



OECD Science, Technology and Industry Working Papers
2008/03

The Geography of Inventive Activities in OECD Regions

Stefano Usai

<https://dx.doi.org/10.1787/230083500260>

Unclassified

DSTI/DOC(2008)3

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

01-Dec-2008

English - Or. English

DIRECTORATE FOR SCIENCE, TECHNOLOGY AND INDUSTRY

THE GEOGRAPHY OF INVENTIVE ACTIVITIES IN OECD REGIONS

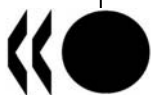
STI WORKING PAPER 2008/3

Statistical Analysis of Science, Technology and Industry

Stefano Usai

JT03256586

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format



DSTI/DOC(2008)3
Unclassified

English - Or. English

STI Working Paper Series

The Working Paper series of the OECD Directorate for Science, Technology and Industry is designed to make available to a wider readership selected studies prepared by staff in the Directorate or by outside consultants working on OECD projects. The papers included in the series cover a broad range of issues, of both a technical and policy-analytical nature, in the areas of work of the DSTI. The Working Papers are generally available only in their original language, English or French, with a summary in the other.

Comments on the papers are invited, and should be sent to the Directorate for Science, Technology and Industry, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

The opinions expressed in these papers are the sole responsibility of the author(s) and do not necessarily reflect those of the OECD or of the governments of its member countries.

<http://www.oecd.org/sti/working-papers>

Copyright OECD, 2008

Applications for permission to reproduce or translate all or part of this material should be made to: OECD Publications, 2 rue André-Pascal, 75775 Paris, Cedex 16, France; e-mail: rights@oecd.org

THE GEOGRAPHY OF INVENTIVE ACTIVITIES IN OECD REGIONS

Stefano Usai*

Abstract

This work reflects an initial analysis employing a pioneering new OECD database; it is among the first systematic attempts to analyse comparatively the distribution of innovative activity across regions in OECD economies with a set of homogenous measures for both input and output in the process of knowledge production and dissemination.

The descriptive analysis shows that there are important differences in the inventive performance of regions in OECD economies, as measured by indicators for one of the key types of intellectual assets (*i.e.*, patents). Inventive performance is concentrated in some regions in continental Europe, in North America and Japan. Highly inventive regions tend to cluster together. This spatial dependence is found to have increased over time. The inventive performance of regions is directly influenced by the availability of human capital and R&D expenditure. Local agglomeration factors (proxied by the density of population) are also found to have a significant impact while some negative effects appear when regions are mainly rural or when they are mainly service-oriented. Cross-country differences point to the importance of national innovation systems which shape the institutional framework within which innovation takes form and diffuses.

* Stefano Usai, CRENoS, University of Cagliari, Via S. Ignazio 78, I-09123 Cagliari, Italy

LA GEOGRAPHIE DES ACTIVITES D'INVENTION DANS LES REGIONS DE L'OCDE

Stefano Usai*

Résumé

Ce travail est le fruit d'une première analyse à partir d'une base de données OCDE nouvelle et novatrice. Il correspond à l'une des premières tentatives d'effectuer de manière systématique des analyses comparatives de la distribution de l'activité d'innovation entre les régions, dans les économies de l'OCDE, et comporte une batterie d'indicateurs homogènes du processus de production et de diffusion du savoir (facteurs de production et produits).

L'analyse descriptive montre que, dans les économies de l'OCDE, l'inventivité des régions, telle que mesurée par les indicateurs de l'un des principaux types d'actifs intellectuels (les brevets, par exemple), n'est absolument pas homogène. L'inventivité se concentre dans quelques régions du continent européen, d'Amérique du Nord et du Japon. Les régions à forte inventivité ont tendance à se constituer en réseaux. On a d'ailleurs constaté que cette dépendance spatiale a augmenté au fil du temps. La disponibilité de capital humain et les dépenses de R-D influent directement sur l'inventivité des régions. On observe également que des facteurs d'agglomération locaux (dont la variable indicatrice est la densité de la population) ont un impact significatif alors que certains effets négatifs se font sentir quand les régions sont principalement rurales ou principalement orientées vers les services. Les disparités transnationales mettent en exergue l'importance des systèmes nationaux d'innovation qui façonnent le cadre institutionnel au sein duquel l'innovation prend forme et se diffuse.

* Stefano Usai, CRENoS, Université de Cagliari, Via S. Ignazio 78, I-09123 Cagliari, Italie.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	7
1. INTRODUCTION	10
2. THEORETICAL AND EMPIRICAL BACKGROUND.....	11
2.1 Theoretical underpinnings	12
2.2 Empirical work	13
3. SOME METHODOLOGICAL AND DATA ISSUES.....	17
4. REGIONAL DESCRIPTIVE ANALYSIS.....	19
4.1 Geographical Analysis	19
4.2 Spatial dependence.....	37
5. ECONOMETRIC ESTIMATION	39
5.1 The KPF Model	40
5.2 The dataset	42
5.3 Robustness checks	46
6. CONCLUSIONS	56
BIBLIOGRAPHY	58

List of figures

Fig. 1 – OECD: PCT, 1998-2000	23
Fig. 2 – OECD: PCT, 2002-2004	24
Fig. 3 - OECD: PCT per million Population, 1998-2000	25
Fig. 4 - OECD: PCT per million Population, 2002-2004	26
Fig. 5 - PCT, Best Performance (30 regions), 2002-2004	29
Fig. 6 - PCT per million population, Best Performance (30 regions), 2002-2004.....	30
Fig. 7 - OECD Regions: PCT per million population variability, 1998-2000	33
Fig. 8 - OECD Regions: PCT per million population variability, 2002-2004	34
Fig. 9 - PCT per million population, Innovation Convergence (OECD Countries) - 1998-2004	36
Fig. 10 - PCT per million population, Innovation Convergence (OECD Regions) - 1998-2004	37

List of tables

Table 1 - OECD: Nations and Regions – PCT, PCT per million Population	21
Table 2 - Macro-Areas – PCT, PCT per million Population	22
Table 3 - PCT OECD by Region – Synthetic indicators.....	31
Table 4 - PCT Europe by Region – Synthetic indicators.....	31
Table 5 - PCT North America by Region – Synthetic indicators	32
Table 6 - PCT Asia/Pacific by Region – Synthetic indicators.....	32
Table 7 - OECD Regions: Coefficient of variation of PCT per capita	35
Table 8 - OECD: PCT per capita, Moran’s I test for spatial autocorrelation.....	38
Table 9 - Europe: PCT per capita, Moran’s I test for spatial autocorrelation.....	38
Table 10 - North America: PCT per capita, Moran’s I test for spatial autocorrelation	39
Table 11 - Asia/Pacific: PCT per capita, Moran’s I test for spatial autocorrelation.....	39
Table 12 - Regressors synthetic statistics	43
Table 13 - KPF estimation.....	45
Table 14 - KPF estimation with interactive dummies	47
Table 15 - KPF estimation with spatial lag of RD.....	49
Table 16 - KPF estimation with distance matrix	51
Table 17 - KPF estimation with Japan and Korea	53
Table 18 - KPF estimation with PCT per worker	55

EXECUTIVE SUMMARY

Introduction

Economic activities, and innovative activities, are not randomly distributed across space but tend to concentrate geographically. Knowledge spill-overs add to other agglomeration forces and locational determinants (local infrastructure, a pool of skilled workers, the presence of other firms, and a relatively larger market). Geography matters for the spatial distribution of intellectual assets and innovation activities in particular, as knowledge flows and specific skills often require proximity to be fully exploited. Indeed, knowledge is often tacit and requires direct interaction, on-the-job training, and the mobility of workers to circulate. The degree of uncertainty surrounding innovation projects can also be reduced by the exchange of information among firms, and the complexity of innovation processes often requires accessing complementary inputs and co-ordinating different aspects of the innovative activity. The presence of upstream and downstream activities further promotes the agglomeration of activities, and in particular of innovation activities that require horizontal and vertical linkages.

This raises a number of questions, which are addressed in this paper: What is the regional profile of inventive activities? How skewed is the cross-regional distribution of inventive activities? What influence (positive, negative) do inventive regions exercise on neighbouring regions? What are the factors that make regions inventive or not – R&D, human capital, influence of neighbouring regions, country-level factors?

The answers are based on a preliminary analysis, both descriptive and econometric, of a new OECD database in which inventors and owners of patents are coded to regions within the 30 OECD countries.

Data issues and descriptive statistics

Given the importance of localised spill-overs for innovation and the diffusion and transmission of intellectual assets in theory, it is important to attempt to measure the extent of such spill-overs, as well as their rate of decay with distance, in practice. This is a difficult task, especially in light of numerous measurement challenges, not only for capturing information on intellectual assets, but also innovative activity itself. The approach taken in this paper uses patents as an indicator of innovative output. There are some advantages in such an approach: for instance, as an invention is needed to apply for a patent, patent statistics do not include the unsuccessful R&D (contrary to R&D statistics). However, it should be stated at the outset that there are some limitations in using patent indicators for such a purpose. This is because the relationship of patents to the overall volume of innovation is not necessarily linear. Some innovators may opt not to employ intellectual property (IP) protection or may choose to rely on other forms of IP protection instead of patents (*e.g.* copyrights or trade secrecy).

The number of Patent Cooperation Treaty (PCT) applications is used as the main indicator of inventive performance. Using patents rather than R&D data has three major advantages: *i*) patents include information about the residence of the inventor and can therefore be grouped regionally; R&D statistics are usually only available for certain regions, or at the national level; *ii*) patents include technological content and can therefore be grouped into industrial sectors; R&D data tend to be aggregated, especially at the regional level; and *iii*) patents are available over long time periods whereas regional R&D data have become available only relatively recently, and not on a continuous basis. The data used come from the

OECD Regional Database of statistics on socio-economic indicators (e.g. demographics, economy and labour market, and social issues) in some 2 014 regions across the 30 OECD member economies. The level of detail chosen for the analysis is the so-called Territorial Level 2 (TL2), where possible, and country level elsewhere.

The paper examines mainly two time periods, 1998-2000 and 2002-2004. Growth in PCT applications is found greatest in Japan and Korea, which are still modest utilisers of this type of applications. Most countries that experienced low PCT values in the first period saw a significant increase in PCT in the second period (Mexico, Poland, Turkey, and Slovak Republic). Countries with strong performances in the first period saw lower growth in the second period (Finland, Norway, Sweden, and Luxembourg). Most European regions showed significant PCT applications growth. Most regions of innovation excellence can be found in Europe and the United States, but Japan also has some high-performance regions.

Geographical proximity of regions also proves to matter, both for high and low performance innovators, especially in Europe and the United States. The results point to four main innovation clusters with strongly significant spatial interactions: the West and East US coastal belts, the “spine” of continental Europe (Switzerland, Southern Germany), and Scandinavia. Asia Pacific regions show a rather different pattern.

Econometric analysis

Finally, a “knowledge production function”, relating inputs to inventive performance at the regional level, has been estimated based on previous works, such as Moreno *et al.* (2005) for Europe and De Crescenzi *et al.* (2007) for Europe and the United States. Inputs include human capital, R&D, agglomeration effects, country level characteristics (institutional factors) and factors of spatial correlation. The percentage of the population having completed a tertiary education degree is used as a proxy for human capital. The clustering patterns of the PCT applications indicators are reflected to some extent in the geographical distribution of the human capital indicator, but are not nearly as pronounced in the distribution of the latter as it tends to be relatively high for most of the territory. The same is true when looking at the patterns of the R&D indicator (the percentage of GDP spent on R&D). Thus, it appears that while human capital and R&D are very important determinants of innovative performance, other factors contribute to the spatial correlation of regions.

The inventive performance of a region can be related to the share of skilled workers as the concentration of skilled workers in a region should enable externalities to materialise since direct communication enables flows of information and tacit knowledge. At the same time such a concentration can give rise to important pecuniary externalities in the job market for talented and experienced workers and researchers. The density of economic exchanges and contacts is assumed to act as a catalyst for agglomeration effects. Innovation systems have been found to matter for innovation performance at the aggregate level and regional innovation institutions and strategies are likely similarly important in stimulating and diffusing innovation. Finally, with local spill-overs, innovation generated in one region may spill-over and help knowledge formation in other regions, especially nearby ones. Indeed, the production of knowledge in a region may depend not only on its own research efforts, but also on the knowledge stock available in the whole economy and on its ability to exploit it. Many factors, including those external to the region, can have an impact on technological activity, such as trade and investment flows and common markets for skilled labour and final goods.

Preliminary empirical evidence on the regional dimension of innovation confirms the importance of R&D and human capital for innovation. Local agglomeration factors (proxied by the density of population) are also found to have a significant impact on the innovation indicator (PCT applications). There are some negative effects from being a rural region, but there is no effect from being an urban region. On the

contrary we found a negative effect from being a capital region, probably due to its specialisation in services rather than manufacturing. Institutional factors are found to have some impact, generally positive in regions with a lot of innovative activity, and small or negative in regions with low innovative activities. This is likely to reflect the general efficiency of research, influenced by economy-wide institutional settings as well as the productivity effects of the diffusion of knowledge throughout the economy. Relatively higher GDP also has a positive impact, although the capacity of R&D and human capital to drive further innovation appears to slow down with GDP. In Europe there is also some positive effect on a region's PCT applications from R&D undertaken in nearby regions which could point to some interregional knowledge spill-overs and a geographical proximity effect.

THE GEOGRAPHY OF INVENTIVE ACTIVITIES IN OECD REGIONS

1. Introduction¹

Our research line starts out from the belief that technological activity is the main engine of growth and that its functioning has an important geographical component. The localisation of innovative activity may, in fact, depend on some factors and externalities which are local in nature since one of the main ingredients of innovation, knowledge, can be more easily transmitted through interpersonal contacts.

This paper is the first systematic attempt to analyse the processes of knowledge creation and diffusion comparatively across regions in OECD countries, using a set of homogenous measures for both input and output in the process of knowledge production and diffusion.

The main objectives are as follows:

- To describe the patterns of innovation across regions in some OECD countries in order to determine what are the main factors behind the process of localisation of such activities.
- To estimate a Knowledge Production Function (KPF) at the regional level for OECD regions.
- To assess the role of R&D efforts, localised knowledge and absorptive capacity in enhancing regions' technological capability.
- To assess if the pattern, extent and pace of the absorptive capacity depend on geographical proximity (measured in several ways) but also on other economic and institutional similarities.

These aims are pursued through a new database concerning social, economic and innovative phenomena in OECD regions, set up by OECD and improved upon by CRENoS (Centre for North-South Economic Research). The new database is thoroughly described and investigated with a wide set of statistical tools in order to understand the distribution pattern of innovation at the regional level. Moreover, the novelty of the database is exploited along the way pioneered by Jaffe and recently renewed by Crescenzi *et al.* (2007). Another important original aspect of this study regards the use of a specific spatial econometric strategy, which allows for rigorous testing procedures in the search for the best specification. Firstly, we check the need to introduce external spillovers in the estimation of a regional KPF. Once this need is assessed, the statistically correct specification of the KPF is searched by means of the usual testing strategy of spatial econometrics. In the process, we study the geographical scope of spillovers, in order to evaluate and measure the possible existence of a spatial decay effect. Additionally, we explore the relevance of sharing a common institutional background in facilitating spillovers across regions.

1. This research is a by-product of a report for the OECD. This paper is partially the result of a research effort made by several researchers at CRENoS and in particular Raffaele Paci and Fabiano Schivardi who have contributed during the setting of the research agenda and with several insightful suggestions. I am also particularly grateful to Matteo Bellinzas, Barbara Dettori and Giuliana Caruso for their excellent research assistance. Finally I would like to thank Dominique Guellec and Colin Webb for their help throughout the process of database building and research.

The descriptive analysis shows that there are important differences in the innovative performance of regions in OECD countries. Such performance is clustered in some areas in continental Europe, in North America and to a lesser extent in Japan. Spatial dependence across regions is quite stable through time in almost all OECD countries except for the dispersed macro-region of Asia - Pacific.

In spite of some regularities which characterise the three macro-areas of Europe, North America and Asia - Pacific, we find that there are important differences across countries which probably indicate the relevance of national innovation systems within regional performance.

The estimation of a model which links innovative output to its main inputs shows that this process is at work across regions in OECD countries. Some particular results are worth noting: R&D is a crucial determinant and so are human capital and agglomeration economies. Moreover, the performance of neighbouring regions is an important determinant of innovative performance. Institutional factors, which vary country by country, are relevant too. The effect of R&D is not monotonous but it decreases as the country becomes richer. These results are differentiated with respect to two macro-areas: Europe and North America. All in all, the KPF model seems to be able to explain innovative performance both in Europe and in North America even though some interesting differences emerge.

This paper has the following structure. The next section presents a review of the main theoretical and empirical underpinnings on which the analysis of the distribution of innovation is based. The third section deals with the main issues concerning the data and the measurement of innovative activity. It provides a justification for the choice of patents in general and the Patent Cooperation Treaty (PCT) ones in particular. The fourth section is the core of the empirical analysis, where maps, tables and figures are used to illustrate the main characteristics of the geographical distribution of innovation across regions in OECD countries. In particular, spatial statistical techniques are implemented in order to assess the existence of phenomena of spatial dependence, their strength and extent. All such phenomena are analysed with respect to two time periods, and some comments of the trend at work are given. We also compare the regional distribution of patents with that of R&D efforts and of human capital, which introduces the econometric estimation of section five based on the Knowledge Production Function (KPF). Section five presents the KPF model, the estimation strategy and the main results. Such results are then tested with respect to a number of robustness checks. Section six concludes the work.

2. Theoretical and empirical background

The last three decades have witnessed a resurgence of interest in the localisation of innovative activity. This resurgence has occurred within the context of a more general attention to the issue of the spatial distribution of economic activity, following the seminal contribution of Krugman on economic geography (Krugman, 1991). It is by now undisputed that economic activity is not randomly distributed in space, even after controlling for spatial characteristics: rather, activities tend to agglomerate spatially (Ellison and Glaeser, 1997). This phenomenon occurs also for innovation activities. Audretsch and Feldman (1996) further argue that “the propensity for innovative activity to cluster is more attributable to the role of knowledge spillovers and not merely to the geographic concentration in production”.

This section reviews and explores the theoretical and empirical underpinnings on which the descriptive and econometric analyses on OECD regions that follows is built. We start from the theoretical literature on the agglomeration of the innovation activity that offers an insight into why this activity tends to be clustered spatially. The basic answer is that innovation activity benefits from a series of inputs external to the innovating entity. Such external inputs can take the form of either pecuniary externalities or knowledge spillovers (KS in what follows). We then move on to review the empirical work that uses patents to study the geographical dimension of the innovation process. We focus in particular (but not exclusively) on the research based on the “Knowledge Production Function” (KPF, Griliches, 1979; Jaffe,

1986). We investigate what spatial dimension would be “correct” for the analysis of spatial interactions in the innovation process. We also review the literature that focuses on the specific channels through which knowledge flows take place. In particular, we consider the role of different actors in the innovation process and on the characteristics of the local structure that foster innovation.

2.1 *Theoretical underpinnings*

From a theoretical point of view, the traditional reasons behind agglomeration, beyond the case of innovation, can be traced back to the work of Marshall (1890) more than one century ago:

“When an industry has thus chosen a locality for itself, it is likely to stay there long: so great are the advantages which people following the same skilled trade get from neighbourhood to one another. The mysteries of trade become no mysteries; but are as in the air, and children learn many of them unconsciously.... Employers are apt to resort to any place where they are likely to find a good choice of workers with the special skill which they require. .. The advantages of variety of employment are combined with those of localised industries in some of our manufacturing towns, and this is a chief cause of their continued economic growth”.

Marshall had mostly manufacturing activities in mind. Subsequently, the economic literature has investigated why spatial clustering is particularly important for innovation.² This question has been addressed by different fields of the economics of innovation.

Urban and regional economics consider the spatial aspects of technological innovation and address the question of why regions differ in their innovation potential. Following Davelaar (1991), one can distinguish two main approaches: the first stresses structural factors, such as differences in endowments of innovative inputs; the second focuses on the elements that make a given environment more conducive to innovation for given inputs. In particular, researchers have considered the characteristics that help knowledge circulate more easily among firms. Aydalot (1986) and other researchers at GREMI (*Groupe de Recherche Européen sur les Milieux Innovateurs*) create the notion of *milieux innovateur*, defined as a network of relations that reduces the risk connected with the innovation activity and facilitates tacit knowledge diffusion. Storper (1992), following this research path, later proposes the notion of technological district, where firms manage to combine the advantage of specialisation and flexibility thanks to a local cluster where dynamic technological learning is the key to success. More recently, Antonelli (2008) stresses the role of pecuniary externalities in both the generation and exploitation of technological knowledge which leads to the emergence of local technological systems of innovation.

Industrial economics and the economics of technological change stress the importance of geographical proximity for innovation. The work on endogenous growth (Romer, 1986; Krugman, 1991) has formalised the idea that innovation is characterised by increasing returns; moreover, such returns are in part external to the firm, due to the non-exclusive and non-rival character of knowledge (Arrow 1962). The idea, again, is that innovation requires inputs external to the firm, in terms of specific skills, knowledge flows etc. Such inputs require geographical proximity to be fully exploited. This is mainly due to three reasons (Jacobs, 1969; Feldman, 1994), namely:

- Knowledge includes an important tacit component that cannot be easily codified and therefore requires direct interaction, on-the-job learning and workers’ mobility to circulate.
- The high degree of uncertainty surrounding innovation activity may be reduced by the exchange of information among firms.

2. See, for example, Arthur (1990) and Swann (1996).

- The complexity of innovation activity requires the co-ordination of different capabilities and benefits by the possibility of accessing sophisticated complementary inputs.

The “*new economic geography*” (NEG), originated by the work of Fujita (1988), Krugman (1991) and Venables (1996), studies the economic forces that can sustain a large permanent imbalance in the distributions of economic activities in a general equilibrium framework with monopolistically competitive firms. This literature, surveyed by Ottaviano and Thisse (2004), distinguishes between centrifugal (*i.e.* promoting dispersion) and centripetal (*i.e.* promoting agglomeration) forces in determining the spatial distribution of activities. The first general result is the so-called “home market effect”, according to which, when transportation costs are sufficiently low, imperfectly competitive sectors tend to locate in the largest market and serve smaller markets via export. This effect is strengthened by two forces. First, labour mobility contributes to further increase the size of the largest markets, as workers move there and contribute to local demand. Second, the location of an industry increases the convenience to localise in the same place for upstream and downstream industries, due to input-output linkages. These reinforcing effects imply that, in general equilibrium, even small differences in initial conditions among regions can give rise to large differences in equilibrium allocation of economic activities. While not explicitly motivated by the study of innovation activities, the insights of the NEG clearly apply to them as well. In particular, the imperfectly competitive structure is a feature of technological change, as well as the horizontal and vertical linkages with other innovative activities.

2.2 Empirical work

Theoretical work helps understanding why innovation activities are strongly clustered spatially, given the importance of KS. Of course, the extent of such spillovers and the rate at which they decay with distance are open to empirical investigation. Some of the most important contributions in the field have been concerned exactly with investigating this question, which is fraught with difficulties. The first important issue is how to measure innovative output, which is dealt with in the next section. Many different indicators have been used in the literature. Given our interest in patents, in what follows we will focus mainly on papers using patents as the indicator of innovative output.

The literature on patents proceeded mainly along two paths. A first strand tried to directly measure the extent of knowledge flows, using patent citation as evidence of such flows. Jaffe, Trajtenberg and Henderson (1993) were the first to use patent citations as a proxy of knowledge spillovers. They found that nearby inventors have a much higher propensity to cite each other’s patents compared to those located far away from each other. This conclusion has been corroborated by subsequent work on patent citations (Jaffe *et al.*, 1993, 1999, 2002; Hall *et al.*, 2001; Lukatch and Plasman, 2003; Criscuolo and Verspagen, 2006; Maurseth and Verspagen, 1999, 2002; Paci and Usai, 2008; Le Sage *et al.*, 2007)³. The robust finding that emerges from this body of work is that knowledge is localised to a large extent. Unlike this literature, our work is not based on patent citations.

A second set of studies uses the KPF to measure the intensity and spatial extent of KS. This literature is directly related to our work and will therefore be reviewed in more detail, even though not exhaustively. The KPF approach was introduced by Griliches (1979) to study the functional relationship between knowledge inputs and outputs. Griliches applied it at the firm level; subsequent literature has generalised both the unit of analysis (an area, a sector...) and expanded the set of inputs that might enter the production of innovation (R&D employee/expense, university or firm-based, the spatial extent of such indicators, urban density and diversity etc.). This approach is very flexible and has been extensively used to analyse

3. One should mention the existence of a parallel line of analysis on knowledge flows, measured through relational data other than patent citations, such as network of co-inventors (see Maggioni *et al.* 2007).

how the surrounding economic activity enters a unit's KPF, using spatial econometric techniques to assess the rate of spatial decay of KS.

Most studies, particularly in the early literature, are based on US data. Jaffe (1989) explores the existence of geographically mediated spillovers from university research to commercial innovation. He uses a modified Cobb-Douglas model to relate the dependent variable (number of patents assigned to corporations) to two inputs: industry R&D and university research. The paper analyses the US States in the 70s; patents are grouped into 5 broad classes, in order to be matched with the academic department's technical areas. Results show a significant effect of university research on corporate patents. University research appears to have an indirect effect on local innovation by inducing industrial R&D spending; on the contrary, the econometric test on university equation shows that there is no evidence to support the hypothesis that private R&D has any sort of effect on university research. These results are broadly confirmed by Acs, Audretsch, Feldman (1992) using innovation counts rather than patents (considered a better proxy of innovation output) as a measure of innovative output.

Audretsch and Feldman (1996) examine the extent to which industrial activity clusters spatially and link this geographic concentration to the existence of knowledge externalities. The main hypothesis of the paper is that knowledge externalities are stronger in industries where "new economic knowledge" plays a greater role, and innovative activity will tend to cluster in such industries. The model takes into account transportation costs, dependence upon natural resource inputs and capital intensity/importance of scale economies. According to their results, the propensity for innovative activity to cluster will tend to be higher in industries where new economic knowledge is predominant - that is where industry R&D, university research and skilled labour are most important. These industries also tend to exhibit greater geographic concentration of production.

Anselin, Varga and Acs (1997) test the role of universities and the spatial extent of its geographic effects in the stimulation of technological innovation. The model analyses the innovation counts and takes into account employment in high technology research laboratories (for industrial R&D activity) and scientific and engineering expenditure at universities and colleges (for university research). Results show significant and positive coefficients for both private R&D and university research (similar values at state level; industry R&D elasticity five times higher than university research at metropolitan areas level). The introduction of a spatial model (at MSAs level) improves the model fit: private R&D effect falls outside state borders, while university research influence remains positive and significant. They also find that the specialisation indexes for high technology employment and employment in business services have a positive impact on innovative output. Acs, Anselin and Varga (2002) perform a similar exercise using patent data as a measure of output to test reliability of the patent data as a proxy for regional innovative activity. All the parameters representing knowledge flows in the region yield significantly higher values for patents than for innovations, an indication that patents might over-emphasise the effects of localised interactions. Furthermore, the influence of local university research spillovers is under-represented if compared to the effects of R&D spillovers when patents are used as a dependent variable. They conclude that patents provide a reliable, albeit imperfect, measure of innovative activity.

The common feature of all studies reviewed above is the focus on pure (knowledge) spillovers rather than local externalities as a whole. The possibility that local externalities may also have a pecuniary nature is usually ruled out *a priori*. Breschi and Lissoni (2001) provide a critical review of the risks of such an approach, which treats knowledge spillovers as homogeneous entities. The latest studies have attempted to investigate the main general mechanisms of the process of creation and diffusion of innovative knowledge rather than just looking for localised KS. Such studies have been implemented on the US case (Varga *et al.*, 2005, Carlino *et al.*, 2007) and on Europe (Bottazzi and Peri, 2003; Greunz, 2003a; Moreno, Paci and Usai, 2005, Rodriguez Pose and Crescenzi, 2007 and Tappeiner *et al.* 2008). All in all, these

contributions find that technological spillovers (both pure and pecuniary) may exist both within and across regions.

In particular, Bottazzi and Peri (2003) estimate the effect of R&D externalities in generating innovation for 86 European Regions in the 1977–1995 period. Their findings show that spillovers are localised and take effect within a 180-minute travel radius. However, the size of these spillovers is small: “doubling R&D spending in a region would increase the output of new ideas in other regions within 300 km only by 2–3%, while it would increase the innovation of the region itself by 80–90%”.

Greunz (2003) investigates inter-regional KS across European sub-national regions, trying to assess if geographical and technological proximity influences the creation of new knowledge within European regions. The econometric model considers a regional KPF that allows for extra regional innovation-generating inputs. Results show that “interregional knowledge spillovers exist between geographically close regions and between regions displaying similar technological profiles”. Although “technological proximity and geographical proximity coincide to a certain extent”, KS result as being driven mainly by private business. When KS take effect at the national level, the country borders appear to hold up inter-regional spillovers.

Moreno *et al.* (2005) find that the external spillovers decay over space, for 17 European countries (EU 15 plus Norway and Switzerland), and that they occur mainly across regions within a country rather than across nations. This implies that national innovation systems, with their institutions and policies, are important determinants of regional performance. Finally, they find that technological proximity alone is not relevant in determining technological spillovers across regions. Nevertheless, technological closeness when associated with geographical proximity is proven to matter in defining the strength and extent of spillovers.

Rodriguez-Pose and Crescenzi (2006) use a model, for the first time referring to the EU-25 regions, that combine R&D, spillovers, and innovation system approaches, allowing discrimination “between the influence of internal factors and external knowledge and institutional flows on regional economic growth”. Results highlight the importance of proximity (spillovers show strong distance decay) and interaction between local and external research in maximising innovation capacity.

Other studies at the national level in the same line of research are Autant-Bernard (2001) for the French departments, Andersson and Ejeremo (2003) for Swedish functional regions, Fischer and Varga (2003) for Austrian political districts and Buesa *et al.* (2006) for the Spanish regions.

The only paper that analyses different continents at the regional level is Crescenzi *et al.* (2007), who study what extent of the innovation gap between the United States and Europe can be attributed to differences in the spatial organisation of innovative activities (as opposed to differences in the levels of inputs). They find that the United States seems to take better advantage of the localised character of innovation activities, due to higher mobility of capital, labour force and knowledge, in turn related to a higher degree of economic integration. It is worth noting that their model, unlike ours, is based on the use of a different dependent variable for the two macro-areas (EPO patents for Europe and USPTO patents for the United States). Moreover, they focus on the dynamics of patent activities rather than on differences in their level, making it problematic to compare their results with others in the literature.

Finally, in her doctoral thesis, Cihan (2005) analyses a KPF for 19 OECD countries based on USPTO patents, but at the country rather than the regional level. The results show that the total stock of domestic innovations and human capital positively and significantly affect knowledge production.

The last strand of literature reflects on which characteristics of the local structure influence innovative activity. Following the previous Marshallian citation, the literature has been debating the relative importance of concentration, of having “innovators” in the same field of activity clustered together (specialisation economies), or having a sort of diversification in production (urbanisation economies). In a seminal contribution, Glaeser, Kallal, Scheinkman, and Shleifer (1992) estimate the effects of alternative sources of externalities on local economic growth. They find strong evidence (in a cross-section of U.S. cities) that specialisation and the scale of activity have negative effects on growth, while productive variety is positively related to the subsequent growth. Similar evidence has emerged in studies regarding other OECD countries.⁴ Due to data limitations, work in the urban growth literature has been based on employment growth regressions, relying on the assumption that increases in productivity, related to innovation, result in proportional employment gains through shifts in labour demand. Cingano and Schivardi (2004) claim that this assumption is empirically hard to defend, as local characteristics are likely to shift labour supply as well. A more direct gauge of innovative output is needed to correctly measure KS. They use local TFP growth, showing that with this indicator the results of the employment regressions are reversed: they find substantial positive effects of both specialisation and city-size on local TFP growth, while diversification does not play any role. This result suggests that the KPF approach should correctly identify KS: in fact patents, as well as TFP, should be considered a direct measure of innovative output.

The debate on the effects of concentration and diversity on innovation is still open. Feldman and Audretsch (1999) test the role of specialisation (Marshallian externalities) vs. diversity (Jakobian externalities) on the creation of innovative activity measured at the sector level. The authors test whether the number of innovations from a given sector in a state owes more to the city specialisation in that specific sector or to the presence, within the state, of other industries whose science base is related to that of the specific sector. The most striking, and probably unexpected, result is that there is no evidence of specialisation externalities, whilst diversity externalities are at work in the case of US metropolitan areas. In other words, in the United States innovation in a specific sector exhibits strong spatial clustering independently of the sectoral distribution of manufacturing activity in the same sector. In other words, diversity fosters innovation more than specialisation, contrary to the findings of Cingano and Schivardi (2004).

Some contrasting results were found by Paci and Usai (2000) for European regions where there is a positive association between the spatial distribution of technological activity and production specialisation. This result has been confirmed with a different statistical analysis and a larger sample of countries by Greunz (2003b). These works are based on an improved Feldman-Audretsch estimation model firstly applied to the Italian local labour systems by Paci and Usai (1999).⁵ These two contributions find spillovers arising from production specialisation and from production and innovation diversity. Moreover, Paci and Usai (1999) provide some evidence that spillovers may not be constrained by administrative regional borders. Thus, the divergence of results between the European and American cases despite the usage of similar methodology suggests a notable difference in the functioning of the local innovation systems in United States and Europe.

These divergent results are also challenged by the outcome of the analysis by Massard and Riou (2002), who applied the analysis to French *départements* and do not find any evidence of either

4. See, for example, Bradley and Gans (1998) for Australia, Paci and Usai (2002, 2008) for Italy, Combes (2000) for France, and Van Oort *et al.* (2005) for the Netherlands.

5. An important result for the case of Italy is found by Laursen *et al.* (2007) who analyse the role of Social Capital on Innovative activity. The authors point out that “social capital can be considered a geographically bound public good that reduces both search costs between firms and between firms and other organisations”, thus easing the innovation production activity in such locations. The empirical analysis of a sample of Italian manufacturing firms at regional level finds that “being located in regions characterised by a high level of social interaction leads to a higher propensity to innovate” and [...] “positively moderates the effectiveness of externally acquired R&D on innovation”.

specialisation or diversity externalities. Finally, Moreno *et al.* (2006) find that specialised innovative clusters do exist and are getting stronger in Europe. This important result contrasts with the common findings of the literature for production clusters, according to which they are continuously eroded by an ongoing delocalisation process. A possible interpretation is that firms' strategic activities, like innovation, are still greatly influenced by locally bounded interactions with similar firms, with which positive localisation externalities, mostly purely technological, are at work.

Along this vein, the latest contribution of Carlino and Hunt (2007) examines the effects of local economic characteristics on the rate of innovation (as measured by patents) in over a dozen industries. According to their analysis, the availability of human capital is perhaps the most important factor explaining the invention rate for most industries. Moreover, they find evidence of increasing returns with respect to city size (total jobs) for many industries and more modest effects for increases in the size of an industry in a city. Crucially, they confirm the difference between US and EU, since they find that specialisation among manufacturing industries is not particularly helpful. Nevertheless, they find the opposite for specialisation among service industries. In other words they show that sectors may benefit to varying degrees from different local sources of R&D (academia, government labs, and private labs).

Following the contribution of Ó Huallicháin (1999), Carlino *et al.* (2007) consider the role of urban density in determining patent intensity – the per capita invention rate. They conclude that patent intensity is positively related to the density of employment in US metropolitan areas. Most importantly, they find evidence of increasing returns to scale in the invention process, but holding density constant, these returns are exhausted at a modest city size - certainly below 1 million in population. Similarly, they find evidence of diminishing returns to density. Their results also support the view that more competitive local market structures are more conducive to innovation: the industrial and technology mix is important in explaining the variation in patent intensity across cities, whilst no significant effect is found for their measures of industrial and technological specialisation. Finally, local R&D inputs, especially human capital, contribute to higher patent intensities.

3. Some methodological and data issues

Several contributions in the past have made extensive use of patent statistics in order to analyse the spatial distribution of innovation activity. In particular, in the case of European regions, Breschi (2000) and Caniels (2000) have provided an articulated and extensive analysis of the spatial distribution of innovation in European regions until the 90s, whilst Paci and Usai (2000) have tried to address the same issue of agglomeration of innovation and production for a smaller set of countries. In the case of the United States, there is a long tradition of studies based on patents as an innovation indicator, starting from Jaffe (1986) to the latest contribution by Carlino and Hunt (2007).

These precedents remind us that the use of patents as indicators of innovative activity may imply some inconveniences and shortcomings which ought to be kept in mind while interpreting the outcome of the analysis, both descriptive and econometric. Since the extensive review by Griliches (1990), economists have debated the issue of measuring innovative activity and technological progress, even though no universal solution has been proposed. Based on the concept of knowledge production function, two types of indicators are usually identified: technology input measures (such as R&D expenditure and employees) and technology output measures (such as patents and new product announcements).

The main drawback of those indicators is that they do not distinguish between firms' efforts for invention and innovation *and* imitation activities. Moreover, they do not take into account informal technological activity and, as a consequence, may underestimate the amount of innovative activity. On the contrary, patents represent the outcome of the inventive and innovative process which has some novelty and usability features. Moreover, since patenting implies relevant costs for the proponent, it is reasonable

to assume that patented innovations have an expected value higher than those costs. This in turn implies that patented innovations, especially those extended in foreign countries, are expected to have economic value, however heterogeneous. Nevertheless, we should remember that there may be inventions which are never patented as well as patents which are never developed into real products and processes.

With respect to the object of our research,⁶ that is, to study innovative activity across regions, sectors and time, patent statistics seem particularly suitable. First of all, given that inventions mean the creation of new ideas or knowledge and that innovations can be defined as “development, commercial utilisation of the new products or processes” [Basberg (1987)], patents cover both of these aspects. Compared to R&D data, they also exhibit some useful properties which are summarised below:

- a) They provide information on the residence of the inventor and proponent and can thus be grouped regionally (potentially at different territorial levels, starting from zip codes), whilst R&D statistics are available only for some regions or at the national level.
- b) They record the technological content of the invention and can, thus, be classified according to the industrial sectors, whilst R&D data is usually aggregated, especially at the regional level.
- c) They are available year by year for a long time span and this allows for a dynamic analysis. On the contrary regional R&D data is available only for recent years and, for some countries, discontinuously.

This paper thus provides an analysis of the performance of OECD member countries where innovation is measured by patent production. The analysis is based on the latest available international comparable data on patents made available by OECD. Patent indicators are designed to reflect trends in innovative activities across a wide range of regions among OECD member countries. The databank on innovation is associated with the OECD Regional Database (RDB) which provides quantitative information on socio-economic issues in some 2 014 regions within 30 OECD member countries. The database includes regional statistics on three major topics (demographics, economy and labour market, social issues) and a derived indicator on the typology of regions, which distinguishes urban, rural and intermediate regions.

The main indicator for innovation activity is identified as the number of Patent Cooperation Treaty (PCT) applications instead of the traditional indexes based on either the European Patent Organisation (EPO) or the United States Patent Office (USPTO), or the Japanese Patent Office (JPTO). The latter data are, as a matter of fact, rather problematic when applied to international comparisons outside their respective macro areas. Since patents at EPO, USPTO and JPTO protect innovation within their respective geographical area, they are preferred by domestic firms, and thus their quota overestimates their innovative capability with respect to foreign firms.

PCT data is selected from a couple of newly built databases recently introduced in the research of innovative activity thanks to the OECD: the Triadic Patent Family dataset (TPF) and the PCT itself. These two datasets contain patent data which do not suffer (or suffer less) from the usual home bias effect of the EPO, USPTO, JPTO data.

6. Note that since 2000 there is an important EU initiative called the European Trend Chart on Innovation which provides several indicators on innovation (based on input and output data and on the Community Innovation Survey) at the regional level and a synthetic measure of them. Unfortunately, the time and the sectoral dimension of such a database are rather limited as its geographical scope is based on Europe. Nevertheless, for the time being such a database is going to become an increasingly crucial point of reference for the analysis in this field.

In particular, the PCT database consists of patent applications and it provides a unified procedure for filing international patent applications to protect inventions in each of its 139 Contracting Countries.⁷ Its procedure is an intermediate step between the priority application and filing for patent protection abroad and it is increasingly used to file international patent applications. This expansion is strongly correlated with the number of contracting states which has doubled since the mid-1990s.

The TPF database is made up of innovations patented at the EPO, JPO and USPTO that share one or more priorities.⁸ In terms of statistical analysis, they improve the international comparability of patent-based indicators, as only patents applied for in the same set of countries are included in the “family”: home advantage and influence of geographical location are therefore eliminated. Moreover, patents included in the family are typically of higher value: patentees only take on the additional costs and delays of extending protection to other countries if they deem it worthwhile.

PCT data are, nevertheless, preferred to TPF data for several reasons clarified below.

The TPF database is difficult to regionalise given that each family refers to multiple patents. On the contrary, PCT data is available at the regional level, thus giving the information required for a detailed report on innovation.

Another important feature regards the real value of the innovative process, and its comparability at the international level. As the PCT procedure is costly and constitutes a step before the national patent award, it is assumed that PCT data is mostly made up of valuable inventions (those which ensure higher profits). The TPF definition is tighter than the PCT one, its procedure is more expensive, and it ensures the selection of fairly valuable inventions only. Therefore, the PCT database ensures a satisfactory balance between selecting valuable inventions and ensuring a detailed representation of the innovation process. In particular, since PCT refer to about 140 contracting partners, the home bias effect is negligible. With TPF, the fact that applications refer only to EPO, JPTO and USPTO makes the indicator potentially biased, since it excludes some key emerging economies such as China, India, Russia and Korea.

The choice of one of the two alternative available datasets is not critical, however, since the rankings of PCT and TPF per capita at the national level are very similar except for a few countries such as Japan, Finland and Korea.

4. Regional descriptive analysis

4.1 Geographical analysis

The maps, figures and tables below refer to two main indicators for innovative performance: the absolute value of PCT and PCT per capita (or better, per million population). This allows to take into account the high heterogeneity of regions with respect to population size. All data refer to two periods made up of three years, in order to smooth out yearly peaks (1998-2000 and 2002-2004). We present an analysis of the short-term dynamics between these two periods. The choice of the territorial units is the Territorial Level 2 (TL2) whenever possible, that is, for 23 out of 30 nations. For the remaining cases (Denmark, Iceland, Ireland, Luxembourg, Mexico, New Zealand and Turkey), the availability of information restricts the analysis at the country level. Some macro-areas (Europe, North America and

7. According to WIPO, the procedure under the PCT has great advantages for the applicant, the patent offices and the general public: *i)* the applicant has up to 18 months more than in a procedure outside the PCT to reflect on the desirability of seeking protection in foreign countries, *ii)* the search and examination work of patent offices can be considerably reduced or virtually eliminated thanks to the international search report; *iii)* since each international application is published together with an international search report, third parties are in a better position to formulate a well-founded opinion about the patentability of the claimed invention.

8. Patent families are derived from priority application (first filing to a patent office). A single priority may lead to several patents or a single patent may include several priorities..

Asia/Pacific) are selected in order to perform a deeper investigation and to make a few comparisons across continents.

Table 1 provides a summary of the statistics concerning the main indicator for innovative performance and its regional counterpart. The first columns provide some general information concerning the regions in each country. Columns two and three, for instance, report the number of regions and the level of variable disaggregation in each country. The maximum number of regions is that of the United States with its 51 states, while the minimum is that of Belgium with three regions. As mentioned before, seven countries' data are reported at the national level. Columns four and five describe the average region within each country in terms of surface area and population size. These two dimensions are quite important to interpret correctly both the geographic analysis of the PCT distribution and the econometric analysis which follows. The sample of regions is extremely heterogeneous, since regions may range from an average size of almost 1 million km² in Australia to 6 000 km² in Switzerland. The regions exhibit a similarly wide range, though less marked, in demographics: the most populous regions, the Japanese prefectures, have almost 12 million inhabitants on average, whilst the smallest ones, in Norway, have around 600 000 people. Looking at macro-areas, North America is characterised by vast territories while the Asia/Pacific basin by quite high population levels, particularly in Japan and Korea. In Europe, the Scandinavian peninsula has the largest territory but an average population density at the regional level which does not exceed 20 inhabitants per km².

The subsequent columns focus on PCT values. At the OECD level, the aggregate flow of PCT in 2002-2004 was of 308 163 granted patents, an increase of around 27% from the 1998-2000 level (around 242 000 in absolute terms). Such growth regards PCT per capita too, which increases by about 24%, from 71.98 to 89.15.

Table 1. OECD: Nations and regions – PCT, PCT per million population

OECD		Level	average region (size km ²)	average region (population 2002-04)	PCT			PCT per million population		
Nation	Regions				1998-2000	2002-2004	var %	1998-2000	2002-2004	var %
Australia	8	TL2	962,919	2,484,040	4,466	5,618	25.8%	78.65	94.23	19.8%
Austria	9	TL2	9,319	902,893	1,856	2,589	39.5%	77.40	106.20	37.2%
Belgium	3	TL2	10,173	3,458,922	2,210	2,332	5.5%	72.06	74.91	4.0%
Canada	12	TL2	766,934	2,639,336	5,579	6,992	25.3%	61.14	73.59	20.4%
Czech Republic	8	TL2	9,859	1,276,004	216	215	-0.5%	7.00	7.02	0.3%
<i>Denmark</i>	<i>1</i>	<i>TL0</i>	<i>43,098</i>	<i>5,389,733</i>	<i>2,396</i>	<i>2,975</i>	<i>24.2%</i>	<i>150.12</i>	<i>183.99</i>	<i>22.6%</i>
Finland	5	TL2	67,629	1,042,780	3,959	3,571	-9.8%	255.50	228.30	-10.6%
France	22	TL2	24,726	2,735,994	12,069	14,800	22.6%	68.56	81.96	19.5%
Germany	41	TL2	8,708	2,012,461	35,355	42,702	20.8%	143.54	172.51	20.2%
Greece	4	TL2	32,907	2,756,083	136	122	-10.3%	4.17	3.69	-11.5%
Hungary	7	TL2	13,290	1,447,386	442	459	3.8%	14.54	15.10	3.8%
<i>Iceland</i>	<i>1</i>	<i>TL0</i>	<i>102,696</i>	<i>290,661</i>	<i>87</i>	<i>124</i>	<i>42.5%</i>	<i>103.96</i>	<i>142.20</i>	<i>36.8%</i>
<i>Ireland</i>	<i>1</i>	<i>TL0</i>	<i>69,797</i>	<i>3,993,767</i>	<i>576</i>	<i>794</i>	<i>37.8%</i>	<i>51.26</i>	<i>66.27</i>	<i>29.3%</i>
Italy	21	TL2	14,349	2,745,044	4,421	5,998	35.7%	25.89	34.68	34.0%
Japan	10	TL2	37,758	12,758,000	23,658	49,560	109.5%	62.25	129.49	108.0%
Korea	7	TL2	14,209	6,835,549	3,844	9,914	157.9%	27.47	69.06	151.4%
<i>Luxembourg</i>	<i>1</i>	<i>TL2</i>	<i>2,586</i>	<i>449,733</i>	<i>145</i>	<i>95</i>	<i>-34.5%</i>	<i>112.27</i>	<i>70.41</i>	<i>-37.3%</i>
<i>Mexico</i>	<i>1</i>	<i>TL0</i>	<i>1,959,248</i>	<i>101,970,271</i>	<i>245</i>	<i>404</i>	<i>64.9%</i>	<i>0.84</i>	<i>1.32</i>	<i>57.0%</i>
Netherlands	4	TL2	8,471	4,054,683	7,528	7,788	3.5%	158.67	160.06	0.9%
<i>New Zealand</i>	<i>1</i>	<i>TL0</i>	<i>277,039</i>	<i>4,002,267</i>	<i>784</i>	<i>1,019</i>	<i>30.0%</i>	<i>68.14</i>	<i>84.87</i>	<i>24.6%</i>
Norway	7	TL2	43,928	650,180	1,482	1,341	-9.5%	111.08	98.21	-11.6%
Poland	16	TL2	19,543	2,387,900	189	239	26.5%	1.64	2.09	27.5%
Portugal	7	TL2	13,135	1,491,029	71	50	-29.6%	2.33	1.60	-31.3%
Slovak Republic	4	TL2	12,259	1,345,067	79	86	8.9%	4.88	5.33	9.2%
Spain	19	TL2	26,631	2,210,705	1,623	2,118	30.5%	13.54	16.81	24.2%
Sweden	8	TL2	55,168	1,119,863	7,584	5,506	-27.4%	285.32	204.86	-28.2%
Switzerland	7	TL2	5,898	1,052,053	3,760	5,148	36.9%	174.95	233.01	33.2%
<i>Turkey</i>	<i>1</i>	<i>TL0</i>	<i>769,603</i>	<i>70,228,333</i>	<i>196</i>	<i>400</i>	<i>104.1%</i>	<i>0.98</i>	<i>1.90</i>	<i>92.7%</i>
United Kingdom	37	TL2	6,573	1,610,381	12,233	12,041	-1.6%	68.96	67.36	-2.3%
United States	51	TL2	119,774	5,702,560	104,803	123,161	17.5%	125.20	141.16	12.7%
Total	324		94,757	3,556,366	241,992	308,163	27.3%	71.98	89.15	23.8%

Source: OECD Regional Database; Note: in italics national values.

DSTI/DOC(2008)3

The country with the highest number of PCT in the two periods is the United States, whilst Portugal lies at the other extreme with the lowest amount. There are some interesting trends to point out, which imply some convergence in innovative performance at least at the national level. Firstly, there is a remarkable growth of PCT production (per capita and in absolute terms) in Japan and Korea (+102.8 and +151.4, respectively). Secondly, most countries that registered low values in the first period (Mexico, Poland, Turkey, Slovak Republic) also show consistent PCT growth. On the contrary, some countries that performed very well in innovation production in the first period experienced a relative worsening of their innovation rate, especially the Scandinavian countries and Luxembourg in central Europe. Most of continental Europe improved its innovative production capacity, as well as some Mediterranean countries (with the exception of Greece).

Maps 1, 2, 3 and 4, which represent the quantile distribution of the variables under examination for the two periods across 2000, provide an interesting picture and a first insight into the cluster distribution of innovation, and help to develop a preliminary idea of the correlation structure across regions and countries.

The first map shows the absolute values of PCT in the OECD countries. Analysing this map, one finds some major innovative agglomerations (most Western and North American territories, a South-Eastern belt that starts from Texas and ends up at Washington State, Japan and Continental Europe). The visual representation does not change much in the second period (Figure 2): although there is a big increase in Australia, the general distribution of the indicator still highlights the same main clusters.

The third and the fourth map show the per capita values of PCT in the OECD countries in the two periods. Compared to the former, this indicator should be a better measure for taking into account the distinct innovation potential given by different population size. The maps show the geography of innovation in OECD regions, highlighting some clusters centred in Continental and Scandinavian Europe, Western and Eastern North-American states and Japan.

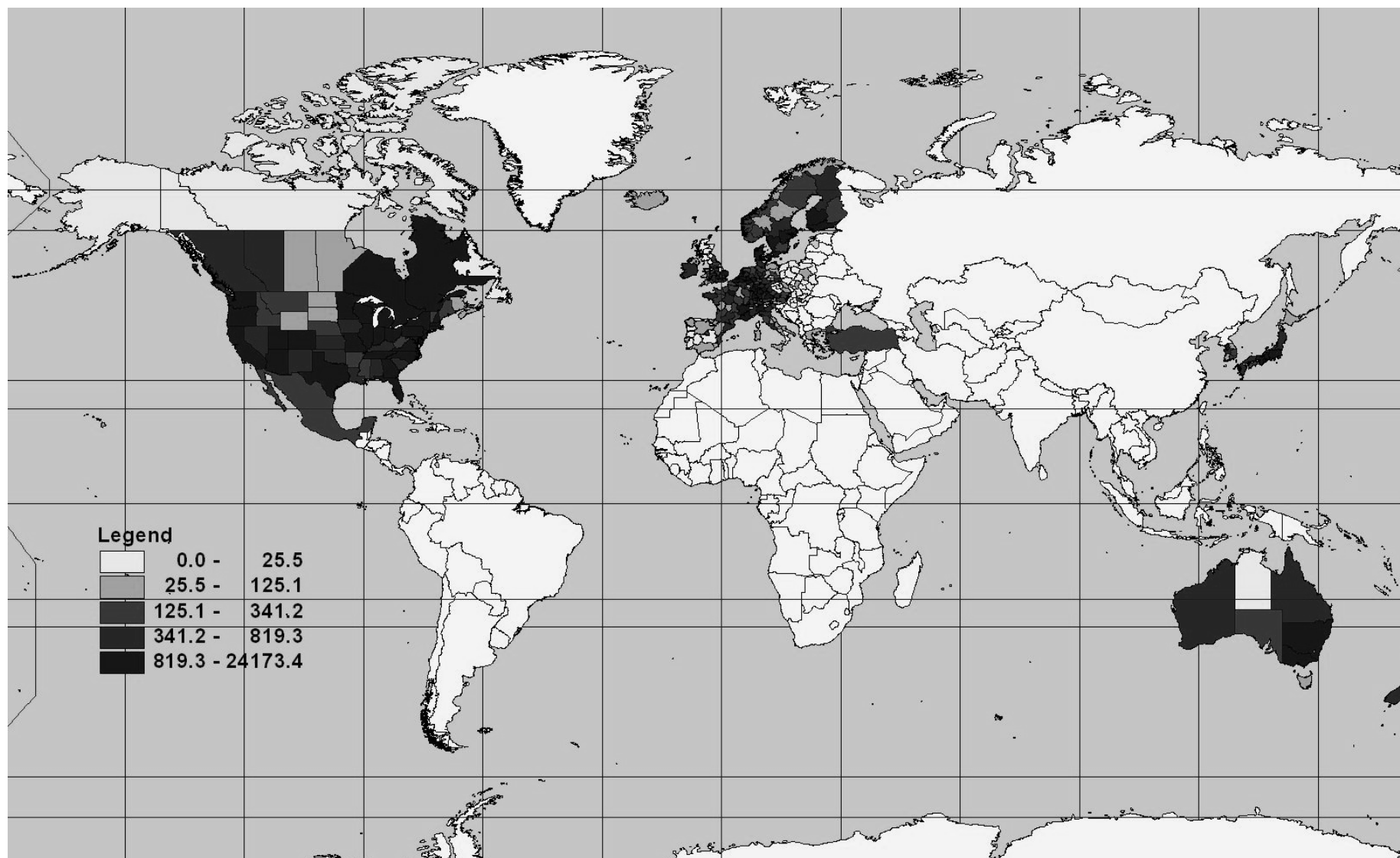
The analysis of these maps is made easier and clearer by Table 2, which provides some information also for some aggregation at the macro-area level: Europe, North America and Asia-Pacific.

Table 2. Macro-areas– PCT, PCT per million population

EUROPE		PCT			PCT per million Population		
<i>Macro-Area</i>	<i>Regions</i>	<i>1998-2000</i>	<i>2002-2004</i>	<i>var %</i>	<i>1998-2000</i>	<i>2002-2004</i>	<i>var %</i>
Europe	234	98613	111495	13.06	63.46	70.32	10.81
N. America	64	110628	130557	18.02	90.70	102.53	13.04
Asia-Pacific	26	32752	66111	101.85	55.68	110.57	98.60

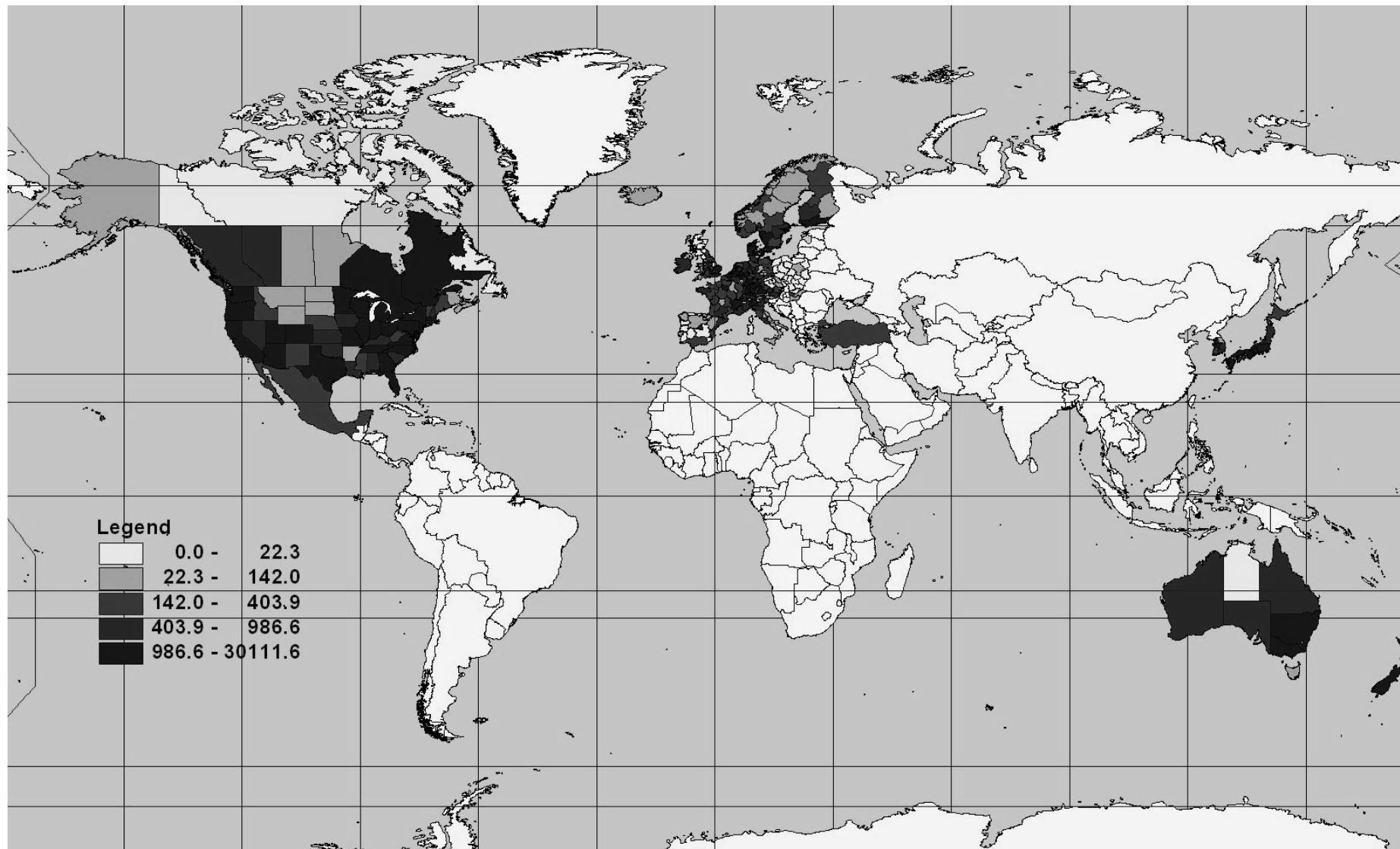
Source: OECD Regional Database.

Figure 1. OECD: PCT, 1998-2000



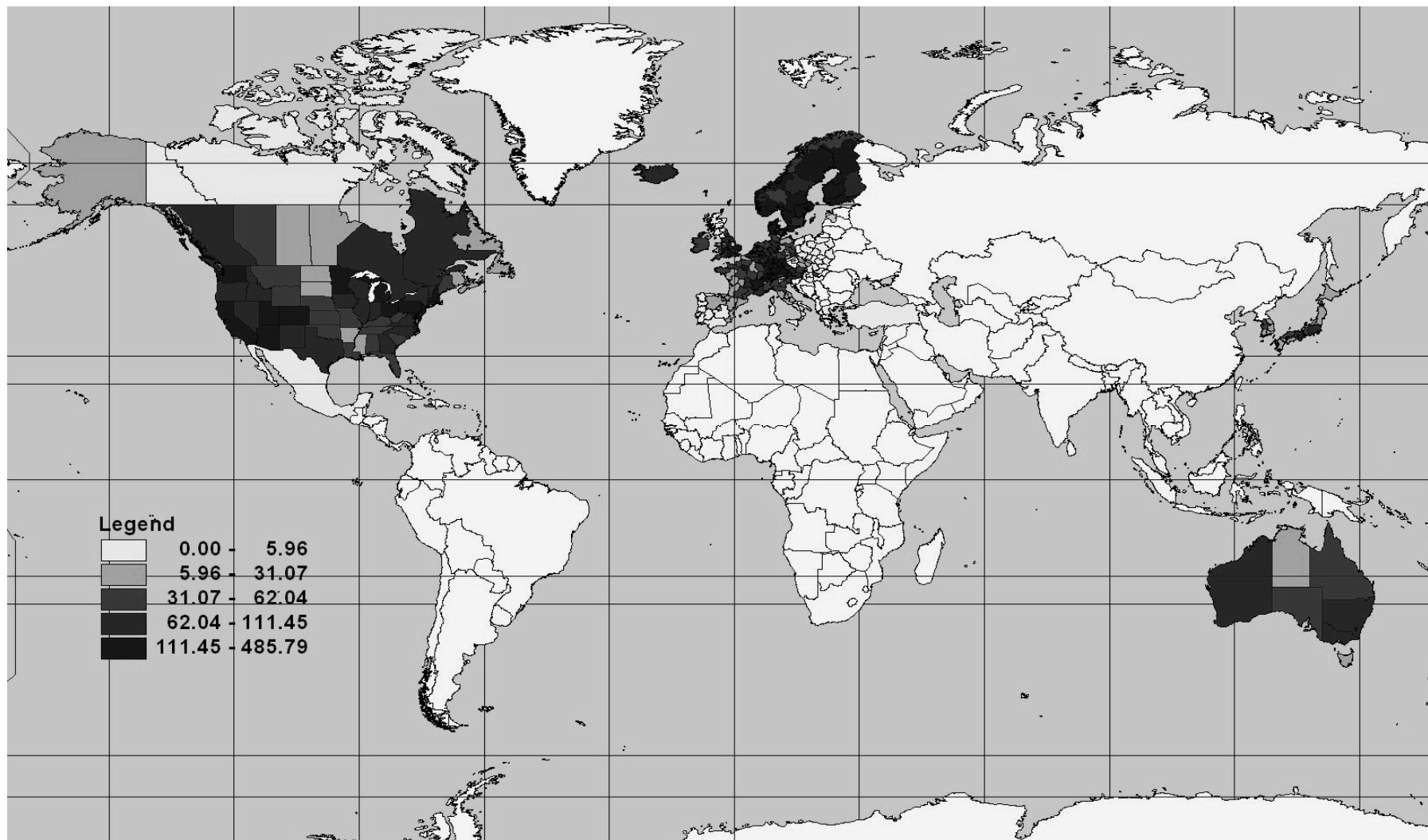
Source: OECD Regional Database.

Figure 2. OECD: PCT, 2002-2004



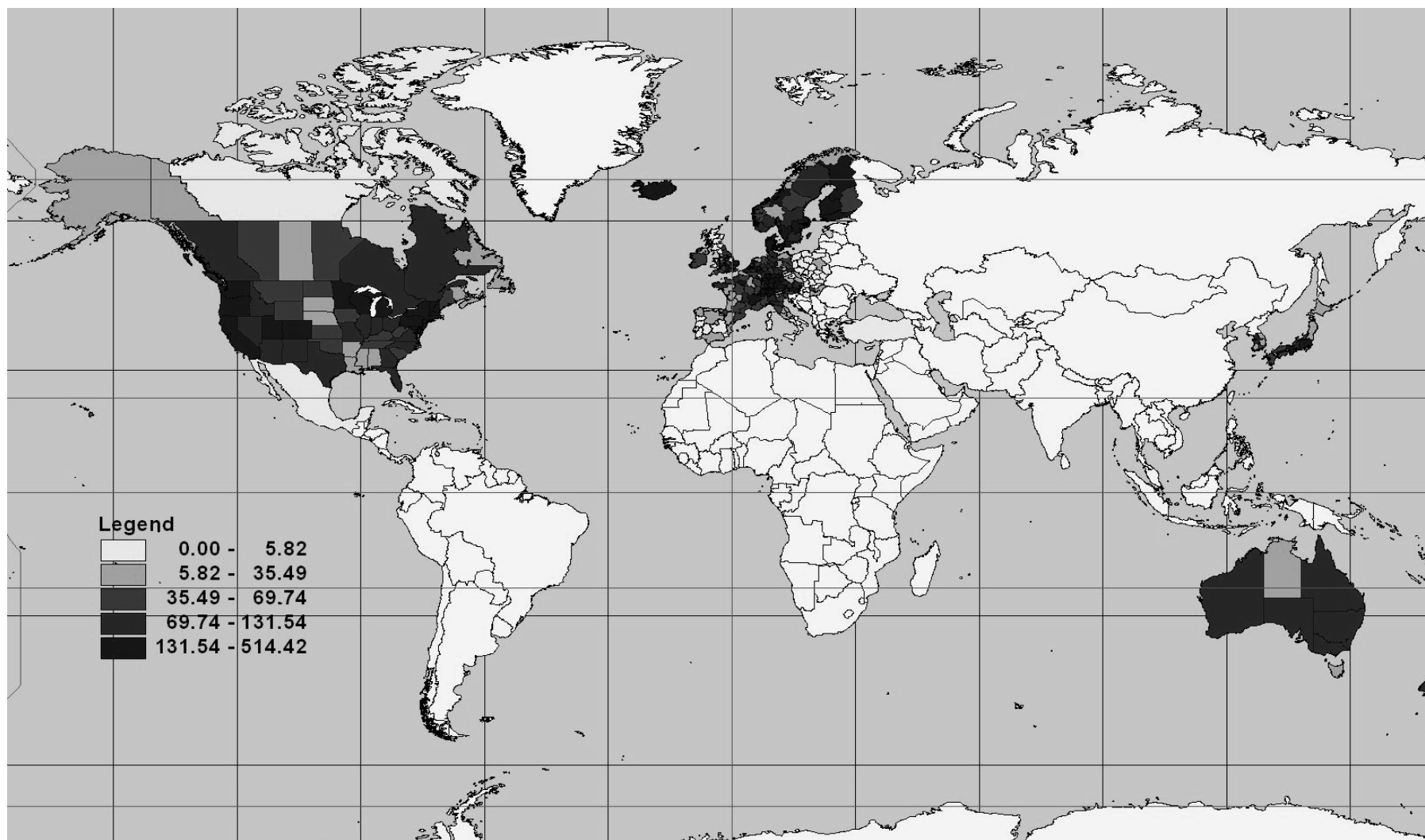
Source: OECD Regional Database.

Figure 3. OECD: PCT per million population, 1998-2000



Source: OECD Regional Database.

Figure 4. OECD: PCT per million population, 2002-2004



Source: OECD Regional Database.

Examining tables and maps together, let us first focus on European countries and regions. In the European area, the increase of PCT per capita is similar to the OECD average, whilst the variation rate of PCT in absolute value is lower. At the country level, a strong negative trend is evident in Sweden and Portugal; on the contrary Austria, Iceland, Italy, Switzerland, Ireland, and Turkey register remarkable growth. At the regional level, we note the positive trend of Austria in the Wien, Steiermark and Vorarlberg regions, of Italy in the Toscana, Lazio and Marche areas and of Switzerland in the Nordwestschweiz region. PCT per capita in Europe shows a large high-performance cluster, which starts from Rhone-Alpes (in France), passes through all Swiss regions, the Vorarlberg region in North Austria, and ends at the South-central part of Germany (Oberbayern, Freiburg, Stuttgart, Rheinhessen-Pfalz, Mittelfranken, Karlsruhe, Oberpfalz, Darmstadt, Tübingen, Unterfranken, Oberfranken, and Schwaben). Close to this agglomeration are those of Düsseldorf, Köln, and Braunschweig. These top performance regions are surrounded by other high performance states: this wider agglomeration could be named as the Continental Europe cluster. Detached from this cluster, one finds the capital regions of Berlin and Paris (Ile de France).

Sweden, Norway (except for the regions of Nord-Norge, Hedmark Og Oppland and Vestlandet) Finland and Denmark show top-high innovation performance, suggesting the presence of a sort of Scandinavian cluster.

Great Britain also shows an English innovation cluster, with top performance regions located mainly in the South East.

On the contrary, the low-performance regions are located at the European borders, in the West (Spain and Portugal), in the South (South-Central Italy, Greece and Turkey), in the East (Poland, Czech Republic, Hungary, Slovak Republic) and finally in the North (Scotland).

In the second period, 2002-2004, some interesting trends appear. First of all, most of Italy shows a growth of PCT per capita, especially the centre-north, which appears to extend the Continental cluster to the South. Secondly, the border states of Iceland and Ireland improve their position in the quantile distribution of the regions, showing high performance values. Thirdly, there is a light increase of PCT per capita in the eastern European regions which are closer to the Continental cluster. Finally, a general worsening of the innovative activity emerges in the Scandinavian cluster.

As far as the North American macro-area is concerned, the strongest increase is recorded in Mexico, although its top values are still very low compared to Canada and the United States. Canada shows a significant increase both in absolute and per capita terms, especially in the provinces of Prince Edward Islands and Nova Scotia, whereas the highest increases in the United States are recorded in Florida, Nevada, Oregon and North Dakota.

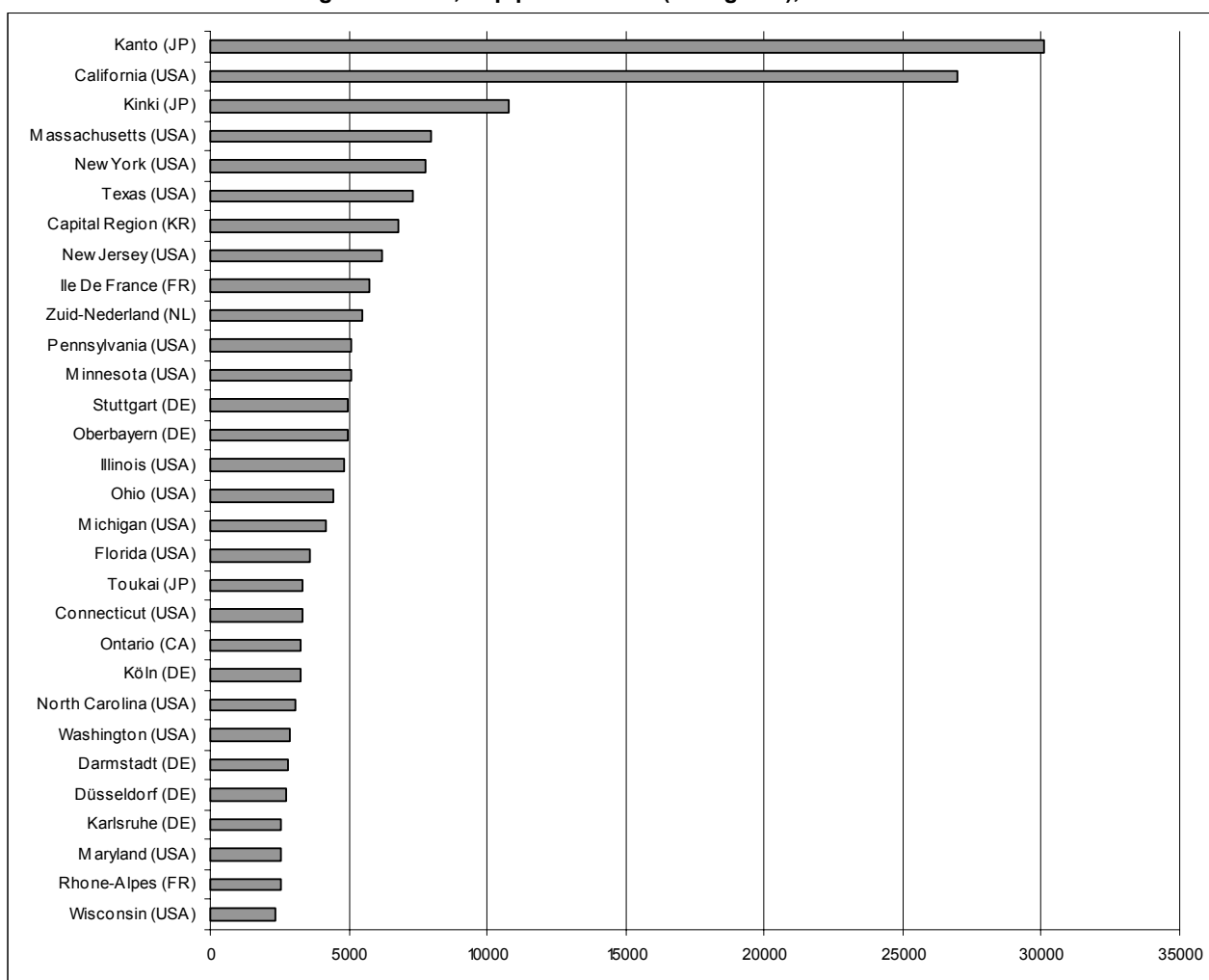
In the first period, a cluster of top performance states materialises in the Eastern US regions (Massachusetts, Delaware, Connecticut, Minnesota, New Hampshire, New Jersey, Maryland, Pennsylvania, Ohio, Rhode Island, Michigan) surrounded by high innovation performance states: this wider geographical zone, that includes Eastern Canadian states, could be considered as a unified belt which may generate some advantages on innovation production and transmission. In the West, we register some high innovative performances too: the top performance states (California, Colorado, Washington, Utah, Arizona) are surrounded by more high performance states, generating a Western innovation belt. As before, this belt goes from the United States (Texas) to Canada (British Columbia). Within these two belts, there are states which record a medium-low performance of PCT per capita. Nevertheless, it seems that the American belts (especially the Western one) tend to extend their influence to the central states along time.

Finally, the Asia-Pacific macro-area, which comprises Japan, Korea, Australia and New Zealand, is the most heterogeneous from a geographical point of view since it includes areas which are quite remote. Nevertheless, we try to assess the presence of a common trend among such diverse and distant countries. The Asia-Pacific area counts the lowest number of regions, but it presents the highest variation rate of PCT. Taking into account the absolute PCT value, the increase is 102%, while for PCT per capita the variation rate is 99%. Japan and Korea show the most significant growth, in particular with the regions of Kanto, Toukai and Kinki in Japan and with Jeolla Region and Gyeongnam in Korea.

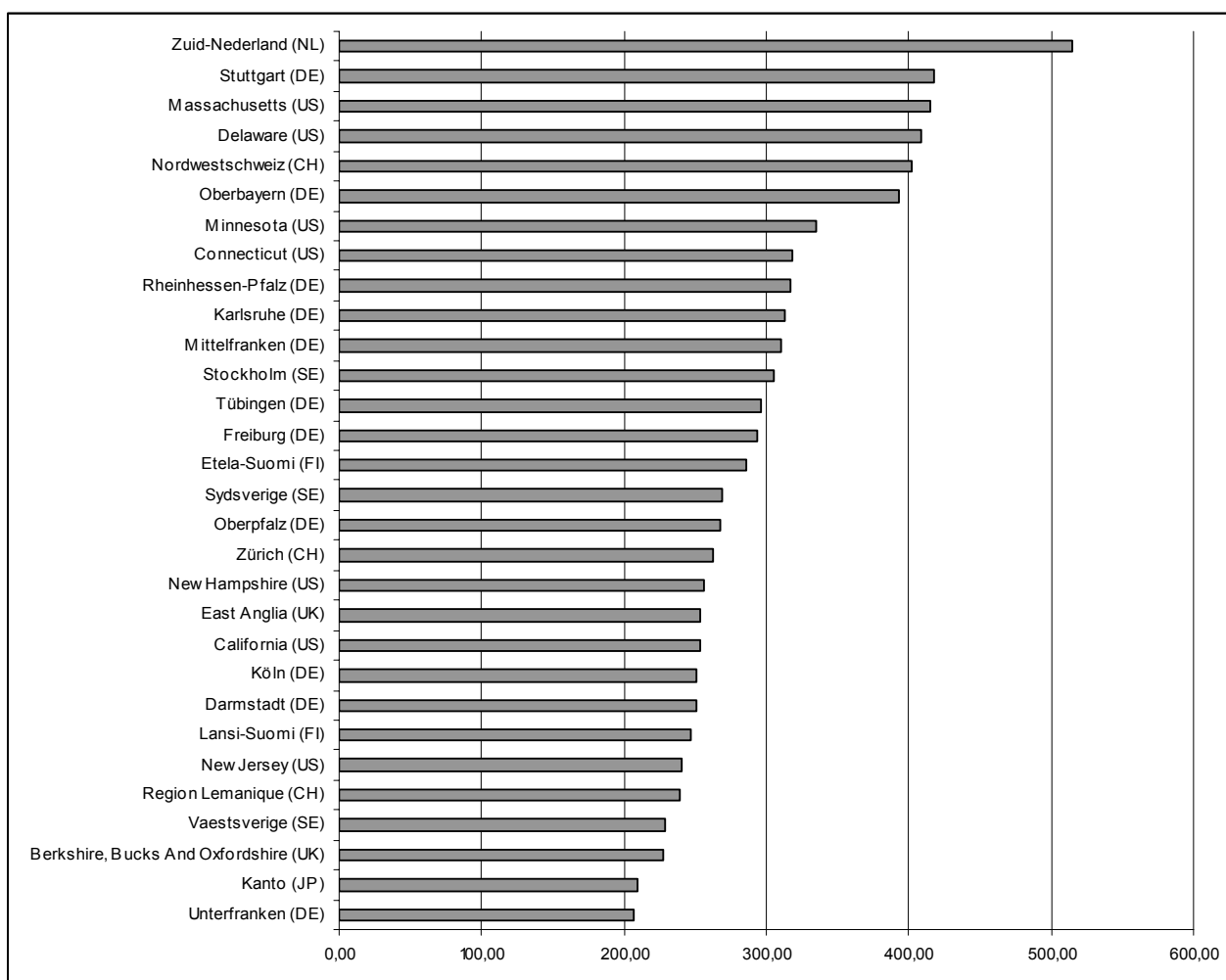
The distribution of PCT per capita in the first period denotes high performance values for the Japanese regions of Toukai and Kinki, in New Zealand and in New South Wales, Victoria and Western Australia. The top regions for innovative production are those of Kanto and the Australian Capital Territory. Considering innovative production, data appears to indicate the presence of a significant cluster in Japan, centred in the region of Tokyo, which seems to consolidate in the second period.

The analysis at the regional level is extended through Figures 5 and 6, which provide the best 30 performances in terms of absolute and per capita values in the second period. The best region in per capita PCT is Zuid-Nederland in the Netherlands (with 514.42 PCT). Looking at the national context, the best performers are Germany with eight regions (Stuttgart, Oberbayern, Rheinhessen-Pfalz, Karlsruhe, Mittelfranken, Tübingen, Freiburg, Oberpfalz), and the United States with five regions (Massachusetts, Delaware, Minnesota, Connecticut, New Hampshire). The other top performing regions are Nordwestschweiz and Zürich in Switzerland, Etela-Suomi in Finland and Stockholm and Sydsverige in Sweden.

Figure 5. PCT, Top performance (30 regions), 2002-2004



Source: OECD Regional Database.

Figure 6. PCT per million population, Top performance (30 regions), 2002-2004

Source: OECD Regional Database.

The analysis so far has shown that the phenomenon under examination is rather uneven at the national level. A certain amount of variability is however evident also at the regional level, which is more closely investigated through the set of tables below. Tables 5 and 6 provide some synthetic indicators about PCT in absolute values and in per capita for the three macro-areas and for the OECD area as a whole.

We find that the minimum value is zero. In the OECD area, in 2002-2004, there were 11 zero regions, which represented 5.25% of the total OECD area. They are: two regions of Portugal (Alentejo and Região Autónoma Dos Açores), two in Spain (Ciudad Autónoma De Ceuta and Ciudad Autónoma De Melilla), six in the United Kingdom (Cumbria, Lincolnshire, North Eastern Scotland, Eastern Scotland, South Western Scotland, Highlands and Islands) and one in Korea (Jeju).⁹

9. These regions are excluded from the econometric analysis below.

In both periods, the OECD regions with the highest value are in Europe: Stockholm in Sweden in 1998-2000 (458.79 PCT per capita) and Zuid-Nederland in the Netherlands in 2002-2004 (514.42 PCT per capita).

In the Asia-Pacific, the maximum value of PCT per capita in 2002-2004 is exhibited by the Kanto region in Japan (with a value of 209.30) while in the North American area, Massachusetts shows the maximum value with 415 PCT per capita. Overall, the average values in each macro area grow between the two time periods for both types of indicator.

As regards the variability, the degree of disparity (determined by the coefficient of variation) in the regional distribution of innovative activities across OECD countries increases when measured by PCT absolute values, whilst it decreases when measured by PCT per capita. However, such phenomena are not homogeneous across macro-areas.

In particular, variability in PCT absolute values decreases in North America. The cross sectional comparison shows that Europe has a lower degree of concentration than the other areas. Such comparisons should be taken with care, as the size of the geographical units varies across areas. As for the per capita indicator, we find that the coefficient of variation decreases significantly in the whole OECD area (from 1.13 to 0.99) and in the Asia-Pacific macro-area (from 0.62 to 0.46), while it decreases slightly in the two other macro-areas.

Table 3. PCT OECD by region – Synthetic indicators

30 nations; 324 regions				
	PCT		PCT_per million population	
	<i>1998-2000</i>	<i>2002-2004</i>	<i>1998-2000</i>	<i>2002-2004</i>
AVERAGE	747	951	71.98	89.15
STAND. DEV	1861	2600	81.42	87.93
CV	2.49	2.73	1.13	0.99
MIN	0	0	0.00	0.00
MAX	24173	30112	485.79	514.42

Source: OECD Regional Database.

Table 4. PCT Europe by region - Synthetic indicators

23 nations; 234 regions				
	PCT		PCT_per million population	
	<i>1998-2000</i>	<i>2002-2004</i>	<i>1998-2000</i>	<i>2002-2004</i>
AVERAGE	421	476	63.46	70.32
STAND. DEV	748	861	83.20	89.19
CV	1.77	1.81	1.31	1.27
MIN	0	0	0.00	0.00
MAX	4908	5752	485.79	514.42

Source: OECD Regional Database.

Table 5. PCT North America by Region - Synthetic indicators

3 nations; 64 regions				
	PCT		PCT per million population	
	1998-2000	2002-2004	1998-2000	2002-2004
AVERAGE	1729	2040	90.70	102.53
STAND. DEV	3293	3737	83.90	90.82
CV	1.91	1.83	0.92	0.89
MIN	0	0	0.00	0.93
MAX	24173	26948	392.65	415.46

Source: OECD Regional Database

Table 6. PCT Asia/Pacific by Region - Synthetic indicators

4 nations; 26 regions				
	PCT		PCT per million population	
	1998-2000	2002-2004	1998-2000	2002-2004
AVERAGE	1260	2543	55.68	110.57
STAND. DEV	2797	5979	34.42	51.11
CV	2.22	2.35	0.62	0.46
MIN	0	0	0.00	0.00
MAX	14006	30112	120.14	209.30

Source: OECD Regional Database

We further analyse the variability of the phenomenon under examination by looking at the difference among regions within national borders. This is done through the boxplot¹⁰ reported in the following two tables for the two periods. In particular, the figures below allow to investigate on the dispersion and skewness of PCT per capita distributions both in the whole of the OECD area and for each single nation.

In both periods, the distributions of Portugal, Poland, Greece, the Slovak Republic, Hungary and the Czech Republic show low dispersion. It is because of the very low values in all their regions - near zero in each country cited above. On the contrary, the German and Swedish distributions show high dispersion; furthermore, those distributions show an increase between the first and the second period. In 2002-2004, high dispersion is also recorded for the Finnish and the Dutch distribution.

The symmetry or asymmetry of the distributions is defined by the distance between the first and the third quartile from the median. In 2002-2004 the most symmetric distributions are those of Finland (which was not quite symmetric in 1998-2000), the United States (which confirms the previous situation in 1998-2000), Italy and the United Kingdom. The most asymmetric cases are those of Switzerland, Australia, the Netherlands and Korea in the 2002-2004 period.

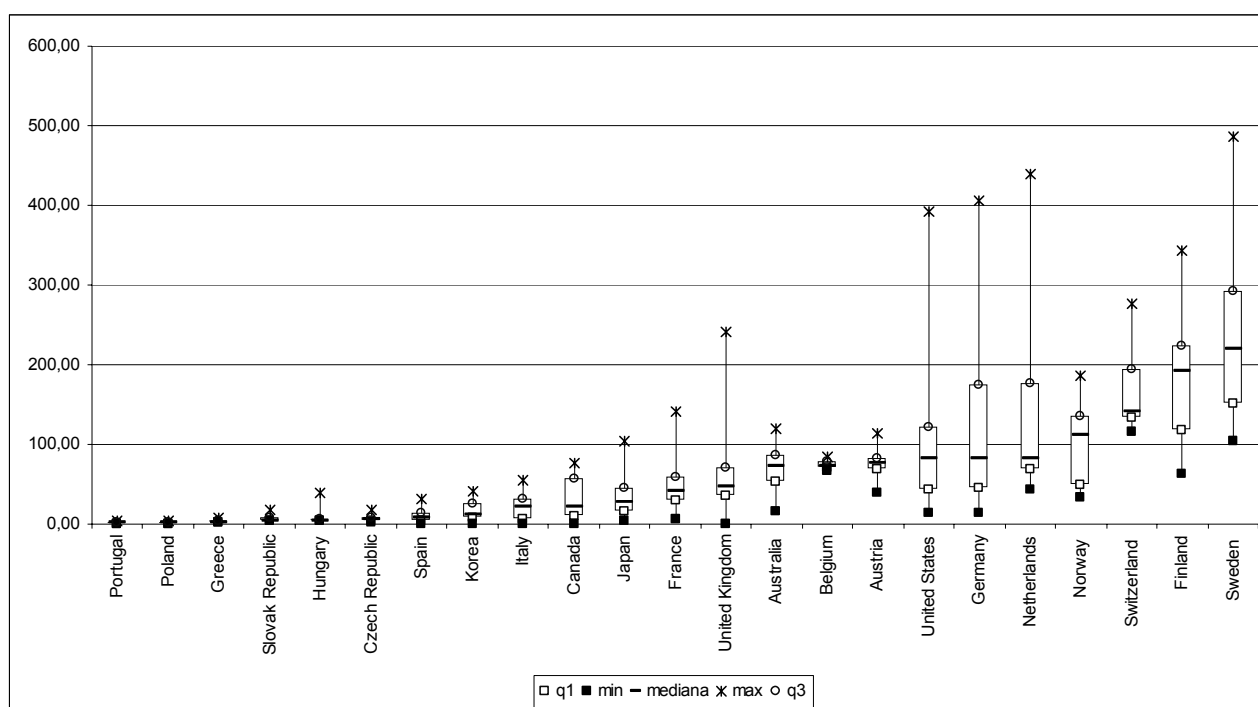
The minimum and the maximum values indicate the distribution variation range. Considering the symmetric distributions mentioned before, in the 2002-2004 period there is a very high variation range for

10. The Box Plot is a histogram-like method to display data. The spacing between the different parts of the box helps to indicate dispersion and skewness and to identify outliers. Another name for the boxplot is "five-number summary" which refers to the five statistics provided, that is the median, the upper and lower quartiles, the minimum and maximum values. The box itself contains the median 50% of the data. The upper edge (hinge) of the box indicates the 75th percentile of the data set, and the lower hinge indicates the 25th percentile. The range of the middle two quartiles is the inter-quartile range. The line in the box indicates the median value of the data. If the median line within the box is not equidistant from the hinges, this implies that the data is skewed. The end of the vertical lines or "whiskers" indicate the minimum and maximum data values, unless outliers are present, in which case the whiskers extend to a maximum of 1.5 times the inter-quartile range.

Finland and the United Kingdom: the first has a maximum value of 286.31 (Etela-Suomi) and a minimum value of 48.17 (Ita-Suomi); the second has a maximum value for both periods in East Anglia, while the minimum values are 0.

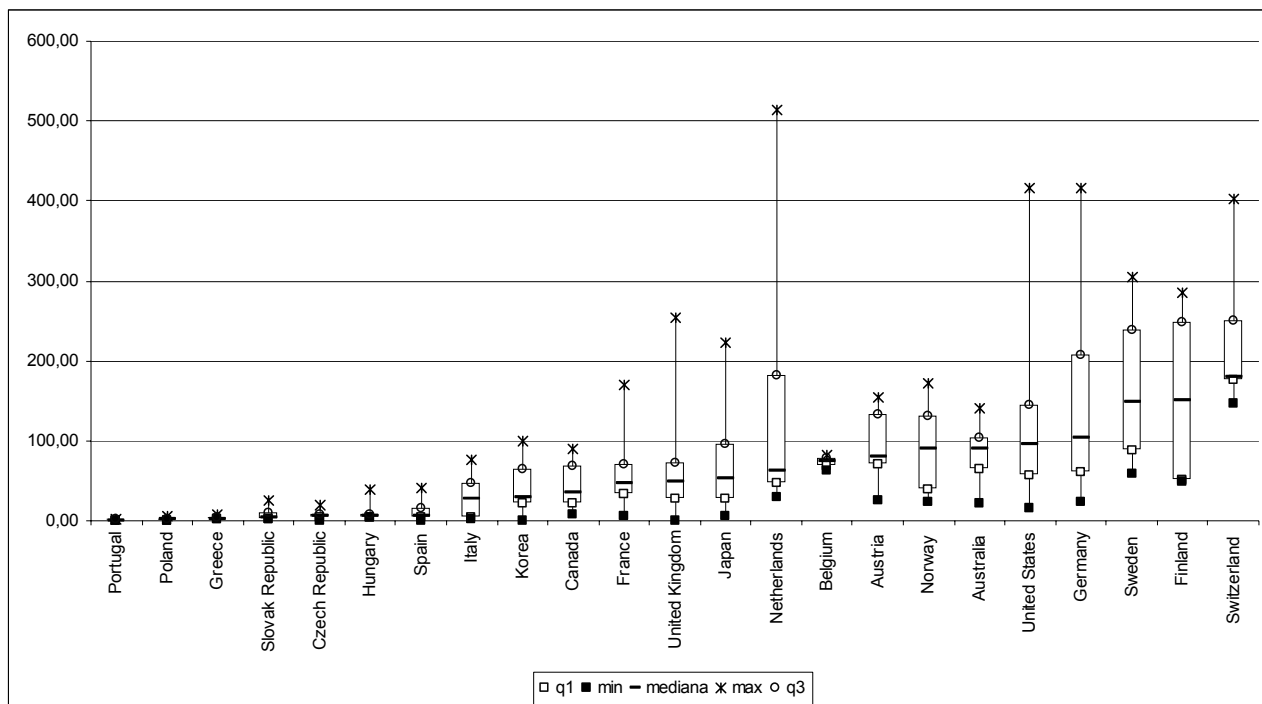
Some other interesting indications emerge from national distributions: Scandinavian countries show a high dispersion in both periods, probably due to the different climatic and density parameters among their regions; high dispersion is also registered for the United Kingdom, which varies its innovation performances between innovation excellences (England regions) and poor ones (Scotland); a great difference stands between the United States and Canadian PCT dispersions: as already seen in the PCT maps, Canadian states have less internal differences than US states; finally, a great dispersion is registered in Germany, for both periods. These distribution characteristics are expected, considering the high number of regions analysed. Asymmetry and dispersion increases significantly in the Netherlands along time.

Figure 7. OECD regions: PCT per million population variability, 1998-2000



Source: OECD Regional Database.

Figure 8. OECD regions: PCT per million population variability, 2002-2004



Source: OECD Regional Database.

Another way to investigate the dispersion of PCT per capita across regions within each country is to provide a quantitative measure which integrates the visual analysis performed with the box plot. In Table 7, the coefficient of variation for each single country for the two periods allows to see that the country with the highest dispersion is Hungary in the first period and the Netherlands in the second period (with quite a remarkable increase). The lowest level is shown by Belgium in both periods, probably because of the small number of regions involved (just three). Another nation with low variability is Switzerland, where the only outlier is the region of Nordwestschweiz. Variability is remarkably stable in most countries (Italy, Germany, Japan, Sweden and the United States to name just a few), but it changes quite significantly in some other countries. For instance, dispersion increases substantially in Austria, the Czech Republic, Finland, the Netherlands, the Slovak Republic and Spain. A decrease is recorded in Canada, Korea and Poland.

Table 7. OECD regions: Coefficient of variation of PCT per capita

<i>Nation</i>	<i>1998-2000</i>	<i>2002-2004</i>
Australia	0.48	0.47
Austria	0.26	0.42
Belgium	0.10	0.12
Canada	0.90	0.72
Czech Republic	0.66	0.79
Finland	0.51	0.63
France	0.64	0.66
Germany	0.77	0.76
Greece	0.62	0.70
Hungary	1.26	1.19
Italy	0.75	0.80
Japan	0.89	0.88
Korea	0.90	0.82
Netherlands	0.99	1.20
Norway	0.54	0.60
Poland	0.85	0.75
Portugal	0.90	0.87
Slovak Republic	0.93	1.15
Spain	0.87	1.05
Sweden	0.51	0.53
Switzerland	0.31	0.36
United Kingdom	0.89	0.94
United States	0.82	0.79

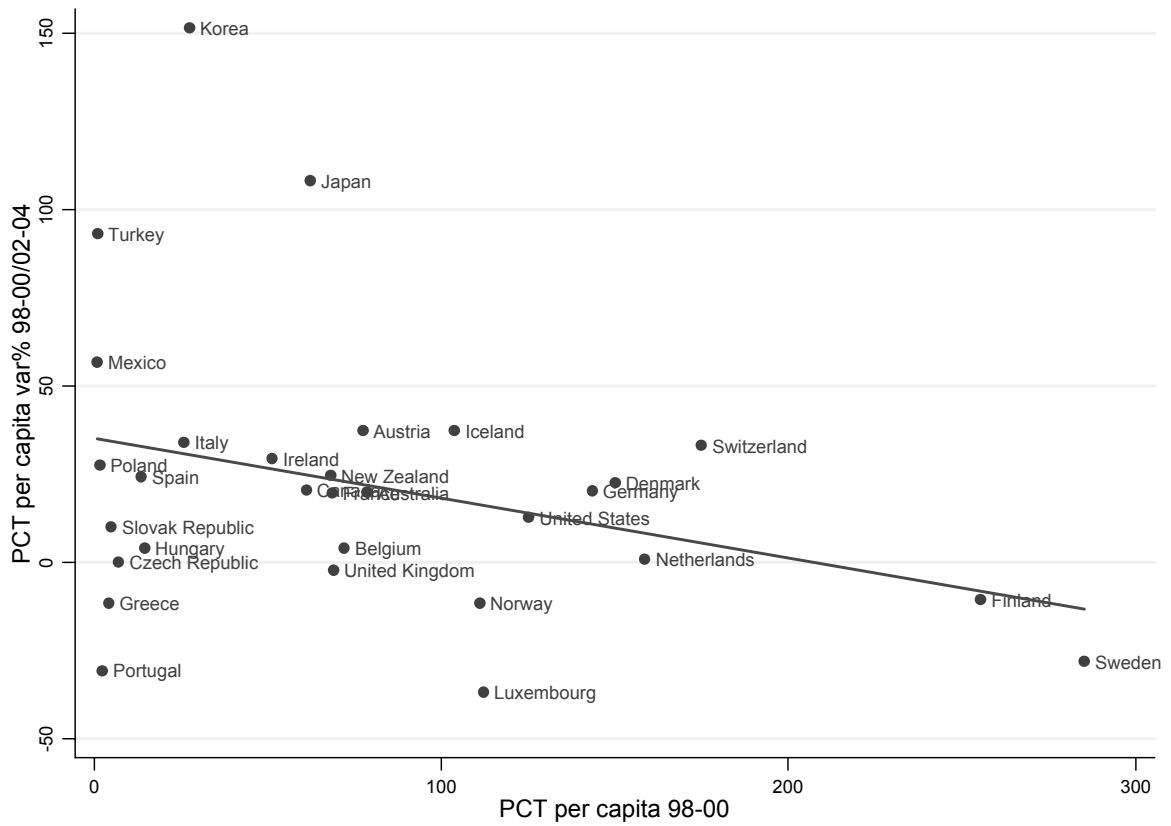
Source: OECD Regional Database.

It could be interesting to assess the degree of regional convergence of innovation levels along time. Using the dispersion graph in the figures below, the per capita PCT variation rate between the two periods is related to the variable's initial level at the national level (Figure 9) and at the regional level (Figure 10).

The first graph provides some evidence in favour of convergence among countries, since the initial value is negatively related to subsequent growth. Hence, in the long term, one can see countries that begin with low innovation values catching up with those that started with high values. For example countries with initially high values such as Germany (143.54), Denmark (150.11), the Netherlands (158.66) or Switzerland (174.95) show modest growth rates (respectively +0.18, +2.25, +0.88, +3.11). Others, which begin with a high value, display negative growth, like Switzerland with 285.30 PCT per capita in 1998-2000 and a growth rate of -28.19, or Finland with 255.48 PCT and then a decrease of -10.65, or Luxembourg (starting value 112.12 and a growth rate -36.99), or Norway (111.08 and -11.58). There are, vice versa, countries with low initial values and high growth rates. Japan and Korea are a case in point: with initial values of 62.25 and 27.48, respectively, they grew to 108.02 and 151.37.

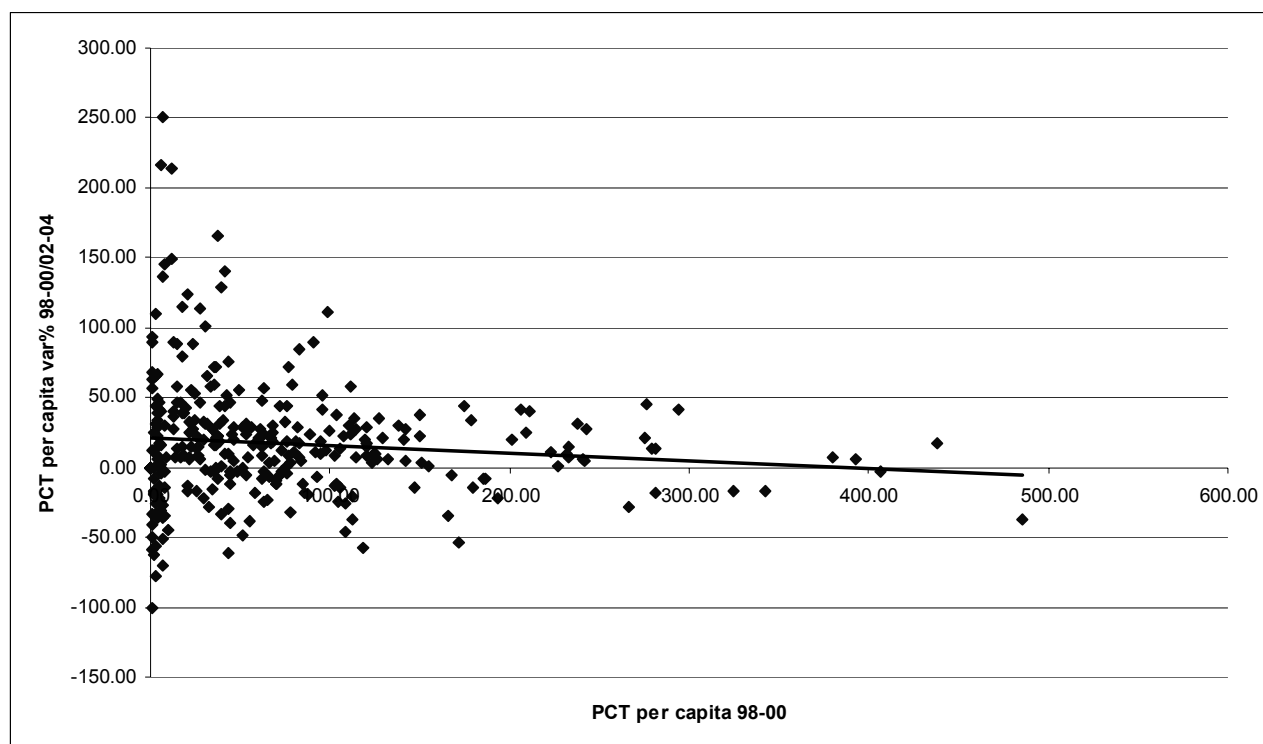
Nevertheless, this stylised fact is put in doubt when one analyses the second graph in Figure 19 which shows the same relationship at the regional level. When using the more disaggregated geographical units, convergence is much less evident even though the negative relationship is still present.

Figure 9. PCT per million population, Innovation Convergence (OECD countries) 1998-2004



Source: OECD Regional Database.

Figure 10. PCT per million population, Innovation Convergence (OECD regions) 1998-2004



Source: OECD Regional Database.

4.2 *Spatial dependence*

Spatial dependence refers to spatial similarity or dissimilarity of neighbouring regions and can be studied in order to assess the geographical extent of different phenomena. The most widely used index to measure spatial dependence, and to test the hypothesis of no-clustering for spatially distributed variables, is the Moran's I index. This index is calculated by taking into account the value assumed by the variable under analysis at different locations (regions) which are mapped through either a contiguity matrix or a distance matrix. The former refers to the presence of a common border between a pair of regions. This concept of contiguity is extended to different orders of adjacency.¹¹ The latter refers to the distance of each pair of regions. As a result, distance matrices are not binary and they usually refer to the inverse of the geographic distance among regions' centroids. The inference is based upon the analysis of the standardised z-value of Moran's index, which also indicates the sign of the geographical clustering (if any).

Tables 8 to 10 present the results for contiguity weight matrices based on adjacency (for first, second and third order of contiguity) and for the distance matrix (except for the OECD case and the Asia-Pacific

11. A pair of regions is of second order contiguity when the regions do not share a border directly, but they share a border with a region which is contiguous to both of them. Contiguity matrices are binary, that is they are made of zero's and one's..

case since we cannot appropriately discriminate between distances which are taken across land and those which are taken across sea).¹²

European and North American regions show a strong positive spatial association for the three orders of contiguity for both periods, indicating a geographic clustering for high and for low values. In other words, high/low values for one location are more similar to the high/low values of its neighbours than it would be the case under spatial randomness. This spatial pattern is confirmed by results for the distance-based statistics. We may also notice that the value of our statistics is much higher for Europe than for North America. This result may be due to structural differences but it can be also attributed to the differences in size for regions in the two macro-areas: regions are quite small in Europe whilst they are very large (at the state level) in North America (see Table 1).

Asia-Pacific regions show a different behaviour, with no significant value for Moran's I, nearly for all periods. This is no surprise, since on the one hand Asia-Pacific regions are very distant from each other except for Korea and Japan; on the other hand the number of regions is quite low.

Table 8. OECD: PCT per capita, Moran's I test for spatial autocorrelation

Period	Matrix	Moran's I	z-value	Prob
1998-2000	contiguity 1	0,518	11,956	0,000
	contiguity 2	0,422	11,182	0,000
	contiguity 3	0,312	8,289	0,000
2002-2004	contiguity 1	0,520	11,995	0,000
	contiguity 2	0,381	10,100	0,000
	contiguity 3	0,243	6,470	0,000

Source: OECD Regional Database.

Table 9. Europe: PCT per capita, Moran's I test for spatial autocorrelation

Period	Matrix	Moran's I	z-value	Prob
1998-2000	contiguity 1	0,555	11,150	0,000
	contiguity 2	0,469	11,299	0,000
	contiguity 3	0,314	7,704	0,000
	distance	0,109	11,616	0,000
2002-2004	contiguity 1	0,555	11,155	0,000
	contiguity 2	0,415	10,014	0,000
	contiguity 3	0,235	5,790	0,000
	distance	0,150	15,842	0,000

Source: OECD Regional Database.

12. For the same reason, distance matrices are not used in the regressions below when they refer to the whole of the OECD countries but are taken into account when they refer to Europe and North America where most distances are across land.

Table 10. North America: PCT per capita, Moran's I test for spatial autocorrelation

Period	Matrix	Moran's I	z-value	Prob
1998-2000	contiguity 1	0,317	3,553	0,000
	contiguity 2	0,202	2,827	0,005
	contiguity 3	0,223	3,049	0,002
	distance	0,195	6,009	0,000
2002-2004	contiguity 1	0,368	4,089	0,000
	contiguity 2	0,215	2,991	0,003
	contiguity 3	0,211	2,897	0,004
	distance	0,187	5,780	0,000

Source: OECD Regional Database.

Table 11. Asia/Pacific: PCT per capita, Moran's I test for spatial autocorrelation

Period	Matrix	Moran's I	z-value	Prob
1998-2000	contiguity 1	0,320	1,945	0,052
	contiguity 2	0,545	2,552	0,011
	contiguity 3	0,164	0,473	0,636
2002-2004	contiguity 1	0,029	0,371	0,711
	contiguity 2	0,385	1,854	0,064
	contiguity 3	-0.323	-0,657	0,511

Source: OECD Regional Database.

The significant values of the Moran indicator for Europe and North America are confirmed by the analysis of the LISA (Local Indicator of Spatial Autocorrelation) indicator and the so-called Moran scatterplot.¹³ Such analysis allows to discriminate among four types of spatial association: two forms of “positive” spatial association, *i.e.*, association between similar values, either high performing regions surrounded by other high performing regions, or low performers surrounded by other low performers; and two forms of “negative” spatial association, *i.e.*, association between dissimilar values. The extent of the “mix” of pairs between these four types of association provides an indication of the stability of the spatial association throughout the whole sample. It may also suggest the existence of different clusters of association in different parts of the world.

The Moran scatterplot confirms the visual analysis provided above, since positive highly innovative clusters are found in Europe and North America. Other clusters, but among regions which are underperforming, appear in Canada and in the Iberian peninsula, in the former Eastern block and in the South of Europe.

5. Econometric estimation

This section describes the econometric model, the variable and the estimation strategy used to investigate the mechanisms and determinants of the process of creation and diffusion of innovative knowledge using spatial econometrics techniques. Later results of the estimation are discussed and interpreted in the light of the descriptive analysis above.

13. Details and maps of this analysis are available from the author.

5.1 *The KPF model*

In section 2 the KPF has been presented as the ordinary tool to investigate the capacity of a firm or a region to produce innovations. It is an imperfect tool but its almost universal use makes its implementation a useful way to compare results in different countries and time periods. Most importantly, its application at the territorial level (instead of firms) allows one to use this tool to evaluate the presence of local externalities of different forms. Although we are not able to discriminate perfectly among such different externalities (in particular between pecuniary and knowledge ones), we try to attribute them to some general categories such as technology inputs, external spillovers, institutional factors and agglomeration effects.

The basic KPF relates the innovative output in region i to R&D input in the same region. We refine this specification by introducing a set of further factors related to the economic and institutional environment, so that the general form of our basic KPF is given by:

$$I_i = RD_i^{\hat{\alpha}_1} IF_i^{\hat{\alpha}_2} e_i \quad (1)$$

where I is a proxy for innovative output, RD indicates the primary input, that is investment in Research and Development, IF , that is Internal Factors, is a vector of variables that reflects the economic and institutional additional determinants and e represents a random i.i.d. error term capturing other unobservable determinants of innovative output, as well as random shocks.

The additional factors included in IF are meant to control for systematic effects that may be present in the innovation process and may enhance or hinder the production of new knowledge. First, in the light of the theoretical and empirical literature, we consider the proportion of people with the necessary competences and skills. In this respect, the concentration of skilled workers in one place is a mechanism through which positive effects may materialise, as direct communication enables flows of information and other tacit components of knowledge (Feldman and Florida, 1994).

Secondly, following Ciccone and Hall (1996), we hypothesise that agglomeration processes may be catalysed with higher density of economic exchanges and contacts. Audretsch and Stephan (1996) argue that the same R&D efforts may result in a higher level of innovative activity in large urban areas rather than in rural areas because of agglomeration economies.

Thirdly, regions, and especially countries, may show remarkable differences in the institutions and strategies for technological change. Such institutions and strategies form the so-called regional or national innovation system, which determines how favourable an environment is for technological change and diffusion.

Finally, the theoretical and empirical literature suggests that the production of knowledge in a region may depend not only on its own research efforts, but also on the knowledge stock available in the whole economy and on its ability to exploit it. In other words, knowledge may spill over from other regions (especially nearby ones). Many factors, external to the region, can act as determinants of technological activity, channelled through trade flows, external investments, imports of machinery, common markets for skilled labour and final goods. Moreover, pecuniary externalities may be at work, thereby shifting externalities at the firm level to higher territorial levels. Our general framework given in (1) may be consequently modified to introduce an additional vector of external factors EF , reflecting the fact that innovation generated in one region may spill over and help knowledge formation in other regions:

$$I_i = RD_i^{\delta_1} IF_i^{\delta_2} EF_i^{\delta_3} e_i \quad (2)$$

This extended model is able to capture the spatial dependence of the innovation phenomena, visually apparent in the graphs in section four, which was quantified and statistically documented through the Moran tests. We should, therefore, expect the estimates of the coefficients of EF to be significant.

Instead of estimating directly model (2), we start with model (1) because we need firstly to understand the best way, from a technical point of view, to introduce spatial dependence in our regression.

Therefore, we begin by assuming that any new knowledge produced by a region, denoted by I, in a given period is related to its R&D efforts in previous periods (RD) and to a vector of internal factors IF = (HK, DENS, NAT, DU, DR, DCAP and DGDP) according to a Cobb-Douglas technology which can be written in logarithmic form as follows:

$$\begin{aligned} \log I_{i,t} = & \beta_1 \log RD_{i,t-s} + \beta_2 \log DENS_{i,t-s} + \beta_3 \log HK_{i,t-s} + \\ & + \sum_{c=1}^n \delta_c NAT_{ic} + \gamma_1 DU_{i,t-s} + \gamma_2 DR_{i,t-s} + \gamma_3 DCAP_{i,t-s} + \gamma_4 DGDP_{i,t-s} + \varepsilon_{i,t} \end{aligned} \quad (3)$$

where the coefficients can be interpreted as elasticities. It is worth noting that independent variables are lagged, since we expect inputs to take some time to produce effects on the output, a very plausible assumption in the case of patents, as they come at the end of a lengthy innovative process. Moreover, this procedure can avoid problems of endogeneity.

The dependent variable I is the number of patents per capita, per million inhabitants more exactly, as it is common in most studies of regional or national KPF.

As for the independent variables, the input of innovative activities (RD) is measured by the share of gross domestic product invested in research and development activities. Among the other potentially relevant internal forces;¹⁴ we introduce density of population (DENS) as a proxy for the agglomeration forces and the availability of human capital (HK) at the regional level, with the index of tertiary education, defined as the percentage of population completing a tertiary education degree.

Further we insert a set of dummies. First of all we include several dummies which designate regions with respect to their specific nature: rural (DR) or urban (DU).¹⁵ Then we try to further differentiate regions with respect to their main specialisation by introducing a dummy (DCAP) which is equal to one when it refers to a capital region and zero otherwise. This is because capital regions are often specialised in service activities and their innovative activity does not show up in patent production which is mostly concentrated in manufacturing. Our set of dummies is finally completed with a dummy (DGDP) which discriminates rich regions with respect to less developed regions.¹⁶ We believe that the former may enjoy several advantages in terms of technological opportunity. Rich countries can, for example, exploit the stock

-
14. Different variables have been used in the literature to proxy other internal factors and agglomeration economies, such as employment in the business sector and high technology employment (Anselin *et al.*, 1997), the relative importance of large firms in the geographical area (Varga, 2000), the quota of manufacturing firms (Moreno *et al.*, 2005).
15. It is important to notice that regions are classified with respect to three modalities: rural, urban and intermediate. Such a classification is provided by the OECD regional database.
16. We do not use the complete series for GDP in order to avoid potential problems of collinearity with the RD and HK variables. We therefore prefer to introduce a dummy in order to use this piece of information. Regions take value one when their GDP per capita is above the average and zero otherwise.

of knowledge and experience accumulated in the past, as in the endogenous model *à la* Romer (1986). The use of a dummy is again motivated by the fact that it is difficult to measure differences in technological opportunities.

Moreover, we include a set of national dummies (NAT) in order to control for any national fixed effect, such as institutional and other structural factors, which may affect either the innovative activity or the propensity to appropriate its results by patenting.

Our empirical exercise directly addresses interregional externalities in the generation of innovation through the use of spatial econometrics techniques. We believe that our regressions may suffer from spatial autocorrelation but we cannot take this for granted. Further, we can not anticipate which form it takes. In other words we need a set of Lagrange multiplier tests, that is LM-ERR and LM-LAG, which are computed for a physical contiguity matrix (W) which gives rise to a binary and symmetric matrix with elements equal to 1 in the case of two neighbouring regions and 0 otherwise. This is done not only to assess the existence of spatial dependence in the specification, but also to ascertain its possible form: a substantive or a nuisance model.¹⁷ The former is the one which is always selected through tests and it is specified as follows:

$$\log I_{i,t} = \beta_1 \log RD_{i,t-q} + \beta_2 \log DENS_{i,t-s} + \beta_3 \log HK_{i,t-s} + \beta_4 W \log I + \sum_{c=1}^n \delta_c NAT_{ic} + \gamma_1 DU_{i,t-s} + \gamma_2 DR_{i,t-s} + \gamma_3 DCAP_{i,t-s} + \gamma_4 DGDP_{i,t-s} + \varepsilon_{i,t} \quad (4)$$

where W is the contiguity matrix defining linkages across regions. The variable represented by the term $W \log I$ is therefore the spatial lag for the innovation output, in other words, it is a weighted measure of patents in the regions with which region i has contacts. We interpret an influence of this variable on the endogenous one as evidence of interregional spillovers of the knowledge located outside the region, whereas the lack of significance of β_4 would indicate that the production of new knowledge is generated internally. This spatial lag term has to be treated as an endogenous variable and Maximum Likelihood (ML) has to be used since OLS estimators are biased and inconsistent.

5.2. The dataset

This paper is based on the latest available international comparable data on patents made available by OECD. The indicators for innovation activity have been identified as the number of Patent Cooperation Treaty (PCT) applications which are available for seven years from 1998 to 2004. Data for the dependent variable refer to the second period (2001-2004) whilst data for the explanatory variables refer to the first period (1998-2000).

Other indicators at the regional level are included in the OECD Regional Database (ORDB), which provides quantitative information on socio-economic issues. Unfortunately, there are several missing values for some important indicators in a handful of countries. In particular, we could not gather regional data for R&D for Japan and Korea. This makes our preliminary analysis basically focus on the comparison of the European and the North American areas. Although this aspect may limit the representativeness of our sample, it should be remembered that PCT data for Japan show a possible undervaluation of its

17. The nuisance model represents a second way to incorporate spatial autocorrelation into the knowledge production function by specifying a spatial process for the disturbance term. Although unbiased, the OLS estimators will be no longer efficient. The spatial error model for the regional KPF would be expression (3) as it stands, with the error term denoted as: $\varepsilon_i = \lambda W \varepsilon_i + \eta_i$ where η_i is asymptotically distributed as $N(0, \sigma^2)$, ε follows a first-order Markov process and λ is the spatial autoregressive coefficient for the error lag. In the case of spatial error autocorrelation, OLS parameter estimates are inefficient, whereas in the presence of spatial lag dependence parameters become not only biased, but also inconsistent (Anselin, 1988 and Anselin and Florax, 1995).

innovative capacity, which would have biased our analysis (even though national dummies are meant to partially control for such cases). In the final part of this section, nevertheless, we attempt to estimate data for the two countries of the Far East and results are mainly maintained.

It is also worth noting that no regionalised data for patents is present for some countries among those included in the analysis, namely Denmark, Iceland, Ireland, Luxembourg, and New Zealand. Data are not regionalised also for Mexico and Turkey but they are not included in the econometric analysis since other pieces of information are also missing. This is not a problem for countries such as Iceland and Luxembourg which are rather small, whilst there is certainly a loss of precious detailed information on the other countries, especially in the case of countries with significant size and population such as Mexico, Turkey and New Zealand (even though they are low producers of innovations).

One more important observation regarding data refers to the UK case where the number of regions for which the econometric analysis is performed (12) does not coincide with that provided in the descriptive analysis. This is due to the fact that R&D data are only available at this higher level of disaggregation. Finally it should be remembered that some regions (11, listed in the section above) are excluded from the analysis since they display zero values and therefore their logarithm was not feasible.

The econometric analysis is therefore made with 271 observations referring to 25 nations out of a potential of 30 countries.

Since we are estimating a cross-section and data refer to several years, we decided to smooth out possible transient effects and to approximate long-run values by using each variable as an average of three years' data. More specifically, since the production of knowledge takes time, we assume a time lag between R&D expenditure and the innovation yield. As a result, variable I refers to the period 2001-2004, whereas DENS and RD refer to 1998-2000. It should be noted that as far as RD is concerned, this information is gathered with different frequency in different countries and therefore it is not necessarily a three-year average for all countries. More detail on the database is given in the appendix.

The table below shows some synthetic statistics for the main regressors.

Table 12. Regressors synthetic statistics

	I (PCT_pc)	RD	DENS	HK
AVERAGE	80.53	1.46	288.91	12.40
STAND. DEV	90.12	1.17	728.11	5.37
CV	1.12	0.80	2.52	0.43
MEDIAN	53.21	1.09	106.13	12.77
MIN	0.31	0.10	0.02	1.86
MAX	514.75	6.23	5935.72	26.93

Source: OECD Regional Database.

As is clear from Table 12, all the variables show quite high variability. In particular we see that the highest dispersion, as anticipated in Table 1, concerns the density of population which shows a CV slightly higher than 2.5 with a range of values going from a minimum of 0.02 to a maximum of almost 6 000. The dependent variable has a high variability, too, but lower than that shown by the RD indicator. Human capital is the most uniform phenomenon.

Econometric results

Main econometric results are summarised in Table 13 which shows results of the KPF for the whole sample of OECD regions in the first two columns.

The general reading is that our model works remarkably well, since almost 90% of the variance of the dependent variable is explained by the set of explanatory variables. However, as was expected, the OLS estimation reported in column one suffers from spatial autocorrelation, since the tests clearly reject its absence at the 1% level of significance, indicating misspecification of the model and potentially some kind of externalities across regions. Following the "classical" specification search approach adopted in the spatial econometrics literature, given that the value of the LM-LAG test is much greater than the LM-ERR, the estimation of the spatial lag model by Maximum Likelihood (ML) is the preferred specification.

Column two, therefore, proposes the ML estimation of equation (4). Results are interesting. In particular we find that the elasticity of patents with respect to R&D expenditure when estimating equation (3) with OLS is around 0.45 and statistically significant. This result confirms the consensus found in the literature. The elasticity goes from 0.2 to 0.9 in the United States (Jaffe, 1989; Acs *et al.* 1994; Anselin *et al.*, 1997), from 0.25 to 0.8 in the European case (see Moreno *et al.* 2005 and Bottazzi and Peri, 2003). With respect to previous contributions, we exploit a larger, more disaggregated and updated database.

The other main factor which contributes to innovative performance, the human capital, is positive and significant with an impact which is more than double than that of RD (0.99).

Local agglomeration factors, measured by density of population, has some impact but a much lower level since the elasticity is around 0.07.

As for the dummies which take into account the nature and the environment of the regions, urban or rural, only the latter's coefficient is significant with the expected negative sign. The capital dummy is, as expected, negative, signalling the presence of a specialisation (mainly in services) which is not conducive to the type of innovative activity that gets patented. The coefficient of the dummy for high income countries is significant and, as expected, positive.

As for the institutional factors, coefficient for the national dummies are in general significant with different values and signs depending on the country performance. They can be interpreted as a reflection of the general efficiency of research influenced by country-wide institutional settings as well as the productivity effects of the knowledge diffused throughout the country.

Finally, as regards the external factors, we find that the lagged value of the dependent variable, which measures interregional linkages, is positive and significant with an elasticity of 0.18 (almost one-third of internal effects of regional R&D). Moreover, the significant value of the LR tests implies the statistical adequacy of the spatial lag models.

Table 13. KPF estimation

Variables	OECD		Europe		North America	
	OLS	ML	OLS	ML	OLS	ML
Log (RD)	0.486 (0.000)	0.446 (0.000)	0.498 (0.000)	0.461 (0.000)	0.548 (0.000)	0.479 (0.000)
log (HK)	1.094 (0.000)	0.991 (0.000)	1.072 (0.000)	0.886 (0.000)	1.061 (0.262)	1.086 (0.008)
log (DENS)	0.070 (0.092)	0.073 (0.045)	0.054 (0.438)	0.059 (0.320)	0.069 (0.182)	0.076 (0.093)
W log (I)		0.182 (0.000)		0.229 (0.000)		0.153 (0.016)
<i>Controls</i>						
Rural dummy	-0.201 (0.050)	-0.202 (0.026)	-0.142 (0.280)	-0.130 (0.248)	-0.236 (0.197)	-0.279 (0.080)
Urban dummy	0.099 (0.452)	0.049 (0.679)	0.268 (0.104)	0.230 (0.103)	-0.271 (0.243)	-0.342 (0.092)
Capital dummy	-0.543 (0.003)	-0.419 (0.010)	-0.515 (0.019)	-0.338 (0.073)	-0.815 (0.440)	-0.821 (0.018)
GDP dummy	0.810 (0.000)	0.652 (0.000)	0.935 (0.000)	0.713 (0.000)	0.466 (0.078)	0.375 (0.103)
NAT dummies	yes	yes	yes	yes	yes	yes
Obs	271	271	201	201	61	61
R ² -adj	0.889	0.906	0.901	0.920	0.679	0.747
LIK	-208.184	-194.813	-156.856	-143.291	-38.144	-35.489
AIC	480.367	455.625	369.713	344.583	94.288	90.978
SC	595.635	574.495	462.205	440.379	113.286	112.087
Moran's I	4.074 (0.000)		3.619 (0.000)		1.656 (0.098)	
LM-ERR	0.002 (0.968)	0.090 (0.764)	0.401 (0.526)	0.065 (0.799)	0.013 (0.909)	0.143 (0.706)
LM-LAG	20.551 (0.000)		22.653 (0.000)		3.990 (0.046)	

Note: probability values in parentheses.

In the same Table 13, columns 3 to 6 are devoted to the estimation of the model separately for Europe and North America. This exercise allows us to assess the differences in elasticities in North America and Europe, as suggested by Crescenzi *et al.* (2007). Moreover we can evaluate if interregional spillovers are homogenous across OECD countries or if there are significant differences between the two macro-areas.

Results show that the regions in these two macro-areas use similar technologies for the production of knowledge even though some interesting specific features appear. Let us focus on columns 4 and 6 since in both cases it is necessary to estimate the spatial lag model. The first thing to notice is that the model fits the European case much better than the North American one, since the power of explanation of the regression is far higher in the former case (around 0.92) rather than in the latter (0.74). In spite of this difference, there is a remarkable similarity of the whole set of coefficients, especially with respect to R&D efforts (even though it is slightly higher in Europe). More evident differences appear with respect to the other two coefficients under examination, that is HK and DENS. HK is positive and significant in both Europe and North America, but with a higher value in the latter case. As for the agglomeration effects, this variable shows a positive sign in both cases but it loses its significance in Europe while this result is kept in the North American case. As for the dummies, we find that the rural dummy keeps its significance in the North American model together with the capital region, whilst the GDP dummy keeps its sign but its significance is only at the 10% level. On the contrary, the rural and urban dummies are both insignificant in the EU model, where the capital region dummy is no longer significant and only the GDP dummy is significant.

Spatial dependence is always present in both models and it is eliminated by the lag of the dependent variable. The coefficient is slightly higher in Europe (0.22) with respect to the OECD result, whilst it is slightly smaller in North America (0.15). To some extent, this difference can be related to the fact that regions in Europe are much smaller than in North America, where potential spillovers are more likely within regions than among regions.

5.3 *Robustness checks*

This final part of the section implements some robustness checks of the main results shown and discussed in the previous paragraphs.

In Table 14 we start by assessing the robustness of the role of the main determinants of innovative activity, that is R&D efforts and human capital, with respect to the GDP dummy which separates the sample into two parts: rich and poor regions. This modification of the model is motivated by the belief that the role of these factors may experience some decreasing return phenomena at a certain point in their accumulation. The interactive dummies should therefore have a negative sign if the effect of RD and HK decreases as GDP augments. The results shown in column one imply that the impact of R&D and HK is relatively lower when the GDP per capita levels are higher. Both interactive terms are negative and significant for the OECD model. This can be interpreted as evidence of the fact that at higher levels of GDP, RD and HK's ability to increase innovative output tends to slow down. The model is less robust when it is implemented for the two macro-areas. In these cases only does the interaction of GDP with RD in the EU case keep its significance. In the North American case the interactive dummies introduce a distortion in the whole model and their use is found inappropriate.

Table 14. KPF estimation with interactive dummies

Variables	OECD		Europe		North America	
	OLS	ML	OLS	ML	OLS	ML
Log (RD)	0.571 (0.000)	0.586 (0.000)	0.600 (0.000)	0.619 (0.000)	0.768 (0.292)	0.605 (0.332)
log (HK)	1.087 (0.000)	0.953 (0.000)	0.969 (0.000)	0.780 (0.000)	1.020 (0.474)	1.408 (0.253)
log (DENS)	0.100 (0.015)	0.106 (0.004)	0.126 (0.080)	0.114 (0.062)	0.073 (0.171)	0.081 (0.074)
W log (I)		0.176 (0.000)		0.223 (0.000)		0.160 (0.013)
DGDP*log(RD)	-0.104 (0.399)	-0.191 (0.085)	-0.155 (0.291)	-0.262 (0.041)	-0.230 (0.753)	-0.141 (0.823)
DGDP*log(HK)	-0.488 (0.002)	0.359 (0.011)	-0.433 (0.027)	-0.201 (0.246)	0.000 (0.999)	-0.400 (0.747)
<i>Controls</i>						
Rural dummy	-0.203 (0.042)	-0.210 (0.017)	-0.950 (0.462)	-0.102 (0.357)	-0.232 (0.213)	-0.277 (0.081)
Urban dummy	0.078 (0.548)	0.021 (0.854)	0.187 (0.253)	0.152 (0.278)	-0.264 (0.263)	-0.339 (0.093)
Capital dummy	-0.478 (0.007)	-0.377 (0.018)	-0.445 (0.038)	-0.300 (0.105)	-0.784 (0.062)	-0.763 (0.031)
GDP dummy	1.953 (0.000)	1.507 (0.000)	1.895 (0.000)	1.192 (0.002)	0.473 (0.902)	1.455 (0.663)
NAT dummies	yes	yes	yes	yes	yes	yes
Obs	271	271	201	201	61	61
R ² -adj	0.893	0.911	0.905	0.923	0.668	0.750
LIK	-199.977	-187.269	-151.127	-138.559	-37.986	-35.142
AIC	467.955	444.538	362.255	339.117	97.972	94.283
SC	590.427	570.613	461.354	441.520	121.192	119.614
Moran's I	0.660 (0.000)		3.454 (0.001)		1.613 (0.107)	
LM-ERR	0.157 (0.692)	0.016 (0.900)	0.547 (0.460)	0.069 (0.793)	0.000 (0.990)	0.033 (0.856)
LM-LAG	20.693 (0.000)		21.230 (0.000)		4.510 (0.034)	

Note: Probability values in parentheses.

Another extension of the model is given below. We start from model (4), where the production of knowledge in a region depends not only on its own research efforts and internal factors, but also on the knowledge available in other regions. The knowledge available in the neighbouring regions is proxied by their innovation output measured through their patents. However, following the idea of Moreno *et al.* (2005), we consider spillovers generated by research efforts, which allow us to analyse the sensitivity of our previous results to other proxies for inter-regional knowledge flows as follows:

$$\log I_{i,t} = \beta_1 \log RD_{i,t-q} + \beta_2 \log DENS_{i,t-s} + \beta_3 \log HK_{i,t-s} + \beta_4 W \log RD + \sum_{c=1}^n \delta_c NAT_{ic} + \gamma_1 DU_{i,t-s} + \gamma_2 DR_{i,t-s} + \gamma_3 DCAP_{i,t-s} + \gamma_4 DGDP_{i,t-s} + \varepsilon_{i,t} \quad (5)$$

where the term $W \log RD$ is a spatial lag for the innovation input. The method of estimation is OLS, since there are no endogeneity problems.

In Table 15, results of the estimation of equation (5) are reported, and they prove extremely interesting. RD performed in nearby regions is significant and its value is almost half the elasticity of internal RD. Spatial autocorrelation is however still present and therefore should be eliminated by means of the usual spatial lag model. Results, which are not shown, prove that the spatial lag of the RD variable does not keep its significance when the lag of the dependent variable is inserted. The latter is a stronger proxy for inter-regional knowledge flows.

The same model is estimated for the EU and the NA cases, too. The results deserve some attention. For the European case we find that the R&D efforts of neighbouring regions are relevant from a statistical and quantitative point of view. On the contrary, for North America we still find a positive coefficient, but we cannot reject the hypothesis that is not statistically different from zero. Again, however, in both models spatial autocorrelation is still detected and the LM tests suggest the use of the spatial lag model. As above the spatial lag of the dependent variable prevails with respect to the spatial lag of RD.¹⁸

18. In the two models (OECD and EU) and we try to insert the spatial lag of second order (neighbours of neighbours) to see if the relationship extends beyond the first ring of regions. Results show that the second lag is positive and significant too (confirming Moreno *et al.*, 2005) but spatial autocorrelation does not completely disappear. The insertion of a third lag of RD did not give significant coefficient and spatial autocorrelation is still detected.

Table 15. KPF estimation with spatial lag of RD

Variables	OECD	Europe	North America	
	OLS	OLS	OLS	
Log (RD)	0.603 (0.000)	0.633 (0.000)	0.627 (0.000)	0.507 (0.000)
log (HK)	1.064 (0.000)	0.940 (0.000)	0.964 (0.000)	1.011 (0.033)
log (DENS)	0.089 (0.031)	0.118 (0.926)	0.126 (0.072)	0.057 (0.277)
W log (RD)	0.253 (0.006)	0.312 (0.010)	0.289 (0.160)	0.214 (0.200)
W2 log (RD)			0.280 (0.051)	
DGDP*log(RD)	-0.155 (0.209)	-0.180 (0.217)	-0.162 (0.261)	
DGDP*log(HK)	-0.483 (0.002)	-0.424 (0.028)	-0.393 (0.041)	
<i>Controls</i>				
Rural dummy	-0.201 (0.041)	-0.092 (0.471)	-0.641 (0.613)	-0.245 (0.178)
Urban dummy	0.062 (0.627)	0.163 (0.311)	0.151 (0.343)	-0.283 (0.220)
Capital dummy	-0.434 (0.014)	-0.396 (0.062)	-0.415 (0.048)	-0.858 (0.034)
GDP dummy	1.923 (0.000)	1.818 (0.000)	1.690 (0.000)	0.513 (0.054)
NAT dummies	yes	yes	yes	yes
Obs	271	201	201	61
R ² -adj	0.897	0.908	0.909	0.683
LIK	-195.657	-147.144	-144.877	-37.154
AIC	461.314	356.289	353.755	94.307
SC	587.388	458.691	459.460	115.416
Moran's I	3.306 (0.001)	3.379 (0.001)	3.290 (0.001)	1.069 (0.285)
LM-ERR	0.347 (0.556)	0.566 (0.452)	0.466 (0.495)	0.123 (0.726)
LM-LAG	14.480 (0.000)	15.139 (0.000)	12.472 (0.000)	2.724 (0.099)

Note: Probability values in parentheses.

Another interesting test for our results rests on the application of a different kind of matrix in order to measure geographical proximity. In other words, we use a distance matrix instead of the contiguity one. The former is made of values where each cell reports the inverse of distance for each pair of regions. Thus, this weight matrix has only non-zero elements for each observation pair that is assumed to interact. For this reason it has been used only for the analysis referring to continental macro-areas: Europe and North-America, since here distances are to some extent homogenous. They are not used when the link between regions can be classified in two very different categories: terrestrial and maritime. Both matrices, nevertheless, rely on the idea that only geographical proximity matters in the interaction across regions.

Table 16. KPF estimation with distance matrix

Variables	Europe		North America
	OLS	ML	OLS
Log (RD)	0.600 (0.000)	0.677 (0.000)	0.548 (0.000)
log (HK)	0.969 (0.000)	0.624 (0.001)	1.061 (0.262)
log (DENS)	0.126 (0.080)	0.075 (0.229)	0.069 (0.182)
W log (I)		0.012 (0.000)	
DGDP*log(RD)	-0.155 (0.291)	-0.209 (0.103)	
DGDP*log(HK)	-0.433 (0.027)	-0.169 (0.340)	
<i>Controls</i>			
Rural dummy	-0.950 (0.462)	-0.088 (0.433)	-0.236 (0.197)
Urban dummy	0.187 (0.253)	0.189 (0.183)	-0.271 (0.243)
Capital dummy	-0.445 (0.038)	-0.232 (0.223)	-0.815 (0.044)
GDP dummy	1.895 (0.000)	1.032 (0.012)	0.466 (0.078)
NAT dummies	yes	yes	yes
Obs	201	201	61
R ² -adj	0.905	0.922	0.679
LIK	-151.127	-139.517	-38.144
AIC	362.255	341.034	94.288
SC	461.354	443.436	113.286
Moran's I	7.125 (0.000)		2.852 (0.004)
LM-ERR	0.780 (0.377)	0.069 (0.793)	1.355 (0.244)
LM-LAG	21.236 (0.000)		0.043 (0.836)

Note: Probability values in parentheses.

Results in Table 16 show that the OLS estimation indicates that spatial autocorrelation has to be corrected in order to get accurate estimates only in the EU case. This result, again, can be attributed to the different geographical setting in the two continents: Europe with relatively small regions and North America, on the contrary, with rather big regions. The model is, therefore, corrected only in the EU case with the spatial lag models reported in column three. Main results prove robust. The only main difference being the fact that now the coefficient of the spatial lag of the dependent variable is much lower since it applies not only to neighbouring regions but to close and far away regions at the same time. Nonetheless, elasticities have the expected sign and significance.

Another important check is the one which is performed thanks to an enlargement of our sample of countries in order to include Japan and Korea, two leaders in the technological competition in many sectors. This operation is made possible through the estimation of some indicators which are not directly provided in the OECD Regional Database. In particular, we estimate HK for Korea, where we used the national average for each region, and RD for both countries. In this case, again, the average national quota of GDP spent in R&D was used for all the regions.

Results in Table 17 show that OLS estimation needs to be corrected with the ML one. Results are surprisingly stable both with respect to the sign and their significance. In particular, it is worth noting that the elasticities of RD and of agglomeration economies are now slightly bigger, possibly due to the inclusion of regions where these two phenomena are particularly important. One more interesting difference is that the dummy for the capital region becomes only marginally significant in the ML model probably due to the fact that the Tokyo and the Seoul regions in these two countries are central in their respective national innovation systems. The global goodness of fit is still remarkably high with almost 90% of the variance explained with our regressors.

Table 17. KPF estimation with Japan and Korea

Variables	OECD	
	OLS	ML
Log (RD)	0.556 (0.000)	0.574 (0.000)
log (HK)	1.114 (0.000)	0.954 (0.000)
log (DENS)	0.093 (0.030)	0.098 (0.009)
W log (I)		0.185 (0.000)
DGDP*log(RD)	-0.113 (0.378)	-0.203 (0.074)
DGDP*log(HK)	-0.411 (0.011)	-0.293 (0.039)
<i>Controls</i>		
Rural dummy	-0.203 (0.045)	-0.228 (0.010)
Urban dummy	0.084 (0.511)	0.016 (0.885)
Capital dummy	-0.358 (0.042)	-0.250 (0.106)
GDP dummy	1.757 (0.000)	1.333 (0.000)
NAT dummies	yes	yes
Obs	287	287
R ² -adj	0.878	0.902
LIK	-222.798	-206.251
AIC	517.596	486.502
SC	649.338	621.903
Moran's I	4.007 (0.000)	
LM-ERR	0.234 (0.629)	0.049 (0.824)
LM-LAG	28.261 (0.000)	

Note: Probability values in parentheses.

Finally, the model is replicated with a different version of the dependent variable, that is PCT per worker (instead of PCT per capita) in order to take into account differences in labour force participation across regions. This indicator is sometimes preferred to the one above, since it refers to those people who are actually involved in the real economic activity and who are therefore potential inventors and innovators. The estimation refers to the model with interactive dummies and it is replicated for EU and NA in order to assess robustness of main elasticities.

Table 18. KPF estimation with PCT per worker

Variables	OECD		Europe		North America	
	OLS	ML	OLS	ML	OLS	ML
Log (RD)	0.531 (0.000)	0.542 (0.000)	0.564 (0.000)	0.580 (0.000)	0.840 (0.238)	0.686 (0.263)
log (HK)	1.068 (0.000)	0.930 (0.000)	0.963 (0.000)	0.764 (0.000)	0.592 (0.670)	0.949 (0.432)
log (DENS)	0.110 (0.008)	0.166 (0.002)	0.146 (0.042)	0.137 (0.027)	0.074 (0.154)	0.082 (0.067)
W log (I)		0.146 (0.000)		0.188 (0.000)		0.133 (0.019)
DGDP*log(RD)	-0.059 (0.630)	-0.134 (0.227)	-0.120 (0.415)	-0.219 (0.090)	-0.296 (0.678)	-0.208 (0.736)
DGDP*log(HK)	-0.488 (0.002)	-0.371 (0.009)	-0.402 (0.040)	-0.193 (0.269)	0.233 (0.866)	-0.119 (0.922)
<i>Controls</i>						
Rural dummy	-0.189 (0.056)	-0.199 (0.250)	-0.081 (0.530)	-0.093 (0.408)	-0.250 (0.170)	-0.288 (0.064)
Urban dummy	0.041 (0.750)	-0.009 (0.936)	0.124 (0.446)	0.093 (0.509)	-0.239 (0.300)	-0.301 (0.128)
Capital dummy	-0.484 (0.007)	-0.394 (0.013)	-0.464 (0.031)	-0.332 (0.076)	-0.791 (0.054)	-0.781 (0.024)
GDP dummy	1.908 (0.000)	1.511 (0.000)	1.789 (0.000)	1.167 (0.003)	-0.220 (0.953)	0.646 (0.843)
NAT dummies	yes	yes	yes	yes	yes	yes
Obs	270	270	201	201	61	61
R ² -adj	0.897	0.905	0.909	0.918	0.661	0.741
LIK	-198.248	-187.226	-151.044	-140.113	-36.415	-33.841
AIC	462.496	442.451	362.087	342.226	94.829	91.682
SC	581.244	564.798	461.186	444.628	118.049	117.013
Moran's I	3.583 (0.000)		3.300 (0.001)		1.505 (0.132)	
LM-ERR	0.012 (0.912)	0.099 (0.753)	0.281 (0.596)	0.033 (0.856)	0.003 (0.956)	0.044 (0.833)
LM-LAG	20.691 (0.000)		18.788 (0.000)		4.197 (0.041)	

Note: Probability values in parentheses.

Results in Table 18 are, again, quite reassuring in terms of robustness of most of the elasticities. Especially, the elasticities for the main factors, that is, RD and HK are almost the same, while the coefficient related to agglomeration economies is slightly higher. As for the dummies, some results are maintained, that is those related to the capital dummy and to the GDP dummy, whilst others are not, namely those referring to the rural dummy. The two regressions for the two macro-areas results mirror those obtained with the previous dependent variable.

6. Conclusions

This report is the first systematic attempt to analyse comparatively the distribution of innovative activity across regions in OECD countries with a set of homogenous measures for both input and output in the process of knowledge production and dissemination.

It describes the patterns of innovation across regions in some OECD countries in order to determine the main factors behind the process of localisation of such activities. Moreover, a Knowledge Production Function (KPF) model has been estimated at the regional level for OECD regions in order to assess the role of R&D efforts, localised knowledge and absorptive capacity in enhancing regions' technological capability. This model has also been used to assess whether the pattern, extent and pace of the absorptive capacity depend on geographical proximity (measured in different ways), but also on other economic and institutional similarities.

The descriptive analysis shows that there are important differences in the innovative performance of regions in OECD countries. Such performance is clustered in some areas both in continental Europe, in North America and, to a lesser extent, in Japan. The presence of spatial dependence is also detected with some specific tests. This spatial dependence is increasing with time and the consequent dispersion of innovation activities is rising in almost all OECD except North America. Some convergence is detected at the country level but not at the regional level.

In spite of some regularities which characterise the three macro-areas of Europe, North America and Asia-Pacific, we find that there are important differences across countries which probably indicate the relevance of a national innovation system which is above regional performance.

The estimation of the KPF shows that the model which links innovative output to main inputs is at work across regions in OECD countries. Some particular results deserve highlighting. R&D is a very important determinant and human capital is even more so. Agglomeration economies measured by population density also prove significant, even though its elasticity is relatively lower. The performance of neighbouring regions is an important determinant of innovative performance, which suggests the existence of important diffusion effects across space. Innovative performance is not favoured in rural regions and it is more facilitated in rich areas. Institutional factors which differentiate each country are still relevant. The impacts of RD and HK are not necessarily monotonous, but they decrease as the country becomes richer.

These results are differentiated with respect to two macro-areas for which it was possible to estimate two separate models: Europe and North America.

In the former, the main results which were found for OECD hold. On the contrary, in North America we find a higher impact of human capital and a significant impact of agglomeration economies, which become non significant in some models in Europe. In both cases we find that diffusion of innovation goes across regions, but this phenomenon appears weaker in North America than in Europe.

It is very difficult to draw some policy lessons from this report which has not focussed on specific strategic issues. Nevertheless, some important aspects can be pinpointed. The importance of the regional

level is the first important result to highlight. Secondly, it is worth noting that not only internal efforts are relevant but also those of neighbouring regions, especially in Europe. Thirdly, agglomeration externalities may have different effects in different institutional and social settings. Finally, institutions at the country level can still play a crucial role in determining the base of innovative performance for regions and countries.

As for future work, given the exploratory nature of this analysis, many research avenues are worth pursuing. One interesting path is the analysis at the sectoral level of the KPF model, in order to assess the characteristics of some of the externalities at the local industrial level. In particular, this kind of analysis would allow one to discriminate among specialisation and diversity externalities, a discrimination which is crucial in order to gain some useful policy implications. Another important avenue for future investigation is the one referring to the analysis at a lower level of territorial disaggregation in order to assess the robustness of our results with respect to this dimension.

BIBLIOGRAPHY

- Acs, Z., Anselin, L. and Varga A. (1997), "Local Geographic Spillovers between University Research and High Technology Innovations", *Journal of Urban Economics*, Vol. 42, Issue 3, pp. 422-448.
- Acs, Z., Anselin, L. and Varga A. (2002), "Patents and innovation counts as measures of regional production of new knowledge", *Research Policy*, Vol. 31, Issue 7, pp. 1069-1085.
- Acs, Z.J., Audretsch D.B. and Feldman M.P. (1992), "Real Effects of Academic Research: Comment", *The American Economic Review*, Vol. 82, No. 1, pp. 363-367.
- Antonelli, C. (2008), "Pecuniary Externalities: the Convergence of Directed Technological Change and the Emergence of Innovation Systems", *Brick working papers* 3/2008
- Arrow, K.J. (1962), "The Economic Implications of Learning by Doing", *The Review of Economic Studies*, Vol. 29, No. 3, pp. 155-173.
- Arthur, B.W. (1990), "Silicon Valley Locational Clusters: When Do Increasing Returns Imply Monopoly", *Mathematical Social Research*, Vol. 19, pp. 235-51.
- Audretsch, D. and Feldman M. (1996), "R&D Spillovers and the Geography of Innovation and Production", *American Economic Review*, 86, 631-640.
- Aydalot, P. (1986), *Milieux innovateurs en Europe*, GREMI, Paris.
- Basberg, L. (1987), "Patents and the Measurement of Technological Change: A Survey of the Literature", *Research Policy*, Vol. 16, pp. 131-41.
- Bottazzi, L. and Peri G. (2003), "Innovation and spillovers in regions: Evidence from European patent data", *European Economic Review*, 47, pp. 687-710.
- Bradley, R. and Gans J.S. (1998), "Growth in Australian cities", *Economic Record*, Vol. 74, pp. 266-278.
- Breschi, S. (2000), "The Geography of Innovation: A Cross-Sector Analysis", *Regional Studies*, 34, pp. 213-29.
- Buesa, M., Heijs J. Martinez-Pellitero M. and Baumert T. (2006), "Regional systems of innovation and the knowledge production function: the Spanish case", *Technovation* 26, pp. 463-472.
- Caniels, M. (2000), *Knowledge spillovers and economic growth*, Edward Elgar, Cheltenham.
- Carlino, G.A. and Hunt R.M. (2007), *Innovation across U.S. industries: the effects of local economic characteristics*, Working Papers 07-28, Federal Reserve Bank of Philadelphia.

- Carlino, G.A., Chatterjee S. and Hunt R.M. (2007), “Urban Density and the Rate of Innovation”, *Journal of Urban Economics*, Vol. 61, pp. 389-419.
- Cihan, C. (2005), *An Empirical Analysis of Knowledge Production Function: What Differs Among The OECD Countries Including Turkey*, Doctoral Thesis, Sydney University.
- Cingano, F. and Schivardi F. (2004), “Identifying the Sources of Local Productivity Growth”, *Journal of the European Economic Association*, Vol. 2, No. 4, pp. 720-742.
- Combes, P.P. (2000), “Economic Structure and Local Growth: France, 1984–1993”, *Journal of Urban Economics*, Volume 47, Issue 3, pp. 329-355.
- Crescenzi, R., Rodriguez-Pose A. and Storper M. (2007), “The territorial dynamics of innovation: a Europe–United States comparative analysis”, *Journal of Economic Geography*, 7, pp. 673–709.
- Criscuolo, P. and Verspagen B. (2006), *Does it matter where patent citations come from? Inventor versus examiner citations in European patents*, Eindhoven Centre for Innovation Studies, Working Paper 05/06.
- Davelaar, E.J. (1991), *Regional Economic Analysis of Innovation and Incubation*, Avebury Publishing Company, Brookfield, VT.
- Dernis, H. and Khan M. (2004), *Triadic Patent Families Methodology*, STI Working Paper 04/2, OECD, Paris.
- Dernis, H., Guellec D. and van Pottelsberghe B. (2001), *Using Patent Counts for Cross-Country comparisons of Technology Output*, OECD, Paris.
- Ellison, G. and Glaeser E.L. (1997), “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach”, *Journal of Political Economy*, vol. 105, no. 5.
- European Commission (2006), European “Innovation Scoreboard 2006, Comparative Analysis of Innovation Performance”, paper prepared by the Maastricht Economic Research Institute on Innovation and Technology (MERIT) and the Joint Research Centre (Institute for the Protection and Security of the Citizen) of the European Commission, INNO-METRICS.
- European Commission (2006), *Innobarometer on cluster’s role in facilitating innovation in Europe*, The Gallup Organisation, Brussels.
- Feldman, A. (1999), “Innovation in cities: Science-based diversity, specialization and localized competition”, *European Economic Review*, Vol. 43, Issue 2, pp. 409-429.
- Feldman, M.P. (1994), *The Geography of Innovation*, Springer, Berlin.
- Fujita, M. (1988), “A monopolistic competition model of spatial agglomeration: A differentiated product approach”, *Regional Science and Urban Economics*, Vol. 18, 87-124.
- Glaeser, E.L., Kallal H.D. Scheinkman J.A. and Shleifer A. (1992), “Growth in Cities”, *The Journal of Political Economy*, 100, 6. pp. 1126-1151.
- Greunz, L. (2003), “Geographically and technologically mediated knowledge spillovers between European regions”, *The Annals of Regional Science*, Springer, Berlin.

- Griliches, Z. (1979), "Issues in Assessing the Contribution of Research and Development to Productivity Growth", *Bell Journal of Economics*, 10, 92-116.
- Griliches, Z. (1979), "Issues in Assessing the Contribution of Research and Development to Productivity Growth", *Bell Journal of Economics*, 10, pp. 92-116.
- Hall, B., Jaffe A.B. and Trajtenberg M. (2001), *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools*, NBER, WP 8498.
- Jacobs, J. (1969), *The Economy of Cities*, Jonathan Cape, London.
- Jaffe, A.B. (1986), "Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value", *American Economic Review*, vol. 76, pp. 984-1001.
- Jaffe, A.B. (1986), "Technological Opportunity and Spillovers of R&D: evidence from Firms' Patents, Profits and Market Value", *American Economic Review*, 76, 984-1001.
- Jaffe, A.B. (1989), "Real Effects of Academic Research", *The American Economic Review*, Vol. 79, No. 5, pp. 957-970.
- Jaffe, A.B. and Trajtenberg M. (1999), "International knowledge flows: evidence from patent citations", *Economics of Innovation and New Technologies*, 8, pp. 105-136.
- Jaffe, A.B. and Trajtenberg M. (2002), *Patents, Citations and Innovations: A Window on the Knowledge Economy*, MIT Press, Cambridge, MA.
- Jaffe, A.B., Trajtenberg M. and Henderson R. (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics*, 108, pp. 577-598.
- Krugman, P. (1991), "Increasing Returns and Economic Geography", *Journal of Political Economy*, vol. 99(3), pages 483-499, June.
- Le Sage, J.P., Fischer M. and Scherngell T. (2007), "Knowledge spillovers across Europe. Evidence from a Poisson spatial interaction model with spatial effects", *Papers in Regional Science* (forthcoming).
- Lukatch, R. and Plasmans J. (2003), "Measuring Knowledge Spillovers in the New Economy Firms in Belgium Using Patent Citations", *Global Business and Economics Review*, 5, pp. 75-99.
- Maggioni, M. A., Nosvelli M. and T. E. Uberti, (2007), "Space versus networks in the geography of innovation: A European analysis", *Papers in Regional Science*, vol. 86(3), pages 471-493, 08
- Marshall, A. (1890), *Principles of Economics*, Macmillan, London.
- Massard, N. and S.Riou (2002), *L'impact des structures locales sur l'innovation en France: spécialisation ou diversité?*, *Région et Développement*, 16, pp. 111-136.
- Maurseth, P.B. and Verspagen B. (2002), "Knowledge spillovers in Europe: a patent citations analysis", *Scandinavian Journal of Economics*, 104, pp. 531-545.

- Maurseth, P. B. and Verspagen B. (1999), *Europe: One or Several Systems of Innovation? An Analysis Based on Patent Citations*, in Fagerberg J., Guerrieri P. and Verspagen B. (eds), *The Economic Challenge for Europe – Adapting to Innovation Based Growth*, Edward Elgar, Cheltenham, pp. 149-74.
- Moreno, R., Paci R. and Usai S. (2005), “Spatial spillovers and innovation activity in European Regions”, *Environment and Planning A*, 37, pp. 1793-1812.
- Moreno, R., Paci R. and Usai S. (2006), “Innovation Clusters in the European Regions”, *European Planning Studies*, vol 14, no 9.
- Ó hUallicháin, B. (1999), “Patent Places: Size Matters”, *Journal of Regional Science*, Vol. 39, No. 4, pp. 613-636.
- OECD (2003), *Overview of Recent Trends in Patent Regimes in the United States, Japan and Europe*, DSTI/STP/TIP(2003)6, OECD, Paris.
- OECD (2007a), *Compendium of OECD Work on intellectual property (IP)*, OECD, Paris.
- OECD (2007b), *Compendium of Patent Statistics*, OECD, Paris.
- Ottaviano, G. P. and Thisse, J.F. (2003), *Agglomeration and economic geography*, Handbook of Regional and Urban Economics, in: J. V. Henderson & J. F. Thisse (ed.), *Handbook of Regional and Urban Economics*, edition 1, Vol. 4, chapter 58, pp. 2563-2608, Elsevier.
- Paci, R. and Usai S. (1999), “Externalities, Knowledge Spillovers and the Spatial Distribution of Innovation”, *Geojournal*, Volume 49, Number 4, pp. 381-390.
- Paci, R. and Usai S. (2000), “Technological Enclaves and Industrial Districts: An Analysis of the Regional Distribution of Innovative Activity in Europe”, *Regional Studies*, Taylor and Francis Journals, vol. 34(2), pp. 97-114.
- Paci, R. and Usai S. (2002), *Externalities and local economic growth in manufacturing industries*, in Fingleton B. (ed.), *European Regional Growth*, Springer-Verlag, Berlin.
- Paci, R. and Usai S. (2006), “Agglomeration economies and growth-The case of Italian local labour systems, 1991-2001”, CRENoS working papers, 2006/12, forthcoming in *Revue d’Economie Industrielle*.
- Paci, R. and Usai S. (2008), Knowledge flows across European regions, *Working Paper CRENoS 2007/04*, forthcoming in *Annals of Regional Science*.
- Rodriguez-Pose, A. and Crescenzi R. (2006), *R&D, Spillovers, Innovation Systems and the Genesis of Regional Growth in Europe*, BEER paper No. 5.
- Romer, P.M. (1986), “Increasing returns and Long Run Growth”, *Journal of Political Economy*, 94, pp. 1002-37.
- Swann, G.M.P. (1996), “Technology Evolution and the Rise and Fall of Industrial Clusters”, in *Revue Internationale de Systemique*, 10(3), pp. 285-302.

Storper, M. (1992), “The Limits to Globalisation: Technology Districts and International Trade”, in *Economic Geography*, 68(1), pp. 60-93.

Tappeiner, G., C. Hauser and J. Walde, (2008), “Regional knowledge spillovers: Fact or artifact?”, in *Research Policy* 37, pp. 861–874

Van Oort, F.G., van Soest D.P. and Gerking S. (2005), *Dynamic information externalities and employment growth in the Netherlands*, in: R. Boschma & R. Kloosterman (eds.), *Learning from clusters. A critical assessment from an economic-geographical perspective*, pp. 303-332, Springer Verlag, Berlin.

Venables, A.J. (1996), “Equilibrium locations of vertically linked industries”, *International Economic Review*, 37, pp. 341-359.

WIPO (2007), *Statistics on Worldwide Patent Activities*, WIPO Patent Report, Switzerland, Geneva.