UNIVERSITÀ CATTOLICA DEL SACRO CUORE
MILANO

Dottorato di ricerca in politica economica
ciclo XXIV
S.S.D: SECS-P/06

Knowledge Spillovers, Externalities and Regional Economic Growth in the EU: Theories and Empirical Evidences

Tesi di Dottorato di Giovanni Guastella
Matricola: 3710485

Anno Accademico 2010/11

UNIVERSITÀ CATTOLICA DEL SACRO CUORE
MILANO

Dottorato di ricerca in politica economica
ciclo XXIV
S.S.D: SECS-P/06

Knowledge Spillovers, Externalities and Regional Economic Growth in the EU: Theories and Empirical Evidences

Coordinatore: Ch.mo Prof. Campiglio Luigi

Tesi di Dottorato di Giovanni Guastella
Matricola: 3710485

Anno Accademico 2010/11

Acknowledgements

To my Dears. You know who you are.

> ”Any pride I might have held in my conclusions was perceptibly lessened by the fact that I knew that the solution of these problems had almost always come to me as the gradual generalization of favourable examples, by a series of fortunate conjectures, after many errors.”
>
> -Hermann von Helmholtz-

This work represents the outcome of three years of research which have been spent in part in Italy, at the Department of Economics and Social Sciences of the Catholic University and in part in the Netherlands, at the Utrecht School of Economics as well as at the Department of Economic Geography of the Utrecht University. Any positive experience I have made during these years would have not been possible without the financial support from the Catholic University through a PhD Scholarship.

My first thank goes to Francesco Timpano and to Frank van Oort, the professors who have supervised my work in these years respectively in Italy and in the Netherlands. They have not only introduced me to their respective fields of interest, but have also led my research toward the most up-to-date issues in the academic and policy debate. I think I have enormously benefited of the friendly relationship I have been able to establish with both and that is the reason why I also would like to thank them for their genuine friendship. Thanks to that, office-hours discussions have frequently turned into *less formal* meetings at bars, pubs, restaurants, especially during conferences, and these meetings have been for *us* the occasion to exchange ideas and viewpoints, more than for *them* the occasion to correct the (many) errors and mistakes I have made. Goes to them my sincere gratitude also for having continuously inspired, motivated and challenged me with great enthusiasm, dedication and expertise. The many conferences and international events they encouraged me to attend have been great opportunities to receive important feedbacks as well as destructive criticisms which obliged me to carefully review what (I thought) I knew and to work harder and better to expand my knowledge.

During the PhD I have also benefited from the contribution of other professors and scholars. I am indebted with Enrico Fabrizi for his kind help with the understanding of important statistical issues. Thanks to him I have been able to glance books and papers the content of which was completely obscure to me before. I also want to thank Marc Schramm from the Utrecht University, with whom I had very useful discussions during my period in the Netherlands. I owe my gratitude to Roberto Basile for the many comments and suggestions I received from him and for whatever I have learned in the course of his lectures held at the University in Piacenza. In the end I would like to thank the two appointed referees for providing useful suggestions in the interpretation of conclusions.

Although the final version of this work has been kindly reviewed several times by my supervisors, by colleagues and by many other people as well, whose work is gratefully acknowledged, it was me collecting the data, running regressions, reporting results and writing references. And I am pretty sure that the final work still contains many errors, lacks and general omissions for which I should accordingly be considered the only responsible person.

As these three years have not been spent only by reading papers and collecting data, I would like to express my gratitude to all the friends who accompanied me during this experience. Among others let me mention Mario, Kim and Gabriele with whom I shared the office and much more, as well as Olimpia, Riccardo, Paola and Giuseppe for having made this experience not as bad as one may think it.

During these years I have received unconditional support from my family, to which I owe my sincere gratitude for having continuously been on my side, rejoicing for me and sustaining me with warm affection. I need to apologize with my girlfriend, Caterina, because the choice of a PhD has forced me to stay faraway from her, physically and too many times also mentally. The regret caused by such a long period of separation is only in part compensated by the wish of a near future in which we would be happy together.

Piacenza
February 27th 2012

Contents

Chapter I

Introduction

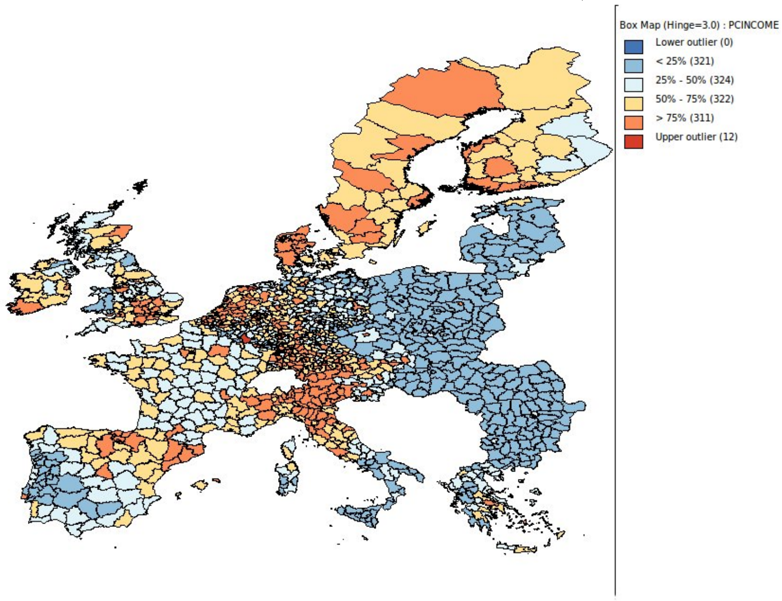
> ”Knowledge is the only competitive advantage of our times, it grows through open interaction with others.”
> -Ronald Coase-

In the introductory summary of the latest report on economic, social and territorial cohesion (EC [35]) smart, inclusive and sustainable growth is highlighted as the key objective of the Europe 2020 strategy. The principle according to which growth should be inclusive represents itself the core of the Cohesion Policy, aimed at reducing disparities between more and less developed regions. Inclusive further refers to the dissemination of the common market opportunities in lagging regions in order to boost the catch-up process of these. Meantime regional growth is also targeted as smart, provided that innovation and knowledge are addressed as the most influential factors promoting competitiveness.

To a certain extent the attributes of *smart* and *inclusive* might be considered as antithetical. As it is noted by Sharp [92] competitiveness and cohesion, respectively pinpointing the objectives of a smart and inclusive growth, both pertain the regional catch-up. However competitiveness is more concerned with economic growth in the European Union, wholly considered as a big player in the global arena, while cohesion is more with the development in lagging regions, to be achieved mainly by getting rid of existing structural barriers. However the extent to which policies pushing competitiveness by encouraging investments in smart-growth-oriented strategies might also benefit cohesion it still appears as an open issue. If it is, in fact, possible to agree that a productivity increase in lagging regions can benefit both the policy objectives, on the other side, a productivity increase in leading regions will likely improve the competitiveness of only these regions, making disparities to grow and, therefore, hampering economic cohesion.

As a matter of fact income is not evenly distributed across European regions. Although the cohesion policy has partially contributed in past years to promote higher productivity growth in lagging regions, yet unbalances look persistent and, in addition, they present a clear territorial character. The figure I.1 shows how the per-capita income[1] is distributed across NUTS III regions in the European geographical space. A clear core-periphery pattern is distinguishable, richer regions concentrating in the curved area covering London , Paris, the Benelux area and down to southern Germany and Northern Italy. So far the target of convergence is far from being reached and this naturally questions the effectiveness of policy actions intended for lagging regions.

Figure I.1: Spatial distribution of per-capita income - 2008 (Source: Eurostat)



More in general the issue concerning the effectiveness of policies intended to reduce territorial disparities have been recently brought at the hearth of an open academic and policy debate, opposing the so-called *spatially blind* approach proposed by the World Bank (World Bank [101]) to the *place-based* approach advocated by the EU's Barca Report (Barca [16]) and by two recent OECD publications (OECD [80] [79]). According to the World Bank chief economist Indermit Gill[2] the economic activity will be naturally unbalanced because of productivity gains which materialize at the local level as a consequence of agglomeration economies. What it is possible to observe in reality is that, in fact, the world is not flat. On the opposite it is and it will continue to be *spicky*[3] and policies oriented toward the reduction of territorial disparities, by promoting shifts from the *unbalanced* equilibrium, are not only ineffective but also inefficient provided that important resources tare subtracted to the competitiveness objective. By the opposite view, Garcilazo and Oliveira Martins[4], among the proposers of the place-based approach to regional policy, emphasize the need for territorial policies aimed at giving equal op-

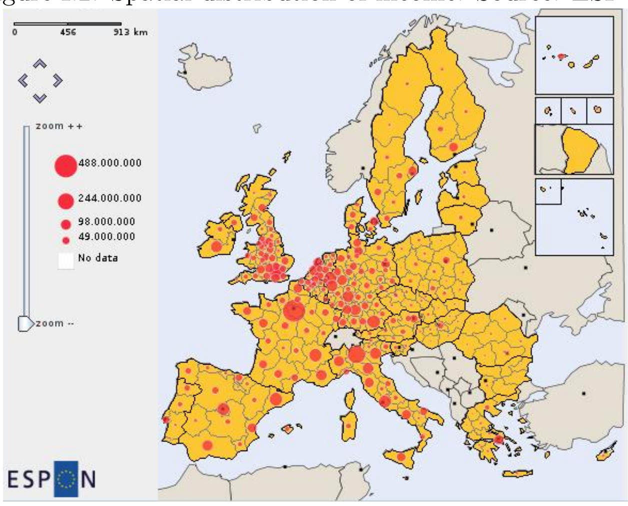
[1]Income is defined as Gross Domestic Product in PPS. Year 2008.
[2]http://www.voxeu.org.
[3]The definition has been coined by Richard Florida by using spikes to represent densities in maps.
[4]http://www.voxeu.org/index.php?q=node/5827.

portunities to all citizens by eliminating the barriers that limit the growth potential at the regional level. In arguing so they claim that agglomeration is neither necessary nor sufficient for growth and, indeed, the evidence on agglomeration and regional growth is rather mixed: not all the peripheral regions are underdeveloped while only some agglomerated regions are well-performing. This, in turn, also suggests that the benefits of agglomeration are bounded and, above a certain point, dis-economies and congestion start to erode the benefits of agglomeration.

Europe is not flat

By looking at European data[5] it immediately appears that income is highly concentrated is some particular regions. In 2008 the 87.46% of the income (GDP) has been produced in the top 25 regions (NUTS II) and the same regions did produce 86.21% of the income in 1995, suggesting income concentration has also increased over time. In addition there is evidence that these top regions are co-located in the central part of the Europe, as it is shown in figure I.2.

Figure I.2: Spatial distribution of income. Source: ESPON



By focusing on regional growth the evidence is more controversial. Higher growth rates have been registered in less developed and more peripheral regions as it is also discussed in the Commission Report (EC [35]). Growth in the periphery is closely linked to the productivity increase in these regions, achieved mainly through the restructuring of economic systems and the consequent shift toward more productive sectors. Nonetheless the data suggest that the same top 25 regions account for the 90% of change in the level of EU GDP between 1995 and 2008, reinforcing the perception that core EU regions also represent the engine of European growth. According to the Commission Report (EC [35]) growth in these regions was boosted by innovation.

Provided that the European Economy is characterized by such large differences, that of coupling the policy targets of cohesion and competitiveness might reveal as a

[5] EU25 excluding Denmark. Source: Eurostat.

vain policy strategy. On the one side investments facilitating the restructuring of lagging economies do contribute only a little to the overall EU growth. On the other side investments in the innovative capacity of regions can considerably sustain the future growth of the EU but, admittedly, these investments are likely to benefit more in leading regions. Should European policy-makers abandon the cohesion objective, accordingly? Not necessarily. The World Development Report answer to such policy issue focuses on economic integration between leading and lagging regions, to be achieved not only by infrastructural improvements but also and even by institutional enhancements. Considering this viewpoint there is no space left to territorial development policies.

The aim of the work

This work aims at demonstrating that space matters. It does for activities concentration, as maps and figures show us. But it does also for innovation and human capital. And since innovation and human capital are the engines of regional productivity, it does for growth. And for policy. In an attempt to answer the previous question this work suggests that the reason for the persistence of income disparities between leading and lagging regions in Europe has to be researched not in the concentration of economic activities nor in the presence of agglomeration forces, but in the clear gap in innovation and human capital which characterizes the periphery of Europe.

The evidence reported in the very first part of the work demonstrates that the concentration of economic activities is actually neither necessary nor sufficient for long-run regional economic growth in Europe. Growth is, on the contrary, induced by the accumulation of human capital and by the creation of knowledge, both inducing significant externalities. To what extent these externalities may also benefit least developed and more peripheral regions, it depends on the mechanisms through which knowledge spreads. These mechanisms are indeed carefully examined in the remaining of the work. In greater detail three issues are investigated.

The first issue is to what extent regional economic development is driven by either increasing or decreasing returns and, more specifically, what are the most important factors of increasing returns. An empirical framework is adopted which extends the traditional model of regional convergence (Barro [19]) by incorporating information on the presence of increasing returns related to agglomeration, human capital and innovation. Improving on the existing literature the empirical formulation adopted in this allows for non-linearly shaped patterns relating these variable to regional growth. The pictures which comes out from the empirical results is quite complex. It appears that a process of convergence exists only between lagging regions, while growth in more developed regions is clearly related to the accumulation of innovation and human capital and, more in general, of knowledge. Moreover the detected non-linearities suggest that there is no clear contribution of agglomeration on regional growth while the positive effects of innovation and human capital are subject to relevant thresholds.

The second issue is the clustering of innovative activities. In an attempt to explain the spatial concentration of innovative firms the *Geography of Innovation* literature (Audretsch and Feldman [10]) has highlighted the role of *Localized Knowledge Spillovers*

(LKS), intended as involuntary transfers of knowledge. Spillovers cause social returns from investments in knowledge to be higher than private returns and are supposed to be localized as a consequence of knowledge stickiness. Hence the presence of LKS might increase the productivity of investments in knowledge in presence of co-localization of innovative firms. Estimation of the degree of spillovers localization turns out to be important to determine, at the regional level, to what extent investments in some (core) regions might benefit other regions (in the periphery). Using a spatially extended Knowledge Production Function approach it is shown that evidence of interregional spillovers in Europe weakens and even disappears once other factors are considered in order to account for the path-dependent and place-dependent nature of innovation. Accordingly spillovers, if any, are probably extremely localized and thus do not cross regional borders. More important the access to markets seems to be the driving force behind the clustering of regional innovative firms.

The third issue is the role of knowledge infrastructures in diffusion of innovations. A large literature based on the seminal work by Jaffe [56] has investigated the contribution of universities to the creation and diffusion of knowledge at the regional level and the important implications of this for regional development. The work improves on this literature by considering the role of Knowledge Intensive Business Services (KIBS), characterized as a second knowledge infrastructures beside universities and public research institutions. An empirical model is developed which explains the intensity of regional innovative activity as a function of the amount of knowledge in the regional economy. In defining the latter the influence of research investments by private firms is considered as well as the contribution of universities and KIBS. The empirical evidence indicates that the productivity of research investments is specially high in those regions where scientific universities are located and, by the opposite, the contribution of KIBS turns more important in regions where there are no scientific universities.

A global look at the empirical results which will be presented in the next chapters unfolds three most important facts. The first is that the higher growth in least developed region, which can be ascribed to the industrial restructuring, is important but it will be not sufficient for the catch-up in the long-term. Growth, in the long-run, is determined by investments in human capital and knowledge, which produce more than proportional returns in terms of income and productivity. The existence of important threshold effects implies that regions need to fill a technological gap in order to catch-up. The second is that such a technological gap is not as easy to be filled. Knowledge externalities, the essence of the more than proportional returns from investments in innovation, are extremely localized and are more likely to be the effect of the spatial concentration of innovative firms more than the cause. Nonetheless aggregate regional innovative activity appears to be driven by market forces as well, probably as innovative firms may want to locate in regions where they can get the highest market value from innovation. Accordingly improvements in physical infrastructures and access to markets might characterize pathways to improve the regional innovative capacity. The third is that knowledge is not only produced through investments in research. Of course investments in research are important for firms to internalize the global knowledge, as that produced by universities and public research institutions. However local knowledge is also important and knowledge intensive services considerably contribute to the creation of a local knowledge base. Thus the promotion of economic restructuring of lagging regions toward a knowledge-based economy should deserve special attention to KIBS and, more in general, to the service-driven innovation which can be easily accessed by firms without requiring large and structured investments in research.

Place-based Policies for a smarter growth

Can EU regional growth be smarter? Yes it can by improving the regional innovative capacity especially in those regions presenting innovation gaps. This is necessary to allow these regions to enter in the increasing returns phase of development and to sustain their growth after the *convergence push* will vanish. The target cannot however be achieved by just promoting higher regional integration and institutional improvements as innovation is strongly path-dependent and place-dependent and filling the gap requires a deep understanding of the forces that, at the territorial level, have so far impeded the full development of a knowledge base. By the opposite a spatially targeted design of policy seems the most reasonable approach to provide local-policy-makers with the right instruments to remove the barriers to the development of knowledge and innovation.

Chapter II

Increasing Returns, Decreasing Returns and Regional Economic Convergence in the EU

II.1

Non Technical Summary

Regional economic development is driven by the accumulation of production factors. More traditional factors like labour and physical capital are accumulated under the law of diminishing returns. This, in turn, allows less developed regions to better perform. Recent branches of theoretical and empirical literature have paid attention to the role of increasing returns in an attempt to explain the persistence in regional economic disparities. Increasing returns are commonly attributed to either the accumulation of non traditional inputs such as human and knowledge capital or the presence of local externalities generated by the spatial concentration of economic activities. In this work the economic performance of 186 European regions is analysed by using the ordinary growth regression approach. An empirical specification which simultaneously accounts for the presence of both decreasing and increasing returns is derived. The study is intended to examine the extent to which regional development originates from the (un)balance between convergence, driven by diminishing returns and divergence, boosted by increasing returns. Results indicate that the accumulation of traditional inputs leads the economic development of less favoured areas while the presence of increasing returns plays a more crucial role in developed regions. Furthermore the use of a non-linear specification for the growth equation highlights evidence of important threshold effects in entering the stage of development characterized by increasing returns. Regional development process is accordingly depicted as a far more complex process than what the simple dualism between increasing and decreasing returns may help to figure out, with very important implications for policy.

II.2

Introduction

Models of regional economic growth have traditionally emphasized the hypothesis of diminishing returns to labour and physical capital to interpret the evidence of an higher productivity in less developed regions. Such scale-related productivity decline is expected to conduce, in turn, to an income convergence between regions in the economy. The growth regression framework proposed by Barro [19] is considered as the workhorse of the empirical literature testing the hypothesis of regional convergence. In short the annual average growth rate of per-capita income over a certain period is regressed on the initial income level. The relation is expected to be negative and a significant value of the estimate corroborates the theoretical hypothesis according to which all regions will converge to the same per-capita income level in the long run (Barro and Sala-i-Martin [18]).

As Martin and Sunley [75] note, the neo-classical approach presents several shortcomings. At the theoretical level the hypothesis of diminishing returns seems to be a very restrictive one. At the empirical level the estimated value of the so-defined *speed of convergence*, the rate at which disparities annually decrease, is found to be quite small (around 2%) and the amount of regional growth which is left unexplained by the model is also very high. Making a step ahead, endogenous growth theories (Romer [91], Lucas [72]) have extended the neo-classical growth model releasing the assumption of decreasing returns in production and, as a consequence, have actually solved much of these shortcomings.

However, when the growth rate is endogenously determined within the model, the prediction about the long-run equilibrium completely differs. While economic convergence, in either its absolute or conditional form, is the equilibrium associated to the hypothesis of decreasing returns, divergence is predicted in presence of increasing returns. The understanding of the extent to which the regional development is driven by either decreasing or increasing returns thus appears as a key issue, especially in Europe.

In Europe regional convergence is expected to take place as a result of not only the integration process, but also and even as a consequence of the large investment programs granted under the Cohesion Policy to boost growth in less developed regions. There are several studies that, indeed, have empirically investigated convergence at the European regional level taking into consideration possible sources of increasing returns. In the work by Ertur and Koch [37] the role of human capital accumulation is emphasized. Similarly, Rodriguez-Pose and Crescenzi [90] study the effect of innovation on growth through investments in research and development. Some studies have also focused the attention on agglomeration economies, like in the case of Bosker [23] and, more in general, on interregional spillovers (Dall'erba and Le Gallo [33]). In all of the mentioned studies there is evidence that, notwithstanding the convergence process, regional growth is affected by the presence of increasing returns in the production.

In the present work a similar approach is adopted and regional growth is studied by using the growth regression framework. Alongside the standard convergence hypothesis the existence of increasing returns is also accounted for in the model specification and tested upon a sample of 186 regions in the period 1995-2007. Building on the existing theoretical and empirical literature, three main determinants of increasing returns are identified, namely the orientation of the regional economy toward innovation, the importance of human capital and skilled workers in the production and, lastly, agglomeration economies. However, differently from the existing empirical literature, these determinants are concurrently related to regional growth.

The model specification further allows for non-linearities in the relations. This, in turn, permits to evaluate the contribution of both decreasing and increasing returns in the different stages of development. The results indicate that a process of economic convergence drives regional growth in less developed areas more than in already developed ones where, by the opposite, production is characterized by increasing returns. More specifically the agglomeration externalities positively contribute to regional growth in only very agglomerated regions and the positive effects of innovation and human capital are noticeable only over a thresholds of, respectively, regional innovative capacity and presence of skilled workers in the economy. These evidence have some very important policy implications. It is shown that regional development is determined by the composition of several factors and that the contribution of each varies along the development path of the region. Accordingly regions follow different development trajectories and, thus, the *"one size fits all"* policy approach to regional development proves to be inappropriate. On the contrary, more attention is claimed toward more place-based approaches.

The remaining of the chapter is organized as follows. In the next section the theoretical and empirical literature on the relation between increasing returns and regional growth is reviewed. The various determinants of increasing returns are discussed and, for each, the issue of non-linearity is addressed. In section three the dataset is presented and three synthetic measures for the determinants of increasing returns are derived by using multivariate data analysis. The empirical model and the results are presented in section four. Follow conclusion.

II.3

Sources of Increasing Returns

The convergence debate has been dominated for decades by the Barro-type regression paradigm (Barro [19]). Such an empirical framework is directly derived from the neoclassical growth model described by Solow [95] in which, under the hypothesis of perfect competition, homogeneous agents and diminishing marginal returns, it is shown that economies follow a path toward a steady-state per-capita income level. The far away from the steady-state, the higher the rate at which the economy grows. Provided that economies have similar structural characteristics, they are expected to converge toward similar steady-state income levels. The empirical test is based on a cross-country or cross-region regression of per-capita income growth rate over a given time period on the initial level of per-capita income. A negative and significant coefficient related to the initial income is perceived as evidence of convergence.

In a series of articles Quah ([88], [89]) has criticized such an approach to the empirical test of the convergence hypothesis, arguing that the approach proves inadequate to explain the persistence or, in same cases, the time-increase in the level of per-capita income disparities, despite the evidence of convergence. Likewise it is argued that, notwithstanding the higher growth in poorer areas, economies not necessarily converge toward the mean of the distribution. By the opposite, the long-run income distribution might be characterized by bi-modality.

Among the theoretical hypothesis behind the Solow-Barro framework, the one on the diminishing marginal returns of factor inputs has been pointed as the most unrealistic. More specifically, recent branches of literature have emerged releasing the assumption of diminishing returns and predicting non-converging long-run scenarios. This is the case of the New Growth Theory[1] (NGT) and of the New Economic Geography[2] (NEG) as well. Models belonging to the first of the two branches of literature emphasize the importance of production factors like human capital and knowledge capital which are capable to determine increasing returns to scale in the economy. In models belonging to the second branch, increasing returns are associated with the presence of pecuniary externalities arising from the spatial concentration of economic activities. For both, the predictions about long-run equilibrium are similar. Economies will diverge and the long-run distribution of per-capita income will be characterized by club-convergence[3] as well as by core-periphery patterns[4].

Consequently this more recent literature proves to be useful in explaining the empirical evidence of bi-modality suggested by Quah [88]. In what follows this literature will be reviewed paying special attention to how the hypothesis of increasing returns is, on the one side, incorporated in the theoretical modelling framework and, on the other side, empirically tested.

Human Capital and the Knowledge Economy

The contribution of human capital to economic growth has been highlighted in the work by Mankiw et al. [73]. Using cross-country data the authors find that human capital can actually explain a large part of between-countries variation in the rate of economic growth. The evidence of a better performance in regions well-endowed of human capital

[1] For a comprehensive review of the literature see the work by Martin and Sunley [75].

[2] See Krugman [62].

[3] Galor [45] extensively discusses the implication of different theoretical growth models on the convergence hypothesis.

[4] An example of theoretical model of endogenous growth integrating NEG is provided by Baldwin and Forslid [15]. Consistently with more generic models of NEG, the long-run equilibrium is characterized by core-periphery patterns.

is ascribed to the prominent role of knowledge. Knowledge is, in fact, embedded in people and not necessarily shows a decreasing marginal productivity. On the contrary, more people working together entail an easier exchange of ideas, experiences and good practices as well, resulting in higher productivity. Accordingly formal models of endogenous growth based on knowledge (Romer [91] and Lucas [72]) assume its marginal productivity to be increasing by allowing knowledge-related externalities to grow with the stock of knowledge. The outcome associated to the predicted model equilibrium is distant from the convergence predicted by the Solow model as, conversely, knowledge can continuously increase generating persistent disparities between the economies.

The empirical test to assess the contribution of human capital on regional growth is based on an extension of the growth regression which includes a measure of human capital. In the study by Lòpez-Rodriguez et al. [71] a survey of the literature is provided together with a critical assessment of the measurement problems. At the EU regional level there is evidence that the long-run equilibrium level of the regional economy is strongly influenced by human capital. Tondl [98] argues that differences in human capital endowments are responsible for the persistence of the disparities between less developed European regions in the south and more developed northern regions. A similar conclusion is indicated also in the study by Badinger and Tondl [14] and by Paci et al. [81] as well.

In a recent contribution, Basile [20] has found evidence that the effect of human capital on regional growth turns positive only after a certain threshold of human capital is passed and that the same effect is larger in region surrounded by neighbouring regions with high levels of human capital as well. According to the author, such an evidence of a non-linear effect is consistent with some theoretical models, as for example that developed by Azariadis and Drazen [12], in which social returns from human capital investments (externalities) appear only after a certain threshold of human capital is passed.

Innovation

Knowledge is not only embedded in people. The part of it which can be codified and formalized materializes in new products and processes. At the heart of the endogenous models of growth based on innovation (Aghion and Howitt [2]), it lies the hypothesis that these new products and processes give the firm a monopolistic power into the market. Increasing returns thus come from innovative activity which, in turn, is the result of specific investments made by the profit-maximizing firm. As a consequence, the growth pattern of the region might be importantly shaped by the relative efforts put by firms in the activities of research and development.

Fagerberg and Verspagen [38] have tested this hypothesis empirically on a sample of European regions, assuming that the the technological gap, measured by mean of R&D-related indicators, explains the persistence of disparities in per-capita GDP. It is shown that the inclusion of R&D in the specification contributes to improve the model fitting and to explain the regional variation in per-capita income growth. A similar framework is also used by Fagerberg et al. [39], who provide analogous evidence but based on a different sample of regions. In a more recent past other studies have investigated the issue using larger samples of regions and more up-to-date datasets as well. In the study by Rodriguez-Pose and Crescenzi [90], grounded on the sample of all the regions of EU25, it is found evidence of convergence, with a clear positive contribution of innovation to regional growth. Likewise Sterlacchini [96] and Verspagen [99], among others, reach to the same conclusion.

At both the theoretical and the empirical levels there are however arguments suggesting that innovation non-linearly relates to growth. Technological change is in fact influenced partly by new innovations and partly by imitations and it is likely thus to be higher in regions with an already significant knowledge base. As it is claimed by Cohen and Levinthal [29] , not only the probability to realize a new innovation but also the probability to successfully replicate an existing innovation positively depends on the level of investments in research. To some extent, it can be argued that R&D investments are necessary to innovate and also represent a pre-condition to imitate (Fagerberg et al. [39]). Shifting this argument to the regional growth and convergence debate, evidence is expected to reveal a slower technological catch-up in less technologically developed regions. Equally, in presence of a wide technological gap, some regions might not catch-up at all. As a matter of fact empirical studies have found evidence of non-linearity and threshold effects in the relation between growth and innovation. For instance such a result is indicated by Fagerberg and Verspagen [38], Sterlacchini [96] and Crescenzi [32]. Very recently, the hypothesis that a lower technological gap facilitates the absorption of new innovation has been included into a model of regional growth which, consequently, predicts club-convergence (Alexiadis and Tomkins [3]). The evidence in the paper supports the theoretical hypothesis.

Agglomeration

Agglomeration economies are at the origin of NEG models (Krugman [62], Krugman and Venables [64]). Externalities arise in presence of multiple co-location of economic activities and are characterized as pecuniary externalities. More precisely they are related to labour market pooling. Manufacturing goods are produced under a Dixit-Stiglitz monopolistic competition framework with scale economies and, hence, the higher is the concentration of economic activities in the area, the higher will be profits for each single firm. The long-run equilibrium is determined by two forces: agglomeration economies boost divergence and high transportation costs promote spreading. Given an initial even distribution of economic activities across regions/countries and high transportation costs, once the latter start declining, it becomes more and more convenient for firms to co-locate in one area to benefit from agglomeration economies.

The original Krugman's framework has been re-adapted by many scholars attempting to accommodate the study of specific cases. Among others, the Krugman and Venables

[63] model is an example of NEG model which interprets the process of European integration and the related decline in transportation costs consequent to the abolition of trade barriers between member states. Empirically, the predicted core-periphery pattern in the spatial distribution of economic activities seems capable to explain the geographical shape of the production in Europe. In their exploratory spatial data analysis of production and income in EU regions, Le Gallo and Ertur [67] provide robust evidence of the regional concentration of activities and of a core-periphery pattern as well. A first attempt to measure the effect of agglomeration economies on regional performance has been made by Ciccone [28], relating total factor productivity to employment density, a standard measure of agglomeration. The effect of agglomeration is positive and sizeable but the analysis, in this specific case, is not further extended to regional growth. The effect of agglomeration economies on regional growth is conversely studied by Bosker [23] for a sample of 208 EU16 regions over the period 1977-2002, differentiating the internal, within the region, effect from the external, between regions, effect. It is found that, for both, the effect is negative. More densely populated regions have lower growth rates and being located nearby other densely populated regions also has a negative impact on growth. Interpretation of this result is straightforward. The negative effects of agglomeration, for instance dis-economies caused by either congestion or high house prices, are, on average, larger than the benefits of agglomeration. As one cannot assume that the agglomeration effect is continuously negative, a natural question arises on what is the critical level of agglomeration at which dis-economies start prevailing on economies. An issue which, according to Bosker [23], is not easy to disentangle.

A Comprehensive Framework

Different attempts have been made to develop empirical models which include testable hypothesis on the effect of innovation, human capital and agglomeration on regional growth. Most of the works surveyed in this section focus on each single determinant of increasing returns, and none of them has considered all the determinants simultaneously. This is probably the consequence of the lack of a theoretical background pinpointing the way externalities from the accumulation of innovation and human capital and externalities from the concentration of economic activities relate to each other. One possible interpretation of this relation lies in the concept of *knowledge spillovers*. Knowledge, in theoretical models and in the reality as well, is classified in two broad categories, explicit and tacit, the second being transmitted exclusively via face-to-face contacts and frequent interactions (Von Hippel [100]). Knowledge externalities are therefore likely to be bounded in space although there is no reason to believe that they cannot cross the regional administrative boundaries and spatial externalities could be accordingly ascribed to localized knowledge spillovers more than to simple agglomeration.

Building on this perspective different studies have applied spatial econometric techniques to the regional growth equation interpreting the evidence in view of spillovers between neighbouring regions (Lopez-Bazo et al. [70], Le Gallo et al. [66], Badinger et al. [13], Erthur and Koch [37], Dall' Erba and Le Gallo [33], Guastella and Timpano [53]). It is however worth noting that localized externalities due to knowledge spillovers are not related to NEG models provided that the latter only account for, as already remarked, pecuniary externalities. Spatial econometric extensions of the growth regression at the regional level thus only in part account for externalities, unless agglomeration economies are not explicitly included.

This work contributes to the existing empirical literature by proposing an unified framework in which the hypothesis of convergence related to diminishing returns is tested jointly with the hypothesis of increasing returns related to the accumulation of human capital and innovation and to the regional concentration of economic activities. Furthermore, spatial heterogeneity and spatial externalities are separately considered.

II.4

Data

All the data used in this work come from the Eurostat regional database. The sample under study is composed by all the regions belonging to countries in the EU25 group. Regions are defined based on the NUTS classification and, for all the countries but Belgium, Greece, Germany and the UK, for which the level I has been taken as reference, the level II is used. The choice to rely on the statistical level I for the four aforementioned countries is motivated by the availability of some of the data at only this level. More generally, for the same countries, the statistical level I seems to be more important than the level II for the definition of relevant administrative units. Overall, the regional classification used here is very close to that used by the OECD[5] in the definition of the territorial level T3.

By following the theoretical literature presented in the previous section, it is derived a list of relevant variables which can proxy the presence of increasing returns at the regional level[6]. Variables are described in table II.1

Admittedly, most of these variables show high correlation between them and the possible collinearity prevents the use of them all in a regression framework. Such correlations are detected by using factor analysis, on the base of which four factors are obtained. Correlations of these factors with base variables are summarized in the table II.2. All together the four factors explain 73.4% of the total variance in the data.

The first factor is highly correlated with *KIS* and *HRST*. High sores in this factor thus indicate a service-based regional economy with a production system prominently

[5] For more information on the territorial classification adopted by OECD please refer to the following documentation: http://www.oecd.org/dataoecd/35/60/42392313.pdf.

[6] In the growth regression framework investments and population change are usually included as controls. However, given the cross-section nature of the dataset, the inclusion of these variables might have produced simultaneity bias in the estimates. According to Grossman and Helpman [52] investments tend to follow GDP growth more than the opposite and, moreover, Fagerberg and Verspagen [39] have shown that differences in physical capital accumulation do not explain regional variation in per-capita GDP. Likewise Fagerberg and Verpagen [38] show that population growth at the regional level is driven by migration flows which, in turn, depend on the economic opportunities in the destination region.

Table II.1: Description of Variables

variable	description
RED	percentage of research expenditure made by both private firms and public institutions located within the region relative to the regional Gross Domestic Product (average in years 1997-1999)
PA	number of applications for patents made at the European Patent Office divided by the number of inhabitants of the region (average in years 1997-1999)
KIS	share of workers in Knowledge Intensive Business Services relative to the total number of workers in all NACE activities (average in years 1997-1999)
HTM	share of workers in High and Medium-High Tech Manufacturing relative to the total number of workers in all NACE activities (average in years 1997-1999)
HRST	percentage of regional population employed in Science and Technology (average in years 1997-1999)
ROAD	total number of kilometres which compose the road network of the region (year 2000) divided by the area of the region in square kilometres
INTERNET	percentage of households having access to internet (average in years 2007-2009)
EMPD	employment density, measured as the ratio between the the number of employees (average in years 1997-1999) and the area of the region in square kilometres

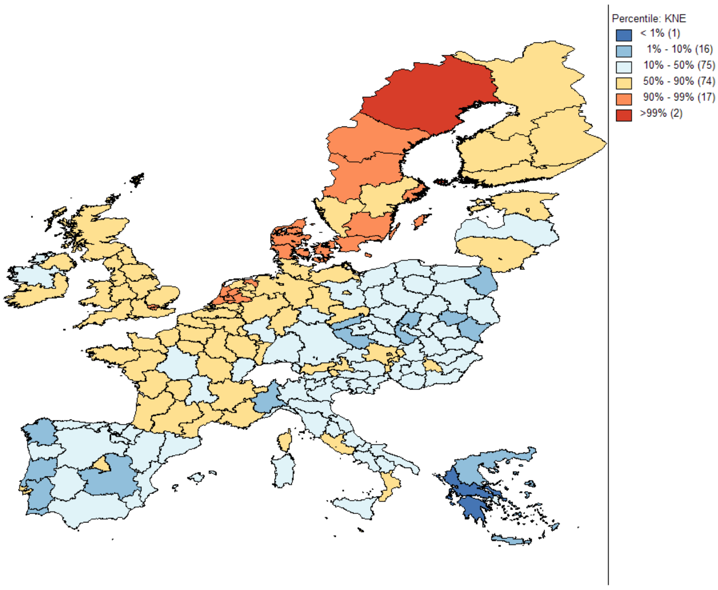
Detailed definitions of NACE activities are provided in the appendix.

Table II.2: Factor Analysis - Varimax Rotation

	Factor 1	Factor 2	Factor 3	Factor 4
RED	0.354	0.846		-0.200
PA	0.385	0.733		
KIS	0.896	0.208	0.267	-0.278
HTM		0.555		0.103
HRST	0.637	0.381	0.288	
ROAD	0.300		0.505	
INTERNET	0.834	0.398	0.116	0.357
EMPD			0.993	
Proportion	0.284	0.239	0.179	0.032
Cumulative	0.284	0.523	0.702	0.734

oriented to knowledge. The high correlation of the factor with the *INTERNET* variable also indicates that the production in high-scoring regions is grounded on a good ICT network infrastructure. For this reason the name of knowledge economy (*KNE*) is attributed to this factor. Its spatial distribution is shown in the figure II.1 and it appears that regions reporting the highest scores are spatially concentrated in the nort-western part of Europe and mostly in Scandinavia.

Figure II.1: Spatial distribution of KNE - percentiles



The second factor is highly correlated with *RED*, *PA* and *MHT*. To this factor it is attributed the name of innovation (*INNO*) as high-scoring regions are characterized by a large use of innovative inputs, both in terms of labour and investments, and a large production of innovative output as well. The spatial distribution of this factor is shown in the figure II.2. It is characterized by a generic core-periphery structure centred on the region of Baden-Wurttenberg. High scores in less central areas are also recorded in the Swedish region of Vastsverige, in East England and, to a lower extent, in Paris and in the Dutch region of Noort Brabant.

The third factor shows high correlation with the *EMPD* variable and with the *ROAD* variable. Accordingly, high scores pinpoint agglomerated regions and the name attributed to the factor is agglomeration (*AGG*). The spatial distribution of this factor, in figure II.3, has, however, a pattern which is different from the expected core-periphery one. It does not surprise that very high scores are registered by the capital regions in the majority of the member states. Nonetheless, according to the indicator, some of the more agglomerated regions appear to be in the eastern part of the Europe, especially in Poland and Chez Republic. On the contrary Spanish and French regions, they are accounted as non agglomerated.

Figure II.2: Spatial distribution of INNO - percentiles

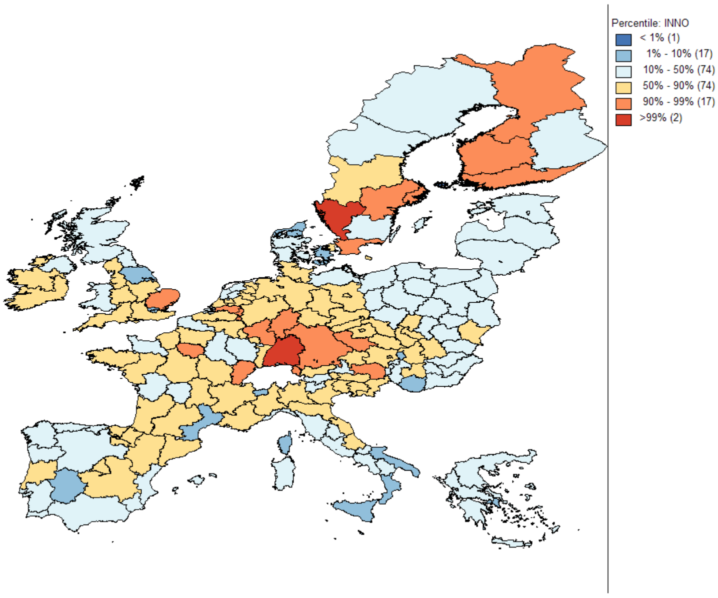
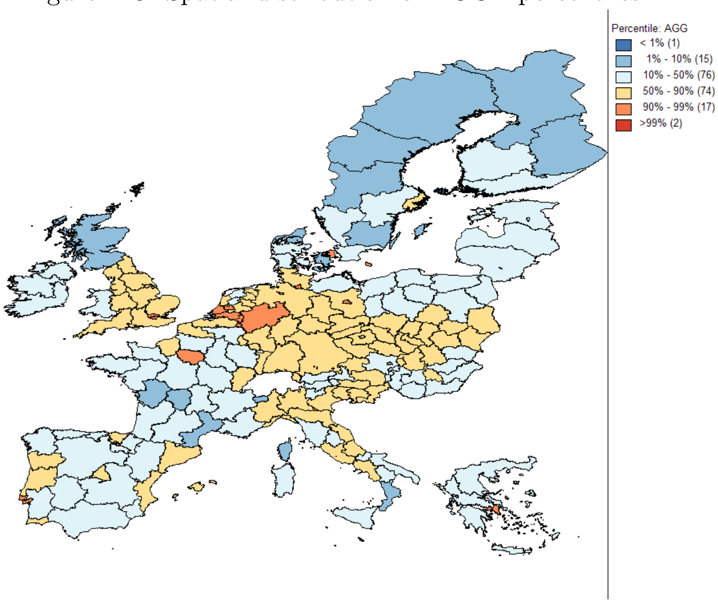


Figure II.3: Spatial distribution of AGG - percentiles



II.5

Empirical Model and Results

The empirical analysis starts by estimating the standard growth equation for the sample of regions, and further adding the three measures derived before to the model (equation II.1). The per-capita Gross Domestic Product[7] (Y_i) of the region is used to measure the regional output and the period under study is that from the year 1995 (t) to the year 2007 ($t+T$).

$$\frac{1}{T}\log\left(\frac{Y_{i,t+T}}{Y_{i,t}}\right) = \alpha + \beta\log\left(Y_{i,t}\right) + \gamma_1 AGG_i + \gamma_2 KNE_i + \gamma_3 INNO_i + \varepsilon_i \tag{II.1}$$

As it is usual, the β coefficient is expected lower than zero. This implies that, as a consequence of the diminishing returns, poorer economies have higher growth rates. On the opposite, the values of γ_1 γ_2 and γ_3 are expected to be positive, so that higher regional growth might be related to the presence of increasing returns. Estimates using this linear specification are summarized in table II.3. In the first three columns of the table estimates have been reported for the models with each factor added separately. In the last column, the three factors are included jointly.

Table II.3: Growth Regression - OLS Estimates

	(1)	(2)	(3)	(4)
Intercept	0.122^{***}	0.161^{***}	0.125^{***}	0.186^{***}
	(0.015)	(0.017)	(0.016)	(0.019)
log(gdp)	-0.010^{***}	-0.014^{***}	-0.011^{***}	-0.017^{***}
	(0.002)	(0.002)	(0.002)	(0.002)
agg	0.001			0.002^{**}
	(0.001)			(0.001)
kne		0.005^{***}		0.006^{***}
		(0.001)		(0.001)
inno			0.001	0.003^{**}
			(0.001)	(0.001)

Notes to table II.3:
SE in parenthesis.
***, ** and * indicate significance at 1%, 5% and 10% confidence levels.

The estimated coefficient related to the initial income is always correctly sloped and highly significant. Its value ranges from -0.017 to -0.010, coherently with previous results in the empirical literature on European regions. Differently, the coefficients on the three factors have positive slopes but, exception made for the factor which interprets human capital and the knowledge economy, they are not significant when considered alone. Nonetheless, they turn out to be strongly significant when considered together in the fourth column. Among all of them, the coefficient related to *KNE* is the largest in magnitude.

[7] Milions of Euro at 2000 prices.

The issues of spatial dependence and non-linearity are further introduced into the analysis. A first attempt is made by estimating a Spatial Error Model (SEM) specification of the equation II.1 with the interaction terms between the logarithm of the initial income and each of the three factors. The choice of the SEM is made on the base of a battery of tests for spatial dependence on the residuals obtained from estimates reported in the column (4) of table II.3. The results of spatial autocorrelation diagnostic tests and of the SEM estimates are reported in the appendix and will not be discussed here. The choice is motivated by the evidence that the SEM specification with interaction terms, although it appears very effective in accounting for spatial relations between units, it also shows weaknesses in accounting for non linearities. Instead, a more flexible semi-parametric specification, firstly applied to the study of regional growth by Basile [20], is preferred. Covariates are introduced as smooth terms into the model formulation and the resulting Generalized Additive Model (GAM) is estimated with the methodology suggested by Wood [103].

Differently from Basile [20], however, spatial relations are taken into account by either including a spatial trend into the model or by using Moran Eigenvectors approach. The choice implies that the empirical model is basically specified as a non-spatial model, to which spatial heterogeneity and spatial relations are added only in a second step. Thus, no a-priory assumptions are made concerning the contribution of interregional externalities to the regional growth. The spatial trend is added to the model as a smooth spline of the geographical coordinates. This seems to be the most suitable choice to handle spatial heterogeneity in a GAM framework, since the same methodology (smooth splines) is used to account for both non linearity and spatial relations. Instead, the Moran Eigenvectors approach (Griffith and Peres-Neto [49]) entails the inclusion of suitable eigenvectors extracted from the contiguity matrix so that any spatial dependence present in OLS residuals[8] is moved into the model (Bivand et al. [21]). It is worth noting that both the approaches, differently from many others spatial regression approaches, permit to include a spatial structure directly into the deterministic part of the model, and not in its random part.

The result of the GAM model are summarized in the table II.4. The simplest model is estimated excluding the spatial component *(a)* from the model specification. It follows the model with the spatial trend *(b)* and that with the spatial filter *(c)*. Significance of each smooth term is evaluated through the value of the related F statistic reported in the table. In all the three models the smooth terms are strongly significant. In the model with the spatial trend, the $s(x, y)$ terms, identifies the smooth term relative to, jointly, latitude and longitude. Finally in the model with spatial filter, the filtering methodology has identified eighteen eigenvector. For the sake of simplicity the related coefficients and statistics have not been reported.

Goodness of fit is assessed by looking at the values of the adjusted R^2, at the per-

[8] The procedure works in two steps. In the first the eigenvectors are selected which minimize the residual autocorrelation of the linear model with the inclusion of covariates. In the second the eigenvectors are included in the non linear model specification.

centage of the deviance explained and at the GCV score[9]. Moreover ANOVA tests have been carried out comparing each of the two models with the non spatial model. The results clearly indicate that, in both cases, the inclusion of spatial effects improves the model's fit. Provided that indicators show that the model with the spatial trend best fit the data, such formulation has been further extended by including the share of agricultural worker *(d)* and a series of country dummy *(e)* to account for regional structural characteristics which are unobserved in the model. In both cases the model fitting significantly improves leaving unchanged the significance of smoothed terms.

Relative to the only model *(e)*, results are presented in figure II.4 in the form of a multiple plot to allow easier interpretation of the effect of non-linearity. Each plot separately scatters the smoothed predicted value on the vertical axis against the original value on the horizontal axis. The value on the vertical axis has a straightforward interpretation. It indicates the predicted contribution of the variable to regional growth. For this reason, all the values on the four different vertical axis have been reproduced on the -.02/.10 range, which permits to compare results.

[9] In this case the lowest it is the value the better the model fits.

Table II.4: Growth Regression - Non-linear Models

	(a)		*(b)*		*(c)*		*(d)*		*(e)*	
					Coefficients					
$Intercept$	0.024^{***}	(0.0008)	0.024^{***}	(0.001)	0.024^{***}	(0.001)	0.014^{**}	(0.006)	0.045^{***}	(0.005)
$ln(agri)$							-0.005	(0.003)	0.001	(0.002)
					F-statistics					
$s(gdp)$	20.046	[0.000]	16.200	[0.000]	29.580	[0.000]	14.611	[0.000]	11.960	[0.000]
$s(agg)$	4.229	[0.041]	2.818	[0.006]	3.459	[0.065]	2.046	[0.041]	2.891	[0.005]
$s(kne)$	8.550	[0.000]	11.149	[0.000]	35.388	[0.000]	6.483	[0.000]	15.814	[0.000]
$s(inno)$	3.067	[0.007]	3.383	[0.004]	4.259	[0.000]	3.432	[0.004]	17.081	[0.000]
$s(x,y)$			6.638	[0.000]			7.048	[0.000]	2.010	[0.019]
$SpatialFilter$					YES					
$CountryDummy$									YES	
					Goodness of Fit					
$Adj.R^2$	0.443		0.746		0.721		0.757		0.892	
$Deviance_e xp.$	48.50%		80.90%		76.90%		82.3%		91.9%	
$GCV \cdot 100$	0.1415		0.0791		0.0792		0.0078		0.0034	
$ANOVA(\chi^2)$			0.000		0.000		0.000		0.000	

Notes to table II.4:
SE in parenthesis, probabilities in square brackets
***, ** and * indicate significance at 1%, 5% and 10% confidence levels

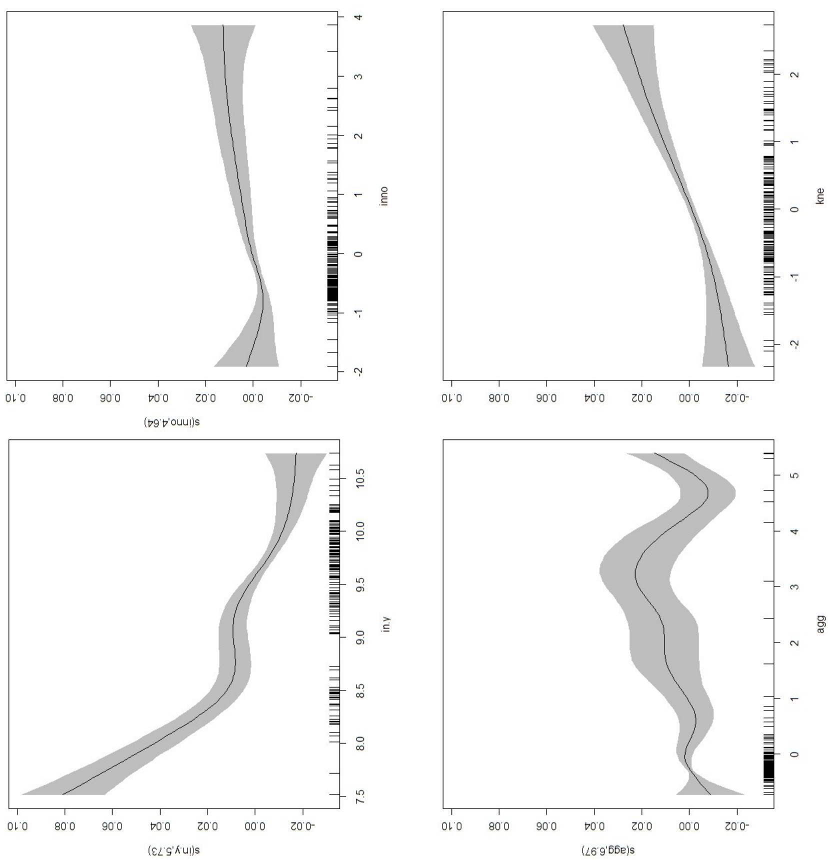


Figure II.4: Plot of predicted smoothed values against variables - Non-linear model with Spatial Trend

By looking at the figure II.4 it is possible to note that the contribution of initial income on growth shows a clear negative relation with the observed values of the initial income. Thus the convergence hypothesis is, on general, verified at the empirical level. A deeper look into the initial income plot, however, indicates that the rate of convergence, graphically identified as the slope of the curve (in absolute values), is higher in regions with a lower initial income level. In greater detail, the income distribution in the initial period seems to be characterized by a strong bi-modality. In the horizontal axis all the observations, each indicated by a small line in the axis, seem to concentrate around two major poles. The part of the curve relative to the observations in the first group of regions, likely the regions of eastern countries and, more generally, of the periphery, looks more sloped if compared to the part of the curve relative to the group of leading regions. The plot moreover shows that, for the majority of these leading regions, the value of the curve stands below the level of zero in the vertical axis. This means that for regions with very high levels of the initial income an increase in income itself has a negative impact on growth and, coherently with the convergence hypothesis, that rich regions have lower growth rates.

Interpretation of the agglomeration plot is more challenging. First glance look allows to detect the characterizing feature of this plot, a strong polarization in the neighbourhood of the value of zero. Only ten regions, in fact, have a score in this factor higher than one and these is likely to be the case of some capital regions. Relatively to the only group of non capital regions, the relation between agglomeration and its contribution to growth looks inverse-U shaped. For low values of agglomeration, its increase has a positive effect on regional growth while, for already densely agglomerated areas an increase in agglomeration produces negative effects on growth. The evidence reinforces the hypothesis on the presence of agglomeration dis-economies or, at least, cast serious doubts on the validity of the opposite hypothesis, according to which agglomeration is good for growth. Finally it is worth noting that the value of $s(agg)$ is higher than zero only in a very small interval on the distribution of agg.

The plot of the innovation factor is characterized by two most important features. The first is the U-shaped pattern, which shows the existence of a first important threshold effect. The second is that the predicted value of the contribution of innovation to growth ($s(inno)$) turns to be positive only after a certain value of the variable, which value represents the second threshold effect. Thus, for very low levels of innovative capacity, an increase in it would have no positive effects on growth. Only when the innovative capacity of a region exceeds the first threshold effect, marginal increases in innovation make the contribution of innovation to growth increase. Scoring higher than the first threshold in innovation is necessary but not sufficient for having positive effects on growth. These effects are present in only regions scoring higher than the second threshold in the innovation variable. The result is consistent with previous evidence found by Sterlacchini [96] for European regions using R&D as a proxy for innovation.

Finally, the interpretation of the knowledge economy factor is the most intuitive. The effect of a marginal increase of the factor is always grater than zero although, and again, the predicted effect on growth becomes positive after a given threshold. The value of $s(kne)$ ranges in between -.02 and .02, which means that, among all the sources of increasing returns, human capital is the one that principally contributes to growth.

II.6

Discussion and Conclusion

The assumption of non-linear patterns in growth-drivers used in this work allows a deeper understanding of the regional convergence. Overall it is found that there is convergence, as the higher the income level, the lower the contribution of income to growth. Nonetheless, the per-capita income distribution appears characterized by bi-modality and, in addition, regions in the two groups converge at different speeds. First group is principally made by regions with a per-capita income in 1995 lower than 10000 euros at 2000 prices. This roughly corresponds to regions eligible for Objective 1 funds under the Cohesion policy[10]. An higher speed of convergence characterizes this group. In the second group, made of regions with an income level in 1995 higher than approximately 13500 euros (at 2000 prices), the speed of convergence is lower. For each region in this group the contribution of income to growth is lower than zero. Growth in these regions, if any, is thus driven not by convergence.

Among the three theoretical hypothesis concerning the way regional growth relates to the presence of increasing returns in the regional economy, explanations grounding on knowledge, human capital and innovation are the most supported by the empirical evidence. On the contrary the contribution of agglomeration seems to be, overall, negative, with an inverse U shape. These results suggest that the way agglomeration relates to regional growth is far more complex than the theory would predict. In grater detail the only positive effect of agglomeration exhibits in very agglomerated regions and, among others, in capital regions are. For all the other regions agglomeration does not rule and, on the opposite, the evidence suggests that, above certain levels of agglomeration, dis-economies become predominant. Innovation and knowledge positively relate to regional growth although with important thresholds to be taken into account. Existence of these thresholds might be associated with the increasing importance of externalities in the path to the formation of a regional environment in which knowledge and technology actually play a prominent role. Innovations are built upon existing knowledge, by investing in research but also by using knowledge accumulated from previous researches and previous experiences. Likewise knowledge exchange between skilled workers takes place only if the pool of workers is sufficiently large at the regional level.

The evidence discussed in the chapter have important policy implications. By looking at regional growth in the recent past it is clear that the *convergence push* in least developed regions, to be ascribed partly to trade liberalization and partly to policy intervention, has worked out very fine. Less developed and more peripheral regions

[10] Actually a larger number of regions can benefit from the eligibility to Objective 1 funds.

have experienced an exceptionally high rate of income growth. This was, however, insufficient to fill the gap with respect to best-performing regions in Europe. As it is highlighted in the Fifth Cohesion Report (EC [35]) this growth has been driven by industrial restructuring with not too much attention on themes like knowledge and innovation. The idea in this work goes exactly the opposite way. Industrial restructuring is not sufficient for long-run growth which is on the contrary determined by accumulation of knowledge, human capital, technology. With this respects less developed regions in Europe face important growth opportunities deriving from the economic integration with probably some of the most knowledge-based economies in the world. In order to allow disparities to diminish over time it is necessary to fill the technology and knowledge gap that impedes these regions to enter the phase of growth characterized by increasing returns.

As a consequence, policies which target low-income regions might result ineffective as far as they are not explicitly intended to remove obstacles to growth. And, for reasons explained above, it could be worth sharpening policy attention on human capital and innovation also by orienting current policy instruments to the creation of a regional knowledge base. To some extent this is the main objective of the Smart Specialization strategy, summarized in a recent report by the European Commission (EC [36]). According to the report, the Smart Specialization strategy should serve to coordinate efforts by public and private institutions for the identification of strategic knowledge areas at the regional level in an attempt to maximize the benefits originating from these efforts. For the specific purpose of identifying regional strategic knowledge areas, the strategy aims at promoting the specialization of leading regions in generic technologies opposed to specialization in sector-specific applications of these technologies in other regions. The same report (EC [36]) however recognizes that most of the commitment to Research and Technological Development for the period 2007-2013 comes from already economically and technologically developed regions, with the perspective that these investments will help reinforcing the virtuous cycle of knowledge creation and growth. Benefits in less developed regions might be accordingly low.

In promoting successful innovation regional policies, local policy-makers should start thinking about the characterizing features of the regional knowledge base, in terms of localization of universities, research centres and higher education institutions. Altogether these are the most important drivers of knowledge, although their contribution to growth will remain considerably scarce unless the activities carried out by these institutions are not coordinated with the local business environment. From this viewpoint, an efficient allocation of resource should promote projects mainly oriented to creating and/or strengthening local research and education institutions as well as to promoting the dissemination of knowledge between institutions and private firms.

II.7

Appendix to Chapter II

Definitions

List of activities included in the definition of Knowledge Intensive Business Services: Post and Telecommunications, Computer and related activities, Research and development, Water transport, Air transport, Real estate activities, Renting of machinery and equipment without operator, and of personal and household goods, Financial intermediation, except compulsory social security, Activities auxiliary to financial intermediation, Education, Health and social work, Recreational, cultural and sporting activities.

List of activities included in the definition of medium/high-tech and high-tech manufacturing: Aerospace, Pharmaceuticals, Computers, Office machinery, Electronics-communications, Scientific instruments, Electrical machinery, Motor vehicles, Chemicals, Other transport equipment, Non-electrical machinery.

Additional Tables

Table II.5: Spatial Autocorrelation Diagnostics

	d=700km	d=500km	d=300km
$Moran's I$	0.3709 [0.000]	0.4167 [0.000]	0.4853 [0.000]
$logLik(SLM)$	575.41	578.05	574.97
$logLik(SEM)$	581.47	586.45	586.54
$logLik(SDM)$	582.23	582.33	582.75
$-LR(SDM - SLM)$			15.5643 [0.004]
$-LR(SDM - SEM)$			-7.5819 [0.108]

Notes to table II.5:
Probabilities in square brackets

Table II.6: Growth Regression - Spatial Error Model

	(1)	(2)	(3)	(4)
Intercept	0.185***	0.183***	0.192***	0.190***
	(0.026)	(0.026)	(0.026)	(0.026)
log(*gdp*)	-0.017***	-0.017***	-0.018***	-0.017***
	(0.003)	(0.003)	(0.003)	(0.003)
agg	0.001*	0.009	0.001*	0.001*
	(0.001)	(0.013)	(0.001)	(0.001)
kne	0.008***	0.008***	0.048***	0.009***
	(0.001)	(0.001)	(0.015)	(0.001)
inno	0.003***	0.003***	0.003***	0.029*
	(0.001)	(0.001)	(0.001)	(0.018)
log(*gdp*) $\cdot$ *agg*	-0.001			
	(0.001)			
log(*gdp*) $\cdot$ *kne*		-0.004***		
		(0.002)		
log(*gdp*) $\cdot$ *inno*				-0.003
				(0.002)
λ	0.719***	0.716***	0.723***	0.716***
	(0.051)	(0.052)	(0.051)	(0.052)

Notes to table II.6
SE in parenthesis
***, **, * indicate significance at 1%, 5% and 10% condifence levels.

Chapter III

On the specification of Heterogeneity in the Spatially Extended Knowledge Production Function

III.1

Non Technical Summary

The use of spatial econometric methods in the estimation of Knowledge Production Functions (KPF) at the regional level is usually motivated by evidence of spatial concentration detected in either dependent and independent variables or in the error term. Theoretical arguments for such a concentration of innovative activities on space are grounded on the tacit character of knowledge. Accordingly, some knowledge can be exchanged only through face-to-face contacts and, hence, within short distances. Nonetheless the spatial clustering of innovative activities is likely to be caused by factors other than knowledge spillovers, and the omission of these factors from the model specification may lead to wrong conclusions, qualifying as spillovers what actually is the effect of other, omitted information. In this the standard KPF model specification is extended with the inclusion of a smooth function of geographical coordinates (spatial trend), attempting to control for unobserved spatial heterogeneity. The resulting Generalized Additive Model is therefore estimated with semi-parametric methods. Evidence reveal that once spatial heterogeneity, captured by the spatial trend, is accounted for, evidence of spillovers weakens and even disappear.

III.2

Introduction

The Griliches-Jaffe [56] formulation of the Knowledge Production Functions (KPF) has become a leading approach to analyse the extent to which knowledge externalities are geographically localized at the regional level. The increasing popularity of spatial econometric methods has undoubtedly contributed to this. At the empirical level, the choice of spatial econometric models in a cross-regional framework is very usually motivated by the detection of spatial autocorrelation - commonly in model residuals - while, on the theoretical side, the evidence of interregional externalities provided by spatial econometric models are motivated in light of the Localized Knowledge Spillovers (LKS) theory (Audretsch and Feldman [10]).

The KPF framework can be summarized as a linear relationship between regional patent applications, the output of innovative activity, and R&D expenditures by private firms and universities, both in the region and in neighbours. A positive and significant estimate of the coefficient related to research expenditure in neighbours is interpreted as evidence of LKS. Such evidence might result misleading, however, as a consequence of the estimation bias due to problems of endogeneity and of unobserved spatial heterogeneity as well.

Interpretation is straightforward. Innovative activity at the regional level is not only the outcome of specific investments. Innovative output is is influenced by region-specific characteristics such as, for instance, the industry-mix, market opportunities, the innovative environment, the social capital and many other factors related to history of the region and to its technological development path. Some of those can be observed, some others cannot. All of them are, however, likely to affect the productivity of R&D investments and the related decisions by firms on the amount to be spent in research, causing R&D to be endogenous[1] (Bottazzi and Peri [24]). In addition, if one or more of these unobserved variables are spatially autocorrelated, this will cause residuals to be also spatially autocorrelated. The evidence might not, however, be ascribed to LKS.

Using data on high-tech patenting activity of 200 European NUTS II regions in 2005-2006, a patent equation is estimated using a negative binomial model. Spatial heterogeneity is introduced by adding to the specification a non-linear trend as a smooth function of geographical coordinates, resulting in a semi-parametric specification. The result of estimates clearly shows that the semi-parametric Generalized Additive Model fits the data better than the parametric Generalized Linear Model. It is found that estimates are biased by the omission of information on market potential as well as of spatial heterogeneity. Evidence of LKS, in fact, weakens and even disappear once

[1]This is especially true at least in cross-regional framework. The availability of panel data would allow, in fact, to control in part for this unobserved heterogeneity. Unfortunately longitudinal innovation data are not available for many regions in the EU. Since the exclusion of regions for which data are not available might cause sample selection bias, many scholars are obliged to work with cross-regional datasets. This is actually the case of many papers which will be reviewed in the next section.

heterogeneity is introduced.

In the remaining of this chapter, specification issues are discussed from both a theoretical and an empirical viewpoint in section 2. The empirical model is discussed in section 3. Results are summarized in section 4. Conclusions follow in the last section.

III.3

Space, heterogeneity and Localized Knowledge Spillovers

Within the Geography of Innovation literature, a special attention is paid to knowledge spillovers in an attempt to explain the determinants of geographical clustering of innovative activities [2]. The existence of positive externalities generated by knowledge transfer between organizations and institutions explains the willingness of innovative firms to co-locate in places from which knowledge can be easily accessed. This, in turn, determines regional differences in innovative activity (Jaffe et al.,[57]) and, eventually, in economic performance. This raises the questions of what are to be considered relevant knowledge sources and to what extent knowledge spillovers in research collaboration are actually localized. In considering the first issue, it is acknowledged that firms are the main investors in research, and thus a primary source of knowledge, accompanied by universities (Jaffe, [56]). Concerning the geographical scope of knowledge spillovers, arguments in favour of localized knowledge refer to its character of *stickiness* (von Hippel, [100]). Although the revolution of communication technologies has depressed the cost of transmitting knowledge, some parts of it are transmitted only through face-to-face contacts and frequent interactions. Admittedly, small distances are then not a sufficient condition for knowledge transfer as institutional and organizational barriers (Boschma, [22]) may also prevent knowledge flows.

The Jaffe's [56] formulation of the knowledge production function (KPF) has become a landmark in this stream of literature because, extending the original Griliches' [50] idea by the inclusion of third-parties research, it also allows econometric modelling interregional externalities by mean of spatial econometric methods (see for example Greunz, [48] and Moreno et al., [76]). Empirical evidence using such a framework generally confirm the theoretical hypothesis (Anselin et al. [8], Piergiovanni and Santarelli [83], Barrio-Castro and Garcia-Quevedo [17], Fritsch and Slavtchev [43]). The described approach has been subject to criticisms, however, on both the theoretical and empirical perspectives.

At a broader theoretical and conceptual level Geroski [46] first argued that standard methodologies do not allow to distinguish knowledge flows which are pure externalities from that mediated by the market dynamics. Knowledge spillovers in research collab-

[2] See Audretsch and Feldman [10] for a complete review of the literature. Knowledge spillovers are here defined as involuntary transfers of knowledge between parts. The mechanisms behind this transfer of knowledge are classified in three main categories: labour mobility, university and industry spin-off and research collaboration. Concerning the study of innovation clustering the focus is primarily on research collaboration.

orations, pertaining the transfer of that part of knowledge which can be difficultly be codified, are saw as pure externalities and their contribution to innovation might accordingly be difficult to disentangle. In addition the conceptualization of *knowledge tacitness* itself - more precisely the use of the concept made be economists and growth theorists - has been subject to criticism in the past (Cowan et al. [31]), provided that in this very broad category many different types of knowledge codifications, with different diffusion mechanisms, are included.

Breschi and Lissoni [25], elaborating on these ideas, denote how knowledge tacitness accordingly is far from being the only explanation to the concentration of innovative activities. Following the authors alternative explanations for the spatial clustering of innovative firms may ground on the presence of a local market for technologies as well as of specialized suppliers of technology. A developed market for technology, matching demand and supply of technologies is likely, in fact, to increase the market value of patents, making it more convenient for innovative firms to locate within short distances from markets. The development of these markets however requires coordination of activities involving different actors and operating at different institutional levels, as well as accompanying skills and expertise (Lamoreaux and Sokoloff, [65]). Also the presence of specialized suppliers of technologies is expected to encourage co-location of innovative firms, as a consequence of reduced complexity of the innovative processes. New technologies may in fact easily be acquired in the market rather than produced internally, which usually requires more time and efforts.

At the empirical level, it has been argued that coefficient estimates of the patent-research relationship are biased because of the omission of relevant variables strongly related to both research investments and patenting activity (Bottazzi and Peri, [24]). This is, for example, the case of the market potential of a region, a variable which is likely to affect the productivity of R&D in the region. The market potential, a measure of the market share which can be accessed by within the region, is expected to be positively correlated with patents, because innovative firms might be willing to locate near the market in which to sell their innovations. But a positive correlation is also expected between market potential and investments in research, as long as higher levels of production are associated with higher propensities to invest in research. More important, patents and research investments at the regional level are both positively correlated with the productivity of research investments. On the one hand because higher productivity of inputs means higher levels of outputs and, on the other hand, because higher productivity of innovative investments further attracts investments in research.

Omitted variables and unobserved heterogeneity are issues strictly related to the coefficient bias in linear models and they turn to play an even more important role in spatial econometric models. If a variable is spatially autocorrelated, as it might be the case of the market potential of the region, its omission from the model specification is likely to bias not only the R&D coefficient estimate, but also the coefficient related to the spatial lag of R&D. Similarly, in case of unobservable heterogeneity, if region-specific characteristics are not randomly distributed across the geographical space, their exclusion from the model causes the estimates related to R&D and to its spatial lag to be biased as well. Not by chance, summarizing the motivations for the use of a spatial econometric model, LeSage and Pace [68, pp. 27-30] point at three main causes for the evidence of spatial correlation in the data, namely omitted spatially correlated variables, unobserved spatial heterogeneity and externalities between units. Nevertheless, spatial model estimates do not allow to distinguish spatial autocorrelation due to unobserved spatial heterogeneity and omitted variables from that caused by spatial interactions and, consequently, there is a possibility that we qualify as interregional knowledge spillovers what actually is the effect of unobserved spatial heterogeneity and omitted variables.

The conceptual weaknesses of the LKS explanation of innovation clustering on the one side and the problems with the empirical specification of the KPF extended to include interregional knowledge spillovers on the other side are closely related. Mutual transfer of tacit knowledge via frequent interactions, indeed, is far from being the exclusive motivation for the spatial concentration of innovative activities. In more detail the market potential of the region and other unobservable characteristics influencing the R&D productivity matter as well. In empirical models, however, the attention is usually paid to only knowledge spillovers, through the inclusion of spatial lags of R&D. Accordingly, the omission of regional market potential and of other unobservable characteristics as well, is likely to bias the evidence in favour of LKS. And, as a consequence of this, it is possible to qualify as knowledge spillovers what actually is the effect of other, excluded, information.

III.4

Econometric strategy

The KPF at the regional level has been always described as a linear function between the rate of patenting activity, a measure of the regional capacity to produce innovative output, and the percentage of R&D in the regional GDP, a measure of the innovative efforts made by firms and public institution located within the region (Jaffe, [56]). This base framework has been extended taking into account spatial relations and spatial interactions between regions using spatial econometric techniques (Anselin et al. [7], Anselin [8], Acs et al. [1], Fischer and Varga [40]).

More recently, a number of studies (Autand-Bernard and LeSage [11] for France, Fritsch and Slavtchev [42] and Grimpe and Patuelli [51] for Germany, Barrio-Castro and Garcia-Quevedo [17] and Gumbau Albert and Maudos [54] for Spain, Ponds et al. [86]) for The Netherlands and Bottazzi and Peri [24] for European Regions) have shown an interest in modelling the number of patents in place of the patenting rate, in an attempt to maximize the information content of the patent variable which is, by definition, discrete and positively defined. Geographically localized knowledge spillovers are accounted for, in the majority of cases, by including the spatially lagged R&D. In this work, this recent stream of literature is followed and, accordingly, distributions for count data are used to model patent applications.

For the 200 NUTS II regions in the sample, the dependent variable Y_i is measured as the average count of patent applications to the European Patent Office during years 2006-2007. Being the purpose of this paper that of testing the relevance of the LKS theory, only applications relative to high-tech industries, in which the diffusion of tacit knowledge is expected to play a crucial role for the development of regional innovation (Keeble and Wilkinson [59]) , have been counted.

Among the covariates, private firms R&D (*REDE*) and university R&D (*REDU*) have been included as inputs of the KPF. The spatially lagged R&D (*WREDE*) is expected to capture the role of knowledge spillovers between firms in neighbouring regions. The size effect is controlled for by including population as additional covariate with the coefficient constrained to unity (offset). In order to avoid problems due to the simultaneity bias, all the variables in the right hand side are taken for a period previous to the years 2006-2007.

Neighbouring relations are described by a row-standardized matrix W, constructed by using the great circle distance to define contiguity. The intensity of neighbouring relations is modelled as an inverse function of squared physical distance between neighbours (d_{ij}), such that the generic element of W is defined as in the equation III.1.

$$w_{ij} = \frac{d_{ij}^{-2}}{\sum_j d_{ij}^{-2}}. \tag{III.1}$$

The basic formulation of the count model for the patent equation is described in equation III.2. As discussed in the previous section, estimates of the model in equation III.2 might be biased by the omission of relevant variables and of region-specific characteristics correlated with *REDE* and *WREDE*.

$$\begin{aligned} Y_i &\sim Poisson\,(\mu_i) \\ \mu_i &= \exp\left(\alpha + \beta_1 REDE_i + \beta_2 REDU_i + \beta_3 WREDE_i + \beta_4 MP_i + \beta_5 MHT_i\right) \end{aligned} \tag{III.2}$$

As far as it concerns the omitted variable bias, in this paper an attempt is made to mitigate such a bias by introducing in the specification a measure of the market potential of the region (*MP*) as well as an indicator of the regional industrial structure composition (*MHT*). The first is a measure of market size which is potentially accessible by within the region and, although available only for the year 2006, it has been preferred to the per-capita income, provided that the latter is itself potentially endogenous. The variable has been constructed under the ESPON project[3] and is a proxy *'for the potential for activities and enterprises in the region to reach markets and activities in other regions'*. Accordingly it is completely exogenous, as it is calculated on the base of the distance separating the origin region by other potential accessible regions and is, at the same

[3] http://www.espon.eu/main

time, appropriate because Gross Domestic Product is used as a weight for distances, hence accommodating a scale of the potential market for (not only) innovative firms. The second is measured as the share of employees in high-tech and medium/high-tech industries in the total of manufacturing employment and controls for the regional specific characteristics related to the industrial structure.

Concerning the unobserved spatial heterogeneity, the issue undoubtedly represents the most demanding part of the specification problem. At the European level, in fact, the lack of sufficient regional data impedes to observe relevant variables affecting the patenting activity of firms. Moreover, most of the aspects which influence the innovative activity of firms are even not observable and measurable. This might be the case, for instance, of the presence of regional markets for technologies as well as of specialized suppliers of technologies. To solve the problem it would be ideal to estimate the model by using panel data techniques, thanks to which such an unobserved heterogeneity would drop with the inclusion of fixed effects. Unfortunately, data on innovation at the regional level for the whole Europe are limited.

In absence of an available panel dataset, the issue of unobservable heterogeneity can be addressed with the inclusion of geographical variables. Unobservable characteristics might in fact be viewed as non-randomly distributed in space. On the contrary they can be thought as the result of a path-dependent and place-dependent process aimed at building an innovation-friendly business and institutional environment which can be eventually proxied by geographical characteristics. Among the possible choices, that of geographical dummy variables would be the simplest and, to some extent, most intuitive solution. Nonetheless it would require an ex ante definition of the geographical space into a set of dichotomous variables, which, in turn, presumes knowledge about the way unobservable characteristics are distributed in space. In this work, by the opposite, geography is accounted for in the specification through the inclusion of regional geographical coordinates among the covariates.

Concerning the relation between the innovation and geographical coordinates, this is expected to be non linear, although the degree of non-linearity is, admittedly, unknown a priori. The linear hypothesis cannot be considered appropriate as long as it would imply that the number of patents will increase (decrease) with increasing (decreasing) latitude or longitude. The most common non-linear specification, the quadratic hypothesis, could be considered more appropriate, for instance; nonetheless still not suitable. It would imply, in fact, that the number of patents increases up to a certain longitude threshold after which it starts decreasing. The same should be implied for latitude. This hypothesis would be consistent with a core-periphery pattern distribution of patenting activity across the regions, but could eventually be misleading in the case the pattern shows to be more complex. For this reason the technical choice for the functional form has fallen on the *spline* fitting method, on the base of which the functional form is chosen optimizing the information in the data. An *spline* term, a non linear function of both longitude and latitude, is thus included as a spatial trend in the mean specification of the model.

The degree of non-linearity is selected by an algorithm minimizing the sum of the model residuals, to which a penalty is added. The penalty is defined as a function itself of the second derivative of the non-linear function, in a way that the the higher the non-linearity, the lower the sum of residuals, the higher the penalty. The optimal choice is, accordingly, the result of the balance between the goodness of model fitting and the penalty. With this respect the smooth spline tends to prefer the more flexible non-linear specification for the spatial trend to the linear specification, correcting for the excess of non-linearities in the trend itself.

The resulting model can be characterized as a Generalized Additive Model in which a smoothed trend of geographical coordinates $s\,(x, y)$ is added to a linear parametric specification of the mean function, and is estimated with semi-parametric methods described in Bivand et al. [21, pp. 297-300].

Finally, it is well known that the limit of Poisson distribution is the assumption of equality between the mean and the variance, a condition which might not hold also in this special case. For this reason a Negative Binomial model is employed, admitting a variance different from the mean thanks to the introduction of the overdispersion parameter θ. Within the full sample of 200 NUTS II regions in the database, for only 8 of them a count of high-tech patent applications equal to zero is reported. This excludes the application of econometric procedures suitable to control for the abundance of zeroes. The complete model is described in equation III.3. Lower cases of variables indicate logarithms[4].

$$\begin{aligned} Y_i &\sim NB\,(\mu_i, \theta) \\ \ln\,(\mu_i) &= \alpha + \beta_1 rede_i + \beta_2 redu_i + \beta_3 Wrede_i + \beta_4 mp_i + \beta_5 mht_i \\ &+ s\,(x_i, y_i) \end{aligned} \tag{III.3}$$

The non-linear component $s\,(x_i, y_i)$ describes the trend surface in the geographical space determined by the value of X and Y coordinates and is expected to capture the unobserved spatial heterogeneity. According to the research question posed in the introduction of this work, once the spatial heterogeneity is accounted for in the model, the bias in the coefficient estimates related to *rede* and *Wrede*, respectively β_1 and β_3 is expected to decrease. In particular in the case of *Wrede* such a result would require to reconsider the validity of the LKS theory at the regional level.

[4] This is the result of the choice to use logarithm as link function for the mean specification of the negative binomial model.

III.5

Results

Results are presented separately for the Generalized Linear Model (GLM) and for the Generalized Additive Model (GAM). To begin with the GLM the knowledge spillovers hypothesis is first tested and control variables are added in a second step. Estimates are presented, respectively with and without control variables, in the upper and the lower parts of table III.1. The model is estimated by using a battery of four different contiguity matrices, constructed with the methodology indicated in equation III.1 and allowing the critical distance of the circle to vary. In this way it is possible to test whether results are sensitive to the specification of the matrix elements.

Table III.1: Negative Binomial Model for the Patent Equation - GLM

	d = 300km	d = 500 km	d = 700 km	d = 900 km
Intercept	-13.356***	-13.837***	-14.074***	-14.320***
	(0.153)	(0.175)	(0.200)	(0.240)
BERD	0.768***	0.705***	0.730***	0.733***
	(0.072)	(0.071)	(0.069)	(0.070)
URD	0.731***	0.770***	0.703***	0.685***
	(0.267)	(0.258)	(0.257)	(0.260)
WBERD	0.777***	1.236***	1.448***	1.673***
	(0.128)	(0.157)	(0.184)	(0.227)
θ	1.370***	1.476***	1.489***	1.459***
	(0.146)	(0.159)	(0.161)	(0.158)
logLik	-762.899	-755.107	-754.432	-757.090
AIC	1535.799	1520.214	1518.863	1524.181
LZ I	0.077**	0.099*	0.063***	0.055***
Intercept	-14.801***	-15.049***	-15.234***	-15.536***
	(0.249)	(0.251)	(0.260)	(0.282)
BERD	0.632***	0.589***	0.601***	0.584***
	(0.073)	(0.071)	(0.071)	(0.071)
URD	0.606**	0.615**	0.607**	0.597**
	(0.255)	(0.248)	(0.247)	(0.247)
WBERD	0.492***	0.866***	1.010***	1.211***
	(0.124)	(0.153)	(0.179)	(0.215)
MP	0.018***	0.017***	0.017***	0.017***
	(0.002)	(0.002)	(0.002)	(0.002)
MHT	0.001	-0.001	0.003	0.007
	(0.020)	(0.019)	(0.019)	(0.019)
θ	1.747***	1.852***	1.862***	1.868***
	(0.193)	(0.207)	(0.208)	(0.210)
logLik	-738.391	-732.698	-732.232	-732.406
AIC	1490.781	1479.395	1478.465	1478.812
LZ I	0.044	0.024	0.006	0.013

Notes to Table III.1
SE in parenthesis.
***, ** and * indicate significance at 1%, 5% and 10% confidence levels.

Coefficient estimates in the upper part of table III.1 show the correct slope and are always significantly different from zero. The size of coefficients related to the R&D expenditure carried by both private firms and universities is not affected by a change in the specification of the contiguity matrix. On the contrary the size of the lagged private firms R&D it is; in addition, throughout the models, it seems to be excessively large.

The introduction of controls does not alter the main result, although it clearly reduces the bias in coefficient estimates. In greater detail, the market potential of the region is, as expected, positively related to the innovative activity in the region and the coefficient is always highly significant. The share of workers in high-tech and medium/high-tech manufacturing is, on the contrary, never significantly different from zero. By looking overall, the evidence suggest that regional innovative activity is determined by the regional expenditure in research, and positive spillovers exist both between firms and universities in the region and between firms in neighbouring regions. Furthermore, the result is robust to several specification of the spatial contiguity relations.

In terms of residual diagnostics, the Lin and Zhang [69] modified version of the Moran's I test ($LZ\ I$) for count data model residuals has been applied in order to test for unexplained spatial clustering of regional innovative activity. The statistic is significant (indicating positive spatial clustering) only in the simplest version of the model and with any of the matrices in the battery. It turns out to be not significant after the inclusion of additional covariates in the model. By comparing the goodness of fit of the models in columns 1 to 4 in table III.1, it appears that best fitting is achieved when the contiguity matrix is set on the base of the critical distance of 700 Kms, at least based on the value of the log-likelihood (logLik) and of the Akaike Information Criterion (AIC). Thus the 700 km matrix is utilized in the remaining of the empirical application.

Although the inclusion of additional covariates, controlling for the size of the regional market and the composition of the industrial structure, significantly contributes to the explanation of the spatial clustering of innovations, it is still not possible to exclude that model estimates are biased by the unaccounted spatial heterogeneity. Geographical coordinates are thus included in the model, through a spatial smooth trend. Estimation results are summarized in the table III.2.

The model fitting, at least according to the AIC, considerably improves after the introduction of the spatial trend. The value of the χ^2 statistics tests the null hypothesis of insignificance of the trend value[5]. The hypotesis, as it is possible to note, is strongly rejected by the test. Furthermore it is worth drawing attention to some changes in the model estimates using the GAM compared to the GLS case. Coefficients related to research made by both firms and universities in the region decrease, even though not much. Nonetheless they continue to be significantly different from zero. On the opposite, the coefficient of firms research in neighbouring regions, the measure of knowledge spillovers, sharply decreases and turns to be insignificant. Finally the coefficient of the market potential is not affected by the introduction of the trend, while the coefficient related to the industrial structure continues to be not significantly different from zero.

Summarizing, innovative activity is driven by regional investments in research and their spatial concentration appears not to be explained by localized knowledge spillovers

[5]Under the null, the GAM estimates in table III.2 do not differ from the GLS estimates reported in the third column of the lower part of table III.1

Table III.2: Negative Binomial Model for the Patent Equation - GAM

	Parametric part	
	Coef	SE
Intercept	-14.354***	(0.599)
BERD	0.574***	(0.071)
URD	0.562**	(0.231)
WBERD	0.265	(0.524)
MP	0.016	(0.002)
MHT	-0.016	(0.024)
	Non-parametric part	
	edf	χ^2
s(X,Y)	24.49	77.17 [0.000]
	Model Fit and Residuals	
UBRE		0.293
Dev exp		0.77
logLik		-693.957
AIC		1448.902
LZ I		-0.0270 [0.2225]

Notes to Table III.2
SE in parenthesis. Probabilities in square brackets.
***, ** and * indicate significance at 1%, 5% and 10% confidence levels.

in research collaboration between neighbouring regions. On the contrary, external factors like the market opportunities and other unobservable characteristics of the region are likely to play a very important role. The introduction of these factors in the patent equation captures most of the effect previously attributed to knowledge spillovers in research. In the remaining of the section the robustness of the aforementioned result will be checked by using other different specifications for the patent equation. In more detail, the model is firstly specified assuming that the dependent variable, the count of patents, is characterized by global autocorrelation, the omission of which would bias the results. In a second step a series of geographical dummy variables are introduced into the model to controls for the spatial heterogeneity. Once again the model will be estimated with and without the spatial trend.

Geographical dummy variables are obtained by categorizing the peripherality indicator developed by ESPON into four categories, namely *core*, *upper intermediate*, *lower intermediate* and *periphery* on the base of 25% quantiles of the indicator distribution. Global autocorrelation is accounted for in the model by using spatial filtering methodologies for count data (Grimpe and Patuelli, [51]). The method assumes that the contiguity structure is represented by the eigenvectors of the contiguity matrix. These eigenvectors are thus first estimated and, in a second step, included into the model specification.

Results of the estimates are summarized in table III.3. In the first two columns the model is estimated with spatial filtering. The third and the fourth columns report estimates of the model including the series of geographical dummy variables. Finally, in fifth and sixth columns both geographical dummy and spatial filtering are included. Every model is estimated twice, respectively without and with the spatial trend in an attempt to test the robustness of the research hypothesis.

Table III.3: Negative Binomial Model for the Patent Equation - Robustness Check

	GLM	GAM	GLM	GAM	GLM	GAM
Intercept	-15.459***	-14.334***	-15.051***	-14.626***	-15.261***	-14.581***
	(0.279)	(0.587)	(0.280)	(0.626)	(0.303)	(0.623)
BERD	0.630***	0.575***	0.634***	0.548***	0.657***	0.550***
	(0.072)	(0.071)	(0.070)	(0.071)	(0.071)	(0.070)
URD	0.466*	0.562**	0.659***	0.579**	0.525**	0.573***
	(0.253)	(0.231)	(0.246)	(0.228)	(0.252)	(0.228)
WBERD	1.290***	0.259	0.876***	0.136	1.124***	0.135
	(0.221)	(0.513)	(0.184)	(0.531)	(0.228)	(0.523)
MP	0.017***	0.016***	0.014***	0.015***	0.015***	0.016***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
MHT	-0.014	-0.018	-0.017	-0.016	-0.031	-0.018
	(0.020)	(0.024)	(0.020)	(0.024)	(0.021)	(0.023)
CORE			0.552***	0.849**	0.541***	0.770*
			(0.192)	(0.425)	(0.191)	(0.422)
UP INTER			0.369**	0.577*	0.338*	0.512
			(0.180)	(0.342)	(0.179)	(0.340)
LOW INTER			0.205	0.625**	0.217	0.591**
			(0.174)	(0.279)	(0.174)	(0.275)
FILTER	-2.238**	5.810			-2.013*	5.256
	(1.158)	(5.911)			(1.163)	(6.345)
$s(X,Y)$						
edf		22.71		24.81		23.52
χ^2		64.02***		69.67***		59.72***
AIC	1477.6	1449.9	1476.4	1447.2	1475.9	1448.4
LZ I	0.0017 [0.7064]	-0.0262 [0.2411]	-0.0005 [0.7982]	-0.0288 [0.1883]	-0.0086 [0.8428]	-0.0281 [0.2013]

Notes to table III.3:
SE in parenthesis. Probabilities in square brackets.
***, ** and * indicate significance at 1%, 5% and 10% confidence levels.

Estimation results strengthen the hypothesis that evidence of research spillovers diminish and even disappear once spatial heterogeneity is accounted for in the model. The inclusion of the spatial filter does not alter the main result of GLM estimation in table III.1, as all the coefficients are still rightly sloped and have almost the same magnitude. Exception is made here for the coefficient related to the industrial structure, which now shows a negative sign although it continues to be insignificant, and for the university research coefficient, the magnitude and significance of which slightly decrease. The filter variable is strongly significant, even tough the related coefficient is not interpretable. The introduction of the trend weakness the significance of the only coefficient relative to research spillovers. As in the case of table III.2 most of the results are, in fact, unchanged and the magnitude of the research spillovers coefficient sharply decreases. The same result emerge from the comparison of columns 3 and 4 in table III.3. The inclusion of geographical dummy variables partially capture the spatial heterogeneity. Comparing the GLM estimates in the third column with estimates in table III.1 the research spillovers coefficient is now lower in magnitude, which in turn indicates that geographical dummy have reduced the estimation bias. The coefficient is however strongly significant. Once again it decreases in magnitude and looses significance only after the trend is included. The geographical variables have the right signs and sizes in both the GLM and the GAM, and all of them turn to be significant in the GAM. Finally, the same conclusion hold when both geographical variables and spatial filtering are included.

Comparing the three couples of columns, it is important to notice the lower values of the AIC scored by the GAMs compared to the relative GLM counterparts. To the authors understanding this is a clear indication of the improved fitting of the model with the spatial trend. This indication is largely confirmed by the value of the three χ^2 statistics which compare, in terms of model likelihood, the model with the trend with respect to the model without the trend. Overall, the lowest AIC score is reached by the GAM with geographical variables and without the spatial filter. In the end, none of the estimated models have produced spatially autocorrelated residuals, at least based on the Lin and Zhang [69] statistic.

III.6

Conclusion

Research spillovers have received increasing attention in the empirical literature on regional innovation as they have been indicated as one of the most important vehicle to convey regional growth. In the search of evidence of research spillovers it has been argued that such spillovers are localized as some knowledge is tacit, cannot be codified and requires face-to-face contacts in order to be exchanged.

Using a spatially extended Knowledge Production Function, this paper has examined the contribution of research spillovers to regional innovation in the EU, with a major attention to spillovers between firms in neighbouring regions. It is argued that the contribution of interregional research spillovers is overestimated if some important variable and regional specific characteristics are not accounted for. Results show that once the model specification controls for relevant variables, an in particular for the market potential of the region and for unobserved spatial heterogeneity, the evidence of research spillovers diminishes and even disappears.

The evidence provided in this paper is in contradiction with the majority of results presented in previous empirical works. This does not stand for a contradiction of the LKS theory and, more in general, of the relevance attributed to knowledge spillovers. By the opposite the evidence in this paper suggests that current methods employed to address the issue of interregional spillovers in research collaboration produces, at best, biased results. This, in turn, requires researchers to work out to improve the econometric specification of the standard patent equation. Among the possible considerations, further research should carefully pay attention to what has been noticed years ago by Glaeser et al. [47], according to whom *"intellectual breakthroughs must cross hallways and streets more easily than oceans and continents"*. In fact knowledge spillovers are likely to be extremely localized and the evidence of interregional spillovers should be searched at a lower geographical scale.

The results have also important policy implications. In the author's opinion too much emphasis has been addressed to the LKS theory in explaining the spatial patterns of innovation and attention should be refocused on what motivates innovative investments. At the regional aggregate level the evidence suggests that innovation is led by existing market opportunities and by regional innovative environment as well. Spillovers are thus likely to be the consequence of innovation clustering, rather than the cause. After all it seems more convincing that innovative firms are willing to locate in regions offering the best opportunities of returns from their own research and not where they can get something from others' research.

Chapter IV

Knowledge Creation vs Knowledge Co-Production: Knowledge Intensive Business Services and Innovative Activity in EU Regions

IV.1

Non Technical Summary

Regional economies are continuously evolving toward a tertiarization of production systems. Despite the increasing relevance of services, however, the analysis of innovation at the regional aggregate level has predominantly focused on manufacturing, gathering the attention on the role of R&D expenditure as input in the production process and, in some cases, accounting for research-based knowledge externalities.

In this paper the role of Knowledge Intensive Business Services is studied and their contribution to the regional aggregate innovation is evaluated. The aim is twofold. First is to provide insights on the role covered by KIBS as a second knowledge infrastructure. Second is to examine the extent to which KIBS operate as bridges between the general purpose analytical knowledge produced by scientific universities and more specific requirement of innovative firms.

A role commonly acknowledged to KIBS is in fact that of knowledge transferors. If on the one side it is however clear *to whom* they transfer knowledge - their client firms - on the other it is not as clear *from whom* the knowledge is originally transferred. For this reason a major attention in this work is dedicated to scientific universities considered as a primary source of knowledge. Being this knowledge analytical and highly codified, it probably can be more easily accessed by nearby located firms having higher opportunities of research collaboration and less easily by firms located in different regions. It is argued that the contribution of KIBS as transferors of knowledge from universities to firms is are therefore specially important in the latter case.

The empirical analysis is based on a sample of 200 EU NUTS II regions and the evidence suggests that the contribution of KIBS to regional innovation is considerable. Furthermore it is find that this contribution is more sizeable in regions in which there are not scientific universities.

IV.2

Introduction

The Jaffe's formulation of the Knowledge Production Function (KPF) (Griliches [50], Jaffe [56]) is a widely adopted approach to study the determinants of innovation at the regional aggregate level. To put it shortly, innovative output is determined by the amount of research made by private firms and by universities, allowing the output elasticity to research to be larger in correspondence of the geographical coincidence between private firms and university research. This gain in input productivity is attributed to the presence of localized knowledge spillovers between universities and industries. The attribute of localized to knowledge spillovers, described as involuntary transfer of knowledge between firms and/or institutions, is explained by the sticky character of knowledge (Von Hippel [100]). Accordingly some knowledge is difficult to be codified and can be transmitted only through face-to-face contacts and frequent interactions.

At the empirical level the model has been estimated by using patent applications as a measure of aggregate innovative output and several studies have reported evidence of a positive effect of knowledge spillovers usually attributed to university-industry collaborations (see Anselin et al.[7], Fischer and Varga [41] and Ponds et al. [85] among the most significant studies.). In their survey of the Geography of Innovation literature Audretsch and Feldman [10] report that the use of other measures of input and output (for instance R&D personnel as alternative input and literature-based measures as alternative output) have usually yielded to very close empirical results.

Despite the general agreement about the robustness of the evidence provided by the use of Jaffe's approach in empirical studies, few attention needs to be deserved to some critical issues. The first worth-considering issue is the extent to which the benefits of co-location of universities and industries could be actually ascribed to the presence of knowledge spillovers. At the micro level Mansfield [74] has documented that the relations between universities and firms are mostly market relations which take place in the form of consulting services. Similarly Zucker at al. [104] find evidence that *"the positive impact of research universities on nearby firms relates to identifiable market exchange between particular university star scientists and firms and not to generalized knowledge spillovers"*. Consequently the use of Jaffe's formulation might determine empirical evidence which, failing to account for market externalities, overestimate the effect of knowledge spillovers.

A second important issue relates to the role played by physical distance in the dissemination of knowledge. In both the cases of market mediated knowledge exchange and pure knowledge spillovers, there is no doubt that distance matters. It is claimed that physical distance, per-se, is neither a necessary nor a sufficient condition for learning (Boschma [22]) but, admittedly, learning and interacting is easier if the distance is short. Therefore the probability of undertaking collaborations is expected to decrease with the distance separating universities and firms[1] and, in addition, if firms and universities are located in two different regions it is expected to be less easier for firms to generally access the knowledge produced by universities.

Long distances do not however imply that academic knowledge is inaccessible. As far as it concerns the mechanisms of knowledge dissemination other than direct interaction, a recent literature has acknowledged the role played by so-called Knowledge Intensive Business Services (KIBS) as producers, providers and, more important, transferors of knowledge (Den Hertog [34]). In particular, according to Den Hertog, KIBS might represent a *"point of fusion"* between more generic global knowledge, as it is that produced by scientific universities, and specific needs of local firms.

The investigation of the role played by KIBS in supporting firms' innovation has recently began to receive attention in firm-level empirical studies (see Cainelli et al. [26] [27] as an example of micro-econometric analysis). Meantime less has been already studied about the contribution of KIBS at the regional aggregate level. In this a spatial econometric analysis of the innovative performance relative to a sample of 200 EU regions is presented and two primary hypothesis are tested. The first hypothesis relates to the participation of KIBS in the production of innovation at the regional level. In more detail it is examined the extent to which the regional concentration of KIBS represents a considerable factor in explaining the regional variation in the level of innovative activity. The second hypothesis more specifically relates to the exact role played by KIBS. As an intermediate level knowledge infrastructure, the activity of KIBS is expected to be more influential for manufacturing firms in those regions in which scientific and academic knowledge is locally absent and, consequently, not directly accessible.

In an attempt to measure the scientific and academic knowledge in the region a new variable is constructed which, contrarily to R&D investments made by universities, is expected to be not biased by the existence of market transactions between universities and firms. The variable is first included in the KPF framework and further used to differentiate the sample of regions endowed with scientific knowledge from the remaining of regions which are not. The empirical models are specified accounting for the presence of spatial relations and spatial interactions between neighbouring regions and are eventually estimated by using heteroschedasticity-consistent estimator for spatial models developed by Arraiz et al. [9] and Kelejian and Prucha [60].

The evidence suggest that KIBS do actually contribute to the regional production of knowledge, mainly as co-producers of innovations assisting their client firms and providing them with the necessary soft skills. Furthermore there is evidence that KIBS working in high-tech sectors can be qualified as scientific knowledge transferors, but only in regions where the scientific and academic knowledge is absent. Oppositely, in

[1] Ponds et al. [87] provide robust empirical evidence supporting this theoretical hypothesis.

presence of scientific and academic knowledge in the region R&D investments by private firms continue to be the most productive knowledge input. Overall the results in this work indicate KIBS as a second infrastructure which, together with research spillovers, contribute to the dissemination of scientific knowledge. The remaining of the work is organized as follows. The next section is aimed at defining what KIBS are and what is, at least according to the existing theoretical literature, their contribution to the regional innovation. In the third section the dataset is illustrated, paying special attention the the issue of measuring scientific and academic knowledge. Empirical results are summarized and discussed in section four. Conclusion follow.

IV.3

KIBS and Innovation

The attention to business services has increased over time together with the progressive shift of national and regional economies from manufacturing-based production systems to more service-oriented development paths. According to Shearmur and Doloreux [93] such an increase in attention has been channelled differently from geographers and innovation economists. For the sooner group of scholars the emphasis was on the urban location of High Order Producer Services[2] (HOBS) and, thus, on the role of cities in a dual picture of the regional production separated in manufacturing activities and service firms. From the viewpoint of innovation economists, oppositely, the focus has been more on the role of service firms in the production of knowledge and only to a lower extent on the distinction between manufacturing and services[3]. Admittedly, the two different conceptualisations actually identify, at the empirical level, the same sectors (Wood [102]).

In a recent survey of the KIBS literature Muller and Doloreux [77] highlight three most important characteristics of KIBS, at least based on the existing definitions. A first characterizing feature is the explicit orientation of the provided services to business enterprises and not to private consumers. Secondly there is the implicit transfer of knowledge between the service firm and the clients (i.e. the business enterprises). Finally the provision of the services is realized with the predominant activity of human capital. Accordingly, the role of KIBS is intermediate in nature but, nonetheless, it still appears to be difficult to disentangle what their exact contribution to innovation is made up of.

The contribution of Den Hertog [34] has represented the point of departure in the analysis of the role played by KIBS in the process of innovation. In his work the

[2]As it is noted by Shearmur and Doloreux [93], the term has been in use to identify business firms providing their clients with management and consulting services and so to distinguish them from providers of more general business services.

[3]As a matter of fact the analysis of service innovation has been mainly conducted focusing on the modes of innovation, which are supposed to differ with respect to the manufacturing case. For instance see the studies by Hollenstein [55], Jensen et al. [58] and Corrocher et al. [30]. Some recent literature (Gallouj and Windrum [44]) has indicated the path for an integrated approach to the study of innovation in both manufacturing and services.

"symbiotic nature" of the relation between KIBS and their clients is emphasized and it is accordingly argued that KIBS act as co-producers of innovation together with their client firms. This in turn indicates that KIBS do not innovate themselves but nonetheless play a fundamental role in assisting manufacturing firms in the innovation process. The idea of co-production is further categorized in the work of Den Hertog in three main dimensions, attributing to KIBS the role of facilitators, carriers and sources of innovation. As facilitators of innovation the role of KIBS is that of bare support to the client manufacturing firm, from which the innovation anyhow originates. As carriers of innovation KIBS act as transferors of existing knowledge to the client firm. The KIBS mediation is motivated by the fact that the knowledge source is generally not directly accessible to the client firm. Also in this latter case the innovation process originates within the client firm but now the contribution of KIBS is more remarkable. Finally as sources of innovation KIBS do initiate the innovation process in place of the client firms and further develop the innovation in close collaboration with them.

All these elements qualify KIBS as a second knowledge infrastructure which, alongside universities and public research institutes (which are considered the first knowledge infrastructure) contribute to the diffusion of knowledge. Remarkable differences however exist between the two knowledge infrastructures especially regarding their relative contribution to the creation and dissemination of knowledge and, consequently to innovation in firms. Research is carried out by universities and public research institutes (hereinafter UPRI) systematically, with structured projects and usually long term horizons. Moreover UPRI projects are designed and developed by specific departments in which the process of knowledge creation is highly formalized and extensively based on R&D investments. On the contrary research is approached by KIBS in a less systematic and structured manner which better accommodates their problem-solving objective. Innovation is not realized through R&D investments but instead via frequent interactions with clients, during which knowledge is mutually exchanged in the attempt to apply general purpose technologies to solve firm-specific problems. (Simmie and Strambach [94]). Due to that, university-industry cooperation is more likely to take place with firms in R&D intensive industries, having these firms the know how and the organizational structure which is necessary to engage in cooperation projects with UPRI. Likewise collaboration between KIBS and firms in more likely in the case of SMEs which are surely willing to engage in innovative projects but, at the same time, lack the specific know how to do that[4]. Muller and Zenker [78] describe the innovation taking place with the interaction between KIBS and firms as a process of re-engineering of existing knowledge which, accordingly, does not require that any or both invest in R&D.

In spite of the fact that the knowledge creation process largely differs between KIBS and UPRI, to some extent their relative contribution to innovation can be considered as overlapping. Based on the Pavitt definition of industries (Pavitt [82]) Strambach [97]

[4]Kleinknecht [61] observes that the lack of adequate know how is likely to be among the most important obstacles to innovation in small and medium size firms. For these firms R&D investments might be insufficient for innovation.

distinguishes two types of knowledge. Analytical knowledge is generated on the base of formal models of research and development and within structured research programs. This knowledge is mainly explicit (as in the case of publishing or patenting) but is highly codified and hence difficult to access. Synthetic knowledge is on the contrary generated by applying generic knowledge to specific problems and, consequently is usually considered as tacit. The knowledge developed within UPRI undoubtedly belongs to the first category but, conversely, not necessarily knowledge developed by KIBS in collaboration with their clients belongs to the second. In fact Strambach defines, according to the type of knowledge used, two categories of KIBS, distinguishing R&D consulting oriented KIBS, which use analytical knowledge, from technical and economic oriented KIBS, which use synthetic knowledge[5].

To sum up KIBS are service firms which not only contribute to innovation by transferring knowledge from various sources to manufacturing firms. KIBS actively participate in the creation of new knowledge by interacting with their client firms with the aim of adapting to specific needs some general purpose technologies (Muller and Doloreux [77]). In doing these they act as bridges between generic knowledge and more specific problems of firms. In the specific case of R&D consulting firms KIBS represent a bridge between analytical knowledge developed by UPRI and firms to which such a knowledge is inaccessible.

The present paper aims at contributing to the existing empirical literature, mainly grounded on either case studies or micro-level analysis, by providing insights from the regional aggregate level analysis. Accordingly, the contribution of KIBS to aggregate regional innovation is examined pinpointing two specific hypothesis.

Hypothesis 1 New knowledge is produced not only through internal R&D investments and some firms might prefer to rely on external sources of knowledge to innovate. Alongside the more traditional first knowledge infrastructure, represented by universities and public institutes, also the localized concentration of KIBS, corresponding to a second knowledge infrastructure, is expected to be positively related with the level of regional innovative performance.

Hypothesis 2 University knowledge, being analytical and highly codified, can be more easily accessed by firms located nearby universities, having these firms higher probabilities to engage in research collaborations with universities. At the regional aggregate level the relative contribution of R&D investments is thus expected to be more sizeable in regions where firms and universities are co-located and, contrarily, external knowledge sources like KIBS are expected to contribute more in regions where university knowledge is less accessible for firms. Especially R&D consulting firms are expected to contribute more in the latter case provided that they act as bridges between academic and scientific knowledge and the demand of knowledge by local firms.

[5] Actually Strambach defines a third category of KIBS, oriented to marketing and advertising, which makes use of what he calls *symbolic* knowledge, particularly relevant in the culture and creativity industries.

IV.4

Empirical Approach and Data

Regional innovation is studied by adopting the standard Griliches-Jaffe's (Griliches [50], Jaffe [56]) framework. The regional innovative activity, as measured by patent applications per millions of inhabitants, is related to the amount of expenditure in research made by both firms and universities, measured as shares of Gross Domestic Product (variables in log). This empirical framework is applied to a sample of 200 European NUTS II regions for which the necessary data were available. The dependent variable pa_i is measured as the average for the years 2006-2007 while the covariates are taken in a previous period (average 2003-2005) in order to avoid the estimation bias due to simultaneity between input and output.

As discussed in the introduction of this work, the use of research expenditure in universities might reveal not a good measure for the identification of spillovers related to university-industry collaborations and thus a second additional measure is used. This measure ($rank$) is obtained by counting the number of regional universities ranking in the top 500 positions of the ARWU ranking[6] (Shanghai Ranking) and weighting their relative position. More in the detail the ranking has been constructed based on the classification of only European Universities and the final measure is the sum of the regional universities present in the ranking multiplied by a factor which was set increasing with the relative score of the university. In such a manner it is guaranteed that the presence of an university scoring in the top 100 of the ranking is valued more respect to the presence of an university scoring between 400 and 500[7]. The advantage of such a measure it that ranking is constructed on the base of publications and citations and hence represents a piece of knowledge completely accessible. Differently from university knowledge accessed through R&D, any market transaction between universities and firms is not necessary. Moreover while R&D expenditure might be registered by universities operating also in non scientific fields of research the ranking is based on publications and citations in only top scientific journals, making the measure a better proxy of the amount of scientific knowledge which can be accessed by within the region.

For the goal of hypothesis testing the empirical specification of the model proceeds in two steps. In the first it is tested the hypothesis that KIBS do contribute to aggregate regional innovation (*Hypothesis 1* in the previous section). The basic model specification is extended by including the share of regional employees in KIBS ($kibs$) on the regional employed population (equation IV.1). The empirical model is further re-specified by using only the share of workers in high-tech KIBS ($kibsht$) and market KIBS ($kibsmkt$)[8].

[6] Rankings are available at the website http://www.arwu.org/.

[7] More information on the procedure are available to the authors upon request

[8] The choice to identify only two additional sub-cathegories is made based on the general definition of KIBS used in this paper and in the majority of the literature, according to which KIBS are predominantly business-oriented. The Eurostat definition includes in fact also non-business services in the general definition of KIBS (like financial services which are not-exclusively business oriented), while actually only business services are considered in the definition of high-tech KIBS and market KIBS. For

In the second it is tested that the contribution of both R&D expenditures and KIBS concentration to regional innovation varies across regions based on the presence, within the region, of a top-quality universities, as indicated by the *rank* variable (*Hypothesis 2* in the previous section). Therefore the general model is estimated separately for *university regions* and *non-university regions* (equation IV.2).

More in the detail, the first hypothesis is tested through inference on the b_5 parameter in equation IV.1. The parameter is expected to be greater than zero. Likewise the coefficients b_1 and b_2, respectively related to the amount of private firms R&D and university R&D are also expected to be positively sloped. Finally, the b_4 coefficient, which is related to the market potential of the region[9], is also expected to be positive.

$$pa_i = a + b_1 berd_i + b_2 urd_i + b_3 rank_i + b_4 mp_i + b_5 kibs_i + e_i \tag{IV.1}$$

The expected value of the b_5 coefficient continues to be greater than zero when *kibs* is substituted in the model with either *kibsht* of *kibsmkt*. Especially the expectation on the coefficient related to *kibsht* is positive because of the peculiar role attributed to KIBS in using analytical knowledge to bridge the availability of technological innovation and the successful application of it to solve the problem of firms. The coefficient related to *kibsmkt* is also expected greater than zero but, in this case, because the availability of specialized providers of market services can stimulate the part of innovation based on the so-called *soft skills*.

The second hypothesis is tested by comparing the coefficient estimates in the two groups (regimes) of regions. One regime is characterized by the presence of at least one top-ranked university (which corresponds to the condition $rank > 0$). By the opposite, the second regime groups regions without top-ranked universities. Accordingly, a dummy variable R is created taking non-zero value if at least one of the top-ranked universities is located in the region. The structural formulation is summarized in the equation IV.2; $X = (berd, urd, mp, kibs)$ is a matrix of covariates and $d^{'} = (a, b_1, b_2, b_4, b_5)$ is a vector of related coefficients. The formulation allows the expected value of each coefficient to vary across the two regimes (the values in the d_1 vector represent the differences between the two). In particular the coefficient related to *berd* is expected to be larger in regions in the university regime and the one related to *kibs* in regions belonging to the non-university regime.

$$\begin{aligned} pa_i &= X_i^{'} d + u_i \\ d &= d_0 + d_1 R_i \end{aligned} \tag{IV.2}$$

Spatial autocorrelation in the data is accounted for by using a suitable spatial econo-

a more detailed description of the European classification see the appendix.

[9] Such a measure is included as control variable in the model specification in an attempt to control for the size of the regional market. It refers to the year 2006 and is available at the ESPON project website www.espon.eu.

metric model indicated by a battery of specification tests and, if necessary, using heterochedasticity consistent estimation methods for spatial data which are described in the next section.

IV.5

Results

The empirical analysis of the innovative activity in the sample of EU regions starts with the estimation of the model described in the equation IV.1 with OLS methods. The estimation output is summarized in the table IV.1. In the first column (model *(a)*) the variation in regional innovative activity is explained by the only private and university expenditure in research and, additionally, by the size of the regional market. Both coefficients related to research are positive and significant. Moreover the estimate relative to the private research is more than double of the one relative to university research. As expected also the coefficient related to the market potential is positive and significant. In the second column (model *(b)*) the *rank* variable is added. Its coefficient is positive and significant and the inclusion of this variable does not alter the main results described above. The variable *kibs* is further introduced in model *(c)* and enters into the model with a positive and significant coefficient. It is now noticeable a significant decrease in the coefficient estimates for private and university research, especially in the latter case.

Table IV.1: Linear Model - OLS Estimates

	(a)	*(b)*	*(c)*
Intercept	-4.931***	-4.519***	-7.847***
	(1.119)	(1.121)	(1.300)
berd	0.698***	0.677***	0.590***
	(0.073)	(0.074)	(0.072)
urd	0.319***	0.269***	0.189***
	(0.078)	(0.080)	(0.079)
rank		0.109**	0.092***
		(0.047)	(0.045)
mp	1.634***	1.502***	1.092***
	(0.236)	(0.241)	(0.247)
kis			1.470***
			(0.326)
	Diagnostics on Linear Model Residuals		
$Moran'sI$	0.135 [0.000]	0.137 [0.000]	0.125 [0.000]
RLM_{ERR}	40.807 [0.000]	40.627 [0.000]	45.886 [0.000]
LR_{ERR}	5.460 [0.150]	5.870 [0.210]	7.860 [0.160]
RLM_{LAG}	3.758 [0.050]	4.969 [0.020]	0.070 [0.790]
LR_{LAG}	15.100 [0.000]	14.490 [0.000]	25.569 [0.000]

Notes to table IV.1:
SE in parenthesis. Probabilities in brackets.
***, ** and * denote significance at 1, 5 and 10% confidence levels.

Based on the OLS estimates a series of diagnostic statistics are further computed, first to detect residual spatial autocorrelation in the data and, secondly, to choose the correct spatial model. Spatial autocorrelation in OLS residuals is detected through the $Moran'sI$ statistic, the value of which is always larger than its expected value under the hypothesis of spatial randomness[10]. The choice of the spatial model is based on two groups of tests. On the one side the usual Robust Lagrange Multipliers (RLM) diagnostics (Anselin et al. [6]) based on OLS residuals compare the most simple non spatial model with the lag (LAG) and error (ERR) alternatives[11]. A significant value of the statistic indicates that the relative spatial model is to be preferred to the linear model and the specification with the higher statistic is chosen. On the other side the Likelihood Ratio (LR) tests compare the most general spatial model, the Spatial Durbin model, with the lag and error alternatives, both nested in the sooner (LeSage and Pace [68]). A significant value of the statistic indicates that the most general, Spatial Durbin, model captures the spatial structure of the data better than the relative alternative and the model with non-significant statistic is chosen.

The RLM statistic is always significant when the basic model is compared to the ERR alternative while it is significant for the only models *(a)* and *(b)* when the alternative is the LAG. On the opposite the statistic turns insignificant in model *(c)*, when *kibs* is introduced in the model specification. In any of the observed cases the value of the RLM_{ERR} statistic is always larger than the RLM_{LAG} one and, therefore, the preference is for the error specification. In addition the LR test is always significant when LAG is the alternative and it is never when ERR is the alternative. Also in this case the error specification is therefore preferred. Accordingly, for the remaining of the empirical analysis the Spatial Error Model (SEM) specification is used to account for spatial correlation in the data.

Estimates obtained by using the SEM specification applied to the model in equation IV.1 are presented in the table IV.2. Coefficients in the first column (model *(d)*) are directly comparable with those in the model *(c)* of table IV.1. Both the coefficients related to research have decreased in magnitude while, on the opposite, the coefficient related to the concentration of KIBS and to the market potential variable have increased. The only noticeable difference is that once the spatial error structure of the residuals is taken into account the coefficient for university research turns insignificant, while the coefficient for the university ranking remains significant and of the same magnitude. Finally the estimated error autoregressive parameter λ is always positive and statistically significant.

Turning the attention to the aim of the reesarch question, it is worth noting that the coefficient relative to the regional concentration of KIBS is strongly significant. Thus the concentration of KIBS in the region positively affects the innovative activity of firms in the same region. The result does not change when the concentration of either the only

[10]Under spatial randomness the statistic should show a value of $E(I) = \frac{-1}{N-1}$, where N indicates the number of observations. The p-values associated to the test in which the alternative hypothesis is non-randomness (two-sided test) are obtained under randomization (Anselin [5]).

[11]LAG refers to the Spatial Lag Model, in which a spatially lagged dependent variable in included as explanatory variable. ERR refers to the Spatial Error Model in which the error structure of the linear model is assumed to follow a conditional autoregressive process.

Table IV.2: Spatial Error Model - ML Estimates

	(d)	*(e)*	*(f)*
Intercept	-9.038***	-5.075***	-4.652***
	(1.412)	(1.184)	(1.220)
berd	0.545***	0.530***	0.583***
	(0.068)	(0.072)	(0.070)
urd	0.112	0.170**	0.176**
	(0.072)	(0.072)	(0.072)
rank	0.091**	0.085**	0.079*
	(0.041)	(0.043)	(0.043)
mp	1.210***	1.460***	1.218***
	(0.256)	(0.255)	(0.302)
kis	1.609***		
	(0.384)		
kisht		0.477**	
		(0.206)	
kismkt			0.613**
		6 (0.254)	
λ	0.874***	0.878***	0.877***
	(0.076)	(0.074)	(0.074)
	Diagnostic for Heteroskedasticity		
Spatial BP test	12.229 [0.032]	9.662 [0.085]	13.239 [0.021]

Notes to table IV.2:
SE in parenthesis. Probabilities in brackets.
***, ** and * denote significance at 1%, 5% and 10% confidence levels.

high-tech KIBS (model *(e)*) or the only marketing KIBS (model *(f)*) are considered. The contribution of KIBS continues to be positive and significant and the magnitude of the estimated coefficient is larger in the latter case. When only high-tech and market KIBS are respectively considered the coefficient related to university research gains significant again.

The lowest part of the table IV.2 reports the result of the Breusch-Pagan test for heteroschedasticity adapted by Anselin [4, pp. 121-122] to spatial models. The null hypothesis of homoschedastic errors is always rejected, although only at a significance level higher than the common 5% in the case of model *(e)*.

In testing the second research hypothesis an heteroschedasticity consistent estimator is applied to the spatial error specification of the model in equation IV.2[12]. The model is basically a restricted form of the general Cliff-Ord type of spatial models[13] in which disturbances are considered heteroschedastic. The so called HAC model is formalized by Arraiz et al. [9] and Kelejian and Prucha [60] and implemented in the *sphet* R-package by Piras [84]. It is considered restricted as long as, in order to maintain the SEM structure, the coefficient of the lagged dependent variable is arbitrarily set to zero.

The estimates are obtained by using the Generalized Spatial Two Stage Least Square

[12] Actually the spatial error structure is added to the reduced form of the model in equation IV.2. This allows to simultaneously estimate the model parameters for the two different group of regions while assuming a spatial structure of the random part of the model which is common to the two groups.

[13] Cliff-Ord type models are spatial models in which both a spatially lagged dependent variable is included in the right hand side of the model equation and a spatial structure is assumed for the disturbances. More details on the classification of spatial models are available in LeSage and Pace [68].

Table IV.3: Heteroskedastic Spatial Error Model - GS2SLS Estimates

	(g)		*(h)*		*(i)*	
	uni	not uni	uni	not uni	uni	not uni
Intercept	-8.215***	-9.489***	-4.917**	-5.031***	-4.882***	-4.302***
	(2.505)	(2.247)	(2.286)	(1.541)	(2.259)	(1.691)
berd	0.745***	0.396***	0.719***	0.394***	0.753***	0.468***
	(0.097)	(0.103)	(0.103)	(0.109)	(0.096)	(0.111)
urd	-0.042	0.121	-0.003	0.178	0.014	0.176
	(0.105)	(0.130)	(0.104)	(0.142)	(0.102)	(0.144)
mp	1.354***	1.030***	1.494***	1.387***	1.405***	1.063**
	(0.441)	(0.313)	(0.486)	(0.328)	(0.513)	(0.481)
kibs	1.228***	1.929***				
	(0.480)	(0.629)				
kibsht			0.314	0.573*		
			(0.264)	(0.343)		
kibsmkt					0.349	0.704
	(0.324)	(0.456)				
λ	0.900***		0.900***		0.900***	
	(0.256)		(0.217)		(0.187)	

Notes to table IV.3:
SE in parenthesis. Probabilities in brackets.
***, ** and * denote significance at 1%, 5% and 10% confidence levels.

procedure described by Piras [84] and results are summarized in the table IV.3. As in the more general case the concentration of all KIBS activities is used (model *(g)*), but also the concentration of only high-tech KIBS (model *(h)*) and of market KIBS as well (model *(i)*). In all the three cases coefficients seem to vary between the two groups of university and non-university regions[14]. By looking at the model *(g)* it emerges that the coefficient related to private firms research is positive and significant while, on the contrary, the one related to university research it is not. Also the coefficient for market potential is of the expected sign and significant and turning the attention to KIBS, the related coefficient continues to be positive and significant as well.

As expected, the *berd* coefficient is larger for the group of university regions. Investments in research made by private firms are more productive in terms of innovation in regions where top-ranking universities are localized. By the opposite, the *kibs* coefficient is larger in the sample of regions where top-ranked universities are absent. The result indicates that in absence of a scientific knowledge base publicly available, firms prefer to rely more on external sources of innovation, as provided by KIBS. Therefore the evidence suggests that actually KIBS do constitute a second knowledge infrastructure. Concerning the hypothesis on the capacity of KIBS to bridge scientific knowledge into practical innovative solutions for firms it is worth looking at the model *(h)* in which only high tech KIBS are considered. Again the estimation output indicates that there are differences between the two groups. More in the detail the *berd* coefficient continues to be larger in university regions while the *kibsht* coefficient is larger in non-university regions. Moreover, limited to the group of university regions, the *kibsht* coefficient turns

[14]Unfortunately an exact statistic for the significance of the differences in coefficients is not available for HAC models. While in fact a modified version of the Chow test for coefficient stability is available for general spatial models in case of likelihood-based estimation, the same version of the test cannot be used after IV/2SLS estimation.

insignificant. The evidence thus further suggests that in presence of wide physical distances between firms and universities firms rely on external sources of knowledge also to internalize the available academic and scientific knowledge.

IV.6

Discussion and Concluding Remarks

Universities, in particular scientific ones, undoubtedly represent a valuable source of knowledge for firms. Nonetheless, the access to such knowledge is limited because of the distance, both geographical an organizational, which separates universities and firms. Low geographical distance facilitates the interaction between firms and universities as research collaboration requires frequent and continuous exchange of knowledge. As far as it concerns the organizational distance, universities, contrarily to firms, do research to get general purpose innovations which eventually find application also in fields other than those they were originally engineered for. Sometimes this field is related, sometimes it is a completely different one. Firms, after all, have to solve problems. And, for such a purpose, they generally require very specific knowledge.

Building on a theoretical stream of literature which has drawn the attention upon the so-called Knowledge Intensive Business Services as actors of innovation through knowledge transformation, this paper has examined the contribution of KIBS to the regional innovative activities. In building the empirical framework it is argued that expenditure in research and development is not the only driver of innovation at the regional level as firms, especially SMEs might prefer to rely on external sources of knowledge like KIBS are. More in the detail it is argued the KIBS work not only as generic co-producers of innovation, but also contribute in bridging the distance between the general purpose research carried out by universities and specific applications required by firms to solve problems.

The evidence in the paper confirm the research hypothesis. It is found that the regional innovative output is positively related with the amount of research carried out by private firms and by the presence of external public knowledge as well. Alongside with these two main inputs, private external knowledge, as measured by the share of workers in KIBS, also significantly contributes to explain the variance in regional innovative activity. Thus KIBS actually promote innovation at the regional level mainly working as a second knowledge infrastructure. Moreover there is evidence that such a second knowledge infrastructure is more important in the group of regions in which the first knowledge infrastructure is absent. The analysis further reveals that this greater importance attributed to KIBS in regions where university knowledge is absent is twice motivated. One the one side it is characterized the contribution of KIBS as co-innovators of their client firms, mainly through an interactive learning process in which soft-skills are exchanged. On the other side some KIBS, like R&D consulting firms, act to bridge the analytical research carried out by universities in some regions and more specific needs of firms who are willing to innovate but lack the know-how which is necessary to repurpose the result of analytical research to their needs.

The interpretation of the role of KIBS proposed in the present paper sheds new light on the academic and policy debate on regional innovation. From the academic perspective it is worth noting that while KIBS have already received attention in micro-level analysis, more work remains to be done at the regional aggregate level. Both in terms of theoretical conceptualization of the contribution of KIBS to the aggregate regional innovation and in the development of empirical tools which enable to understand the functioning of innovation processes in services, not only in the manufacturing. As far as policy is concerned, with the continue tertiarization of regional production systems, more attention should be deserved to monitoring the composition of different regional innovation systems. R&D targeting has represented the traditional instrument for the evaluation of innovative capacity at the regional level. This mainly because R&D based indicators allow to identify the structural nature of technological gaps in some less developed regions. However the evidence in this paper has emphasized the role played by knowledge which is not only developed within firms through formal and formalized research activities but is also produced in collaboration with third parties. The role of these third parties might thus become fundamental in the diffusion of knowledge, with the important consequences of contributing to fill the technological gap of least developed regions and of promoting the technological catch-up at the European level.

IV.7

Appendix to Chapter IV

Eurostat Classification of Knowledge Intensive Services

Knowledge-intensive high-tech services: Post and Telecommunications (64); Computer and related activities (72); Research and development (73).

Knowledge-intensive market services: Water transport (61); Air transport (62); Real estate activities (70); Renting of machinery and equipment without operator, and of personal and household goods (71); Other business activities (74);

Knowledge-intensive financial services: Financial intermediation, except insurance and pension funding (65); Insurance and pension funding, except compulsory social security (66); Activities auxiliary to financial intermediation (67).

Other knowledge-intensive services: Education (80); Health and social work (85); Recreational, cultural and sporting activities (92).

Bibliography

[1] Zoltan J Acs, Luc Anselin, and Attila Varga. Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, 31(7):1069 – 1085, 2002.

[2] Philippe Aghion and Peter Howitt. A model of growth through creative destruction. *Econometrica*, 60(2):pp. 323–351, 1992.

[3] Stilianos Alexiadis and Judith Tomkins. Technology adoption and club convergence. Technical report, MPRA WP, 2010.

[4] L. Anselin. *Spatial econometrics: methods and models.* Kluwer, Dordrecht, 1988.

[5] Luc Anselin. Spatial regression analysis in r: A workbook. Technical report, Center for Spatially Integrated Social Sciences, 2007.

[6] Luc Anselin, Anil K. Bera, Raymond Florax, and Mann J. Yoon. Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, 26(1):77 – 104, 1996.

[7] Luc Anselin, Attila Varga, and Zoltan Acs. Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics*, 42(3):422 – 448, 1997.

[8] Luc Anselin, Attila Varga, and Zoltan Acs. Geographical spillovers and university research: A spatial econometric perspective. *Growth and Change*, 31(4):501–515, 2000.

[9] Irani Arraiz, David M. Drukker, Harry H. Kelejian, and Ingmar R. Prucha. A spatial cliff-ord-type model with heteroskedastic innovations: Small and large sample results. *Journal of Regional Science*, 50(2):592–614, 2010.

[10] David B. Audretsch and Maryann P. Feldman. Chapter 61 knowledge spillovers and the geography of innovation. In J. Vernon Henderson and Jacques-François Thisse, editors, *Cities and Geography*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2713 – 2739. Elsevier, 2004.

[11] Corinne Autant-Bernard and James P. LeSage. Quantifying knowledge spillovers using spatial econometric models. *Journal of Regional Science*, 51(3):471–496, 2011.

[12] Costas Azariadis and Allan Drazen. Threshold externalities in economic development. *The Quarterly Journal of Economics*, 105(2):501–526, 1990.

[13] Harald Badinger, Werner Muller, and Gabriele Tondl. Regional convergence in the european union, 1985- 1999: A spatial dynamic panel analysis. *Regional Studies*, 38(3):241–253, 2004.

[14] Hararld Badinger and Gabriele Tondl. *European Regional Growth*, chapter Trade, Human Capital and Innovation: The Engines of European regional growth in the 1990s, pages 215–238. Springer, 2003.

[15] Richard E. Baldwin and Rikard Forslid. The core-periphery model and endogenous growth: Stabilizing and destabilizing integration. *Economica*, 67(267):pp. 307–324, 2000.

[16] Fabrizio Barca. An agenda for a reformed cohesion policy. Technical report, Independent Report, 2009.

[17] Tomás Del Barrio-Castro and José García-Quevedo. Effects of university research on the geography of innovation. *Regional Studies*, 39(9):1217–1229, 2005.

[18] R. Barro and X. Sala-i Martin. *Economic Growth*. MIT Press, 2004.

[19] Robert J. Barro. Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, 106(2):pp. 407–443, 1991.

[20] Roberto Basile. Regional economic growth in europe: A semiparametric spatial dependence approach. *Papers in Regional Science*, 87(4):527 – 544, 2008.

[21] Roger S. Bivand, Edzer J. Pebesma, and Virgilio Gomez-Rubio. *Applied Spatial Data Analysis*. Springer, 2008.

[22] Ron Boschma. Proximity and innovation: A critical assessment. *Regional Studies*, 39(1):61–74, 2005.

[23] Maarten Bosker. Growth, agglomeration and convergence: a space-time analysis for european regions. *Spatial Economic Analysis*, 2(1):91–100, 2007.

[24] Laura Bottazzi and Giovanni Peri. Innovation and spillovers in regions: Evidence from european patent data. *European Economic Review*, 47(4):687 – 710, 2003.

[25] Stefano Breschi and Francesco Lissoni. Knowledge spillovers and local innovation systems: A critical survey. *Industrial and Corporate Change*, 10(4):975–1005, 2001.

[26] Giulio Cainelli, Rinaldo Evangelista, and Maria Savona. The impact of innovation on economic performance in services. *The Service Industries Journal*, 24(1):116–130, 2004.

[27] Giulio Cainelli, Rinaldo Evangelista, and Maria Savona. Innovation and economic performance in services: a firm-level analysis. *Cambridge Journal of Economics*, 30(3):435–458, 2006.

[28] Antonio Ciccone. Agglomeration effects in europe. *European Economic Review*, 46(2):213 – 227, 2002.

[29] Wesley M. Cohen and Daniel A. Levinthal. Innovation and learning: The two faces of r & d. *The Economic Journal*, 99(397):pp. 569–596, 1989.

[30] Nicoletta Corrocher, Lucia Cusmano, and Andrea Morrison. Modes of innovation in knowledge-intensive business services evidence from lombardy. *Journal of Evolutionary Economics*, 19:173–196, 2009. 10.1007/s00191-008-0128-2.

[31] R Cowan, PA David, and D Foray. The explicit economics of knowledge codification and tacitness. *Industrial and Corporate Change*, 9(2):211–253, 2000.

[32] Riccardo Crescenzi. Innovation and regional growth in the enlarged europe: The role of local innovative capabilities, peripherality, and education. *Growth & Change*, 36(4):471 – 507, 2005.

[33] Sandy Dall'erba and Julie Le Gallo. Regional convergence and the impact of european structural funds over 1989-1999: A spatial econometric analysis. *Papers in Regional Science*, 87(2):219 – 244, 2008.

[34] Pim Den Hertog. Knowledge intensive business services as co-producers of innovation. *International Journal of Innovation Management*, 4:491–528, 2000.

[35] EC. Investing in europe's future. fifth report on economic, social and territorial cohesion. Technical report, European Commission, 2010.

[36] EC. Regional policy for smart growth in europe 2020. Technical report, European Commission, 2011.

[37] Cem Ertur and Wilfried Koch. Regional disparities in the european union and the enlargement process: an exploratory spatial data analysis, 1995-2000. *The Annals of Regional Science*, 40:723–765, 2006.

[38] Jan Fagerberg and Bart Verspagen. Heading for divergence? regional growth in europe reconsidered. *Journal of Common Market Studies*, 34(3):431–448, 1996.

[39] Jan Fagerberg, Bart Verspagen, and Marjolein Caniels. Technology, growth and unemployment across european regions. *Regional Studies*, 31(5):457–466, 1997.

[40] Manfred M. Fischer and Attila Varga. Spatial knowledge spillovers and university research: Evidence from austria. *The Annals of Regional Science*, 37:303–322, 2003.

[41] Manfred M. Fischer and Attila Varga. Spatial knowledge spillovers and university research: Evidence from austria. *The Annals of Regional Science*, 37:303–322, 2003.

[42] Michael Fritsch and Viktor Slavtchev. The role of regional knowledge sources for innovation? an empirical assessment. Technical report, Freiberg Working Papers 15, 2005.

[43] Michael Fritsch and Viktor Slavtchev. Universities and innovation in space. *Industry & Innovation*, 14(2):201–218, 2007.

[44] Faïz Gallouj and Paul Windrum. Services and services innovation. *Journal of Evolutionary Economics*, 19:141–148, 2009. 10.1007/s00191-008-0123-7.

[45] Oded Galor. Convergence? inferences from theoretical models. *The Economic Journal*, 106(437):pp. 1056–1069, 1996.

[46] P. Geroski. Markets for technology: Knowledge, innovation and appropriability. In P. Stoneman, editor, *Handbook of Economics of Innovation and Technological Change*, pages 90–131. Blackwell: Oxford, 1995.

[47] Edward L. Glaeser, Hedi D. Kallal, Josï£ · A. Scheinkman, and Andrei Shleifer. Growth in cities. *The Journal of Political Economy*, 100(6):1126–1152, 1992. Press.

[48] Lydia Greunz. Geographically and technologically mediated knowledge spillovers between european regions. *The Annals of Regional Science*, 37:657–680, 2003.

[49] Daniel A. Griffith and Pedro R. Peres-Neto. Spatial modeling in ecology: the flexibility of the eigenfunction spatial analyses. *Ecology*, 87(10):2603–2613, 2006.

[50] Zvi Griliches. Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics*, 10(1):92–116, 1979.

[51] Christoph Grimpe and Roberto Patuelli. Regional knowledge production in nanomaterials: a spatial filtering approach. *The Annals of Regional Science*, 46:519–541, 2011. 10.1007/s00168-009-0355-y.

[52] G. M. Grossman and E. Helpman. *Innovation and Growth in the Global Economy*. MIT Press, Cambridge, MA, 1991.

[53] Giovanni Guastella and Francesco Timpano. Spillover diffusion and regional convergence: a gravity approach. *Regional Science Inquiry Journal*, 2 (2):71–82, 2010.

[54] Mercedes Gumbau Albert and Joaquin Maudos. Patents, technological inputs and spillovers among regions. *Applied Economics*, 41(12):1473–1486, 2009.

[55] Heinz Hollenstein. Innovation modes in the swiss service sector: a cluster analysis based on firm-level data. *Research Policy*, 32(5):845 – 863, 2003.

[56] Adam B. Jaffe. Real effects of academic research. *The American Economic Review*, 79(5):pp. 957–970, 1989.

[57] Adam B. Jaffe, Manuel Trajtenberg, and Rebecca Henderson. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–598, 1993.

[58] Morten Berg Jensen, Björn Johnson, Edward Lorenz, and Bengt Åke Lundvall. Forms of knowledge and modes of innovation. *Research Policy*, 36(5):680 – 693, 2007.

[59] David Keeble and Frank Wilkinson. Collective learning and knowledge development in the evolution of regional clusters of high technology smes in europe. *Regional Studies*, 33(4):295–303, 1999.

[60] Harry H. Kelejian and Ingmar R. Prucha. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157(1):53 – 67, 2010.

[61] A. Kleinknecht. *Firm Size and Innovation. Observations in Dutch Manufacturing Industries.* SEO, Amsterdam, 1989.

[62] P. Krugman. *Geography and Trade.* The MIT Press, Cambridge, Massachussets, 1991.

[63] P. Krugman and Antony J. Venables. Integration, specialization, and adjustment. *European Economic Review*, 40(3):959–967, 1996.

[64] Paul Krugman and Anthony J. Venables. Globalization and the inequality of nations. *The Quarterly Journal of Economics*, 110(4):pp. 857–880, 1995.

[65] Naomi R. Lamoreaux and Kenneth L. Sokoloff. *Inventors, Firms, and the Market for Technology in the Late Nineteenth and Early Twentieth Centuries*, pages 19–60. University of Chicago Press, January 1999.

[66] J. Le Gallo, C. Ertur, and C. Baumont. *European regional growth*, chapter A Spatial Econometric Analysis of Convergence Across European Regions, 1980 1995, pages 99–130. Springer, 2003.

[67] Julie Le Gallo and Cem Ertur. Exploratory spatial data analysis of the distribution of regional per capita gdp in europe, 1980 1995. *Papers in Regional Science*, 82(2):175–201, 2003.

[68] James LeSage and Robert Kelley Pace. *Introduction to Spatial Econometrics.* Chapman & Hall, 2009.

[69] Ge Lin and Tonglin Zhang. Loglinear residual tests of moran's i autocorrelation and their applications to kentucky breast cancer data. *Geographical Analysis*, 39(3):293–310, 2007.

[70] Enrique Lopez-Bazo, Esther Vaya, Antonio J. Mora, and Jordi Surinach. Regional economic dynamics and convergence in the european union. *The Annals of Regional Science*, 33:343–370, 1999.

[71] Jesús López-Rodríguez, J. Andrés Faíña, and José López-Rodríguez. Human capital accumulation and geography: Empirical evidence from the european union. *Regional Studies*, 41(2):217 – 234, 2007.

[72] Robert Lucas. On the mechanics of economic development. *Journal of Monetary Economics*, 22(1):3 – 42, 1988.

[73] N. Gregory Mankiw, David Romer, and David N. Weil. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, 107(2):407–437, 1992.

[74] Edwin Mansfield. Academic research underlying industrial innovations: Sources, characteristics, and financing. *The Review of Economics and Statistics*, 77(1):pp. 55–65, 1995.

[75] Ron Martin and Peter Sunley. Slow convergence? the new endogenous growth theory and regional development. *Economic Geography*, 74(3):pp. 201–227, 1998.

[76] Rosina Moreno, Raffaele Paci, and Stefano Usai. Spatial spillovers and innovation activity in european regions. *Environment and Planning*, 37:1793–1812, 2005.

[77] Emmanuel Muller and David Doloreux. What we should know about knowledge-intensive business services. *Technology in Society*, 31(1):64 – 72, 2009.

[78] Emmanuel Muller and Andrea Zenker. Business services as actors of knowledge transformation: the role of kibs in regional and national innovation systems. *Research Policy*, 30(9):1501 – 1516, 2001.

[79] OECD. How regions grow. Technical report, Organization for Economic Cooperation and Development, 2009.

[80] OECD. Regions matter. Technical report, Organization for Economic Cooperation and Development, 2009.

[81] Raffaele Paci, Francesco Pigliaru, and Maurizio Pugno. *Regional Economic Growth, SMEs and the Wider Europe*, chapter Disparities in Economic Growth and Unemployment Across European regions: A Sectoral perspective, pages 75–102. Ashgate, 2003.

[82] K. Pavitt. Sectoral patterns of technical change: towards a taxonomy and a theory. *Research Policy*, 13:343–373, 1984.

[83] Roberta Piergiovanni and Enrico Santarelli. Patents and the geographic localization of r&d spillovers in french manufacturing. *Regional Studies*, 35(8):697–702, 2001.

[84] Gianfranco Piras. sphet: Spatial models with heteroskedastic innovations in r. *Journal of Statistical Software*, 35:1–21, 2010.

[85] Roderik Ponds, Frank van Oort, and Koen Frenken. Innovation, spillovers and university industry collaboration: an extended knowledge production function approach. *Journal of Economic Geography*, 10(2):231–255, 2010.

[86] Roderik Ponds, Frank van Oort, and Koen Frenken. Innovation, spillovers and universityâĂŞindustry collaboration: an extended knowledge production function approach. *Journal of Economic Geography*, 10(2):231–255, 2010.

[87] Roderik Ponds, Frank Van Oort, and Koen Frenken. The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, 86(3):423–443, 2007.

[88] Danny T. Quah. Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics*, 95 (4):427–443, 1993.

[89] Danny T. Quah. Empirics for economic growth and convergence. *European Economic Review*, 40(6):1353 – 1375, 1996.

[90] Andrés Rodríguez-Pose and Riccardo Crescenzi. Research and development, spillovers, innovation systems, and the genesis of regional growth in europe. *Regional Studies*, 42(1):51 – 67, 2008.

[91] Paul M. Romer. Increasing returns and long-run growth. *Journal of Political Economy*, 94(5):pp. 1002–1037, 1986.

[92] Margaret Sharp. Competitiveness and cohesion. are the two compatible? *Research Policy*, 27(6):569 – 588, 1998.

[93] Richard Shearmur and David Doloreux. Urban hierarchy or local buzz? high-order producer service and (or) knowledge-intensive business service location in canada, 1991-2001. *The Professional Geographer*, 60(3):333–355, 2008.

[94] James Simmie and Simone Strambach. The contribution of kibs to innovation in cities: an evolutionary and institutional perspective. *Journal of Knowledge Management*, 10:26–40, 2006.

[95] Robert M. Solow. A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1):65–94, 1956.

[96] Alessandro Sterlacchini. R & d, higher education and regional growth. uneven linkages among european regions. *Research Policy*, 37(6-7):1096 – 1107, 2008.

[97] S. Strambach. Knowledge-intensive business services (kibs) as drivers of multilevel knowledge dynamics. *International Journal of Services Technology and Management*, 10:152–174, 2008.

[98] Gabriele Tondl. The changing pattern of regional convergence in europe. *Jahrbuch für Regionalwissenschaft*, 91(1):1–33, 1999.

[99] Bart Verspagen. The spatial hierarchy of technological change and economic development in europe. *The Annals of Regional Science*, 45:109–132, 2010.

[100] Eric von Hippel. "sticky information" and the locus of problem solving: Implications for innovation. *Management Science*, 40(4):pp. 429–439, 1994.

[101] WB. Reshaping economic geography. Technical report, World Bank, 2009.

[102] P. Wood. *Consultancy and Innovation.* Routledge: London, 2002.

[103] S. N. Wood. *Generalized Additive Models: An Introduction with R.* Chapman & Hall, Boca Raton, 2006.

[104] Lynne G. Zucker, Michael R. Darby, and Jeff Armstrong. Geographically localized knowledge: Spillovers or markets? *Economic Inquiry*, 36(1):65–86, 1998.