–32

Braczyk, H.-J., Cooke, P., Heidenreich, M. and Krauss, G. (eds.) (1998) Regional innovation systems. The role of governances in a globalized world. UCL Press, London

Breschi, S. and Lissoni, F. (2001) Localised knowledge spillovers vs. innovative milieux: Knowledge "tacitness" reconsidered. Papers in Regional Science 80: 255–273

Caniels, M. and Romijn, H. (2003) What drives innovativeness in industrial clusters? Transcending the debate. Regional Studies Association International Conference, April 2003, Pisa

Cleff, Th. and Rennings, K. (2000) Determinants of environmental products and process innovation – evidence from the Mannheim innovation panel and a follow-up telephone survey. In: Hemmelskamp, J., Rennings, K. and Leone, F. (eds.), Innovation-oriented environmental regulation. Theoretical approaches and empirical analysis. Physica, Heidelberg: 331–347

Coase, R.H. (1960) The problem of social cost. Journal of Law and Economics 3: 1–44

Commission of the European Communities (2003) Communication from the Commission: The role of the universities in the Europe of knowledge. COM(2003) 58 final, Brussels

Commission of the European Communities, (2004) Management of intellectual property in publicly funded research organizations: "Towards European Guidelines". Working Paper (EUR 20915 EN), http://europa.eu.int/comm/research/era/ipr_en.html

Cooke, P. (1998) Introduction. Origins of the concept. In: Braczyk, H., Cooke, P., Heidenreich, M. and Krauss, G. (eds.) Regional innovation systems. The role of governances in a globalized world. UCL Press, London: 2–25

Cooke, P., Boekholt, P. and Tödtling, F. (2000) The governance of innovation in Europe. Regional perspectives on global competitiveness. Pinter, London

Dorfman, N. (1983) Route 128: the development of a regional high technology economy. Research Policy 12: 299–316

Drenth, P.J.D. (2002) Mobility and mobstacles in European science. Conference "Flexible Europe – Mobility as a tool for enhancing research capacity", September 2002, Tallinn, Estonia

Fischer, C., Withagen, C. and Toman, M. (2004) Optimal investment in clean production capacity. Environmental and Resource Economics 28: 325–345

Gale, R., Barg, S. and Gillies, A. (1995) Green budget reform. Earthscan, London

Gouldson, A. and Murphy, J. (1998) Regulatory realities: The implementation and impact of industrial environmental regulation, Earthscan, London

Griliches, Z. (1988) Hybrid corn: an exploration of the economics of technological change. In: Griliches, Z., Technology, education and productivity: Early papers with notes to subsequent literature. Blackwell, New York: 27–52

Hall, P. (1987) The geography of high technology: An Anglo-american comparison. In: Brothie, J., Hall, P. and Newton, P. (eds.) The spatial impact of technological change. Croom Helm, London: 141–156

den Hertog, P., Bergman, E. M. and Charles, D. (eds.) (2001) Innovative clusters: drivers of national innovation systems. OECD, Paris: 405–419

Hoel, M. and Shapiro, P. (2004) Transboundary environmental problems with mobile but heterogeneous populations. Environmental and Resource Economics 28: 265–272

Isik, M. (2004) Incentives for technology adoptions under environmental policy uncertainty: Implications for green payment programs. Environmental and Resource Economics 27: 227–246

Johansson, B. (2004) A menagerie of agglomeration and network externalities. CESIS Working Papers No 2, Centre of Excellence for Science and Innovation Studies, The Royal Institute of Technology, Stockholm

Karlsson, C. and Zhang, W. B. (2001) The role of universities in regional development: endogenous human capital and growth in a two-region model. The Annals of Regional Science 35: 179–197

Klemmer, P., Lehr, U. and Löbbe, K. (1999) Umweltinnovationen. Anreize und Hemmnisse. Innovative Wirkungen umweltpolitischer Instrumente, Band 2. Analytica Verlag, Berlin

Kline, S. J. and Rosenberg, N. (1986) Economic growth – the basis for any society's hope for the future. In: Landau, R., Rosenberg, N. (eds.) The positive sum strategy. National Academy Press, Washington: 1–16

Krugman, P.R. (1991) Geography and trade. Leuven University Press, Leuven

Küppers, G. and Pyka, A. (2002) The self-organisation of innovation networks: Introductory remarks. In: Pyka, A. and Küppers, G. (eds.), Innovation networks. Theory and practice. Edward Elgar, Cheltenham: 3–21

Marin, B. and Mayntz, R. (eds.) (1991) Policy networks. Empirical evidence and theoretical considerations. Campus, Frankfurt

de la Mothe, J. and Paquet, G. (1998) Local and regional systems of innovation. Kluwer Academic Publishers, Boston, Mass.

Pezzey, J. (1992) The symmetry between controlling pollution by price and controlling it by quantity. Canadian Journal of Economics 25: 983–991

Pezzey, J. (2003) Emission taxes and tradable permits – A comparison of views on long-run efficiency. Environmental and Resource Economics 26: 329–342

Pigou, A.C. (1932) The economics of welfare. 4th edition. Macmillan, London

Porter, M.E. (2000) Locations, clusters and company strategy. In: Clark, G. L., Feldman, M.P. and Gertler, M.S. (eds.) The Oxford handbook of economic geography, Oxford University Press, New York: 253–274

Putnam, R. (1993) Making democracy work: civic traditions in modern Italy. Princeton University Press, Princeton, N.J.

Pyka, A., Gilbert, N.G. and Ahrweiler, P. (2002) Simulating innovation networks. In: Pyka, A. and Küppers, G. (eds.), Innovation networks. Theory and practice. Edward Elgar, Cheltenham, UK: 169–196

Rallet, A. and Torre, A. (1998) On geography and technology: proximity relations in localised innovation networks. In: Steiner, M. (ed.), Clusters and regional specialisation. Pion, London: 44–45

Rosenkranz, St. (1996) Co-operation for product innovations. WZB Berlin, Berlin

Sampat, B.N. (2003) The effects of Bayh-Dole on technology transfer and the academic enterprise: a survey of the empirical literature (http://www.vannevar.gatech.edu/papers/bdsurvey.pdf)

Schmitz, H. (1995) Collective efficiency: growth path for small-scale industry. Journal of Development Studies 31: 529–566

Schmitz, H. (1999) Collective efficiency and increasing returns. Cambridge Journal of Economics 23: 465–483

Schrama, G.J.I. and Sedlacek, S. (2003) Environmental and technology policy in Europe. Technological innovation and policy integration. Kluwer, Dordrecht

Schubert, U. (1979) Environmental quality and urban land markets. University Microfilms, Ann Arbor, Mich.

Sedlacek, S. and Schubert, U. (2004) Inter-firm and inter- institution networks as the organizational basis of environment-oriented innovation projects. In: Johansson, B., Karlsson, Ch. and Stough, R.R. (eds.), Entrepreneurship and dynamics in a knowledge-economy. Routledge, in press

Stanners, D. and Bourdeau, P. (1995) Europe's environment – The Dobris assessment. European Environment Agency, Copenhagen, DK

Torero, M. (2000) Analyzing the spillover mechanism on the semiconductor industry in the silicon valley and route 128. Econometric Society World Congress 2000

Ulph, D. (1997) Environmental policy and technological innovation. In: Carraro, C. and Siniscalco, D. (eds.), New directions in the economic theory of the environment. Cambridge University Press, Cambridge: 43–68

Zucker, L.G. and Darby, M.R. (2001) Capturing technological opportunity via Japan's star scientists: evidence from Japanese firms' biotech patents and products. Journal of Technology Transfer 26: 37–58

Zucker, L.G., Darby, M.R. and Armstrong, J.S. (1998) Geographically localized knowledge: spillovers or markets? Economic Inquiry 36: 65–86

Zucker, L.G., Darby, M.R. and Brewer, M. B. (1998) Intellectual human capital and the birth of U.S. biotechnology enterprises. American Economic Review 88: 290–306

Zucker, L.G., Darby, M.R. and Torero, M. (2000) Determinants of embodied technology transfer from stars to firms. Working paper, UCLA, Anderson School

Zucker, L.G., Darby, M.R. and Torero, M. (2002) Labor mobility from academe to commerce. Journal of Labor Economics 20: 629–660

Springer-Verlag and the Environment

WE AT SPRINGER-VERLAG FIRMLY BELIEVE THAT AN international science publisher has a special obligation to the environment, and our corporate policies consistently reflect this conviction.

WE ALSO EXPECT OUR BUSINESS PARTNERS – PRINTERS, paper mills, packaging manufacturers, etc. – to commit themselves to using environmentally friendly materials and production processes.

THE PAPER IN THIS BOOK IS MADE FROM NO-CHLORINE pulp and is acid free, in conformance with international standards for paper permanency.