UNIVERSITÀ DEGLI STUDI DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

Corso di Dottorato in
Economia della Produzione e dello Sviluppo
Ciclo XXIV

Determinants of Knowledge Flows and their Effects on Economic Growth

Relatore:
Prof. Fabio Montobbio

Tesi di dottorato di:
Riccardo Cappelli

Esame Finale Anno 2012

Dedicated to my parents Andrea and Floriana
And my girlfriend Sabrina

ACKNOWLEDGMENT

I would like to thank all people who helped me to develop this thesis.

I want to thank Prof. Dirk Czarnitzki for his precious help during the course of my studies. I want also to thank Prof. Kornelius Kraft and Thorsten Doherr for their help in developing this thesis.

I want to thank Dr. Nicola Matteucci who encouraged me to undertake the PhD studies.

I want to thank Prof. Gioacchino Garofoli and all the people who are engaged in the organisation of the PhD of our faculty.

I want to thank my friends and colleagues at the University of Insubria for their support and the pleasant evenings of good food and wine.

I want to thank Dr. Bettina Peters, Prof. Katrin Hussinger and Prof. Dirk Czarnitzki who hosted me in their universities and research institutions.

I want to thank my parents Andrea and Floriana who have made possible my whole academic formation with their moral and financial support. I want to thank my girlfriend Sabrina for her love which embellishes my life.

I am deeply indebted to my supervisor, Prof. Fabio Montobbio, for his constant support, suggestions and teaching. His enthusiasm for the research and his encouragement helped me to write a more accurate thesis.

Contents

LIST OF ACRONYMS

CIS	Community Innovation Survey
EPO	European Patent Office
ERA	European Research Area
EU	European Union
GDP	Gross Domestic Product
ICT	Information and Communication Technologies
IPC	International Patent Classification
ISTAT	Italian National Institute of Statistics
MIP	Mannheim Innovation Panel
NUTS	Nomenclature des Unités Territoriales Statistiques
OECD	Organisation for Economic Cooperation and Development
OLS	Ordinary Least Squares
OST	Observatoire des sciences et des techniques
PPML	Poisson Pseudo Maximum Likelihood
R&D	Research and Development
US	United States
USPTO	United States Patent and Trademark Office
ZEW	Centre for European Economic Research

Chapter 1

Introduction

The diffusion of knowledge is pivotal in the knowledge-based economy, from the perspectives of both academics and policy makers. In contrast to standard neoclassical theory (Solow, 1956, 1957), endogenous (Romer, 1987, 1990; Grossman and Helpmann, 1991) and evolutionary (Nelson and Winter, 1982; Fagerberg, 1994) theories identify knowledge flows as a key factor explaining economic growth. The EU in its Europe 2020 strategy (European Commission, 2010) underlines the importance of knowledge flows for achieving a “smart” economy, i.e. one where growth is driven by knowledge and innovation. One of the goals of this strategy is to remove the territorial barriers to knowledge flows in order to develop an integrated ERA (European Commission, 2000, 2010).

The term knowledge flows generally refers to the diffusion of knowledge from one institution, the knowledge generating institution, to another, the knowledge receiving institution. Knowledge receiving institutions can use the acquired knowledge to generate new ideas (e.g. invention) and/or to exploit the knowledge in the economy through the development of products. Thus, we can distinguish two aspects to knowledge flow processes: the spread of knowledge, and the effect of the knowledge on innovation and economic output. This thesis examines both of these aspects of the diffusion of knowledge and investigates some areas that so far have been neglected in the literature.

This thesis is based on three empirical studies of knowledge flows: an analysis of the determinants of interregional knowledge flows, and two studies, at different levels of aggregation, on the economic impact of the diffusion of knowledge. One of these analyses the effect of knowledge flows on firms' economic performance, and the other investigates the impact of knowledge flows on regional economic growth. Although the research questions and methodological analyses used differ, all three studies have a common thread, which is the effort to find a suitable measure to capture the peculiarities of the phenomenon of knowledge diffusion.

The concept of knowledge diffusion is not unique and has several facets. Knowledge has an important feature that distinguishes it from traditional assets such as capital and labour. Knowledge is an immaterial good, which may be embedded, for example, in blueprints, in human beings or in organizations (Soete and Ter Weel, 1999). Thus, knowledge can be codified or tacit. Some examples of codified knowledge are the description of an invention in a patent document, or the ideas presented in scientific and journal articles. The concept of tacit knowledge was introduced by Polanyi (1967) and subsequently adopted in evolutionary theories by Nelson and Winter (1982). Some knowledge cannot be codified: this may be because it would be too costly, or because it resides only in a person or an organization. This type of knowledge is also described as sticky (von Hippel, 1994). The literature on knowledge flows mainly uses measures that capture the diffusion of codified knowledge and tend to ignore the diffusion of tacit knowledge. Although some scholars analyse the diffusion of tacit knowledge, studies of both types of knowledge are rare. The studies in this thesis use a set of measures of knowledge flows that capture the diffusion of both codified and tacit knowledge. Given the increasing interest of governments in innovation policies, a better understanding of the mechanisms of knowledge diffusion and the innovative process is fundamental for the formulation of effective policy.

Knowledge flows at the spatial level

One of the focuses of this thesis is the spatial aspect of knowledge flows. The diffusion of knowledge across space has a long tradition in the economic literature. Since the seminal contribution of Marshall (1920), several economists have analysed knowledge flows and the mechanisms which contribute to explain the geographical diffusion of knowledge. The growth pole theory developed by Perroux (1950) and the following studies on local productive systems (Garofoli, 1983), industrial districts (Becattini, 1979) and innovative milieus (Camagni, 1991) underline the importance of geographical proximity in the diffusion of knowledge.

Measuring knowledge flows is a central issue in the empirical analysis which attempt to investigate the geographical dimension of the diffusion of knowledge. Krugman[1] (1991) emphasizes that knowledge flows are "invisible" and cannot be "measured and tracked". This extreme view, which leaves no room for analysis of knowledge flows, was contested by Jaffe et al. (1993), who pointed to patent citations as one means of tracking the diffusion of knowledge in space. The citations in a patent are an indication that the knowledge contained in the cited patent was used to develop the new ideas contained in the citing patent. Patent citations are similar to the references in journal articles, but differ from them in one important aspect. For example, the author of an article may cite other authors for reasons of gratitude to a "master", or friendship, or because a referenced author may be a possible referee. Inventors are not driven by these reasons, and the citing of other patents introduces the risk that the citing patent is not entirely novel and not a patentable invention. Patent citations provide a paper trail of the

[1] Analysing the reasons for the geographical location of the firms in an industry described by Marshall (1920), i.e. labour market pooling, the presence of specialized suppliers and knowledge spillovers, Krugman (1991, p. 53) has argued that economists should focus on the first two of these because "knowledge flows [...] are invisible; they leave no paper trail by which they may be measured and tracked".

diffusion of knowledge among inventors (Jaffe et al., 1993; Peri, 2005; Montobbio and Sterzi, 2011).

Following Jaffe et al. (1993), analyses of the diffusion of knowledge in geographical space mainly use patent citations data. These studies show that the diffusion of knowledge is geographically localized (Jaffe et al., 1993; Maurseth and Verspagen, 2002; Bottazzi and Peri, 2003; Peri, 2005). This means that the diffusion of knowledge is more likely within the territory, i.e. country and region, in which it is generated (Maurseth and Verspagen, 2002; Peri, 2005). For instance, Peri[2] (2005, p. 208) shows that "only 20% of average knowledge is learned outside the average region of origin, and only 9% is learned outside the country of origin". Some studies explore the geographical limit to the diffusion of knowledge: for instance, Bottazzi and Peri (2003), show that the diffusion of knowledge occurs within a radius of 300 km.[3]

To sum up, the works on knowledge flows shows that there are significant barriers preventing knowledge from flowing freely in a geographic space. These barriers include physical distance and the historical, social and institutional differences that characterize different countries and regions. Geographical proximity facilitates contact and communication with the knowledge generator; also, people in the same organization with a common culture and similar values, will be better able and more willing to communicate, will be more trusting and will have some expectation of reciprocity (Agrawal et al., 2008).

Technological advances in the field of communications and increased integration among countries suggest that the diffusion of knowledge is becoming less localized. The development and diffusion of ICT (e.g. fibre optics, satellite communications,

[2] The author uses patent citations data for 113 regions of Europe and North America.
[3] The authors use a patent dataset that covers 86 regions and 13 EU countries.

social networks) have reduced the costs of communicating and eased the exchange of knowledge over long distances. In parallel, the greater integration among countries related to trade and cultural and institutional factors suggest that the impact of territorial boundaries on knowledge flows has decreased. One of the objectives of this thesis research is to analyse whether these "agents of change" have had an effect on knowledge flows.

Although the diffusion of knowledge may be geographically localized, studies that use patent citations data do not take account of the fact that knowledge can spread in various ways and that geographic proximity can have different effects on different knowledge flow channels. Knowledge can be diffused simply by someone reading, e.g. a patent document, or by personal interaction[4] with the inventors. In the first case, the diffusion of knowledge is easy and costless (e.g. using the internet); in the second case, face-to-face contact may be hindered by geographical distance. However, personal interaction also allows tacit knowledge (e.g. know-how) to be transferred. Patent citations are a good measure of the diffusion of codified, but not tacit knowledge. In this thesis we consider measures that takes account of both codified and tacit knowledge flows.

Chapter 2 analyses the patterns of knowledge flows among European regions[5] in the period 1981-2000. This examination takes three directions::

- analysis of the determinants of knowledge flows;
- analysis of the trends of these determinants;

[4] The literature on industrial districts (see e.g.: Becattini, 1979) points out that informal relationships among people and organizations are one reason why knowledge flows occur more easily among firms located in a district.

[5] Our dataset contains data for the EU25 member states plus Norway and Switzerland. A region is defined according to the Eurostat NUTS definition, which, in most cases, corresponds to a lower geographical level than the national (e.g. the 20 Italian administrative regions), but in some cases refers to whole countries (e.g. Luxembourg).

- analysis of the role of European integration on knowledge flows.

The starting point is to confirm whether the diffusion of knowledge among European regions is geographically localized. Several works (Maurseth and Verspagen, 2002; Peri, 2005, Paci and Usai, 2009) show that the diffusion of knowledge among European regions is hampered by geographical distance and national borders. The present analysis extends the previous works using a larger dataset, i.e. a larger number of countries and/or years, and in contrast to existing work, provides a direct comparison of how knowledge diffuses through two channels, which convey different types of knowledge. Using inventor citations, we capture the diffusion of mainly codified knowledge, and using inventor collaborations we capture the diffusion of tacit knowledge. Several scholars show that inventor collaborations are a good measure of the diffusion of tacit knowledge (Almeida and Kogut, 1999; Singh, 2005; Breschi and Lissoni, 2009).

The diffusion of tacit knowledge, is likely to be hampered by geographical barriers more than the diffusion of codified knowledge. Thus, we hypothesize that inventor collaborations will be more localized than inventor citations.

These kinds of analyses provide only statistic pictures of the diffusion of knowledge and do not consider that the patterns of knowledge flows are probably changing due to the observed technological advances and greater integration among countries. Work on evolution of the impact of geographical factors on knowledge flows is scarce and provides mixed results (Johnson et al., 2006; Griffith et al., 2007; Sonn and Storper, 2008; Montobbio and Sterzi, 2012). Among European regions, to our knowledge, only the work by Paci and Usai (2009) analyses the dynamics of the localization of knowledge flows. Paci and Usai compare the geographical distribution of patent citations over two years and show that the impact of physical distance has increased over time, but the impact of national borders has decreased. We extend this work by considering a broader temporal sample, using a refined methodology and exploiting inventor collaborations as a measure of knowledge flows.

This work is one of the few attempts to analyse the impact of European integration on knowledge flows. In particular, it investigates whether the EU enlargement processes following the annexation of Spain and Portugal in 1986 and of Austria, Sweden and Finland in 1995, have increased knowledge flows among the regions of EU country members. The effect of European integration has been studied in the trade flows literature (Baldwin, 1995; Bussière et al., 2008; Carrère, 2006); here, we apply the analysis to knowledge flows because the reduction in the institutional barriers between EU countries suggests there is greater international exchange of knowledge among EU members.

To make these analysis we use a modified gravity model.[6] Gravity models are used traditionally in the literature on bilateral trade between countries (Tinbergen, 1962; Rose, 2001; Micco et al., 2003; Anderson and van Wincoop, 2003), and have found application in the study of knowledge flows (Maurseth and Verspagen, 2002; Picci, 2010). Gravity models derive from Newton's law of gravity that any two bodies attract one another with a force that is proportional to the product of their masses and is inversely proportional to the distance between them. In the case of knowledge flows, the "bodies" are the regions' numbers of patents, and "distance" is represented by some measure of geographical proximity such as physical distance. Econometric procedures are used to estimate the impact of proximity on knowledge flows, indicated by the sign

[6] Another methodology common in the literature is the matching approach (Jaffe et al., 1993). In this work we use the gravity model because, unlike the matching approach, it allows quantification of the impact of geographic proximity on knowledge flows and identification of the dynamics of proximity effects.

of the coefficient. For instance, if the sign of the coefficient related to the physical distance is negative this is an indication of geographically localized diffusion of knowledge because knowledge flows between regions decrease with the geographical distance between the regions. In order to take account of possible sources of bias, which are neglected in the traditional literature, this work uses PPML estimates (Sanots Silva and Tenreyro, 2006).

Knowledge flows and economic growth: regional level

Another line of work on the diffusion of knowledge analyses the relationship between knowledge flows and economic growth. Recent models of endogenous growth (Romer, 1987, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1998), which challenge standard neoclassical approaches (Solow, 1956, 1957), assume that knowledge is a partially excludable good and that economic growth is driven by knowledge flows. Evolutionary theories, such as endogenous theories, also assume that economic growth is explained by knowledge flows (i.e. imitation processes). However, evolutionary theory assumes that the positive impact of knowledge flows on economic growth cannot be considered guaranteed because the knowledge receiving system or region needs adequate technological (Cohen and Levinthal, 1989) and social (Abramovitz, 1986) absorptive capacity.

The impact of knowledge flows on the economic growth of countries or regions is analysed empirically making use mainly of indirect measures of knowledge flows such as stock of knowledge generated by R&D activities of other countries or regions (for an analytic survey see e.g.: Griliches, 1979; Czarnitzki et al., 2006; Hall et al., 2010). This body of work is based on the idea that institutions (territories) appropriate the knowledge produced by other institutions (territories) depending on their proximity.

Various measures of proximity are used such as physical distance (Rodriguez-Pose and Crescenzi, 2008), trade intensity (Coe and Helpman, 1995) and technological proximity (Park, 1995).

"Foreign" (i.e., in other regions) R&D stock is an indirect measure of knowledge flows and may be an indication of potential knowledge exchange between regions. To overcome this limitation, in this thesis we adopt a more direct measure of knowledge flows which considered the diffusion of both tacit and codified knowledge. Although the importance of tacit knowledge for the economic growth is widely recognized in the theoretical literature,[7] in empirical analyses the diffusion of tacit knowledge is neglected. This is mainly due to the unavailability of appropriate data.

Chapter 3 analyses the impact of knowledge capital on the economic growth of Italian regions, in 1995-2007. This work follows two directions:

- the measurement of knowledge capital;
- the impact of knowledge capital on economic growth.

Both aspects are related. An important aspect of analysing the relationship between knowledge and economic growth is measuring the knowledge capital of a system or region. The accumulation of knowledge in a region generally is determined by the processes of generation and diffusion of knowledge. The R&D activities of firms and the scientific output of universities and other research institutes contribute to the knowledge capital of a region. Since it is not possible to completely protect or keep secret R&D results, part of the knowledge is appropriated by other companies, which allows knowledge to flow from one region to another.

[7] E.g. see the literature on the industrial districts (Brusco, 1996; Becattini and Rullani, 1996).

The knowledge generated within a region traditionally is measured by the intensity of R&D (e.g. R&D expenditure as a percentage of GDP), and by number of patents to evaluate "successful knowledge" (Fagerberg et al., 1997; Fagerberg and Verspagen 2002; Sterlacchini, 2008). Both measures are adopted in this work.

Instead of R&D stock of other regions, we measure knowledge flows using a direct measure - backward patent citations. An advantage of using patent citations is that it allows consideration of intraregional knowledge flows.

To take account of the diffusion of tacit knowledge, we introduce a measure of knowledge flows based on the mobility of the inventors. An inventor who moves between firms carries with him a wealth of knowledge, skills and experience, which results in the diffusion of knowledge. At the regional level, this involves both knowledge leakage and acquisition of new knowledge, and the opposite effects in relation to innovation capability. To construct inventor mobility indexes we construct a new dataset of Italian inventors with at least one patent application to the EPO.

To assess the impact of knowledge capital in terms of economic growth we use a model based on the technology-gap approach (Fagerberg et al., 1997), where the rate of growth of GDP per capita is a function of changes to the stock of knowledge, measured as R&D and number of patents for internal innovative activities, and patent citations and inventor mobility for knowledge diffusion processes.

Knowledge flows and economic growth: firm level

The impact of knowledge flows on economic performance has been widely analysed at firm level (for a survey see e.g.: Mairesse and Sassenou, 1991; Los and Verspagen, 2000). Similar to macro level studies, knowledge flows generally are measured as the stock of R&D generated by firms within the same and sometimes other industries (for a survey see: Hall et al., 2010). However, at micro level these measures have some limitations because they do not consider some important sources of knowledge such as university or other research institution.

Chapter 4 analyses the impact of knowledge flows on economic performance at firm level. We exploit a dataset of German firms for the period 2000-2002[8] to:

- measure knowledge flows between firms;
- investigate the impact of knowledge flows on economic performance.

Firms belong to supplier, customer and organizational networks, all of which can be sources of knowledge. The traditional measure of knowledge flows, i.e. the stock of R&D of other firms, can be considered an indirect measure of the knowledge flows from other firms. However, this measure treats all sources of knowledge in the same way and does not take account of the specificity of single channels for the appropriation of knowledge. The effects of knowledge diffusion on innovative or economic output may differ depending on the source of the knowledge. For instance, it can be argued that the knowledge acquired by rivals leads mainly to product imitation, while the knowledge originating from firms working in different sectors has very different effects on innovative and economic outputs.

To overcome these limitations we open the black box of knowledge flows and distinguish among the various sources of knowledge: customers, competitors, suppliers, research institutions. A distinction is made between imitation, i.e., the introduction of products new to the firm, but not the market, and innovation, i.e. the introduction of products new to the market. This allows the effect of the four sources of knowledge on imitation and innovation processes to be analysed.

[8] The data are from the MIP dataset.

We exploit an empirical model in which firm sales from innovation and firm sales from imitation are functions of the different sources of knowledge flows and of the knowledge capital generated by internal innovative activities.

The thesis is organized as follows. Chapter 2 presents the work on the diffusion of knowledge among European regions. Chapters 3 and 4 discuss the work on the economic impact of knowledge flows at the regional (Chapter 3) and firm (Chapter 4) levels. Chapter 5 provides conclusion, offers some final considerations and suggests some possible directions for future research.

Chapter 2

European Integration and Knowledge Flows across European Regions[9]

2.1. Introduction

The diffusion of knowledge is an important engine of economic growth. Models of endogenous growth (Romer 1986, 1990) show that, if at least part of the knowledge produced as result of R&D activities is non-rivalrous and non-excludable, knowledge diffusion generates economic growth. At the same time, there is a large body of literature (e.g. Jaffe et al., 1993; Maurseth and Verspagen, 2002) showing that the diffusion of knowledge is geographically localized. This means that knowledge does not flow freely through the space due to the presence of geographical and institutional barriers. The limited diffusion of knowledge is one of the factors that might explain the lack of convergence across countries or regions and their disparities in terms of economic growth (Grossman and Helpman, 1991; Fagerberg et al., 1997).

Since 1993 (Jaffe et. al., 1993), many studies have analysed the geographical patterns of knowledge diffusion, but little is known about the evolution over time of the impact of geographical and institutional factors on knowledge flows. Over time, decreased transport costs, technological advances and diffusion of ICT and the greater both commercial and political integration among countries, have eased the exchange of knowledge over long distances. We could assume, then, that the diffusion of knowledge is being hindered less by geographical and institutional barriers.

[9] This chapter represents a slightly different version of the paper: Cappelli, R., Montobbio, F., 2012. European integration and knowledge flows across European regions, mimeo.

In this work, we analyse the patterns of knowledge flows between the regions of the EU 27 member states, plus Norway and Switzerland. Using EPO data for the period 1981-2004, we study whether knowledge flows among European regions are less localized than in the past. This analysis is particularly important in view of the innovative policies recently adopted by the EU (Europe 2020 agenda) in order to reduce the technology gap with leader countries and the EU's prioritization of "smart growth", i.e. the idea of Europe as an economy based on knowledge and innovation as key drivers of economic growth (European Commission, 2010). The EU wants to promote the diffusion of knowledge among member countries in order to "complete" the ERA (European Commission, 2000, 2010).

We follow studies in the literature and use patent citations to measure knowledge flows (see e.g. Maurseth and Verspagen, 2002; Bottazzi and Peri, 2003; Paci and Usai 2009; Bacchiocchi and Montobbio 2010). Studies show that knowledge flows between European regions are hindered by physical distance and country borders. In particular, interregional knowledge flows are more likely to occur within nations and decrease with increasing distance between regions. However, these analyses provide only static pictures of the phenomenon considered and whether or not Europe is more integrated in the field of knowledge remains an open question. A first, and perhaps unique, attempt to analyse the diffusion of knowledge within Europe from a dynamic point of view is provided by Paci and Usai (2009). They use the geographic distribution of patent citations for two different periods and compare the impact of geographic distance and national boundaries. Their comparison shows that the impact of national borders is smaller but the impact of geographical distance is larger. In this work, we extend this analysis in two directions: first we consider a broader temporal sample (i.e. we compare

20 different periods), second, we adopt an analytical methodology that takes account of some sources of bias that have been neglected in other studies.

We also analyse the effect of European integration on knowledge flows. The integration among countries reduces the political and institutional barriers between them and reduces the cultural distance between people, thereby easing the diffusion of knowledge among countries and regions. This framework seems particularly useful to describe the situation of Europe, where there is a continuous process of reducing the institutional barriers that divide countries and a growing implementation of EU policies to promote the diffusion of knowledge. Although this line of analysis was suggested by Maurseth and Verspagen (2002), there is no empirical work analysing the impact of European integration on knowledge flows.

This work studies the processes of European integration, from 1981 to 2000 in particular, when two processes of enlargement expanded the EU from 10 to 15 members. The integration effect stemming from EU enlargement has been investigated mainly in the trade flow literature (see e.g. Baldwin, 1995; Bussière et al., 2008; Gil et al., 2008). These works analyse the effect of European institutional integration on bilateral trade both between EU country members and between EU members and non-members. We go further and consider the hypothesis that a reduction in the institutional barriers, in addition to having an effect in terms of trade flows be facilitating the exchange of knowledge between EU regions.

The diffusion of knowledge can occur through different channels, which, in turn, are differentiated by the type of knowledge conveyed. Since the diffusion of knowledge is not a univocal concept and has different facets, in our analysis we consider both patent citations and inventor collaborations. Both measures capture the diffusion of knowledge among inventors, with the difference that inventor collaboration, by definition, requires face-to-face contact, while citation requires no personal contact. In the latter case, knowledge diffusion is related to only the codified component of the knowledge. There are empirical studies showing the pattern of knowledge flows based on one of these measures, but not both. In this work we adopt both measures in order to obtain a more complete picture of the diffusion of knowledge. We also compare the impact of geographical and institutional barriers on inventor citation and collaboration in order to investigate whether these measures are affected equally or whether one dominates the other.

We exploit a modified version of the gravity model and our results are obtained through PPML estimates. Our estimation strategy is in two steps. First, in order to evaluate the evolution over time of the impact of geographical and institutional barriers on the two measures of knowledge flows, we look at the different sub-periods in our sample and perform separate cross-sectional analyses for each period. We check the robustness of our results through pooled time series cross-sectional analysis. Second, we investigate the existence of changes in the diffusion of knowledge due to the process of European integration. In this case, in order to take account of possible heterogeneity bias due to the presence of unobserved factors we perform fixed-effects estimates.

The chapter is organized as follows. Section 2.2 discusses in more detail the literature on the diffusion of knowledge and the theoretical justification for our analysis. Sections 2.3 and 2.4 respectively present the gravity model used in our estimates and the methodology adopted in order to resolve some sources of bias. Section 2.5 presents the data and the variables used, and explains the difference between the two measures of knowledge flows adopted. Section 2.6 presents and discusses the results of our estimates. Section 2.7 offers some final considerations and some suggestions for future work.

2.2 Background to the study

This work examines the patterns of knowledge flows between European regions. Recent endogenous growth (Romer, 1986, 1990) and evolutionary (Nelson and Winter, 1982) theories underline the importance of the diffusion of knowledge for the economic growth of a system or region. However, due to the intangible nature of knowledge, an important and non-trivial aspect of knowledge flows analysis is measurement (Krugman, 1991). Also, the diffusion of knowledge can occur through different channels, which may involve people, companies or other organizations and, more importantly, these channels may convey different types of knowledge (e.g. codified and tacit). Thus, we need to know more about the typology of the knowledge flows investigated; the concept of knowledge flows is not unique, but has several facets. In the following sub-sections we give more detail on the sources of knowledge flows (section 2.2.1) and we review the related literature, distinguishing between analysis of static (section 2.2.2) and dynamic (section 2.2.3) knowledge.

2.2.1 Economic growth and sources of knowledge flows

Unlike neoclassical theories (Solow, 1956, 1957), endogenous growth theories (Romer, 1986, 1990; Lucas, 1988) assume that knowledge is a public good that is partially excludable. This means that, on the one hand, innovation activities carried out by local actors contribute to the technological and economic growth of countries or regions (Romer, 1990; Aghion and Howitt, 1998) and, on the other hand, the diffusion of knowledge is a way to reduce the technological gap and, consequently, to promote economic convergence (Grossman and Helpman, 1991). Similar to endogenous growth theories, the technology gap approach (Fagerberg, 1994; Fagerberg and Verspagen, 2002; Sterlacchini, 2008), which can be seen as an evolutionary theory (Nelson and Winter, 1982), assumes that the diffusion of knowledge is an important means of technology catching-up, although this process is not automatic and requires adequate technological (Cohen and Levinthal, 1989) and social (Abramovitz, 1986) absorptive capacity.

The relationship between knowledge flows and economic growth is analysed using different knowledge measures. For instance, Coe and Helpman (1995), show that the importation of products allows a country to take advantage of the results of the innovative activities of the trading partners and, therefore, has positive effects in terms of economic growth. Other studies use foreign direct investments to capture knowledge flows (Blomstrom and Kokko, 1998; Aitken and Harrison, 1999; Crespo and Fontoura, 2007), although the results in terms of economic benefits vary (Keller, 2004). Finally, several authors measure knowledge flows using "narrowly defined" indicators based on patent data, such as number of patents, patent citations and inventor collaborations (Keller, 2001).

Due to the diversity of the indicators used in the knowledge flows literature, it is useful to make a classification of knowledge based on its characteristics. A general but important classification is provided by Griliches (1979) who distinguishes between two types of knowledge flows which result in externalities: "rent spillovers" and "pure spillovers". Both are externalities in the sense that an institution benefits from the knowledge generated by the R&D activities conducted by another institution. However, "rent spillovers" require a market transaction and occur when the higher quality of an improved or new product is not matched by a commensurate increase in price. "Pure spillovers" occur when a new idea spreads from one institution to another without any market transaction.[10] In this case, knowledge has the characteristics of a public good because it is non-rival and non-excludable. Given these two specific properties, "pure" or knowledge spillovers are crucial for explaining the increasing returns in endogenous growth theories (Romer, 1986, 1990).

Due to the importance of knowledge spillovers, patent citations are probably "the best measure of spillovers in the sense of an externality" (Keller, 2001, p. 48) because other measures might capture some element of market transactions. Patent citations are widely used in the literature on the diffusion of knowledge between countries or regions (Jaffe et al., 1993; Maurseth and Verspagen, 2002; Bottazzi and Peri, 2003; Peri, 2005). The citations included in a patent can be considered references to the knowledge used by the inventors to develop the new ideas contained therein.[11] Thus, patent citations allow the "tracing" of knowledge flows among inventors (Jaffe et al., 1993). However, the use of patent citations as measure of knowledge flows has some shortcomings due to the fact that citations can be added by patent examiners without the inventor being aware of the patent cited. Some studies that compare patent citations data with case study or innovation survey data, show that the patent citations, although imperfect, are a valid measure of "real" knowledge flows (Jaffe et al., 1998, Jaffe and Trajtenberg, 2002; Dugeut and MacGarvie, 2005).

The knowledge contained in a patent can be divided into codified and tacit knowledge. Codified knowledge is represented by the description of the invention

[10] However, it should be noted that distinguishing between these two types of spillovers is not always easy because some channels of knowledge flows traditionally considered "pure spillovers" hide market transactions (Feldman and Kogler, 2010).

[11] E.g., if patent A cites patent B, we can assume that the inventors who developed the idea contained in patent A know and use the idea contained in patent B.

contained in the patent document; tacit knowledge is represented by the information that is not codified (e.g. technical know-how) and is embedded in the inventors of the patent. Since tacit knowledge is transferable only through personal interaction among individuals, technological collaborations can provide a measure of knowledge diffusion that is superior to patent citations.

Several papers show that inventor collaboration as a good measure of knowledge spillovers (Singh, 2005; Picci, 2010). Singh (2005) shows that the relationships among the inventors, identified by their direct or indirect collaboration, explain much of the diffusion of knowledge (measured by patent citations) between firms and regions. Further confirmation of the validity of use of collaboration as a measure of knowledge flows is provided in Breschi and Lissoni (2009).

In line with the above, to measure knowledge flows between European regions we use patent citations and inventor collaborations. In this way, we can verify whether changes in technology and EU integration have had different effects on the two channels of knowledge diffusion.

Note that, although we use data on patent citations and inventor collaborations we do not take into account of "rent spillovers"; these measures refer to only a specific part of knowledge spillovers. In particular, inventor citations and collaboration occur only if both regions - the one generating the knowledge and the one receiving it - perform R&D and apply for patents. Thus, we do not consider knowledge spillovers stemming from informal innovation activities.

2.2.2 Geographic localization of knowledge flows

Empirical analysis based on patent citations shows that the diffusion of knowledge is a geographically localized phenomenon (Jaffe et al., 1993; Maurseth and Verspagen, 2002; Bottazzi and Peri, 2003; Fischer et al., 2009; Bacchiocchi and Montobbio, 2010). These studies make use of different country dataset and methodologies.[12] For instance, the pioneering work of Jaffe et al. (1993), provides evidence for the US on the localized diffusion of knowledge in geographical areas. Using USPTO data and a matching procedure that takes account of the existing geographic concentration of patent activity, they find that a patent is more likely to be cited by other patents originating in the same country, state or metropolitan statistical area. The first empirical evidence for Europe was provided by Maurseth and Verspagen (2002). They use EPO data for 14 European countries and gravity model estimates to show that the likelihood of citations between two patents developed in two different regions is negatively affected by the presence of an international border and by the geographical distance between them. Peri (2005), using National Bureau of Economic Research patent data for 113 regions of Europe and the US and gravity model estimates, shows that only 20% of knowledge goes out from the region of origin on average, and only 9% goes out from the country of origin.

The literature on the relationship between geographical and institutional factors and inventor collaboration is smaller than the body of work on patent citations and is mostly country level (Guellec et al., 2001; Picci, 2010; Montobbio and Sterzi, 2012). Guellec et al. (2001), using EPO data for 29 OECD countries and a gravity model show, that the possibility of collaboration between inventors[13] residing in two different countries decreases as the geographical distance between them increases (expressed as sharing the same territorial border). Picci (2010) uses a series of datasets (EPO, USPTO and other national patent office data) for 42 countries (including OECD countries) and applies a

[12] Different explanations have been provided for the geographically localized diffusion of knowledge. For a survey see e.g.: Breschi and Lissoni, 2001; Audretsch and Feldman, 2004.

[13] They consider different types of collaboration based on the geographical locations of inventor and applicant. Here, we examine only inventor collaborations.

gravity model, to show that the possibility of international collaboration is affected by geographic distance, both physical and in the form of a territorial border.

The above shows that the diffusion of knowledge, measured by patent citations and by inventor collaborations, is geographically localized. The present work contributes by providing a direct comparison of the effect of geographical factors on the two channels of knowledge diffusion.

2.2.3. Evolution over time of geographical and institutional factors

Despite growing interest in the spatial diffusion of knowledge, few studies analyse the evolution over time of the impact of geographical factors on knowledge spillovers (Johnson et al., 2006; Griffith et al., 2007; Paci and Usai, 2009; Sonn and Storper, 2008; Montobbio and Sterzi, 2012) and there is also no consensus: some show that the diffusion of knowledge is occurring in a more localized way than in the past, while others show the opposite. Johnson et al. (2006), using data for all the US inventors with USPTO patents for the period 1975-1999 and using at Tobit model with geographical distance as the dependent variable, a time trend of variable of interest, and a set of control variables (e.g. technological characteristics of the patent), show that the average distance between the citing and the cited patents increases by almost seven miles per year and the average distance between coinventors also increase by four miles per year. Griffith et al. (2007), analyse the changes over time of the propensity for inventor citations to be national, using USPTO data for the period 1975-1995 for 5 countries (US, Japan, France, Germany, UK), and two groups of countries (EU countries and Rest of the World). They apply a duration model that looks at the "speed" of the patents of different countries to cite the same patent, and show that the national border effect decreased during the period investigated. Sonn and Storper (2008), using USPTO data for the period 1975-1997 for the US and matching procedures (e.g. Jaffe et al., 1993), find that inventor citations became more localized at country, state and metropolitan levels.

The work that is closest to ours in relation to the data and methodology used is Paci and Usai (2008), which uses EPO patent citations data for the regions in 17 European countries and makes use of gravity model estimates. They construct two cohorts of citing patents, of patents granted in 1990 and in 1998. For each cohort they consider backward citations (i.e. citations to previous patents), for 1978-1990 for the first cohort and for 1978-1998 for the second cohort. They run two separate estimates and compare the results for the impact of geographical distance and national borders on interregional knowledge flows. Their results show that the geographical distance effect has increased, while the impact of national border has decreased.

The present analysis extends that by Paci and Usai (2009) by comparing the impact of physical distance and country border for 20 periods during 1981-2000 and implementing a methodology that overcomes some of the aspects neglected in the literature such as lack of control for heterogeneity and truncations bias problems. We compare the dynamics of distance and national border effects for inventor citations and inventor collaborations.

We also analyse the effect of European integration on knowledge flows. The literature on trade flows and knowledge flows mainly identifies two reasons for the presumed "death" of distance and institutional barriers: technological advances in the field of transportation and communications, and increased integration among countries.

Both factors are analysed in the literature on trade flows, but work on knowledge flows ignores the effect of integration. Some studies analyse the difference in the diffusion of knowledge among members of a group of countries united by some supranational institutions. For instance, Picci (2010) shows that EU membership positively affects international collaboration. However, this paper considers only static effects and, therefore, does not investigate whether greater integration between countries has an effect on reducing the national barriers to knowledge flows, which would require dynamic analysis that takes account of the factors that over time contribute to the integration of countries.

In the literature on trade flows, works that analyses the dynamic impact of European integration mainly consider the effects of the processes of EU enlargement on trade for the countries involved. Based on this work and the hypothesis that a reduction in institutional barriers may affect knowledge flows, we analyse the impact of integration through the EU enlargement processes that occurred during the period investigated. This allows us to distinguish between the effect on the three types of countries involved, i.e. new members, old members, and non-EU members.

To our knowledge, this analysis is one of the few attempts to test the dynamic impact of European integration on the diffusion of knowledge.

2.3 The empirical model

The econometric model used to analyse knowledge flows among European regions is a modified version of the gravity model. The gravity model has been widely used in work on bilateral trade between countries (Rose, 2001; Micco et al., 2002; Anderson and van Wincoop, 2003), and also has found application in the study of knowledge flows, measured by patent citations (Maurseth and Verspagen, 2002) and by collaborations (Picci, 2010; Montobbio and Sterzi, 2012). In its basic form, the model predicts that the diffusion of knowledge between two regions is directly proportional to the inventive mass of the regions and inversely proportional to the geographic distance between regions. The model can be expressed by the following formula:

$$[2.1] \quad C_{ijt} = G \frac{{P_{it}}^{\alpha} \, {P_{jt}}^{\beta}}{{dist_{ij}}^{\gamma}}$$

where C_{ijt} is the variable capturing knowledge flows (in our case measured by number of citations or collaborations) between regions *i* and *j* at period *t*, *G* is a constant, P_{it} and P_{jt} are total numbers of patents (inventive mass) for the two regions, and $dist_{ij}$ is the geographical distance between the two regions.

To identify citations between patents from different regions, we identified a citation from region *j* to region *i* as occurring when the citing patent as at least one inventor residing in the region *j* and the cited patent has at least one inventor residing in the region *i*. In the case of patents with more than one inventor residing in the same region (*i* or *j*), citations are counted only once.

Since we are interested in "pure knowledge spillovers" (Griliches, 1979), we do not consider citations between two patents developed by a single firm. Self-citations between firms are not considered externalities. We also exclude self-citations between inventors because, by definition, this cannot be considered an exchange of knowledge between individuals (Agrawal et al., 2006).[14]

Inherent in the use of patent citations is a truncation bias problem, due to the fact that we observe only a limited period of the legal life of the patent. This problem is greater for recent patent cohorts. For instance, a patent developed in 1985 may be cited by patents developed in the period 1985-2004, but a patent developed in 2000 can only be

[14] It would be useful in order to isolate the "rent spillovers" to control over time for firm mergers, acquisitions (Maurseth and Verspagen, 2002) and patent licensing (Jaffe et al., 1993), but the availability of adequate data and the effort involved make this task very difficult.

cited by patents developed in the period 2000-2004. Thus, patents from 1985 have 25 years of time to be cited, while patents from 2000 have only 5 years. This time lag difference could be a source of bias in evaluation of the changes in distance or country border effects because the diffusion of knowledge could follow paths that are influenced by time. For instance, it is possible that the "new" knowledge flows, in the first periods, more easily at the local level than beyond.

In order to overcome this problem, we consider only the pairs of patents where the time lag between cited and citing patent is four years or less.[15] For instance, C_{ijt} is the total number of citations contained in patents developed in region *j* (spillover-receiving region) during the period *t-(t+4)* and directed to patents developed in region *i* (spillover-generating region) in period *t*.[16] Thus, we have a sample consisting of a set of cited patents for the period 1981-2000, and a set of citing patents for the period 1981-2004.

The second measure of knowledge flows used in this work is technological collaborations. To identify a collaboration among inventors from different regions, we observe a collaboration between the region *i* and the region *j* if, in a patent developed by more than one inventor, at least one co-inventor is resident in region *i* and at least one co-inventor is resident in region *j*. Similar to the case of patent citations, if a patent has more than one inventor resident in the same region (*i* or *j*) the collaboration is counted only once. For inventor collaborations, there is obviously no truncation problem.

In the empirical studies, variables are added to the basic model in order to take account of regional differences in terms of technological specialization (Maurseth and

[15] We consider different time lags, but the results obtained from our estimates are quite similar (see the appendix 2.D).

[16] Moreover, the "inventive mass" in all the equations with patent citations as dependent variable is adapted in order to take into account of these temporal windows of four years. Thus, the term P_{it} become the total number of patents developed in period *t*, while P_{jt} is the total number of patents developed during the period *t-(t+4)*.

Verspagen, 2002), social, political and institutional differences between regions (Picci, 2010) and geographic factors that may enhance the localization effect determined by physical distance. In this work, we include the following control variables:

- Technological proximity (*Technology*): this controls for the sectoral distribution of patents within the two regions because citations are mostly at the intrasectoral level and, also, there are some combinations of sectors that are more frequently cited than others (Maurseth and Verspagen, 2002; Montobbio and Sterzi, 2012). Therefore, the geographical distance effect could be influenced by the technological specialization of regions. Following the literature (see e.g. Peri, 2005; Montobbio and Sterzi, 2012), this variable is measured by the Jaffe (1986) index, i.e. the uncentred correlation between the vectors expressing the distribution of the patents in 30 technology classes (OST, 2004) for the region *i* and the region *j*, that is: $TP_{ij} = P_iP_j'/[(P_iP_i')(P_jP_j')]1/2$. This variable takes values between 0 (when the vectors are orthogonal) and 1 (when the vectors are identical).

- Common language (*Language*): this variable controls for the language spoken in the two regions. A common language facilitates interpersonal relationships and, thus, facilitates the diffusion of knowledge between regions (Maurseth and Verspagen, 2002). This variable is represented by a dummy that takes the value 1 if the two regions have the same language.

- Common Border (*Border*): this variable controls for whether the regions are neighbours. It determines whether adjacent regions engage in greater exchange of knowledge (Paci and Usai, 2009). It is represented a dummy that takes the value 1 if the two regions have a common border.

- Country border (*National_Border*): this variable controls for whether two regions are located in the same country and takes account of political, social and historical features specific to a nation. These features can facilitate the exchange of knowledge among inventors located in the same nation (Maurseth and Verspagen, 2002; Paci and Usai, 2009). The variable is represented by a dummy that takes the value 1 if the two regions belong to the same country.

- We include a set of dummies (Reg_i and Reg_j) to take account of regional-specific unobservable effects of region i and region j.

In order to reduce the impact of outliers, we express the variable for the inventive mass (P_i and P_j) and distance ($dist_{ij}$) in logarithmic form:

$$[2.2] \quad E(C_{ijt}) = \mu_{ijt} = \exp(\text{constant} + \alpha \ln(P_{it}) + \beta \ln(P_{jt}) + \gamma \ln(\text{dist}_{ij}) + \\ + \varrho \text{ Technology}_{ijt} + \sigma \text{ Language}_{ij} + \Omega \text{ National\_Border}_{ij} + \\ + \varphi \text{ Border}_{ij} + \eta_i \text{ Reg}_i + \Theta_j \text{ Reg}_j)$$

Equation [2.2] estimates the aggregate effects for the whole period investigated and for the different sub periods in order to evaluate changes in effects. We conduct estimations to check the previous results on the dynamics of the distance effect. In particular, we perform these estimates on a panel dataset obtained by pooling annual data and a specification using the same variables but also a set of time dummies ($Year_t$). Moreover, the variable for the distance ($dist_{ij}$) is replaced by a set of variables ($dist_{ijt}$) obtained by interacting the variable distance with the time dummies in order to allow the coefficient of distance to shift yearly. This gives the following specification:

[2.3] $$\mu_{ijt} = \exp(\text{constant} + \alpha \ln(P_{it}) + \beta \ln(P_{jt}) + \gamma_t \ln(\text{dist}_{ijt}) + \\ + \varrho\ \text{Technology}_{ijt} + \sigma\ \text{Language}_{ij} + \Omega\ \text{National\_Border}_{ij} + \\ + \varphi\ \text{Border}_{ij} + \eta_i\ \text{Reg}_i + \Theta_j\ \text{Reg}_j + \varrho_t\ \text{Year}_t)$$

Since changes in the distance effect can also capture changes in the country border effect, we make a further check allowing the variable for national border effect (*National_Border*$_{ij}$) to vary over time. In particular, we replace the *National_Border*$_{ij}$ variable with a set of dummies (*National_Border*$_{ijt}$) constructed by interacting this *National_Border*$_{ij}$ with the time dummies. This gives:

[2.4] $$\mu_{ijt} = \exp(\text{constant} + \alpha \ln(P_{it}) + \beta \ln(P_{jt}) + \gamma_t \ln(\text{dist}_{ijt}) + \\ + \varrho\ \text{Technology}_{ijt} + \sigma\ \text{Language}_{ij} + \Omega_t\ \text{National\_Border}_{ijt} + \\ + \varphi\ \text{Border}_{ij} + \eta_i\ \text{Reg}_i + \Theta_j\ \text{Reg}_j + \varrho_t\ \text{Year}_t)$$

As already mentioned, one of the aims of our analysis is to examine the effect of the European Integration process on interregional knowledge flows. The time period covered by our analysis, 1981 to 2000, includes two enlargement processes. The first was in 1986, following the entry of Spain and Portugal to the EU, and the second was in 1995, following the entry of Austria, Finland and Sweden.

The effect of European integration on reducing national border barriers has been analysed in depth in the bilateral trade literature (see e.g. Spies and Marques, 2009). The impact of European integration is estimated using a dummy variable added to the basic gravity model in order to capture deviations from the volumes of trade predicted by the model. We follow the same methodology and make use of a time varying dummy variable (*EU_both*) which is set equal to 1 for the pairs of regions that are members of the EU. In order to take account of a possible effect on knowledge flows towards non-

EU members we add a time varying dummy (*EU_one*) which is set equal to 1 when only one region is a member of the EU.

To estimate the effect of European enlargement on the knowledge flows between regions we estimate the following equation:

[2.5] $$\begin{aligned}\mu_{ijt} = {} & \exp(\text{constant} + \alpha \ln(P_{it}) + \beta \ln(P_{jt}) + \gamma_t \ln(\text{dist}_{ijt}) + \\ & + \varrho\ \text{Technology}_{ijt} + \sigma\ \text{Language}_{ij} + \Omega\ \text{National\_Border}_{ij} + \\ & + \varphi\ \text{EU\_both}_{ijt} + \omega\ \text{EU\_one}_{ijt} + \varrho\ \text{Border}_{ij} + \eta_i\ \text{Reg}_i + \theta_j\ \text{Reg}_j + \\ & + \varrho_t\ \text{Year}_t)\end{aligned}$$

In our sample we can distinguish between three types of regions based on membership of the EU: *old* refers to regions in countries that were EU member before 1981; *new* refers to regions of countries that joined the EU in the period 1981-2000; *never* refers to regions of non-EU member countries. The variable *EU_both* is equal to 1 when the two regions are *old* or *new,* or *old* and *new*, *EU_one* is equal to 1 when one region is *old* or *new* and the other one is *never*. This allows us to identify whether the aggregate effect of EU membership (*EU_both* and *EU_one*) hides different behaviors in the different subgroups. Because our dataset is at regional level, we can distinguish between the effects of European integration on the diffusion of knowledge within and between countries by breaking down the above variables on the basis of a shared national border. Figure 2.1 shows that there are 12 types of pairs of regions.

	old	new	never
old	intra / extra	extra	extra
new	extra	intra / extra	extra
never	extra	extra	intra / extra

old: regions in countries that were EU member before 1981 (Belgium, Denmark, Germany, Greece, France, Ireland, Italy, Luxembourg, Nederland and United Kingdom)

new: regions of countries that joined the EU in the period 1981-2000 (Spain, Portugal, Austria, Finland and Sweden)

never: regions of non-EU member countries (Bulgaria, Czech Republic, Hungary, Poland, Norway and Switzerland)

intra: regions are located in the same country

extra: regions are located in two different countries

Figure 2.1. Matrix of the combinations between European regions

However, the two measures of knowledge flows we use have some characteristics that need to be taken into account in determining the specification to be used in the estimates. In particular, patent citations capture the diffusion of knowledge from patent inventors to other inventors who developed a patent in a subsequent period. Thus, patent citations measures unidirectional flows between inventors or regions. Collaborations captures the interchange of knowledge between inventors for the generation of a new patent. Thus, inventor collaborations measures bidirectional flows between inventors or regions. This distinction means that in evaluating the impact of European integration on pairs of regions using patent citations rather than inventor collaborations, we can distinguish between the effects on the knowledge generating region and the knowledge receiving region. For instance, for the pairs of *old* and *new* regions, we can distinguish between the diffusion of knowledge from *old* to *new* regions (*EU_both_old_new*) and knowledge flows from *old* to *new* regions (*EU_both_new_old*). In the case of inventor collaborations there are only the bidirectional flows between *old* and *new* regions, thus, we have only one variable (*EU_both_old_new*).

Thus, we have a further specification in which the variables *EU_one* and *EU_both* are replaced with their subgroups. In this regard, see Figure 2.2.

Finally, since it is possible to distinguish between the two phases of EU enlargement (1986 and 1995), we test whether the effect of EU integration is different in the two periods and, consequently, in the two different groups of nations. For instance, the variable *EU_both_old_new* can be divided into *EU_both_old_new_86* and *EU_both_old_new_95*.

Initial variable	*Description*	*Region*	*National Border*	*Final variable*
EU_both	knowledge flows between EU members	knowledge flows between *old* and *new* regions	international flows	EU_old_new EU_new_old
		knowledge flows between *new* regions	intranational flows	EU_both_new_new_intra
			international flows	EU_both_new_new_extra
		knowledge flows between *old* regions	intranational flows	EU_both_old_old_intra
			international flows	EU_both_old_old_extra
EU_one	knowledge flows between EU members and not EU members	knowledge flows between *old* and *never* regions	international flows	EU_one_old_no EU_one_no_old
		knowledge flows between *new* and *never* regions	international flows	EU_one_new_no EU_one_no_new
EU_never	knowledge flows between not EU members	knowledge flows between *never* regions	intranational flows	EU_no_no_intra
			international flows	EU_no_no_extra

Figure 2.2. European integration and sub group of regions

2.4 Methodology

The gravity models in equations [2.2] to [2.5] can be estimated using different econometric specifications. Following a procedure widely used in the literature on international trade, our gravity model can be estimated using OLS on the log-linear version of the previous equations. However, this procedure has some problems which can lead to biased estimates. First, there are pairs of regions that do not have any interchange of knowledge (either citations and/or collaborations), which means a zero value of the dependent variable (C_{ij}). These observations are treated as missing in the estimates which introduces bias in the coefficients estimated. Gravity models also have an inherent problem of heteroschedasticity, which can lead to biased estimates. To address these issues we use a PPML estimator (Santos Silva and Tenreyro, 2006).[17]

To estimate the effect of European enlargement on the knowledge flows between regions we perform pooled cross-section (equation [2.5]) and fixed-effects estimates. The latter are statistically more robust than the former because they control for unobserved heterogeneity (Cheng and Wall, 2005),[18] which can explain the amount of bilateral knowledge flows and, additionally, the probability that two regions are in the same European agreement. However, this procedure has the disadvantage that we cannot estimate the impact of the European integration for regions whose "EU member status" does not change during the period covered by our analysis (e.g. *old* and *never* regions). In fact, the inclusion of pair region dummies implies that only information on time variation in the variables is used to estimate their coefficient values, while

---

[17] Thus, we adopt an econometric specifications where C_{ijt}, our count data dependent variable, follows a Poisson distribution $\left(Pr[C_{ijt}] = \frac{exp^{-\mu_{ijt}} \mu_{ijt}{}^{C_{ijt}}}{C_{ijt}!} \right)$ and the results are obtained through Poisson estimates with robust standard errors.

[18] To take account of the times-series correlation we should use time-varying dummy regions (Baldwin and Taglioni, 2006), but the large number of regions and years investigated makes this calculation difficult.

information on cross-sectional variations is excluded. This mean that we cannot estimate the effect for time invarying variables such as those used to represent the pairs of regions that do not involve at least one *new* region.[19] Thus, the fixed-effects models allow estimates of the European integration effects for only six pairs of regions that involve at least one country that became a new member of the EU. This might be seen as a limitation, but is not because we are interested in evaluating the effect on knowledge flows of greater integration among countries, and this effect is captured by looking at the exchange of knowledge between new EU member regions and other regions (EU members or not).[20] The pair of regions excluded by fixed-effects analysis are shaded grey in Figure 2.3.

	old	new	never
old	intra / extra	extra	extra
new	extra	intra / extra	extra
never	extra	extra	intra / extra

old: regions in countries that were EU member before 1981 (Belgium, Denmark, Germany, Greece, France, Ireland, Italy, Luxembourg, Nederland and United Kingdom)

new: regions of countries that joined the EU in the period 1981-2000 (Spain, Portugal, Austria, Finland and Sweden)

never: regions of non-EU member countries (Bulgaria, Czech Republic, Hungary, Poland, Norway and Switzerland)

intra: regions are located in the same country

extra: regions are located in two different countries

Figure 2.3. Matrix of the combinations between European regions (regions excluded by fixed-effects analysis are shaded grey)

[19] The other time invariant variables in equation [2.5], and thus not considered in the fixed-effects model, are *Border*, *Language* and *National_Border*.

[20] Also, with regard to *EU_both* and *EU_one* variables, in the pooled cross section models we estimate the effect of being part of the EU, while fixed-effects model imply that we estimate the effect of a region joining the EU because information on time invariant pairs of regions (*old_old*, *never_never* and *old_never*) are not used in the estimates.

2.5. Data

In order to analyse the diffusion of knowledge between European regions we use the patent citations and collaboration among inventors. To construct these measures of knowledge flows we use information contained in EPO patents. The patents include detailed information on the content (e.g. title, abstract, priority year, technological classification) of the invention and on applicants and inventors (name and address) that developed them. Address of inventor or applicant allows us to assign a patent to the territory where it was developed. We assign patents to regions based on the inventor's address which is more precise than address of patent application for several reasons. First, the applicant may have registered a patent using the address of a headquarters even though the patent might have been developed in a R&D laboratory located in a different region. Also, since we are interested in the diffusion of tacit and codified knowledge, it is more appropriate to consider the inventor than the applicant because tacit knowledge is embedded in the inventor.

The analysis of knowledge flows for the period 1981-2000 is performed at the level of NUTS2 regions (EUROSTAT, 2007). Our initial dataset contains data on patents with at least one inventor residing in one of the 285 regions of the EU 27 member states, and Norway and Switzerland. However, in our estimates we consider only those regions that have at least one patent in each year of the period in question because if a region has no patents then, by definition, it cannot have a regional knowledge flow. Thus, the final dataset contains patents data from 191 regions (177 regions in 19 countries in the EU27, 7 regions in Norway and 7 regions in Switzerland).[21,22] As discussed above, using inventor citations we can measure unidirectional knowledge

[21] However, the results obtained using the sample with all the regions are very similar (see the appendix Table 2.C1).
[22] See the appendix Table 2.E1 for the list of regions covered by our analysis.

flows from one region to another; inventor collaborations measure only bidirectional flows between two regions. Thus our dataset contains 729,620 observations [191 regions* 191 regions *20 years] for patent citations and 366,720 observations [(((191^2-191)/2) +191)*20] for inventor collaborations.

The geographical distance between two regions is calculated using the great circle distance method on the basis of the geographical coordinates of the centre point of the regions (Maurseth and Verspagen, 2002). In considering knowledge flows within regions, the intra-regional distance is calculated as two thirds of the radius of the regional geographic size, which is presumed to be circular in shape (Hoeckman et al., 2010).

To construct the variable related to technological proximity (*Technology*) we use the OST (2004) sectoral classification because the industrial classification used in the patent documents, i.e. the IPC, does not have a direct connection to the industrial sector for which the invention is developed. This connection is obtained through appropriate conversion tables by IPC to OST.

Finally, the variable that controls for the language (*Language*) is built on the basis of the regional official languages.

2.6. Results

In this section we present and compare the results of the estimates for the two measures of knowledge flows. Sub-section 2.6.1 presents some descriptive statistics looking at the distribution over time of both variables; sub-section 2.6.2.1 presents the results of the estimates of equation [2.2] aimed at checking whether the diffusion of knowledge among European regions is geographically localized. The cross-sectional estimates are made using aggregated data.[23] Then we split the sample into sub periods and provide separate estimates for each sub period in order to assess the evolution over time of the impact of geographical factors (2.6.2.2). Changes in the coefficients of geographical distance and national border allow us to assess whether the diffusion of knowledge is more or less circumscribed in the space than in the past. Sub-section 2.6.3 performs checks for the previous results on estimates carried out using panel data (equations [2.3] and [2.4]). Sub-section 2.6.4 estimates the impact of European integration on the diffusion of knowledge (equation [2.5]).

2.6.1 Descriptive statistics

Figure 2.4 shows the distribution over time of interregional patent citations (right side) and technological collaborations (left side) as percentages of the total. Interregional patent citations have decreased over time (from 91.4% in 1981 to 88.1% in 2000), while interregional collaborations have increased over time (from 33.5% in 1981 to 46.6% in 2000). Figure 2.5 shows the distribution over time of international patent citations and technological collaboration as percentages of the total (regional excluded). The international patent citations (right side) have decreased over time (from 67.4% in 1981 to 58.2% in 2000), while international collaborations (left side) have increased over time (from 11.9% in 1981 to 22.1% in 2000). These figures indicate two aspects of the diffusion of knowledge between regions. One the one hand, inventor collaborations, throughout the period examined, are more localized than inventor citations. On the other hand, these two measures of knowledge flows exhibit different time trends with inventor citations becoming more localized than in the past, and the reverse applying to inventor collaborations. These aspects will be confirmed in succeeding paragraphs.

[23] To determine the number of patents of a region (P_i and P_j) and the number of citations/collaborations between two regions (C_{ij}) for the period considered we add up the yearly values of these variables.

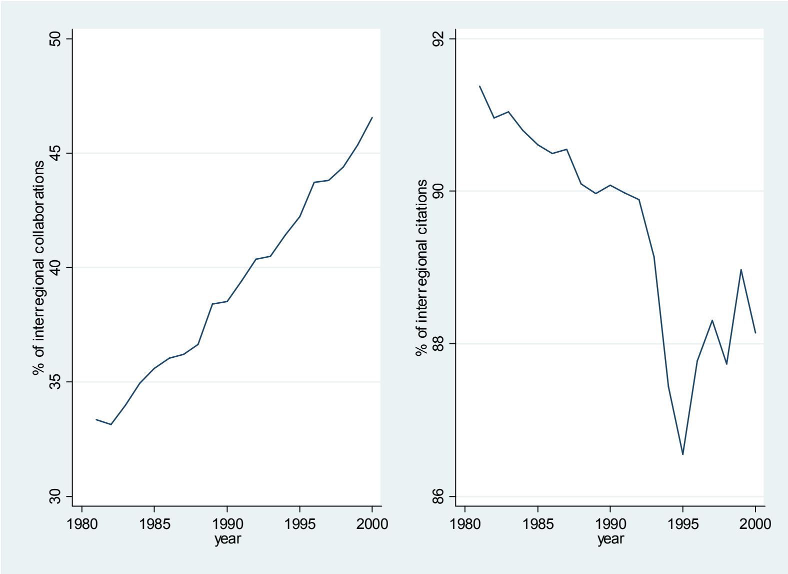


Figure 2.4. Interregional patent citations and collaborations in percentage on total

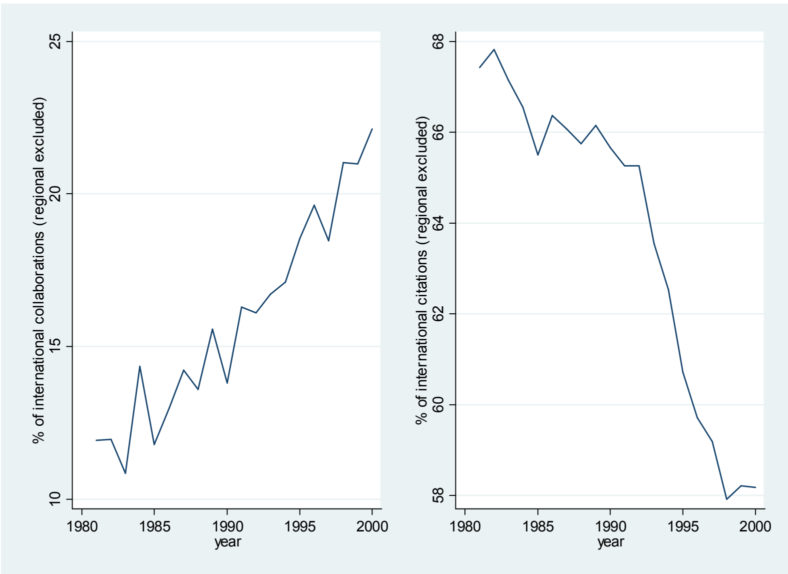


Figure 2.5. International patent citations and collaborations in percentage on total (regional excluded)

2.6.2. Cross-section estimates

2.6.2.1 Cross-section estimates for the whole period

The results of the estimates of equation [2.2] using aggregated data for the whole period analysed (1981-2000) are shown in Table 2.1. Table 2.1 presents each measure of knowledge flows in separate columns: first, in line with the extant literature (Maurseth and Verspagen, 2002; Paci and Usai, 2009), columns 1a and 1b show the results of the estimates that do not consider intraregional knowledge flows (excluding observations for which $i = j$), while, as a robustness check, columns 2a and 2b show the results for estimates that include intraregional knowledge flows. The difference between the number of observations in the first set of columns (1a, 1b) and the second set of columns (1b, 2b) is equal to the number of regions (i.e. 191). The difference in the number of observations between the patent citations (1a and 2a) and the inventor collaborations (1b and 2b) columns are due to the different characteristics of these two measures (unidirectional/bidirectional knowledge flows). The number of observations for the first column of the patent citations (1a) is 36,290 [(191*191)-190], and the number of observations for the first column of inventor collaborations is 18,145 [((191*191)-191)/2] and the number of observations for the second column of patent citations (2a) is 36,481 [((191*191)-190)+191], and the number of observations for inter-inventors collaborations is 18,336 [(((191*191)-191)/2)+191].

In Columns 1a and 1b all the coefficients are statistically significant and their signs are consistent are as expected. The distance (*dist*) effect is negative for both citations and collaborations and especially for the latter. Therefore, the diffusion of knowledge between European regions is weaker with increasing geographical distance. For the variables *Language*, *Border* and *National_Border* the coefficients are positive for both measures of knowledge flows. The diffusion of knowledge is greater if the regions share a common language and the number of citations and collaborations is higher for geographically contiguous regions. Finally, diffusion of knowledge is higher for two regions in the same country. The significance of the variable *National_Border*, irrespective of controlling for geographic proximity (*dist* and *Border*) or technological proximity (*Technology*), can be interpreted as being due to social, institutional or historical reasons which lead to greater knowledge flows within than between countries.

Another interesting result is the difference in the coefficient values for the variables for geographical (*dist* and *Border*), social and institutional (*Language* and *National_*B*order*) proximity for both measures of knowledge flows. The coefficient values of the variables *dist, Border, Language* and *National_Border* for collaboration are greater than for patent citations, meaning that technological collaboration tends to be more geographically localized than patent citation. This is consistent with inventor citations not requiring face-to-face contact. For instance, an inventor can know about the invention cited simply by reading the description contained in the patent document. To sum up, our analysis confirms the hypothesis that geographical and institutional factors are more important for inventor citations than for inventor collaborations.

For the coefficient values of the variable *Technology*, we see that technological proximity is more important for inventor citations than for inventor collaboration. This is consistent with the very many citations that are added by patent examiners, often aimed at limiting inventors' claims to novelty in a technological field. On the other hand, technological complementarities are an important incentive for inventors to collaborate. While absorptive capacity and, thus, a degree of technological proximity are necessary for effective knowledge exchange between inventors, technological complementarities and, thus, a degree of technological distance, allow inventors to learn new knowledge.

Table 2.1. Determinants of knowledge flows (aggregated data for the period 1981-2000) - PPML

Variable	*Citations* (1a)		(2a)		*Collaborations* (1b)		(2b)	
log (P_i)	0.902	***	0.880	***	0.476	***	0.488	***
	(0.017)		(0.022)		(0.093)		(0.110)	
log (P_j)	0.901	***	0.876	***	0.556	***	0.914	***
	(0.024)		(0.028)		(0.082)		(0.263)	
Technology	2.221	***	2.138	***	1.615	***	2.014	***
	(0.047)		(0.075)		(0.213)		(0.195)	
log (dist)	-0.215	***	-0.243	***	-0.939	***	-0.828	***
	(0.011)		(0.015)		(0.057)		(0.060)	
Language	0.226	***	0.225	***	0.505	***	0.398	***
	(0.020)		(0.023)		(0.084)		(0.096)	
National_Border	0.452	***	0.454	***	1.763	***	1.791	***
	(0.023)		(0.024)		(0.111)		(0.128)	
Border	0.180	***	0.152	***	0.705	***	0.733	***
	(0.025)		(0.026)		(0.084)		(0.078)	
region			0.351	***			0.374	***
			(0.068)				(0.081)	
constant	-17.288	***	-16.659	***	-3.016	***	-7.757	***
	(0.436)		(0.489)		(1.098)		(1.458)	
dummy region i	Yes		Yes		Yes		Yes	
dummy region j	Yes		Yes		Yes		Yes	
regional observations	excluded		included		excluded		included	
Log likelihood	-97906.8		-110892.01		-47263.8		-62547.15	
R-squared	0.955		0.926		0.908		0.983	
N. of regions	191		191		191		191	
N. of observations	36290		36481		18145		18336	

Note: ***, ** and * indicate significance at 1, 5 and 10 %, respectively.

As a robustness check, we estimated the above specifications including the observations for intraregional knowledge flows. We also include the variable *region*[24] to take account of the possible existence of regional barriers to knowledge flows. The results of these estimates (columns 2a and 2b) show a significant and positive effect of *region* on both measures of knowledge flows.[25] This means, that knowledge flows are

[24] This dummy variable is set equal to 1 when knowledge flows occur within a region (*i=j*).

[25] Several works underline the importance of accounting for heterogeneity in gravity model estimates (see e.g. Cheng and Wall, 2005; Baldwin and Taglioni, 2006). With cross-sectional data the use of a set of

more likely within regions and, thus, there are regional barriers that contribute to the geographically localized diffusion of knowledge. All the other variables are significant with coefficient values similar to the above results.

2.6.2.2 Cross-sections for different sub periods

The next step is analysis of the evolution over time of the coefficients of the above variables. We break the dataset into five year sub periods and perform five separate cross-sectional analyses (equation [2.2]), i.e. one for each sub period. As above, the number of observations for inventor citations is 36,290 [(191*191)-191], and the number of observations for the first column of inventor collaborations is 18,145 [((191*191)-191)/2]. Table 2.2 presents the results of these estimates. In general, the estimates confirm that geographical factors hinder the diffusion of knowledge among European regions, and the evolution over time is different for patent citations and collaborations. The distance effect increases over time (from -0.14 to -0.21) for citations, but decreases for collaborations (from -1.05 to -0.88). So, the supposed reduction in the distance effect is found only for collaborations, whereas for citations we find the opposite dynamic. At the same time, the national border effect has increased for patent citations (from 0.30 to 0.54) and decreased for technological collaboration (from 1.86 to 1.71). The coefficient of *Border* increases for patent citations (from 0.09 to 0.24), but slightly decreases for the collaborations (from 0.69 to 0.68). The language

dummies for region *i* and another set of dummies for region *j* solves most sources of bias (Baldwin and Taglioni, 2006). When we exclude the regional dummies (Reg_i and Reg_j) from our specifications, we find that the coefficient values of the variables are very different (significance and signs are the same). For instance, taking the first set of columns (1a and 1b) as a reference, the effect of physical distance (*dist*) is underestimated for both measures, i.e. from -0.215 to -0.091 for patent citations and from -0.939 to -0.683 for the inventor citations. Also, for the specifications with intraregional flows (columns 2a and 2b), we find that the regional effect (*region*) is overestimated, i.e. from 0.351 to 0.678 for patent citations and from 0.374 to 2.210 for inventor collaborations. Thus, our results confirm the importance of controlling for heterogeneity using regional fixed-effects.

effect decreases for both measures of knowledge flows, i.e. from 0.30 to 0.17 for inventor citations and from 0.69 to 0.42 for inventor collaborations, meaning the importance of a common language is less. Finally, the importance of technological proximity (*Technology*) increases for patent citations (from 2.09 to 2.28), and decreases for inventor collaborations (from 1.89 to 1.65).

Based on the above results we can say that over time interregional collaboration among European inventors is being affected less and less by geographical proximity and territorial border, while the effect for patent citations is the reverse.

To sum up, technological collaborations support the hypothesis of decreased importance of spatial proximity as a determinant of interregional knowledge flows, and citations identify what the trade literature defines as the “missing globalization puzzle” (Bhavnani et al., 2002).

Table 2.2. Determinants of interregional knowledge flows (sub-periods estimates) - PPML

	Citations				Collaborations			
Variable	1981-1985	1986-1990	1991-1995	1996-2000	1981-1985	1986-1990	1991-1995	1996-2000
log (P_i)	0.372	0.847 ***	0.544 ***	0.704 ***	0.600 ***	0.553	0.772 ***	0.610 ***
	(0.248)	(0.105)	(0.245)	(0.237)	(0.221)	(0.370)	(0.181)	(0.130)
log (P_j)	0.686 **	1.210 ***	1.299 ***	0.825 ***	1.187 ***	0.253	0.481 ***	0.717 ***
	(0.310)	(0.172)	(0.418)	(0.042)	(0.335)	(0.194)	(0.132)	(0.099)
Technology	2.092 ***	2.092 ***	2.286 ***	2.283 ***	1.889 ***	1.683 ***	1.526 ***	1.653 ***
	(0.065)	(0.064)	(0.060)	(0.055)	(0.200)	(0.226)	(0.234)	(0.196)
log (dist)	-0.139 ***	-0.184 ***	-0.256 ***	-0.211 ***	-1.056 ***	-0.971 ***	-0.994 ***	-0.883 ***
	(0.016)	(0.016)	(0.015)	(0.014)	(0.078)	(0.073)	(0.066)	(0.051)
Language	0.309 ***	0.190 ***	0.224 ***	0.179 ***	0.698 ***	0.651 ***	0.513 ***	0.424 ***
	(0.027)	(0.026)	(0.028)	(0.025)	(0.122)	(0.075)	(0.097)	(0.077)
National_Border	0.302 ***	0.443 ***	0.393 ***	0.538 ***	1.864 ***	1.819 ***	1.790 ***	1.711 ***
	(0.027)	(0.027)	(0.032)	(0.028)	(0.154)	(0.139)	(0.130)	(0.102)
Border	0.090 **	0.141 ***	0.164 ***	0.240 ***	0.692 ***	0.758 ***	0.677 ***	0.684 ***
	(0.045)	(0.035)	(0.033)	(0.029)	(0.082)	(0.075)	(0.070)	(0.061)
constant	-12.108 ***	-19.104 ***	-17.896 ***	-14.628 ***	-6.824 ***	-2.284 ***	-3.433 **	-4.658 ***
	(3.126)	(1.929)	(3.893)	(1.238)	(2.561)	(0.226)	(1.529)	(1.196)
Dummy region i	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dummy region j	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log likelihood	-37270.1	-48472.6	-60367.7	-71079.2	-9792.2	-15134.5	-19466.3	-30899.6
R-squared	0.894	0.911	0.914	0.934	0.925	0.914	0.906	0.899
N. of regions	191	191	191	191	191	191	191	191
N. of observations	36290	36290	36290	36290	18145	18145	18145	18145

Note: ***, ** and * indicate significance at 1, 5 and 10 %, respectively.

2.6.3 Panel estimates of the distance effect

To check the previous results on the dynamics of the distance effect, we make pooled cross-section estimates using a panel dataset obtained by pooling annual data. The number of observations for patent citations is 725,800 [((191*191)-191)*20], and the number of observations for inventor collaborations is 362,900 [(((191*191)-191)/2) *20].

The results of equation [2.3] confirm the trends in the cross-sectional estimates. Figure 2.6 reports coefficient values (and the confidence interval at ± 95%) for the variable *dist.*[26] For inventor citations the distance effect (negative values) increases (becomes more negative) over time, while for inventor collaborations it decreases (becomes less negative).

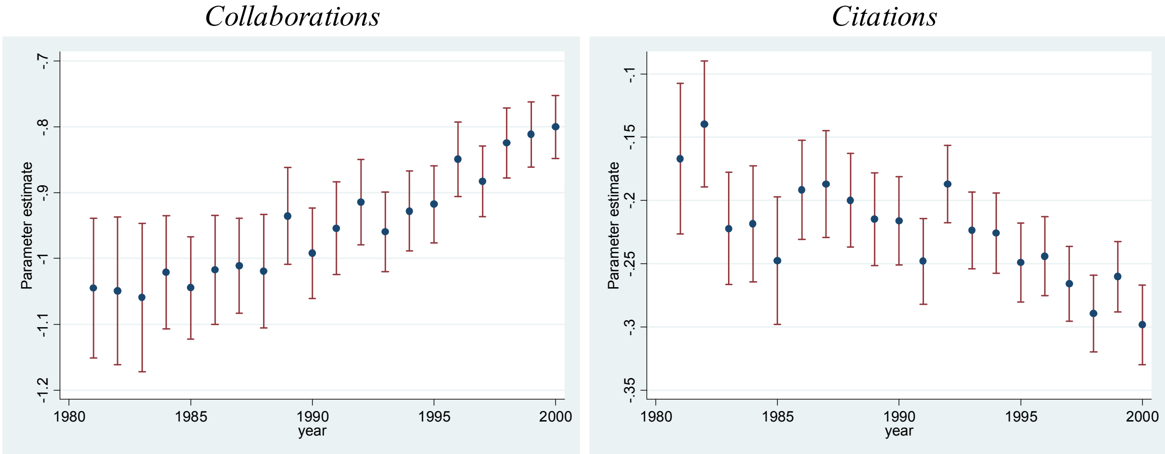
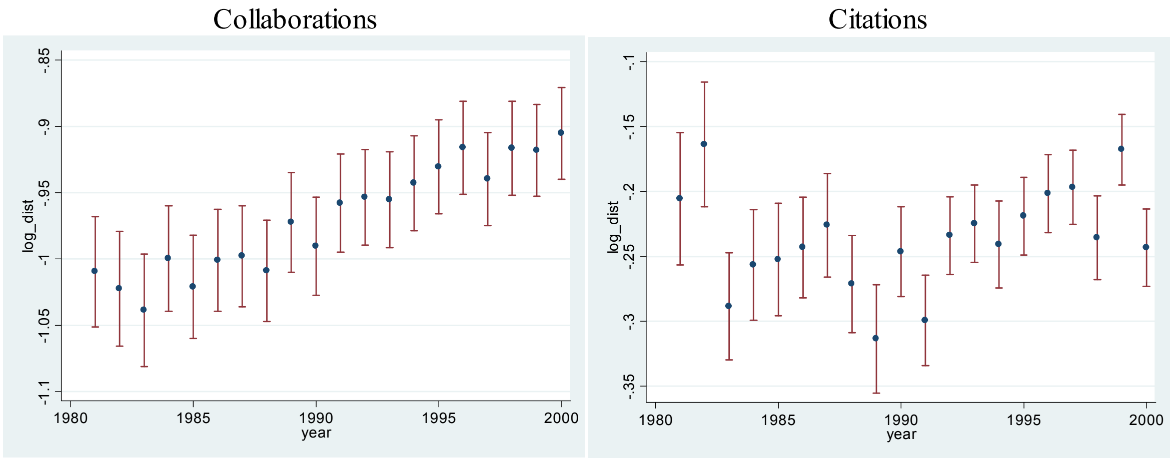


Figure 2.6. Interaction between distance and time dummies: evolution over time of the distance effect

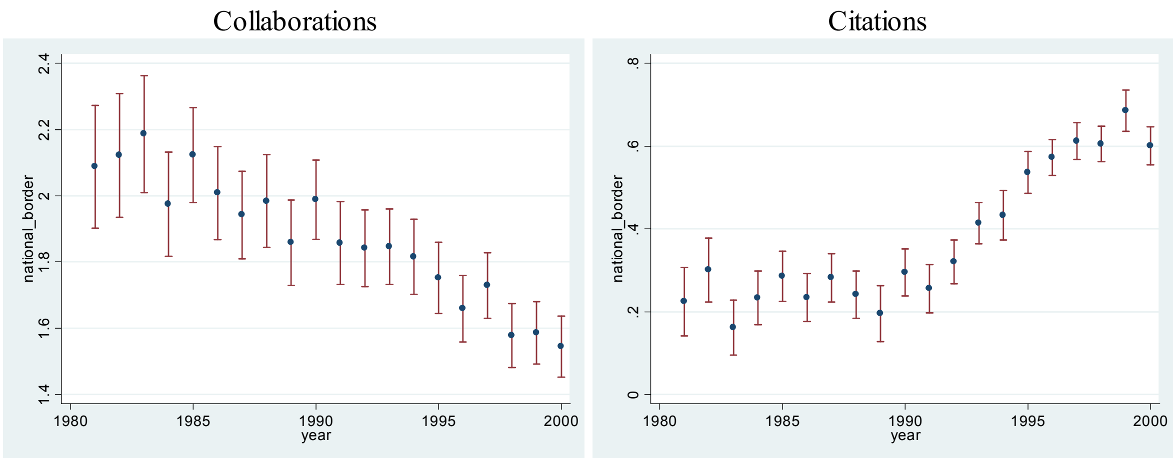
As a further check, we perform an estimate with distance and national border varying over time simultaneously (equation [2.4]). Thus, both distance and national border are interacted with time dummies. Figure 2.7 shows the results for the distance effect (graph a) and for the national border effect (graph b). We can see that the national border effect decreases for

[26] The coefficients of the other variables (not reported here) are all significant and have the same sign as the above estimates.

inventor collaborations, while the reverse happens for inventor citations. These findings are further evidence that inventor citations and inventor collaborations follow two different trends in which the former become ever more geographically localized, and the latter become less and less localized.



a) evolution over time of the distance effect



b) evolution over time of the national border effect

Figure 2.7. Distance and national border are interacted with time dummies: evolution over time of the distance and national border effect

With regard to the distance effect, the decreasing trend for the inventor collaborations is confirmed in Figure 2.7, and the trend for inventor citations follows a U-shaped curve. This means that if we do not control for the national border effect over time, the distance effect captures mainly the increased tendency for EU inventors to cite national patents. Thus, the increased localized diffusion of patent citations is due mainly to the increased home bias[27] effect.

2.6.4 Panel estimates of the impact of European integration

An argument supporting the supposed "death of distance" (Cairncross, 1997) is that reduced transport costs and diffusion of ICT have facilitated knowledge exchange among inventors residing in different regions. However, in addition to technological advances in the field of transportation and communications, the period analysed is characterized by a greater integration among European countries, which should have facilitated the international diffusion of knowledge. In particular, the European integration process has reduced the institutional barriers between countries and, thus, should have increased the knowledge flows between inventors residing in the EU countries.

Since Viner's (1950) contribution, numerous studies (see e.g. Glick and Rose, 2002; Carrère, 2006; Baier and Bergstrand, 2007) have analysed the impact of cross-country or regional trade agreements in terms of trade flows. In the present analysis, we transpose this type of analysis to the field of knowledge flows to examine the impact of the European enlargement processes on knowledge flows. Coherent with the trade literature (see section 2.3), we use a set of dummies to identify the impact of European integration in different groups of regions, i.e. regions in countries that were EU members before 1981 (*old*), regions in countries joining the EU during the period 1981-2000 (*new*) and regions that are not EU members (*never*).

Analysis of the effect of the European integration process on interregional knowledge flows is conducted using equation [2.5]. Table 2.3 presents the results of the pooled cross-

[27] Home bias is a term used in the knowledge flows literature (see e.g. Bacchiocchi and Montobbio, 2010) to indicate that inventor citations are more likely within than between countries.

section and fixed-effects estimates. For the pooled cross-sections estimates, the number of observations for the patent citations (columns 1a, 3a and 5a) is 725,820 [((191*191)-191)*20], and the number of observations for the inventor collaborations (columns 1b, 3b and 5b) is 362,900 [(((191*191)-191)/2) *20]. The differences in the numbers of observations for the pooled cross-section and fixed effects estimates are due to the fact that in the fixed effect estimates the observations for the groups of regions with zero variations over time of the dependent variables (i.e. the two measures of knowledge flows) are dropped.[28] Thus, the number of observations for the patent citations (columns 2a, 4a and 6a) is 540,920 (i.e. 725,800-184,880), while the number of observations for the inventor collaborations (columns 2b, 4b and 6b) is 161,340 (i.e. 362,900-201,560).

Knowledge flows between two different types of regions (e.g. *old* and *new* regions) are captured by two variables (*EU_both_old_new* and *EU_both_new_old*) for patent citations, and one variable (*EU_both_old_new*) for collaborations because, in that case, the diffusion of knowledge is bidirectional and, thus, we do not distinguish between spillovers receiving and spillovers generating regions. Finally, the difference in the number of variables between pooled cross-section and fixed-effects estimates is due to the fact that the latter do not allow estimation of time invarying variables.

From the pooled cross-section estimates (columns 1a and 1b) we observe that European integration increases the knowledge flows between EU regions (*EU_both*), for both inventor citations and inventor collaborations. There is also an effect exerted on third regions because EU integration reduces knowledge flows between EU regions and non-EU regions (*EU_one*).

[28] For inventor citations we have 36,290 [191*191)-191] groups of regions in the pooled cross-sections and 27,046 (36290-9244) groups of regions in the fixed-effects estimates. For inventor collaborations we have 18,145 [((191*191)-191)/2] groups of regions in the pooled cross-sections and 8,067 (18145-10078) groups of regions in the fixed-effects estimates.

The fixed-effects estimates (columns 2a and 2b) show level and significance of these variables. These results underline the need to control for region-pairs fixed effects in order to obtain unbiased estimates of the integration effect (Cheng and Wall, 2005; Carrère, 2006).[29] For the *EU_both* dummy, the coefficient is positive for both measures of knowledge flows, but significant only for patent citations. Thus, it seems that there is an EU integration effect only in the case of inventor citations. The *EU_one* dummy is insignificant for both measures, thus, there are no effects on third countries, of EU integration.

For the different groups of regions in *EU_both* (columns 4a and 4b) the picture of European integration effects is more detailed. For inventor collaborations, we find a positive and significant effect on collaboration between *old* and *new* regions (*EU_both_old_new*). Thus, European integration has increased international collaboration between EU regions but it has no effect on knowledge flows between new EU members (*EU_both_new_new_intra* and *EU_both_new_new_extra*). There are also no effects on knowledge flows between new and non-EU members (*EU_one_new_no*).

With regards to patent citations, we observe a positive and significant effect between *old* and *new* regions in relation to *old* regions citing the patents of *new* regions, but a negative and insignificant effect for *new* regions citing the patents of *old* regions. Thus, EU integration increases international knowledge flows only from new EU members to old EU members. There is a positive and significant effect on international knowledge flows between *new* regions (*EU_both_new_new_extra*) and a negative and significant effect on national knowledge flows between *new* regions (*EU_both_new_new_intra*). Thus, EU integration increases international knowledge flows while decreasing national flows between new EU

[29] To take account of the time-series correlation we would use time-varying dummy regions (Baldwin and Taglioni, 2006), but the large number of regions and years investigated makes it difficult to make this calculation.

members. Finally, the EU integration has no effects on knowledge flows between new and not EU members (*EU_one_new_no* and *EU_one_no_new*).

In a further step, we use disaggregated data to evaluate the effect of single EU enlargement (columns 6a and 6b). For group involving regions joining the EU we split them into two periods: for the year 1986 when the first EU enlargement occurred with the annexations of Spain and Portugal; and for the year 1995 when the second EU enlargement occurred with the annexations of Austria, Finland and Sweden. For inventor collaborations, there is a positive and significant effect confirmed between *old* and *new* regions with each EU enlargement. For patent citations, the aggregate effects of European integration are based only on the second EU enlargement (with the exception of *EU_both_new_new_extra_86).*[30]

To sum up, European integration has had a significant effect on reducing the national barriers to knowledge flows between new and old EU members. However, for patent citations, this effect relates only to the second EU enlargement and is confined to knowledge flows from new to old EU members.

[30] The high values of the coefficients of the variables for knowledge flows between *new* regions in the first EU enlargement (e.g. *EU_both_new_new_extra_86*) are due to the initial low levels of collaborations/patent citations before 1986 and the relatively high increase after 1986.

Table 2.3. European Integration - Poisson regressions (robust standard errors)

	Citations						*Collaborations*					
Variable	(1a)	(2a)	(3a)	(4a)	(5a)	(6a)	(1b)	(2b)	(3b)	(4b)	(5b)	(6b)
EU_both	0.148 ***	0.085 ***					0.071 **	0.058				
	(0.014)	(0.019)					(0.029)	(0.039)				
EU_one	-0.121 ***	0.033					-0.236 ***	0.040				
	(0.016)	(0.039)					(0.044)	(0.090)				
EU_both_old_new			0.005	-0.014					0.263 ***	0.266 ***		
			(0.012)	(0.022)					(0.040)	(0.047)		
EU_both_new_old			0.259 ***	0.220 ***								
			(0.019)	(0.026)								
EU_both_old_old_intra			-0.273 ***						0.868 ***			
			(0.053)						(0.110)			
EU_both_old_old_extra			0.135 ***						0.358 ***			
			(0.032)						(0.075)			
EU_both_new_new_intra			-0.269 ***	-0.200 ***					-0.024	-0.099		
			(0.065)	(0.069)					(0.041)	(0.067)		
EU_both_new_new_extra			0.407 ***	0.193 ***					-0.424 ***	-0.087		
			(0.046)	(0.039)					(0.142)	(0.152)		
EU_one_old_no			-0.011						-0.317 ***			
			(0.037)						(0.053)			
EU_one_no_old			0.070 *									
			(0.037)									
EU_one_new_no			-0.021	-0.006					-0.226 ***	0.048		
			(0.047)	(0.059)					(0.074)	(0.090)		
EU_one_no_new			-0.022	0.078								
			(0.049)	(0.051)								
EU_both_old_old_intra					-0.263 ***						0.953 ***	
					(0.054)						(0.117)	

(continued)

Table 2.3. (continued)

Variable	Citations (1a)	(2a)	(3a)	(4a)	(5a)	(6a)	Collaborations (1b)	(2b)	(3b)	(4b)	(5b)	(6b)
EU_both_old_old_extra					0.142 ***						0.434 ***	
					(0.034)						(0.080)	
EU_one_old_no					-0.010						-0.323 ***	
					(0.037)						(0.055)	
EU_one_no_old					0.074 **							
					(0.037)							
EU_both_old_new_86					0.146 **	0.115					0.880 ***	0.551 ***
					(0.059)	(0.086)					(0.146)	(0.214)
EU_both_new_old_86					0.135 *	-0.101						
					(0.069)	(0.090)						
EU_both_new_new_intra_86					0.401 ***	1.073					0.712 **	0.989
					(0.155)	(1.029)					(0.289)	(0.703)
EU_both_new_new_extra_86					1.688 ***	11.765 ***					1.032	10.732 ***
					(0.453)	(0.708)					(0.677)	(0.744)
EU_one_new_no_86					-0.248 **	-0.320					-1.051 ***	-0.718
					(0.111)	(0.268)					(0.219)	(0.440)
EU_one_no_new_86					-0.336 ***	-0.166						
					(0.095)	(0.265)						
EU_both_old_new_95					-0.005	-0.017					0.236 ***	0.260 ***
					(0.018)	(0.023)					(0.042)	(0.048)
EU_both_new_old_95					0.268 ***	0.224 ***						
					(0.019)	(0.026)						
EU_both_new_new_intra_95					-0.294 ***	-0.202 ***					-0.031	-0.100
					(0.066)	(0.069)					(0.041)	(0.067)
EU_both_new_new_extra_95					0.399 ***	0.193 ***					-0.467 ***	-0.088
					(0.047)	(0.039)					(0.145)	(0.153)

(continued)

Table 2.3. (continued)

	Citations						*Collaborations*					
Variable	(1a)	(2a)	(3a)	(4a)	(5a)	(6a)	(1b)	(2b)	(3b)	(4b)	(5b)	(6b)
EU_one_new_no_95					-0.004 (0.049)	-0.000 (0.059)					-0.102 (0.075)	0.058 (0.091)
EU_one_no_new_95					0.057 (0.052)	0.084 (0.052)						
distance * time dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
P_i and P_j	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Technology	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
National_Border	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Border	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Language	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Model	Pooled	FE	Pooled	FE	Pooled	FE	Pooled	FE	Pooled	FE	Pooled	FE
Number of groups	36290	27046	36290	27046	36290	27046	18145	8067	18145	8067	18145	8067
Number of regions	191	191	191	191	191	191	191	191	191	191	191	191
Number of observations	725800	540920	725800	540920	725800	540920	362900	161340	362900	161340	362900	161340

Note: ***, ** and * indicate significance at 1, 5 and 10 %, respectively.

2.7 Conclusion

This work analysed the evolution over time of the impact of geographical and institutional barriers to the diffusion of knowledge among European regions based on patent citations and technological collaborations. The results show that knowledge flows are geographically localized for both measures and that the impacts of geographical and institutional factors are higher for inventor collaborations than for inventor citations. We showed also that, although national borders are still important barriers to the diffusion of knowledge, their impacts on the two measures of knowledge flows follow different time trends. In particular, the national border effect decreased for technological collaborations and increased for patent citations. On the one hand, inventors tend to collaborate more with other international inventors, but on the other hand, the tendency to cite national inventors increases. The evolution over time of the distance effect, which decreases only for inventor collaborations, confirms that inventor collaborations are becoming less localized, while the reverse is true for inventor citations.

We also analysed whether European integration has an impact on reducing the national barriers to knowledge flows. An important result of our estimates is that European integration favours international collaborations between new EU members and existing EU members. For patent citations, it seems that European integration positively affects the diffusion of knowledge only in the case of the second EU enlargement and only for knowledge generated in new member regions, and spills over to old EU members.

Our analysis shows that knowledge flows between European regions are hindered by geographical and institutional barriers, which led to the implementation, at the end of the last century, of a ERA (Commission of the European Communities, 2001). However, the objectives of this policy are still far from being achieved.

The next step in this research would be detailed analysis of the reasons underlying the different evolution over time of patent citations and technological collaborations and the impact of national borders in order to understand why, despite technological advances and increased country integrations, patent citations are becoming more localized. This line of research could be enriched by disaggregated sectoral level analysis to investigate sectoral trends. Other research could investigate in more detail the EU integration processes and their effects on knowledge flows. The Union goes back to the 1950s and, over time, has been affected by various changes such as creating a free trade area or a single market, where, however, the process of enlargement extending to other countries is a crucial step towards greater integration.

Appendix (Chapter 2)

Appendix 2.A. The construction of the dataset used in our work

The initial dataset was obtained by extrapolating from the KITES dataset the information on patents (EP number, priority year), inventors (name and address), applicants (name and address) and citations. The dataset contains data on patents registered in various patent offices, national or supranational, but for our analysis we consider only the patents registered at the EPO. Patents are assigned to European NUTS 2 regions (Eurostat, 2007) based inventor's address. Patents with more than one inventor are assigned to regions based on the addresses of all inventor addresses. For instance, a patent with two inventors from two different regions is assigned to both regions.

In most cases the connection between inventor's address and NUTS 2 region, is contained in the KITES dataset. However, in order to avoid bias in our estimates, we reduced the number of inventors without a NUTS 2 region assigned. To do this, we merged the KITES and the OECD REGPAT datasets based on EPO pubblication number and inventor name in order to obtain full correspondence between datasets. This allowed us to identify the NUTS 2 region for some of the inventors. For the inventors without a NUTS 2 designation we manually assigned the NUTS 2 region based on inventor's place of residence and postcode. The percentage of inventors without a NUTS 2 is approximately 0.8% of the total and is the same across time periods and countries.[31]

The procedures adopted for the construction of our dependent variables are described in section 2.3.

[31] Although we do not know the inventor's NUTS 2, we know country of residence and details of the patent developed.

Appendix 2.B. Definition, source and descriptive statistics of the variables

Table 2.B1 reports the definitions and sources of the variables used in the analysis.

Table 2.B1. Definition and sources of the variables

Variable	*Definition*	*Source*
Citations *ijt*	Number of patents (at time *t*) of inventors residing in region *i* cited by patents (at time *t-(t+4)*) with at least an inventor residing in region *j*. Patents with more than one inventor residing in the same region (*i* or *j*), citations are counted only once.	KITES/OECD REGPAT
Collaborations *ijt*	Number of patents with at least an inventor residing in region *j* and an inventor residing in region *i*. Patents with more than one inventor residing in the same region (*i* or *j*), collaborations are counted only once.	KITES/OECD REGPAT
P *it*	Number of patents with at least an inventor residing in country *i*.	KITES/OECD REGPAT
P *jt*	Number of patents with at least one inventor residing in country *j*. For patent citations we consider a temporal window of four years (*t-(t+4)*).	KITES/OECD REGPAT
Technology *ijt*	Jaffe (1986) index based on 30 technology classes (OST, 2004). It is an indicator of the technological proximity between region *i* and region *j*.	KITES
National_border *ij*	Time invarying dummy equal to 1 if the two regions are located in the same country.	KITES/OECD REGPAT
dist *ij*	Geographical distance between two regions, calculated using the great circle distance method on the basis of the geographical coordinates of the centre point of the regions. Intra-regional distance is calculated as two thirds of the radius of the regional geographic size.	EUROSTAT/GISCO
Border *ij*	Time invarying dummy equal to 1 if the two regions are neighbours.	Authors' elaborations
Language *ij*	Time invarying dummy equal to 1 if the two regions have the same official language.	Authors' elaborations
region *ij*	Time invarying dummy equal to 1 for intra-regional knowledge flows (*i=j*).	KITES/OECD REGPAT

(continued)

Table 2.B1. (continued)

Variable	*Definition*	*Source*
EU_both *ijt*	Time varying dummy equal to 1 if the two regions are from EU member states.	Authors' elaborations
EU_one *ijt*	Time varying dummy equal to 1 if only one region is from EU member states.	Authors' elaborations
EU_both_old_new *ijt*	Time varying dummy equal to 1 if knowledge flows from an *old* region (i.e. region of countries that were EU member before 1981) to a *new* region (i.e. region of EU entering states).	Authors' elaborations
EU_both_new_old *ijt*	Time varying dummy equal to 1 if knowledge flows from a *new* region to an *old* region.	Authors' elaborations
EU_both_old_old_intra *ij*	Time invarying dummy equal to 1 if knowledge flows within countries that were EU member before 1981.	Authors' elaborations
EU_both_old_old_extra *ij*	Time invarying dummy equal to 1 if knowledge flows between countries that were EU member before 1981.	Authors' elaborations
EU_one_old_no *ij*	Time invarying dummy equal to 1 if knowledge flows from an *old* region to a non–EU member country.	Authors' elaborations
EU_one_no_old *ij*	Time invarying dummy equal to 1 if knowledge flows from a *never* region (i.e. region of non-EU member countries) to an *old* region.	Authors' elaborations
EU_both_new_new_intra *ijt*	Time varying dummy equal to 1 if knowledge flows within EU new member states.	Authors' elaborations
EU_both_new_new_extra *ijt*	Time varying dummy equal to 1 if knowledge flows between EU new member states.	Authors' elaborations
EU_one_new_no *ijt*	Time varying dummy equal to 1 if knowledge flows from a *new* region to a *never* region.	Authors' elaborations
EU_one_no_new *ijt*	Time varying dummy equal to 1 if knowledge flows from a *never* region to a *new* region.	Authors' elaborations
EU_both_old_new_86 *ijt*	Time varying dummy. The variable is "EU_both_old_new", but *new* regions include only regions of countries that joined the EU in 1986 (Spain and Portugal).	Authors' elaborations
EU_both_new_old_86 *ijt*	Time varying dummy. The variable is "EU_both_new_old", but *new* regions include only regions of countries that joined the EU in 1986 (Spain and Portugal).	Authors' elaborations

(continued)

Table 2.B1. (continued)

Variable	*Definition*	*Source*
EU_both_old_new_95 *ijt*	Time varying dummy. The variable is "EU_both_old_new", but *new* regions include only regions of countries that joined the EU in 1995 (Austria, Finland and Sweden).	Authors' elaborations
EU_both_new_old_95 *ijt*	Time varying dummy. The variable is "EU_both_new_old", but *new* regions include only regions of countries that joined the EU in 1995 (Austria, Finland and Sweden).	Authors' elaborations
EU_both_new_new_intra_86 *ijt*	Time varying dummy. The variable is "EU_both_new_new_intra", but *new* regions include only regions of countries that entered the EU in 1986 (Spain and Portugal).	Authors' elaborations
EU_both_new_new_intra_95 *ijt*	Time varying dummy. The variable is "EU_both_new_new_intra", but *new* regions include only regions of countries that joined the EU in the 1995 (Austria, Finland and Sweden).	Authors' elaborations
EU_both_new_new_extra_86 *ijt*	Time varying dummy. The variable is "EU_both_new_new_extra", but *new* regions include only regions of countries that entered the EU in 1986 (Spain and Portugal).	Authors' elaborations
EU_both_new_new_extra_95 *ijt*	Time varying dummy. The variable is "EU_both_new_new_extra", but *new* regions include only regions of countries that joined the EU in 1995 (Austria, Finland and Sweden).	Authors' elaborations
EU_one_new_no_86 *ijt*	Time varying dummy. The variable is "EU_both_new_no", but *new* regions include only regions of states that joined the EU in 1986 (Spain and Portugal).	Authors' elaborations
EU_one_new_no_95 *ijt*	Time varying dummy. The variable is "EU_both_new_no", but *new* regions include only regions of states that joined the EU in 1995 (Austria, Finland and Sweden).	Authors' elaborations
EU_one_no_new_86 *ijt*	Time varying dummy. The variable is "EU_both_no_new", but *new* regions include only regions of states that joined the EU in 1986 (Spain and Portugal).	Authors' elaborations
EU_one_no_new_95 *ijt*	Time varying dummy. The variable is "EU_both_no_new", but *new* regions include only regions of countries that joined the EU in 1995 (Austria, Finland and Sweden).	Authors' elaborations

In the Table 2.B2 are shown the descriptive statistics of the variables used in our analysis (regional observations are excluded).

Table 2.B2. Descriptive statistics (period 1981-2000)

Variable	Citations					Collaborations				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
C_{ij}	725800	0.82	3.75	0	327	362900	0.57	5.82	0	593
P_i	725800	210.18	351.99	1	3742	362900	246.83	401.44	1	3742
P_j	725800	1190.01	1939.71	10	18784	362900	173.53	289.76	1	3742
$\log(P_i)$	725800	4.47	1.41	0	8.23	362900	4.59	1.45	0	8.23
$\log(P_j)$	725800	6.25	1.35	2.30	9.84	362900	4.35	1.35	0	8.23
Technology	725800	0.55	0.18	0	0.99	362900	0.5	0.19	0	1
National_Border	725800	0.11	0.31	0	1	362900	0.11	0.31	0	1
dist	725800	947.27	566.52	6.56	3775.18	362900	947.27	566.52	6.56	3775.18
log(dist)	725800	6.64	0.71	1.88	8.24	362900	6.64	0.71	1.88	8.24
Border	725800	0.02	0.15	0	1	362900	0.02	0.15	0	1
Language	725800	0.18	0.38	0	1	362900	0.18	0.38	0	1
EU_both	725800	0.64	0.48	0	1	362900	0.64	0.48	0	1
EU_one	725800	0.19	0.39	0	1	362900	0.19	0.39	0	1
EU_both_old_new	725800	0.04	0.20	0	1	362900	0.09	0.28	0	1
EU_both_new_old	725800	0.04	0.20	0	1					
EU_both_old_old_intra	725800	0.10	0.30	0	1	362900	0.10	0.30	0	1
EU_both_old_old_extra	725800	0.45	0.50	0	1	362900	0.45	0.50	0	1
EU_one_old_no	725800	0.09	0.28	0	1	362900	0.17	0.38	0	1
EU_one_no_old	725800	0.09	0.28	0	1					
EU_both_new_new_intra	725800	0.00	0.04	0	1	362900	0.00	0.04	0	1

(continued)

Table 2.B2. (continued)

Variable	Citations					Collaborations				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
EU_both_new_new_extra	725800	0.00	0.05	0	1	362900	0.00	0.05	0	1
EU_one_new_no	725800	0.01	0.08	0	1	362900	0.01	0.11	0	1
EU_one_no_new	725800	0.01	0.08	0	1					
EU_both_old_new_86	725800	0.02	0.13	0	1	362900	0.04	0.18	0	1
EU_both_new_old_86	725800	0.02	0.13	0	1					
EU_both_old_new_95	725800	0.03	0.16	0	1	362900	0.05	0.22	0	1
EU_both_new_old_95	725800	0.03	0.16	0	1					
EU_both_new_new_intra_86	725800	0.00	0.02	0	1	362900	0.00	0.02	0	1
EU_both_new_new_intra_95	725800	0.00	0.03	0	1	362900	0.00	0.03	0	1
EU_both_new_new_extra_86	725800	0.00	0.01	0	1	362900	0.00	0.01	0	1
EU_both_new_new_extra_95	725800	0.00	0.05	0	1	362900	0.00	0.05	0	1
EU_one_new_no_86	725800	0.00	0.05	0	1	362900	0.01	0.07	0	1
EU_one_new_no_95	725800	0.00	0.06	0	1	362900	0.01	0.09	0	1
EU_one_no_new_86	725800	0.00	0.05	0	1					
EU_one_no_new_95	725800	0.00	0.06	0	1					

Appendix 2.C. Cross-section estimates for the whole sample and the restricted sample

Table 2.C1 compares the results of the estimates of equation [2.2] obtained for the whole[32] sample of 281 regions and the restricted sample of 191 regions (aggregate data).

Table 2.C1. Citations and collaborations (restricted and whole sample) - PPML

	Citations				Collaborations			
Variable	restricted		whole		restricted		whole	
log (P_i)	0.880	***	1.581	***	0.488	***	1.506	***
	(0.022)		(0.257)		(0.110)		(0.352)	
log (P_j)	0.876	***	1.435	***	0.914	***	0.266	***
	(0.028)		(0.251)		(0.263)		(0.133)	
Technology	2.138	***	2.146	***	2.014	***	2.033	***
	(0.075)		(0.074)		(0.195)		(0.191)	
log (dist)	-0.243	***	-0.246	***	-0.828	***	-0.833	***
	(0.015)		(0.015)		(0.060)		(0.058)	
Language	0.225	***	0.228	***	0.398	***	0.431	***
	(0.023)		(0.023)		(0.096)		(0.096)	
National_Border	0.454	***	0.453	***	1.791	***	1.802	***
	(0.024)		(0.024)		(0.128)		(0.127)	
Border	0.152	***	0.152	***	0.733	***	0.724	***
	(0.026)		(0.026)		(0.078)		(0.076)	
region	0.351	***	0.347	***	0.374	***	0.367	***
	(0.068)		(0.068)		(0.081)		(0.079)	
constant	-16.659	***	-26.467	***	-7.757	***	-9.811	***
	(0.489)		(3.017)		(1.458)		(1.958)	
dummy region i	Yes		Yes		Yes		Yes	
dummy region j	Yes		Yes		Yes		Yes	
regional observations	included		included		included		included	
Log likelihood	-110892.0		-124317.9		-62547.1		-67,821.0	
R-squared	0.926		0.927		0.983		0.983	
N. of regions	191		281		191		281	
N. of observations	36481		78961		18336		39621	

Note: ***, ** and * indicate significance at 1, 5 and 10 %, respectively.

[32] We consider all regions with at least 1 EPO patent during the period analysed, which leaves 4 regions (285-281) without an EPO patent.

Appendix 2.D. Pooled cross-section estimates for different time lags

For patent citations, Figures 2.D1 and 2.D2 show the results for distance and national border effects[33] obtained from three different pooled cross-section estimates for equation [2.3] (Figure 2.D1) and three different pooled cross-section estimates for equation [2.4] (Figure 2.D2). To these estimations we created three different samples on the basis of the temporal lag between the priority years of the cited and citing patents. The temporal lags used are: 0-2 years (lag_0_2); 3-5 years (lag_3_5); 6-9 years (lag_6_9).

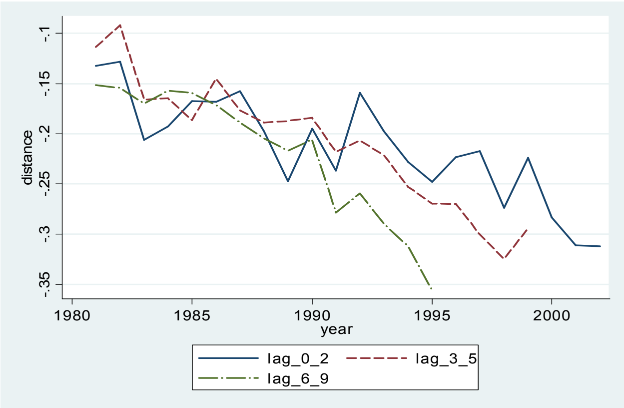
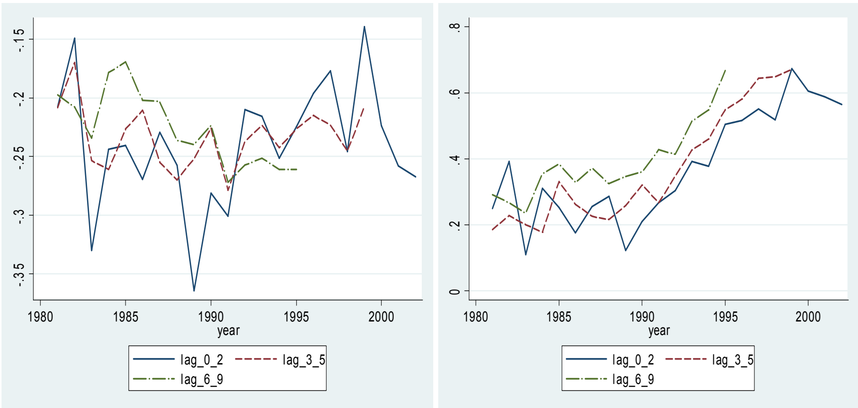


Figure 2.D1. Citations and lag: evolution over time of the distance effect



a) distance effect b) national border effect

Figure 2.D2. Citations and lag: evolution over time of the distance and national border effect

[33] All the coefficients for the distance and national border effect are significant at 1%. The results for the other variables, not shown here, are very similar to those obtained with the dataset with a time lag of 4 years.

Appendix 2.E. List of the European regions

Table 2.E1. List of Nuts 2 regions (Eurostat) used in our estimates

Country	Nuts 2	Name	Country	Nuts 2	Name
Austria	AT11	Burgenland (A)	Greece	GR30	Attiki
	AT12	Niederösterreich	Hungary	HU10	Közép-Magyarország
	AT13	Wien		HU21	Közép-Dunántúl
	AT21	Kärnten		HU23	Dél-Dunántúl
	AT22	Steiermark		HU32	Észak-Alföld
	AT31	Oberösterreich	Ireland	IE01	Border, Midland and Western
	AT32	Salzburg		IE02	Southern and Eastern
	AT33	Tirol	Italy	ITC1	Piemonte
	AT34	Vorarlberg		ITC3	Liguria
Belgium	BE10	Région de Bruxelles-Capitale		ITC4	Lombardia
	BE21	Prov. Antwerpen		ITD1	Provincia Autonoma Bolzano
	BE22	Prov. Limburg (B)		ITD3	Veneto
	BE23	Prov. Oost-Vlaanderen		ITD4	Friuli-Venezia Giulia
	BE24	Prov. Vlaams-Brabant		ITD5	Emilia-Romagna
	BE25	Prov. West-Vlaanderen		ITE1	Toscana
	BE31	Prov. Brabant Wallon		ITE2	Umbria
	BE32	Prov. Hainaut		ITE3	Marche
	BE33	Prov. Liège		ITE4	Lazio
	BE34	Prov. Luxembourg (B)		ITF1	Abruzzo
	BE35	Prov. Namur		ITF3	Campania
Bulgaria	BG41	Yugozapaden		ITF4	Puglia
Czech Republic	CZ01	Praha		ITG1	Sicilia
Germany	DE11	Stuttgart		ITG2	Sardegna
	DE12	Karlsruhe	Luxembourg	LU00	Luxembourg (Grand-Duché)

(continued)

Table 2.E1 (continued)

Country	Nuts 2	Name	Country	Nuts 2	Name
Germany	DE13	Freiburg	Netherlands	NL11	Groningen
	DE14	Tübingen		NL12	Friesland (NL)
	DE21	Oberbayern		NL13	Drenthe
	DE22	Niederbayern		NL21	Overijssel
	DE23	Oberpfalz		NL22	Gelderland
	DE24	Oberfranken		NL31	Utrecht
	DE25	Mittelfranken		NL32	Noord-Holland
	DE26	Unterfranken		NL33	Zuid-Holland
	DE27	Schwaben		NL34	Zeeland
	DE30	Berlin		NL41	Noord-Brabant
	DE42	Brandenburg - Südwest		NL42	Limburg (NL)
	DE50	Bremen	Poland	PL12	Mazowieckie
	DE60	Hamburg		PL22	Slaskie
	DE71	Darmstadt	Portugal	PT17	Lisboa
	DE72	Gießen	Sweden	SE11	Stockholm
	DE73	Kassel		SE12	Östra Mellansverige
	DE80	Mecklenburg-Vorpommern		SE21	Småland med öarna
	DE91	Braunschweig		SE22	Sydsverige
	DE92	Hannover		SE23	Västsverige
	DE93	Lüneburg		SE31	Norra Mellansverige
	DE94	Weser-Ems		SE32	Mellersta Norrland
	DEA1	Düsseldorf		SE33	Övre Norrland
	DEA2	Köln	United Kingdom	UKC1	Tees Valley and Durham
	DEA3	Münster		UKC2	Northumberland and Tyne and Wear
	DEA4	Detmold		UKD1	Cumbria
	DEA5	Arnsberg		UKD2	Cheshire

(continued)

Table 2.E1 (continued)

Country	Nuts 2	Name	Country	Nuts 2	Name
Germany	DEB1	Koblenz	United Kingdom	UKD3	Greater Manchester
	DEB2	Trier		UKD4	Lancashire
	DEB3	Rheinhessen-Pfalz		UKD5	Merseyside
	DEC0	Saarland		UKE1	East Yorkshire and Northern Lincolnshire
	DED1	Chemnitz		UKE2	North Yorkshire
	DED2	Dresden		UKE3	South Yorkshire
	DED3	Leipzig		UKE4	West Yorkshire
	DEE0	Sachsen-Anhalt		UKF1	Derbyshire and Nottinghamshire
	DEF0	Schleswig-Holstein		UKF2	Leicestershire, Rutland and Northamptonshire
	DEG0	Thüringen		UKF3	Lincolnshire
Denmark	DK01	Hovedstaden		UKG1	Herefordshire, Worcestershire and Warwickshire
	DK02	Sjælland		UKG2	Shropshire and Staffordshire
	DK03	Syddanmark		UKG3	West Midlands
	DK04	Midtjylland		UKH1	East Anglia
	DK05	Nordjylland		UKH2	Bedfordshire and Hertfordshire
Spain	ES21	País Vasco		UKH3	Essex
	ES30	Comunidad de Madrid		UKI1	Inner London
	ES51	Cataluña		UKI2	Outer London
	ES52	Comunidad Valenciana		UKJ1	Berkshire, Buckinghamshire and Oxfordshire
	ES61	Andalucía		UKJ2	Surrey, East and West Sussex
Finland	FI13	Itä-Suomi		UKJ3	Hampshire and Isle of Wight
	FI18	Etelä-Suomi		UKJ4	Kent
	FI19	Länsi-Suomi		UKK1	Gloucestershire, Wiltshire and Bristol/Bath area
	FI1A	Pohjois-Suomi		UKK2	Dorset and Somerset
France	FR10	Île de France		UKK3	Cornwall and Isles of Scilly
	FR21	Champagne-Ardenne		UKK4	Devon

(continued)

Table 2.E1 (continued)

Country	Nuts 2	Name	Country	Nuts 2	Name
France	FR22	Picardie	United Kingdom	UKL1	West Wales and The Valleys
	FR23	Haute-Normandie		UKL2	East Wales
	FR24	Centre		UKM2	Eastern Scotland
	FR25	Basse-Normandie		UKM3	South Western Scotland
	FR26	Bourgogne		UKM5	North Eastern Scotland
	FR30	Nord - Pas-de-Calais		UKN0	Northern Ireland
	FR41	Lorraine	Switzerland	CH01	Lake Geneva region
	FR42	Alsace		CH02	Espace Mittelland
	FR43	Franche-Comté		CH03	Northwestern Switzerland
	FR51	Pays de la Loire		CH04	Zurich
	FR52	Bretagne		CH05	Eastern Switzerland
	FR53	Poitou-Charentes		CH06	Central Switzerland
	FR61	Aquitaine		CH07	Ticino
	FR62	Midi-Pyrénées	Norway	NO01	Oslo og Akershus
	FR63	Limousin		NO02	Hedmark og Oppland
	FR71	Rhône-Alpes		NO03	Sør-Østlandet
	FR72	Auvergne		NO04	Agder og Rogaland
	FR81	Languedoc-Roussillon		NO05	Vestlandet
	FR82	Provence-Alpes-Côte d'Azur		NO06	Trøndelag
				NO07	Nord-Norge

Chapter 3

Inventor Mobility and Regions' Innovation Potential[34]

3.1 Introduction

Knowledge is an important engine of economic growth. In contrast to standard neoclassical theories (Solow, 1956, 1957), the technology gap approach (Fagerberg, 1987, 1994) assumes that the rate of economic growth of a region is determined by the rate of growth of its knowledge capital. Regional knowledge capital is generated mainly by firm innovation or scientific activities by universities or other research institutions. Knowledge, unlike traditional goods (i.e. labour and capital), is a non-rivalrous and partially excludable public good, and firms and regions cannot prevent parts of that knowledge from spreading to other firms and regions. It follows that knowledge capital is determined also by the knowledge diffusion processes. Knowledge flows are a key element in explaining the economic growth and the catching-up processes of regions.

Although theoretical and empirical work on economic growth is mainly at country level, analysis at this geographical level ignores economic and technological differences within countries (Fagerberg and Verspagen, 1996; Oughton et al., 2002; Sterlacchini, 2008). Italy shows some remarkable intra-national disparities, with significant economic differences between the northern and central regions and the less developed regions of the country's south. For instance, in the sample used in the current analysis, in 2000, GDP per capita of the richest region (Lombardia) was more than double that of the

[34] This chapter represents a slightly different version of the paper: Cappelli, R., Czarnitzki, D., Doherr, T., Montobbio, F., 2012. Inventor mobility and regions' innovation potential, mimeo.

poorest region (Calabria) and the ratio of R&D expenditure per capita of the region with the highest value (Lazio) and the region with the lowest value (Calabria) was about 12:1.

In this work we analyse the impact of knowledge capital on economic growth in the Italian regions for the period 1995-2007. In line with the technology gap approach, the knowledge capital of a region is a function of both the knowledge generating and knowledge diffusion processes.

Knowledge capital is an important key to understanding regional differences in economic growth rates. Since knowledge is an immaterial good, a fundamental question related to analysis of the relationship between knowledge and economic growth is how to measure regional knowledge capital. The R&D activities of firms and research institutions, in different ways and with different aims, contribute to increasing regional knowledge capital, so it follows that it can be measured using data such as R&D expenditure and number of patents (Fagerberg et al. 1997; Badinger and Tondl, 2003; Sterlacchini, 2008). However, knowledge is a partially excludable good and, thus, the knowledge capital of a region may stem also from the voluntary or involuntary diffusion of the outcomes of R&D activities. In order to take account of knowledge diffusion processes, some scholars use measures of knowledge flows, which traditionally are modelled by including the unweighted or weighted sum of "foreign" R&D activities (for a survey see Hall et al. 2010).

Traditional measures of knowledge flows, i.e. that make use of "foreign" R&D stock, have some shortcomings. First, they are an indirect measure only of the diffusion of knowledge and may simply indicate potential diffusion of knowledge between regions. Second, they do not consider intra-regional knowledge flows. There is a large literature showing that the diffusion of knowledge is geographically localized and that most knowledge flows are within regions (see e.g. Peri, 2005). It follows that "foreign" stock of R&D does not consider an important part of a region's knowledge flows and knowledge capital.

To overcome these limitations, we use patent citations (i.e. backward patent citations)[35] to measure knowledge flows. Patent citations leave a paper trail of the knowledge flows between inventors and between regions. The appropriateness of the patent citations measures is supported by the fact that patent citations are used widely in the literature that explicitly analyses the geographical diffusion of knowledge (Jaffe et al., 1993; Maurseth and Verspagen, 2002; Bottazzi and Peri, 2003; Paci and Usai, 2009).

The measures of stock of "foreign" R&D and patent citations capture mainly the diffusion of codified knowledge. However, an important part of knowledge, such as skills and competences, is not codifiable (e.g. in blueprints) and remains embedded in people (Polanyi, 1967). It can be argued also that as advances in ICT render access to codified knowledge easier and less expensive, the tacit part of knowledge and, thus, the ability to absorb and transform knowledge into new products or ideas, is becoming more important. The diffusion of tacit knowledge is neglected in empirical work on economic growth mainly because of the difficulty to find adequate datasets.

In this work, we take account also of the diffusion of tacit knowledge through the inventor mobility index. Inventor mobility represents a channel of diffusion of tacit knowledge between firms and between regions because inventors who move from firm to firm take with them their skills and abilities. Making use of the information contained

[35] Backward citation is citation in a patent to another patent developed in a previous period.

in EPO patents, we built a new dataset of Italian inventors which allows us to identify whether an inventor moves between firms and between regions.

This work proposes an empirical model in which the different economic growth rates of Italian regions are explained by differences in the rates of growth of their knowledge capital. To measure the knowledge generated within a region we use a set of variables constructed on the basis of the R&D and patenting activities. R&D variables confirm whether intentional R&D activity explains the economic growth of Italian regions while patenting shows whether there is a further effect exerted by successful R&D. To measure the diffusion of knowledge we use patent citations and inventor mobility indexes to identify the impact on economic growth of the diffusion of codified knowledge through patent citations, and of tacit knowledge through inventor mobility.

This chapter is organized as follows. Section 3.2 discusses in more detail the theoretical and empirical literature on the determinants of regional economic growth. Particular attention is paid to the knowledge capital measures adopted in the literature and to our motivation for using the inventor mobility indexes. Section 3.3 presents the empirical model used in our estimates. Section 3.4 describes the data and variables used, with particular attention to the methodology adopted to construct the inventor mobility indexes. Section 3.5 presents and discusses in more detail the data on GDP per capita growth and the geographical mobility of inventors for the Italian regions. Section 3.6 presents and discusses the results of our estimates and Section 3.7 presents some final considerations and proposes some guidelines for future work.

3.2 Background to the study

3.2.1 Knowledge capital and economic growth

The neoclassical growth model (Solow, 1956, 1957) leaves no room for knowledge as a factor capable of explaining the economic growth of regions, since knowledge is considered a non-rival and non-excludable public good, freely available to everyone in time and space. An important implication of these theories is that economic agents do not have an incentive to engage in R&D activities because of the non-appropriability of the knowledge generated by these activities and, thus, of the returns from R&D investment. Also, state and local governments are not motivated to formulate innovation policies because the benefits of these policies are equally distributed among the administered territory and the territories of other governments. It follows that knowledge capital plays no role in explaining economic growth. However, this view is inconsistent with the real functioning of economic systems in which firms are involved in R&D activities, and national or regional policy provides financial support for innovation. Endogenous growth theories (Romer, 1987, 1990; Grossman and Helpman, 1991) overcome these problems by assuming that knowledge is a non-rivalrous and partially excludable good and, in that framework, firms' R&D activities are explained by the possibility of increasing profits, and policies directed to reinforcing these investments are reasonable and practical. Endogenous growth models underline the importance of the spatial aspect for economic growth because the diffusion of knowledge generates positive externalities which enhance the productivity of the whole economic system (Grossman and Helpman, 1991).

The importance of the diffusion of knowledge for economic growth can be found also in the technology gap approach (Fagerberg, 1994). Framed within evolutionary theory (Nelson and Winter, 1982), the technology-gap approach assumes that regional growth is a function of the internal processes of knowledge generation and of the processes of knowledge diffusion among regions.[36] However, technology-gap models assume that the diffusion of knowledge among regions and catching-up processes cannot be taken for granted. To absorb the knowledge generated externally, regions need an adequate level of technical competence (Cohen and Levinthal, 1989) and appropriate characteristics in the form of appropriate social, political and institutional structures (Gerschenkron, 1962; Abramovitz, 1986; Audretsch, 2007). This is coherent with the reality that the diffusion of knowledge is geographically localized as a result of geographical, social and institutional barriers (Jaffe et al., 1993; Maurseth and Verspagen, 2002; Peri, 2005). In turn, the limited diffusion of knowledge explains the lack of convergence between regions and the presence of income polarizations and regional growth clubs (see e.g. Doring and Schnellenbach, 2006, for a survey). Ultimately, differences in regional growth rates can be explained by innovation and imitation processes, neither of which is mechanical, while innovation can lead to differences among regions, and imitation can lead to greater convergence.

From an empirical point of view, one of the challenges of analysing the impact of knowledge capital on economic growth is the quantification of the knowledge present in a system or region. While data on traditional factors, i.e. labour and capital, are easily accessible via national accounting systems, which provide data at different territorial levels, data on knowledge, given its immaterial nature, must be constructed based on

[36] Following Schumpeterian thinking, the technology gap approach assumes that economic growth is a process determined by technological discontinuities not predictable ex ante (Fagerberg, 1994, 2002). The introduction of a radical innovation with its creative destruction, breaks down the existing economic system and provides the opportunity for a jump in the economic growth rate. The new knowledge generated tends to spread beyond the territorial borders where it originated giving the opportunity for imitation.

data on knowledge generating activities and knowledge diffusion channels. Scholars have used different tools to measure knowledge and the next two subsections discuss the literature on measures of knowledge creation processes (section 3.2.2) and measures of knowledge diffusion processes (section 3.2.3.1). In the last subsection (section 3.2.3.2), we argue for the usefulness of introducing the inventor mobility indexes to measure the diffusion of knowledge.

3.2.2. Knowledge creation

To measure the knowledge generated in an economy the literature makes use mainly of data on R&D expenditure or R&D employment. Firms invest in R&D activities in order to develop new products or processes, and the innovations resulting from these activities enhance the firms' knowledge capital. At the same time, many studies underline the importance of the basic knowledge generated by academic and other research institutions to facilitate industry innovation[37] (Mansfield, 1995). The positive impact of R&D activities on regional economic growth is demonstrated in several papers (Fagerberg et al., 1997; Cappellen et al., 1999, Sterlacchini, 2008). For instance, Fagerberg et al. (1997), using regional data for four European countries in the period 1980-1990, shows that the economic growth of a region, expressed as GDP per capita growth, is positively affected by internal R&D activities, expressed as share of R&D employment in the labour force. However, other works show that R&D intensity has a positive impact on regional economic growth only for the more developed regions (Cappellen et al., 1999; Sterlacchini, 2008). For instance, Sterlacchini (2008) using regional data for 12 European countries in the period 1995-2002, shows a positive and significant effect of R&D activities, expressed by share of R&D expenditure in Gross

[37] To be rigorous and in line with our exposition, the use of scientific knowledge by firms can be considered the result of the diffusion of knowledge between research institutions and firms.

Value Added, on economic growth, expressed by the GDP per capita growth, only for regions above a certain threshold of per capita GDP, while the effect for less developed regions is not significant.

Another measure used for regional knowledge creation processes is patent indicators. The use of patents is justified by the fact that R&D is a measure of the inputs used in the innovation processes rather than a measure of innovative outcomes. It can be argued that the resources allocated to R&D activities do not guarantee results in terms of the production of new products or knowledge and, therefore, patents represent a measure of successful R&D.[38] Some scholars estimate the impact of innovative activities on national or regional economic growth using number of patents developed within the territorial border, rather than or in addition to R&D indicators (e.g. Fagerberg, 2002; Crescenzi, 2005; Sterlacchini, 2008)[39].

To measure innovative performance, some scholars suggest using patent citations, i.e. forward citations,[40] instead of the simple number of patents (Trajtenberg, 1990; Hall et al., 2005). The technological and, therefore, economic value of patents has a highly skewed distribution and the simple count of the patents within a territory does not take account of this heterogeneity in values. Instead, citations received captures this heterogeneity in the value of patents.

In this work, we consider data on R&D and patenting to measure regional knowledge generation processes. We make use of data on forward patent citation to control for heterogeneity in patent values.

[38] Patents allow us to measure inventions, but not to capture innovations that correspond to the commercialization of the ideas contained in the patent document. Thus, patents represent an intermediate output of the innovation process which ends with the introduction of a new product in the market.

[39] E.g., Sterlacchini (2008) shows that patenting activities, expressed by number of EPO applications in the population, do not have an additional effect on regional economic growth.

[40] A forward citation is a citation to the patent in another patent developed later.

3.2.3. Knowledge diffusion

3.2.3.1 Traditional measures of knowledge diffusion

Economists are unanimous about the process of knowledge diffusion as an important driver of economic growth. Knowledge can spread across space through various channels which can be distinguished primarily by two characteristics: type of externalities produced, and type of knowledge conveyed. In the first case, according to Griliches (1979), we can distinguish between different channels of knowledge flows depending on whether they require market transactions or not. On the one hand, we have "rent spillovers", which are the result of market transactions where the price of a product does not reflects the real value of the technology contained in the product. On the other hand, we have pure externalities ("pure spillovers"), which are the knowledge flows that occur without a market transaction and derive from the imperfect appropriability of knowledge and from voluntary exchange of knowledge through personal contacts.[41] The channels of knowledge diffusion can also be categorized according to the type of knowledge transmitted, i.e. codified or tacit. The knowledge is codified when an idea or a technology is described in a document (e.g. patent, scientific article), knowledge is tacit when it cannot be codified in documents (e.g. know-how). Tacit knowledge can be transferred easily and at low cost over large distances using ICT, but tacit knowledge requires personal contact and its distribution, therefore, is more localized.

There are multiple ways that knowledge can flow across space and the literature uses different measures of knowledge spillovers. The technique used to measure the diffusion of knowledge among countries or regions consists of the weighted sum of

[41] However, distinguishing between the two types of spillovers is not straightforward since some channels of knowledge flows, traditionally considered as "pure spillovers", may hide market transactions (Feldman and Kogler, 2010).

R&D,[42] which is based on the idea that a region has access to the technology and knowledge created in other regions (measured by R&D) in proportion to some distance (measured by some weighting matrix) between the two regions. Constructing the weighting matrix and, therefore, identifying the proximity between two countries or regions make use of trade (Coe and Helpman, 1995), technological proximity (Park, 1995), foreign direct investments (van Pottelsberghe and Lichtenberg, 2001) and geographical distance (Rodriguez-Pose and Crescenzi, 2008).

However, as discussed above, stock of "foreign" R&D is only a suggestive measure of knowledge flows, and disregards the real exchange of knowledge among people, firms and regions. The measures that make use of stock of "foreign" R&D do not provide information concerning the actual use by an economic system of the knowledge produced elsewhere. In addition, intra-regional knowledge flows, by construction, are not considered. Intra-regional knowledge flows are an important element explaining external economies of agglomeration (Marshall, 1920) and regional economic growth (Lundvall, 1992; Morgan, 1997).

We overcome these limitations using backward patent citations to measure knowledge flows between regions. Backward patent citations are a direct measure of knowledge spillovers. Backward patent citations allow us to observe that in the inventive process leading to a patent an inventor has used knowledge generated by other inventors. In addition, use of backward patent citations allows us to measure the diffusion of knowledge within a territorial border. Although very few studies incorporate backward patent citations in their growth models (Caballero and Jaffe, 1993), several researchers use this tool to measure knowledge flows between countries

[42] The formula used in the calculation of knowledge flows between regions is : $KF_i = \sum_{j \neq i} w_{ji} \; R\&D_j$, where *KF* are knowledge flows to region *i*, *R&D* are the R&D expenditures of the region *j* (*j*≠*i*)*,* and *w* is a weighting matrix.

or regions (Jaffe et al., 1993; Maurseth and Verspagen, 2002; Peri, 2005; Fischer et al., 2009; Montobbio and Sterzi, 2011). These works show that the diffusion of knowledge is geographically localized. For instance, Maurseth and Verspagen (2002), using EPO patent citations data for 112 European regions, show that interregional knowledge flows are more likely to occur within nations, and decrease with increasing distance between regions.

The use of patent citations to measure knowledge flows has been criticized by some scholars because of the patent procedures (Alcacer and Gittelman, 2006; Thompson, 2006). In particular, they argue that citations can be added by patent examiners and by firm applicants, but inventors do not know which patents are cited. This limitation is exacerbated for patents submitted to the EPO and other patent offices which, unlike the US Patent and Trademark Office, do not honour the rule of "duty of candour". Several authors however claim that patent citations is a valid although imperfect measure of knowledge flows (Jaffe et al., 1998; Jaffe and Trajtenberg, 2002; Dugeut and MacGarvie, 2005). For instance, Dugeut and MacGarvie (2005) using CIS data for France show that EPO patent citations are strongly correlated with real knowledge flows.

Patent citations are a good measure of the diffusion of codified, but not tacit knowledge. The diffusion of tacit knowledge is rather neglected in empirical analyses of the relationship between knowledge capital and economic growth because of lack of data. In this work we take account of the diffusion of tacit knowledge through the use of a new dataset on Italian inventor mobility.

3.2.3.2 Inventor mobility

The knowledge capital of firms and regions is strongly linked to their human capital. Human capital accumulates through education and experience (job-training) (Becker, 1962, 1993). However, since a part of the knowledge, such as skills, is embedded in people, mobility of workers is an important channel for the diffusion of tacit knowledge between firms and regions. Almeida and Kogut (1999) show the close link between knowledge flows and labour mobility. Analysing the determinants of diffusion of knowledge (measured by patent citations) between regions in the US semiconductor industry, they show that the mobility of engineers is an important factor explaining the localized diffusion of knowledge within regions. In particular, regions with greater internal mobility of engineers show higher levels of localized diffusion. This stems from the fact that an important part of an invention is represented by the tacit knowledge embedded in engineers. The mobility of workers also creates links between firms through social ties, which involve the worker that moves and the workers in his or her previous firm. These ties favour the diffusion of knowledge among firms and regions (Breschi and Lissoni, 2009). Agrawal et al. (2006), analyse the diffusion of knowledge (measured by patent citations) between US regions, generated by the mobility of inventors. They show that an inventor who moves from one region to another is more likely to cite inventors in the previous region, compared to those who have never lived in that region. The social networks between inventors reduce the frictions in knowledge flows exerted by geographical factors such as physical distance.

In this work we propose the use of the inventor mobility to measure the knowledge flows between regions. The mobility of inventors represents a measure of knowledge spillovers, which, unlike traditional measures of knowledge flows, captures the diffusion of the tacit component of knowledge. When an inventor changes jobs, he or she transfers from the old to the new firm detailed information on the technologies used in the previous employment and also the knowledge, skills and experience embedded in the mobile inventor.

3.3. The empirical model

To conduct our analysis we use an empirical model based on the technology gap approach. The technology gap approach emerged because of the inadequacy of formal neoclassical theories to explain the economic growth of countries. The technology gap approach is an "appreciative theory" (Nelson and Winter, 1982; Fagerberg, 1994) based on empirical studies that abstract from the rigidity of formal mathematical models and assume that the processes of creation and diffusion of knowledge are important for explaining economic growth. In particular, empirical work on economic growth within the technology gap approach is based on three hypotheses (Fagerberg, 1987, 1997). First, regional rate of growth is positively influenced by the rate of growth of the region's knowledge capital. Second, a follower region, i.e. a region with a technology gap with respect to regions at the technological frontier, can increase its rate of economic growth by means of imitation. The possibility of a follower region exploiting the knowledge generated externally depends in its absorptive capacity.

Based on this, our analysis of Italian regions exploits a model in which annual rates of growth of GDP per capita are a function of the following variables: log of GDP per capita in level (*GDP/POP*); log of population (*POP*); log of fixed investments per capita (*INV/POP*); log of R&D expenditure per capita (*R&D/POP*); ratio of number of patent applications and R&D expenditure (*PAT/R&D*). These variables are expressed in lagged values. R&D expenditure per capita, which includes expenditure on R&D in the public and private sectors, is used to measure the change in the region's knowledge stock based on intentional innovative activities in the region. We use the ratio of patents and R&D expenditure to check for an additional effect exerted by successful R&D. GDP per capita in level measures the technological distance of the region from the technology leader regions. This variable captures catching-up based on imitation processes. Catch-up cannot be taken for granted; it depends on set of economic, social and institutional factors that determine the ability to absorb external knowledge. In line with the literature (Fagerberg, 1987, 2002), we use fixed investments to proxy for the country's capacity to exploit external knowledge. Population controls for the size of the region.

Our basic model is represented by the following equation:

$$[3.1]\ \ln\left(\frac{\left(\frac{\mathrm{GDP}}{\mathrm{POP}}\right)_t}{\left(\frac{\mathrm{GDP}}{\mathrm{POP}}\right)_{t-1}}\right) = \alpha + \beta \ln\left(\frac{\mathrm{GDP}}{\mathrm{POP}}\right)_{t-1} + \gamma \ln\,(\mathrm{POP})_{\mathrm{t-1}} + \pi\ \ln\left(\frac{\mathrm{INV}}{\mathrm{POP}}\right)_{t-1} + \\ + \varphi\ \ln\left(\frac{\mathrm{R\&D}}{\mathrm{POP}}\right)_{t-1} + \upsilon\ \left(\frac{\mathrm{PAT}}{\mathrm{R\&D}}\right)_{t-1}$$

We add to this model the variables constructed using patent citations data: backward and forward patent citations. Backward patent citations measure the diffusion of codified knowledge between regions in order to verify whether these kinds of knowledge flows have an impact in terms of regional economic growth. We distinguish between intra-regional (*Back_intra*) and interregional (*Back_inter*) backward patent citations in order to check for a different impact of knowledge flows determined by the origin of knowledge. Forward citations control for heterogeneity in patent values. Since the distribution of patent values is highly skewed, forward patent citations is a more direct measure of the economic value of a patent. We distinguish between intra-regional (*Forw_intra*) and interregional (*Forw_inter*) forward patent citations to verify whether the “geographical dimension” of the invention is relevant. Backward and forward patent citations are divided by number of the patents for the region in order to take account of regional difference in terms of capacity to patent. Once again these variables are expressed in lagged values. This gives the following equation:

$$[3.2]\ \ln\left(\frac{\left(\frac{\text{GDP}}{\text{POP}}\right)_t}{\left(\frac{\text{GDP}}{\text{POP}}\right)_{t-1}}\right) = \alpha + \beta \ln\left(\frac{\text{GDP}}{\text{POP}}\right)_{t-1} + \gamma \ln\,(\text{POP})_{\text{t}-1} + \pi\ \ln\left(\frac{\text{INV}}{\text{POP}}\right)_{t-1} + \\ + \varphi\ \ln\left(\frac{\text{R\&D}}{\text{POP}}\right)_{t-1} + \upsilon\ \left(\frac{\text{PAT}}{\text{R\&D}}\right)_{t-1} + \mu\ \left(\frac{\text{Back\_intra}}{\text{PAT}}\right)_{t-1} + \\ + \Omega\ \left(\frac{\text{Back\_inter}}{\text{PAT}}\right)_{t-1} + \eta\left(\frac{\text{Forw\_intra}}{\text{PAT}}\right)_{t-1} + \theta\left(\frac{\text{Forw\_inter}}{\text{PAT}}\right)_{t-1}$$

We add the inventor mobility indexes to take account of the diffusion of tacit knowledge. We consider a mobile inventor to be an inventor who moves between firms (switch applicants).[43] We distinguish between intra- and inter-regional inventor mobility. Intra-regional mobility (*Mob_intra*) is the number of inventors that switch applicants within the same region. For interregional inventor mobility we have two types of inventor mobility in order to distinguish between inflow of inventors from other regions (*Mob_inflow*) and outflow of inventors to other regions (*Mob_outflow*). The inventor mobility indexes are expressed as the ratio of number of mobile inventors in a period and the regional stock of inventors (*SI*) in the previous period. For instance, the intraregional mobility index of region A at time *t-1* is given by the ratio of number of inventors that move within region A in the period *t-1* and stock of inventors in region A at time *t-2*. Therefore, we have the following equation:

[43] In order to eliminate possible bias in the mobility indexes it would be useful to control over time for firm's mergers and acquisitions, but data availability and the effort involved make this task very difficult.

[3.3] $$\ln\left(\frac{\left(\frac{\mathrm{GDP}}{\mathrm{POP}}\right)_t}{\left(\frac{\mathrm{GDP}}{\mathrm{POP}}\right)_{t-1}}\right) = \alpha + \beta \ln\left(\frac{\mathrm{GDP}}{\mathrm{POP}}\right)_{t-1} + \gamma \ln(\mathrm{POP})_{t-1} + \pi \ln\left(\frac{\mathrm{INV}}{\mathrm{POP}}\right)_{t-1} + \\ + \varphi \ln\left(\frac{\mathrm{R\&D}}{\mathrm{POP}}\right)_{t-1} + \upsilon \left(\frac{\mathrm{PAT}}{\mathrm{R\&D}}\right)_{t-1} + \mu \left(\frac{\mathrm{Back\_intra}}{\mathrm{PAT}}\right)_{t-1} + \\ + \Omega \left(\frac{\mathrm{Back\_inter}}{\mathrm{PAT}}\right)_{t-1} + \eta \left(\frac{\mathrm{Forw\_intra}}{\mathrm{PAT}}\right)_{t-1} + \theta \left(\frac{\mathrm{Forw\_inter}}{\mathrm{PAT}}\right)_{t-1} + \\ + \lambda \left(\frac{\mathrm{Mob\_intra}_{t-1}}{\mathrm{SI}_{t-2}}\right) + \rho \left(\frac{\mathrm{Mob\_inflow}_{t-1}}{\mathrm{SI}_{t-2}}\right) + \tau \left(\frac{\mathrm{Mob\_outflow}_{t-1}}{\mathrm{SI}_{t-2}}\right)$$

In an additional specification interregional inventor mobility indexes are replaced by net inflows of inventors (*Mob_netflow*), i.e. the difference between inflow and outflow of inventors. Therefore, we have the following equations:

[3.4] $$\ln\left(\frac{\left(\frac{\mathrm{GDP}}{\mathrm{POP}}\right)_t}{\left(\frac{\mathrm{GDP}}{\mathrm{POP}}\right)_{t-1}}\right) = \alpha + \beta \ln\left(\frac{\mathrm{GDP}}{\mathrm{POP}}\right)_{t-1} + \gamma \ln(\mathrm{POP})_{t-1} + \pi \ln\left(\frac{\mathrm{INV}}{\mathrm{POP}}\right)_{t-1} + \\ + \varphi \ln\left(\frac{\mathrm{R\&D}}{\mathrm{POP}}\right)_{t-1} + \upsilon \left(\frac{\mathrm{PAT}}{\mathrm{R\&D}}\right)_{t-1} + \mu \left(\frac{\mathrm{Back\_intra}}{\mathrm{PAT}}\right)_{t-1} + \\ + \Omega \left(\frac{\mathrm{Back\_inter}}{\mathrm{PAT}}\right)_{t-1} + \eta \left(\frac{\mathrm{Forw\_intra}}{\mathrm{PAT}}\right)_{t-1} + \theta \left(\frac{\mathrm{Forw\_inter}}{\mathrm{PAT}}\right)_{t-1} + \\ + \lambda \left(\frac{\mathrm{Mob\_intra}_{t-1}}{\mathrm{SI}_{t-2}}\right) + \sigma \left(\frac{\mathrm{Mob\_netflow}_{t-1}}{\mathrm{SI}_{t-2}}\right)$$

Since the potential mobility of an inventor is determined by the number of applicants in the different regions, as a robustness check we make other estimates to control for this potential source of bias. In particular, we add regional applicant shares (*applicant_share*) to equations [3.3] and [3.4] (see Appendix). This variable, which is time varying, is constructed as the ratio of number of patent applicants in the region and total number of applicants in Italy.

3.4. Data

For the empirical analysis we constructed a set of variables for the period 1995-2007 for the 20 Italian regions, i.e. the first level of administrative divisions in the Italian state. A first group of variables is constructed using data from the ISTAT: GDP, R&D

expenditure, fixed investment in capital. Population data are provided by the EUROSTAT database.

The large patent datasets[44] supplied by patent offices make it possible to construct various measures based on patents (number of patents, patent citations, etc.) at country or regional level. However, for the construction of measures related to the mobility in space of inventors, these data suffer some important limitations because of the “who is who” and the “John Smith” problems (Trajtenberg et al., 2006). The former refers to the fact the name of an inventor with two or more patents may be spelled differently on different patents. The latter refers to the same name sometimes referring to different inventors. To overcome these limitations we built a separate dataset using a procedure referred to as “name game” analysis (Trajtenberg et al., 2006; Raffo and Lhuillery, 2009) on PATSTAT data on EPO patent applications. Our name game analysis[45] takes account of inventor’s name, inventor’s address, technological class of the patent, name of the applicant, and co-inventors.

This dataset of Italian inventors allows us to identify all the patents developed by a single inventor over time, i.e. this dataset does not suffer from “who is who” or “John Smith” problems. Using this dataset we identify the inventors who move between firms and also between regions. We consider inventors with at least two EPO patent applications and look at the applicants of these patents. If an inventor, in a given period, has an EPO patent for an applicant and the same inventor, in a later period, appears on an EPO patent with a different applicant name, we assume that this inventor moved

[44] Patent documents provide a variety of information regarding the invention (e.g. description of the invention and its technological class), applicants (names and addresses) and inventors (names and addresses).

[45] In our procedure we do not consider the patents developed by Italian inventors resident outside of Italy. Thus, we do not consider mobility from an Italian region to another country, and vice versa. However, we suppose that this type of mobility is rare and would not influence the results of our analysis. See the appendix for more detail on the “name game” analysis adopted in this work.

from one firm to another (switch applicant) during the two periods. To identify whether the inventor moved within or between regions we look at the inventor's address for the two periods. If the two addresses correspond to localities in the same region, then the inventor moved within the region (i.e. intra-regional mobility). If the two addresses correspond to localities in two different regions, then the inventor moved between regions (i.e. inter-regional mobility).

The stock of regional inventors (*SI*) (used to construct the inventor mobility indexes) is calculated using the perpetual inventory method (we have data from 1978), which gives:

$$SI_{it} = (1 - \delta)\, SI_{i,t-1} + I_{it}$$

where I is the number of inventors in year t and δ is the constant depreciation rate that is set at 5% (see Griliches and Mairesse, 1984, for a more detailed description).[46] We "depreciate" the stock of inventors in order to takes account of the exit of inventors due to retirement.[47]

The number of EPO patent applications of each Italian region is obtained using the inventor's address to assign a patent to a Italian region. In the case of coinventorship with inventors residing in different regions, patent is assigned at each regions. In the case of several inventors residing in the same region patent is counted only once.

We also use a set of variables constructed using both our dataset on inventor mobility and the KITES (Bocconi University) patent database: intra-regional and interregional backward citations, intra-regional and interregional forward citations. The two dataset

[46] Stock corrected for double counting of inventors.
[47] However, the results obtained using the stock of inventors calculated without a depreciation rate are very similar.

are merged.[48] The dataset on inventor mobility provides information on the region of residence of the patent's inventors and the KITES dataset provides information on the patent's citations to other EPO patents.[49,50] We identify a citation from region A to region B when the citing patent has at least one inventor resident in region A and the cited patent has at least one inventor resident in region B. In this case we have an interregional patent citation. An intra-regional patent citations is when the citations are to another inventor in the same region. In the case of several inventors residing in the same region (A or B) citations are counted only once. Coherent with the literature (Maurseth and Verspagen, 2002; Paci and Usai, 2009; Hall and MacGarvie, 2010) we do not consider self-citations among firms. Finally, in order to overcome the truncation bias problem[51] (Hall et al., 2005; Fisher et al., 2009), we count only the citations where the time lag between cited and citing patent is within a temporal window of one year.[52] Thus, a backward citation (intraregional or interregional) is a citation in a patent registered (priority date) at the period t to another patent registered during the period t – $(t-1)$. A forward citation (intraregional or interregional) is a citation received by a patent registered at the period t from another patent registered during the period t – $(t+1)$.

Finally, we use the applicant's address to construct the regional applicant shares. The applicant addresses are provided by the PATSTAT database.

Table 3.1 provides descriptive statistics of the variables used in our analysis.

[48] The merge between the two dataset is made using the EPO publication number in order to obtain full correspondence between datasets.
[49] We exclude patents that do not have at least one Italian inventor.
[50] KITES database contains EPO citations data cleaned from errors (e.g. EPO patents that cite EPO patents with higher EPO publication number) found in the original data provided by the patent office.
[51] Truncation bias is due to the fact that we observe only a limited period of the legal life of a patent and this period differs for each cohort of patents.
[52] We consider a temporal window of one year because the patent database contains patent applications with a priority year that does not exceed year 2008.

Table 3.1. Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
GDP/POP	240	0.0202	0.0051	0.0116	0.0283
POP	240	2874971	2278932	117063	9545441
INV/POP	240	0.0044	0.0012	0.0022	0.0081
PAT/R&D	240	0.3055	0.2409	0.0233	1.7699
R&D/ POP	240	0.0002	0.0001	0.0000	0.0005
GDP per capita growth	240	0.011	0.016	-0.031	0.057
log(GDP/POP) *t-1*	240	-3.949	0.271	-4.481	-3.544
log(POP) *t-1*	240	14.448	1.058	11.667	16.064
log (INV/POP) *t-1*	240	-5.490	0.282	-6.158	-4.818
(PAT/R&D) *t-1*	240	0.305	0.241	0.023	1.767
log(R&D/POP) *t-1*	240	-8.870	0.680	-10.856	-7.655
(Forw_inter/PAT) *t-1*	240	0.019	0.085	0	1.000
(Forw_intra/PAT) *t-1*	240	0.007	0.024	0	0.250
(Back_inter/PAT) *t-1*	240	0.002	0.020	0	0.250
(Bacw_intra/PAT) *t-1*	240	0.007	0.034	0	0.500
applicant_share *t-1*	240	0.050	0.075	0.001	0.342
Mob_intra *t-1* / SI *t-2*	240	0.015	0.016	0	0.156
Mob_inflow *t-1* / SI *t-2*	240	0.002	0.009	0	0.078
Mob_outflow *t-1* / SI *t-2*	240	0.002	0.005	0	0.053
Mob_netflow *t-1* / SI *t-2*	240	0.001	0.010	-0.053	0.078

Note:values expressed in €millions

3.5. Regional differences in GDP per capita and inventor mobility

The empirical analysis focuses on per capita GDP growth in Italian regions, during the period 1995-2007. Italian regions are characterised by remarkable differences in per capita GDP. Table 3.2 shows per capita GDP for the Italian regions in 1995, 2001 and 2007. In 1995 the highest per capita GDP was in Val d'Aosta[53] (about 28,892 euro) and the lowest was in Calabria (about 11,326 euro). There are marked differences in GDP per capita between the northern (Emilia Romagna, Friuli Venezia Giulia, Liguria, Lombardia, Piemonte, Trentino Alto Adige, Veneto and Val d'Aosta) and central (Lazio, Marche, Umbria and Toscana) regions, and the regions in the south of Italy

[53] Val d'Aosta is also the region with the lowest level of population. See the appendix for more detail.

(Abruzzo, Basilicata, Calabria, Campania, Puglia, Molise, Sicilia and Sardegna). Table 3.2 shows that levels of per capita GDP are lower in all the regions in the south of Italy than in the north and central regions for the three years analysed.

Table 3.2. GDP per capita of Italian regions (years 1995, 2001 and 2007)

Region	1995	2001	2007
Abruzzo	16165	18361	18133
Basilicata	12227	14681	15431
Calabria	11326	13403	13974
Campania	12115	13602	13918
Emilia Romagna	24341	27152	27269
Friuli Venezia Giulia	20755	23763	24354
Lazio	22412	24561	25544
Liguria	18873	21860	22064
Lombardia	25633	27975	28177
Marche	18686	21350	22241
Molise	13607	15563	16570
Piemonte	21665	23532	23984
Puglia	12148	14052	14184
Sardegna	14392	16181	16679
Sicilia	12279	13878	14319
Toscana	20686	23309	23778
Trentino Alto Adige	24994	26884	27058
Umbria	18326	20644	20511
Val d'Aosta	28892	27460	28271
Veneto	22470	24973	25558

Note: values expressed in €; ISTAT data (chain values – reference year: 2000).

Table 3.3 presents average annual (compound) growth rates in per capita GDP for the Italian regions, in 1995-2007 and the two sub periods 1995-2001 and 2001-2007. In 1995-2007 the highest GDP per capita growth was in Basilicata (about 1.96%), and the lowest was in Val d’Aosta, which showed negative growth (about -0.18%). It is interesting that the less developed regions in the south of the country show the highest GDP growth rates. After Basilicata, Calabria (about 1.77%) and Molise (about 1.66%) are ranked respectively second and third for per capita GDP growth. In the period analysed, it can be seen that there was a process of catching-up by the less developed regions of south Italy. When we look at the two sub periods it is clear that in the second period (2001-2007), there was a slowdown in GDP per capita growth in the Italian regions compared to the first period (1995-2001).

Table 3.3. Annual (compound) growth of per capita GDP for the Italian regions: 1995-2007 in percentages

Region	Period		
	1995-2007	1995-2001	2001-2007
Abruzzo	0.96	2.15	-0.21
Basilicata	1.96	3.10	0.83
Calabria	1.77	2.85	0.70
Campania	1.16	1.95	0.38
Emilia Romagna	0.95	1.84	0.07
Friuli Venezia Giulia	1.34	2.28	0.41
Lazio	1.10	1.54	0.66
Liguria	1.31	2.48	0.15
Lombardia	0.79	1.47	0.12
Marche	1.46	2.25	0.68
Molise	1.66	2.26	1.05
Piemonte	0.85	1.39	0.32
Puglia	1.30	2.46	0.16
Sardegna	1.24	1.97	0.51
Sicilia	1.29	2.06	0.52
Toscana	1.17	2.01	0.33
Trentino Alto Adige	0.66	1.22	0.11
Umbria	0.94	2.01	-0.11
Val d'Aosta	-0.18	-0.84	0.49
Veneto	1.08	1.78	0.39

Source: authors' elaboration of ISTAT data (chain values – reference year: 2000).

The main focus of the analysis is on inventor mobility as the source of knowledge flows. Table 3.4 shows total numbers of inventors and inventor mobility for the Italian regions in the period 1995-2007. Column 1 shows the geographical distribution of Italian inventors, i.e. Italian inventors with at least one EPO patent application.[54] The

[54] Number of inventors corrected for double counting of inventors, at country and at regional level.

total number of Italian inventors is 35,706.[55] The highest number of inventors (11,022) is in Lombardia and the lowest (30) is in Molise. Columns 2 to 5 show inventor mobility. As discussed above, we consider inventors with more than one EPO patent application to define inventor mobility as events in which an inventor has a patent application for an applicant in a period, but for another applicant in a later period (switch applicant). Geographical mobility of the inventor (intra- or inter-regional) is determined by his/her residences (regional level) observed in the two patents. Column 2 shows intra-regional mobility, i.e. the total number of cases where two patents developed by the same inventor have two different applicants, but the inventor's region of residence does not change. Total intra-regional mobility is 5,692. The region with the highest value for intra-regional mobility is Lombardia with 2,246 cases, while the region with the lowest value for intra-regional mobility is Molise with 1 case. Columns 3 and 4 present interregional mobility, i.e. the total number of cases where two patents developed by the same inventor have two different applicants and the inventor's region of residence is different between the two patents. Column 3 shows interregional inflows, i.e. the total number of cases where an inventor who switches applicants enters the region from another region. Total interregional inflows are 332. The region with the highest value for inventor inflows is Lombardia with 84 cases, while Basilicata, Calabria, Molise and Sardegna are the regions with the lowest value of inventor inflows with 1 case each. Column 4 shows interregional outflows, i.e. the total number of cases where an inventor who switches applicants moves from one region to another.[56] The region with the highest value for inventor outflow is Lombardia with 89 cases; the

[55] Number of inventors in Italy does not equate with the sum of the inventors in each region because inventors who patented in more than one region are counted only once (i.e. double counting correction).
[56] Total inventor outflow is equal to total inventor inflow because we consider inventor mobility between Italian regions but within the country of Italy.

region with the lowest value for inventor outflow is Val d'Aosta (none of its inventors moved). Column 5 shows interregional net inflows, i.e. the difference between inflow and outflow of inventors. The region with the highest value of net inflow is Emilia Romagna with a value of 20 (48 cases of inflow and 28 of outflow); the region with the lowest value of net inflow is Piemonte with a value of -12 (37 cases of inflow and 49 of outflow).

Table 3.4. Italian inventor Mobility, period 1995-2007

Region	Number of inventors	Intra-regional mobility	Interregional inflow	Interregional outflow	Interregional net inflow
Abruzzo	415	27	9	10	-1
Basilicata	67	7	1	2	-1
Calabria	165	33	1	1	0
Campania	699	92	8	11	-3
Emilia Romagna	4893	787	48	28	20
Friuli V. Giulia	1066	107	11	16	-5
Lazio	2440	327	34	30	4
Liguria	921	94	8	10	-2
Lombardia	11022	2246	84	89	-5
Marche	808	71	6	7	-1
Molise	30	1	1	1	0
Piemonte	4965	793	37	49	-12
Puglia	518	39	3	6	-3
Sardegna	210	18	1	3	-2
Sicilia	795	58	3	7	-4
Toscana	2409	370	31	26	5
Trentino A. Adige	539	43	4	3	1
Umbria	347	35	6	5	1
Valle d'Aosta	69	4	7	0	7
Veneto	3909	540	29	28	1
Italy	35706	5692	332	332	0

Source: authors' elaborations on PATSTAT data.

Table 3.5 presents the combination of inflow and outflow of inventors between the Italian regions. Reading from left to right we can see the inflow of inventors to a region from each of the other 19 regions; reading from top to bottom we can see the outflow of inventors from a region to each of the other 19 regions. For instance, the value in the cell for the intersection of the first column (Piemonte) and the second row (Val d'Aosta) is 7. This is the number of cases of inventors entering Val d'Aosta from Piemonte, i.e. inventor inflow to Val d'Aosta from Piemonte. The value for the number of cases of inventors moving from Piemonte to Val d'Aosta is also 7, i.e. inventor outflow from Piemonte to Val d'Aosta. The value in the cell at the intersection of the second column (Val d'Aosta) and the first row (Piemonte) is 0. This is the number of cases of inventors moving to Piemonte from Val d'Aosta, i.e. inventor inflow to Piemonte from Val d'Aosta. The number of cases of inventors moving from Val d'Aosta to Piemonte, i.e. inventor outflow from Val d'Aosta to Piemonte, is also 0 .

The highest value of inflows/outflows between regions is 22. There are 22 cases of inventors moving from Lombardia to Emilia Romagna, thus, inventor inflow to Emilia Romagna from Lombardia is 22 and inventor outflow from Lombardia to Emilia Romagna is 22.

Table 3.5. Inflows and outflows of inventors by region, period 1995-2007

	Piemonte	Val d'Aosta	Liguria	Lombardia	Trentino A. A.	Veneto	Friuli V. G.	Emilia R.	Toscana	Umbria	Marche	Lazio	Abruzzo	Molise	Campania	Puglia	Basilicata	Calabria	Sicilia	Sardegna	Total Inflow
Piemonte	-	0	2	12	0	0	8	5	1	1	1	2	1	0	1	0	1	0	0	2	37
Val d'Aosta	7	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
Liguria	2	0	-	3	0	0	0	0	1	0	0	0	0	0	0	2	0	0	0	0	8
Lombardia	13	0	4	-	1	11	2	12	14	2	1	8	6	0	4	3	0	0	3	0	84
Trentino A. A.	0	0	0	1	-	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	4
Veneto	2	0	0	14	1	-	0	6	1	0	0	3	0	0	0	0	0	0	2	0	29
Friuli V. G.	7	0	0	1	0	1	-	0	0	0	1	1	0	0	0	0	0	0	0	0	11
Emilia R.	8	0	0	22	0	10	0	-	5	1	2	0	0	0	0	0	0	0	0	0	48
Toscana	6	0	0	15	1	1	1	1	-	0	0	3	0	0	1	0	0	0	1	1	31
Umbria	0	0	0	1	0	0	0	0	0	-	0	5	0	0	0	0	0	0	0	0	6
Marche	0	0	0	1	0	1	0	2	0	0	-	2	0	0	0	0	0	0	0	0	6
Lazio	1	0	3	11	0	2	4	1	4	1	1	-	2	0	2	1	0	0	1	0	34
Abruzzo	1	0	0	2	0	0	1	0	0	0	1	2	-	0	2	0	0	0	0	0	9
Molise	0	0	1	0	0	0	0	0	0	0	0	0	0	-	0	0	0	0	0	0	1
Campania	1	0	0	4	0	0	0	0	0	0	0	2	0	1	-	0	0	0	0	0	8
Puglia	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	-	0	1	0	0	3
Basilicata	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	-	0	0	0	1
Calabria	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	-	0	0	1
Sicilia	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	-	0	3
Sardegna	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	-	1
Total outflow	49	0	10	89	3	28	16	28	26	5	7	30	10	1	11	6	2	1	7	3	332

Source: authors' elaborations on PATSTAT data.

3.6. Results

3.6.1 Pooled OLS estimates

To investigate the relationship between knowledge capital and per capita GDP growth of Italian regions in 1995 to 2007 we implement a set of four OLS estimates for each of the equations in Section 3.3. The number of observations for each estimate is 240 (12 for each of 20 regions). In order to control for various sources of bias, we add a set of time dummies and use clustered standard errors in the estimates. Time dummies are included to control for time effects. Clustered standard errors control for likely correlation between observations in a region. The usual assumption is that the errors are independently and identically distributed, but in many cases this assumption is violated. In these situations of within region correlation, OLS estimates are unbiased, but lead to incorrect statistical inference tests of significance. Table 3.6 presents the results of our estimates.

The first column of Table 3.6 (Model 1) shows the results of the estimates of equation [3.1]. Lagged GDP per capita (*GDP/POP*) is significant and has a negative sign. Thus, during the period examined there was a process of catching-up, which means that regions with lower levels of GDP per capita, ceteris paribus, show higher growth rates. There is a positive effect of R&D activities (*R&D/POP*) on GDP growth rates, although the coefficient is significant only at 10%. We also find a positive and significant effect of the variable for patent applications (*PAT/R&D*). This means that successful innovation activity, i.e. which results in a patent application to the EPO, contributes positively to regional growth. The other variables controlling for country size (*POP*) and fixed investment (*INV/POP*) are not significant.

The second column (Model 2) shows the results of estimates of equation [3.2] which adds the variables constructed using patent citations (backward and forward) data to the variables in the previous equation ([3.1]). Backward patent citations confirm the existence of intra-regional or interregional knowledge flow impacts on regional economic growth. As discussed above (see section 3.3), intra-regional and interregional forward citations control for heterogeneity in patent values. We find that the coefficients of the backward citations, both intra-regional (*Back_intra/PAT*) and interregional (*Back_inter/PAT*) are not significant. Thus, knowledge flows, captured by backward citations, do not explain a change in GDP growth among Italian regions. These results can be explained by the fact that backward patent citations capture mainly diffusion of codified knowledge. The knowledge codified in documents, such as patents, although generated in one region, is accessible to all other regions equally and, thus, cannot explain regional differences in growth.

For forward patent citations, we find a significant and positive effect for interregional forward citations (*Forw_inter/PAT*). Thus, regional differences in patent value help to explain regional difference in rates of economic growth. However, since intra-regional forward patent citations (*Forw_intra/PAT*) are not significant, only patents with an "interregional dimension" contribute to explaining these differences.

The third (Model 3) and fourth (Model 4) columns of Table 3.6 show the results of the OLS estimates respectively for equations [3.3] and [3.4]. By means of these equations we can verify the impact on regional economic growth of the diffusion of tacit knowledge through the channel of the inventor mobility. In the third column (Model 3), we have one variable for intraregional mobility (*Mob_intra/SI*) and two variables for interregional mobility - one for inflow of inventors (*Mob_inflow/SI*) and one for outflow of inventors (*Mob_outflow/SI*). As expected, outflow of inventors has a negative effect on regional economic growth, while inflow of inventors has a positive effect. However, the coefficient of the inflow of inventors is not significant. The coefficient of intra-regional mobility is negative but not significant. The fourth column (Model 4) includes one variable for intraregional mobility (*Mob_intra/SI*) and one variable for interregional mobility (*Mob_netflow/SI*), to show the effects of inflow of inventors and outflow of inventors. Net inflow of inventors has a positive and significant effect on regional economic growth. The effect of intraregional mobility is negative, but insignificant.

The above results indicate that knowledge that "travels" within the inventor contributes to explaining the changes in GDP growth among Italian regions. Note that intra-regional inventor mobility has no effect on regional economic growth. This may be due to the lock-in problem (Bathelt et al., 2004), i.e. that knowledge that flows between the firms in the same region is too similar to have an effect on improving firms' economic performance. For externally acquired knowledge to improve innovative and economic performance requires that it has some complementarity with the knowledge already owned (Boschma et al., 2009).

Tab. 3.6. Estimation results (OLS, Cluster standard error)

Variable	Model 1		Model 2		Model 3		Model 4	
log(GDP/POP) ***t-1***	-0.042	**	-0.043	**	-0.043	***	-0.043	**
	(0.016)		(0.016)		(0.015)		(0.015)	
log(POP) ***t-1***	-0.000		-0.000		0.000		0.000	
	(0.001)		(0.001)		(0.001)		(0.001)	
log (INV/POP) ***t-1***	0.011		0.012		0.012		0.125	
	(0.008)		(0.009)		(0.009)		(0.009)	
(PAT/R&D) ***t-1***	0.016	***	0.017	***	0.016	***	0.015	***
	(0.006)		(0.006)		(0.005)		(0.005)	
log(R&D/POP) ***t-1***	0.007	*	0.008	**	0.007	**	0.007	**
	(0.004)		(0.003)		(0.003)		(0.003)	
(Forw_inter/PAT) ***t-1***			0.011	***	0.009	**	0.008	**
			(0.004)		(0.003)		(0.004)	
(Forw_intra/PAT) ***t-1***			0.031		0.029		0.028	
			(0.029)		(0.030)		(0.030)	
(Back_inter/PAT) ***t-1***			-0.002		-0.002		-0.002	
			(0.050)		(0.054)		(0.052)	
(Back_intra/PAT) ***t-1***			0.030		0.034		0.038	
			(0.059)		(0.062)		(0.061)	
Mobility								
Mob_intra ***t-1*** / SI ***t-2***					-0.009		-0.018	
					(0.046)		(0.042)	
Mob_inflow ***t-1*** / SI ***t-2***					0.107			
					(0.118)			
Mob_outflow ***t-1*** / SI ***t-2***					-0.293	***		
					(0.102)			
Mob_netflow ***t-1*** / SI ***t-2***							0.159	**
							(0.064)	
Time year dummies	Yes		Yes		Yes		Yes	
Number of observations	240		240		240		240	
R-squared	0.508		0.518		0.528		0.526	

Note: ***, ** and * indicate significance at 1, 5 and 10 %, respectively.

3.6.2 Fixed-effects estimates

As a robustness check we implement a set of four fixed-effects estimates for each of the equations in Section 3.3. In particular, we take account of the potential bias due to unobserved heterogeneity, through the introduction in each of the above equations of a set of regional dummies. Table 3.7 presents the results of our estimates.

Table 3.7 Column 1 (Model 1) shows the results of the estimates of equation [3.1]. The significant and negative sign of lagged GDP per capita (*GDP/POP*) is confirmed. However, the value of the coefficient is higher than the value of the coefficient obtained from the pooled OLS estimates (see Table 3.6). Population (*POP*) has a negative and significant effect, which again contrasts with the pooled OLS. Also, the value of the coefficient is higher than the value of the coefficient obtained from the pooled OLS estimates. Thus, region size helps to explain economic growth, i.e. regions with smaller populations, ceteris paribus, show higher growth rates. The positive and significant effect of successful R&D (*PAT/R&D*) is confirmed by the fixed-effects estimates. However, in contrast to the pooled OLS, we find a negative and insignificant effect of R&D (*R&D/POP*). Finally, the coefficient of fixed investment (*INV/POP*) is positive and not significant, as in the pooled OLS.

Column 2 (Model 2) shows the results of the estimates of equation [3.2]. Similar to the pooled estimates, we find that backward citations, both within (*Back_intra/PAT*) and between (*Back_inter/PAT*) are not significant. Also, the not significant effect for intraregional forward citations (*Forw_intra/PAT*) is confirmed. However, in contrast to the pooled estimates, intraregional forward citations (*Forw_inter/PAT*) is not significant.

Column 3 (Model 3) presents the results of the estimates of equation [3.3]. For interregional mobility, as in the pooled estimates, we find a negative and significant effect for outflow of inventors (*Mob_outflow/SI*) and a positive, but not significant effect for inflow of inventors (*Mob_inflow/SI*). The negative, but not significant effect for intraregional mobility (*Mob_intra/SI*) is confirmed.

Tab. 3.7. Estimation results (fixed-effects, cluster standard error)

Variable	Model 1		Model 2		Model 3		Model 4	
log(GDP/POP) ***t-1***	-0.406	***	-0.409	***	-0.396	***	-0.393	***
	(0.035)		(0.035)		(0.033)		(0.033)	
log(POP) ***t-1***	-0.276	***	-0.280	***	-0.263	***	-0.257	***
	(0.061)		(0.065)		(0.069)		(0.066)	
log (INV/POP) ***t-1***	0.016		0.015		0.010		0.011	
	(0.019)		(0.019)		(0.017)		(0.017)	
(PAT/R&D) ***t-1***	0.019	***	0.018	***	0.017	**	0.016	**
	(0.059)		(0.006)		(0.007)		(0.007)	
log(R&D/POP) ***t-1***	-0.002		-0.002		-0.005		-0.006	
	(0.004)		(0.004)		(0.008)		(0.007)	
(Forw_inter/PAT) ***t-1***			0.003		-0.002		-0.003	
			(0.005)		(0.008)		(0.007)	
(Forw_intra/PAT) ***t-1***			-0.015		-0.016		-0.018	
			(0.025)		(0.026)		(0.026)	
(Back_inter/PAT) ***t-1***			0.032		0.032		0.031	
			(0.046)		(0.049)		(0.049)	
(Back_intra/PAT) ***t-1***			0.005		0.009		0.014	
			(0.053)		(0.059)		(0.059)	
Mobility								
Mob_intra ***t-1*** / SI ***t-2***					-0.005		-0.013	
					(0.053)		(0.055)	
Mob_inflow ***t-1*** / SI ***t-2***					0.103			
					(0.131)			
Mob_outflow ***t-1*** / SI ***t-2***					-0.265	***		
					(0.071)			
Mob_netflow ***t-1*** / SI ***t-2***							0.159	**
							(0.067)	
Time year dummies	Yes		Yes		Yes		Yes	
Regional dummies	Yes		Yes		Yes		Yes	
Number of observations	240		240		240		240	
R-squared	0.641		0.643		0.650		0.649	

Note: ***, ** and * indicate significance at 1, 5 and 10 %, respectively.

Column 4 (Model 4) shows the results of the estimates of equation [3.4]. The positive and significant effect of net inflow of inventors (*Mob_netflow/SI*) is confirmed.

To sum up, the fixed-effects estimates show that during the period examined there was a process of catching-up. With regard to the relationship between knowledge capital and economic growth, only internal successful R&D activities, and interregional knowledge flows through the channel of inventor mobility, explain the economic growth of Italian regions.

3.7 Conclusion

This work has analysed the impact of knowledge capital on the economic growth of Italian regions. Following the technology gap approach, regional knowledge capital is determined by the knowledge generation and knowledge diffusion processes. In relation to knowledge diffusion, we distinguish between the diffusion of codified knowledge through the channel of patent citations, and the diffusion of tacit knowledge through the channel of inventor mobility.

Our results show that the knowledge generated within a region through R&D activities, has a positive impact on economic growth, especially if the activities lead to patents (successful R&D). With regard to knowledge flows, backward patent citations, either intra-regional or interregional, does not explain regional economic growth. Interregional inventor mobility helps explain regional economic growth, but intra-regional inventor mobility does not explain economic growth.

Overall, our results show that regions can achieve economic growth based on internal R&D efforts and the processes of knowledge diffusion. However, for knowledge flows to contribute to regional innovative and economic performance it must be tacit and must have been generated in another region. Tacit knowledge, which resides in the skills and knowledge embedded in people, is crucial for the innovative processes of firms. However, it is likely that the knowledge that is transferred within a region, from one firm to another, is already owned by the knowledge receiving firm and, therefore, does not increase the receiving firm's knowledge capital. However, when inventors move between regions, there is a higher likelihood that the knowledge transferred will be complementary knowledge that is new to the receiving firm.

One of the problems in the literature on the relationship between knowledge capital and economic growth is how to measure the knowledge capital of a system or region. Our work contributes to this literature by considering a more direct measure of knowledge flows and also accounting for the diffusion of both codified and tacit knowledge. This work constitutes the first attempt to consider explicitly inventor mobility as a measure of knowledge flows in empirical analysis of the determinants of regional economic growth.

Future research should be directed to improving the analysis in this work. First, we need more refined inventor mobility indexes (i.e. a new algorithm for the "name game" analysis). The mobility indexes in this work do not take account of inventors who move beyond the country border and then return to Italy, and vice versa. Second, we need more data on Italian regions to allow us to control for regional differences in terms of social, political and institutional structures. Third, we need to cope with potential endogeneity of measures such as R&D, patenting and inventor mobility, through the implementation of more sophisticated estimation procedures.

Appendix (Chapter 3)

Appendix 3.A. Population, GDP and GDP per capita of the Italian regions

Table 3.A1. Population, GDP and GDP per capita of the Italian regions (year 1995)

Region	GDP per capita	GDP	Population
Abruzzo	16.165	20276387	1254352
Basilicata	12.227	7456266	609799
Calabria	11.326	23387285	2064883
Campania	12.115	68958001	5691818
Emilia Romagna	24.341	94810310	3895014
Friuli Venezia Giulia	20.755	24555363	1183131
Lazio	22.412	115515470	5154261
Liguria	18.873	30973379	1641159
Lombardia	25.633	227515530	8875974
Marche	18.686	26819394	1435302
Molise	13.607	4479578	329214
Piemonte	21.665	92316874	4261168
Puglia	12.148	49273275	4055934
Sardegna	14.392	23762948	1651147
Sicilia	12.279	61452905	5004913
Toscana	20.686	72510740	3505309
Trentino Alto Adige	24.994	22549317	902177
Umbria	18.326	14911076	813664
Val d'Aosta	28.892	3370389	116653
Veneto	22.470	98868367	4400073

Note: GDP and GDP per capita values expressed in €thousands; ISTAT data (chain values – reference year: 2000).

Appendix 3.B. Robustness check for regional applicant shares

Table 3.B1 show the results of the OLS and fixed-effects estimates of equations [3.3] (Model 3) and [3.4] (Model 4) with the addition of the regional applicant shares (*applicant_share*). The signs and the significance of the mobility indexes showed in Tables 3.6 and 3.7 are confirmed.

Table 3.B1. Estimation results (OLS and fixed-effects estimates; cluster standard errors)

	Model 3				Model 4			
Variable	OLS		FE		OLS		FE	
log(GDP/POP) ***t-1***	-0.044	***	-0.404	***	-0.044	**	-0.401	***
	(0.015)		(0.032)		(0.016)		(0.033)	
log(POP) ***t-1***	0.000		-0.271	***	0.000		-0.266	***
	(0.001)		(0.069)		(0.001)		(0.065)	
log (INV/POP) ***t-1***	0.011		0.010		0.012		0.011	
	(0.009)		(0.016)		(0.009)		(0.016)	
(PAT/R&D) ***t-1***	0.015	***	0.015	**	0.015	***	0.015	**
	(0.005)		(0.007)		(0.005)		(0.007)	
log(R&D/POP) ***t-1***	0.007	**	-0.006		0.007	**	-0.007	
	(0.003)		(0.008)		(0.003)		(0.007)	
(Forw_inter/PAT) ***t-1***	0.008	**	-0.002		0.007	*	-0.003	
	(0.004)		(0.008)		(0.004)		(0.007)	
(Forw_intra/PAT) ***t-1***	0.029		-0.016		0.028		-0.018	
	(0.030)		(0.026)		(0.031)		(0.026)	
(Back_inter/PAT) ***t-1***	-0.000		0.033		-0.000		0.033	
	(0.055)		(0.049)		(0.054)		(0.050)	
(Back_intra/PAT) ***t-1***	0.033		0.008		0.036		0.014	
	(0.063)		(0.058)		(0.061)		(0.059)	
applicant_share ***t-1***	0.007		0.176		0.005		0.183	
	(0.013)		(0.148)		(0.012)		(0.141)	
Mobility								
Mob_intra ***t-1*** / SI ***t-2***	-0.011		-0.003		-0.019		-0.011	
	(0.047)		(0.054)		(0.043)		(0.055)	
Mob_inflow ***t-1*** / SI ***t-2***	0.105		0.109					
	(0.118)		(0.131)					
Mob_ouflow ***t-1*** / SI ***t-2***	-0.299	***	-0.265	***				
	(0.103)		(0.072)					
Mob_netflow ***t-1*** / SI ***t-2***					0.159	**	0.163	**
					(0.063)		(0.068)	
Time year dummies	Yes		Yes		Yes		Yes	
Regional dummies	No		Yes		No		Yes	
Number of observations	240		240		240		240	
R-squared	0.528		0.651		0.526		0.650	

Note: ***, ** and * indicate significance at 1, 5 and 10 %, respectively.

Appendix 3.C. Brief description of the name game analysis

To construct the mobility indexes used in this work we built a dataset of Italian inventors in order to resolve the "who is who" and "John Smith" problems (Trajtenberg et al., 2006). The dataset was obtained by applying a cleaning procedure to the PASTAT dataset, based on the following information included in patents:

- name of inventor;
- address of inventor;
- technological class of the patent;
- name of applicant,
- co-inventors.

The procedure consists of five steps[57]:

Step 1: we grouped all the inventors with the same name (first and last name) using the key "invname". To solve the problem of inventors whose names are misspelled, we used a metaphone matching procedure combined with a heuristic procedure. At the end of this step, all inventors with the same name were considered as unique persons (see example below). Thus, "invname" is our first definition of inventor. Since this definition suffers some limitations, the following steps were adopted in order to refine the inventor definition.

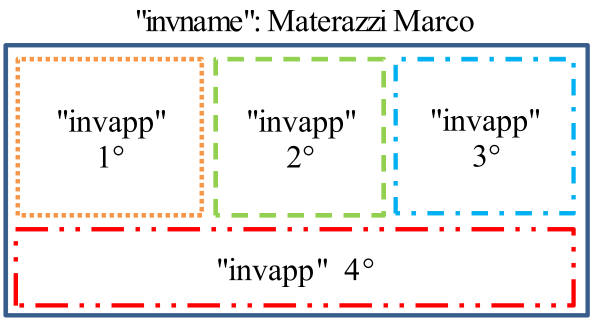
Example of invname

"invname" 1°:
all the italian inventors
with the name Marco Materazzi

[57] This procedure is developed by Thorsten Doherr.

Step 2: the previous definition of inventor ("invname") is restricted through the procedures applied in this step. All inventors of an applicant with the same name ("invname") are the same person. These inventors grouped by the key "invapp" (see example below). Therefore, we have a set of groups, i.e."invapp", which are subsets of "invname".

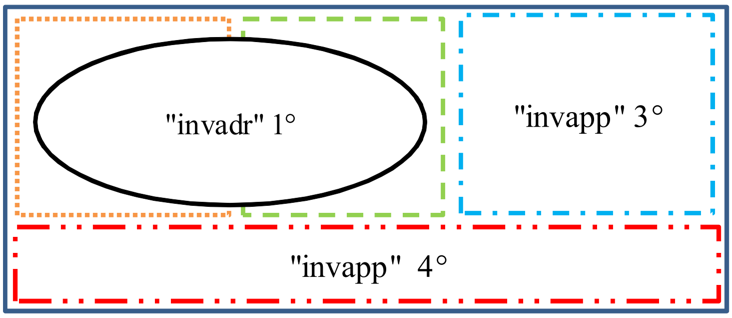
Example of invapp



for instance:
"invapp" 1°: Materazzi Marco, Barilla SpA
"invapp" 2°: Materazzi Marco, Buitoni SpA
"invapp" 3°: Materazzi Marco, De Cecco SpA
"invapp" 4°: Materazzi Marco, Ferrero SpA

Next, we use all the addresses of an inventor. All inventors with the same name ("invname") and the same addresses are an unique persons. Because an inventor can have multiple addresses, associated by patenting for the same applicant but with different home addresses, this rule connects inventors of different applicants, using the group key "invadr". Thus, inventors with the same name and address (but different applicant) are the same person ("invadr") (see example below). Thus, we have a set of groups, i.e. "invadr", which are subsets of "invname", but larger than the "invapp" sets.

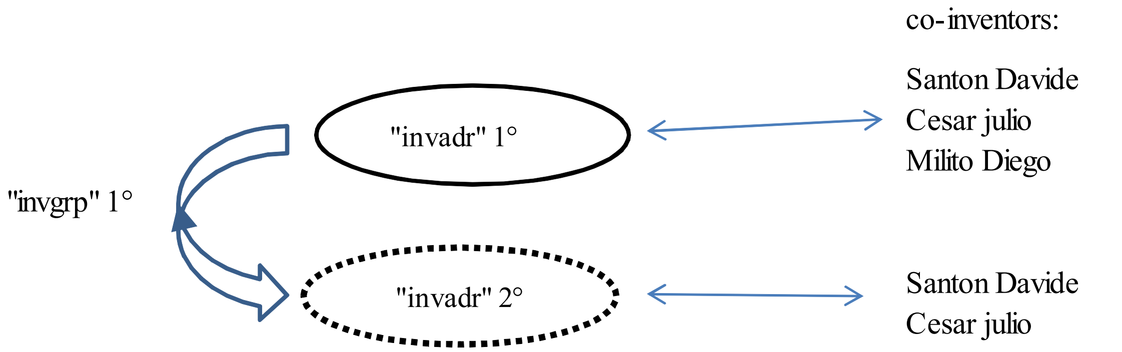
Example of invadr



for instance:
"invapp" 1°: Materazzi Marco - Barilla SpA -Via Meazza 3, Milano
"invapp" 2°: Materazzi Marco - Buitoni SpA -Via Meazza 3, Milano
"invadr" 1°

Step 3: we scrutinized inventors' collaborations. An inventor can have patents with a group of inventors (coinventorship). We assume that different inventors ("invadr") who collaborate with some members of this group are unique persons. Thus, we use coinventorship to connect inventors with the same name ("invname"), but different home addresses ("invadr") (see example below). Thus, we have a set of groups, i.e. "invgrp", which are supersets of "invadr" but subsets of "invname".

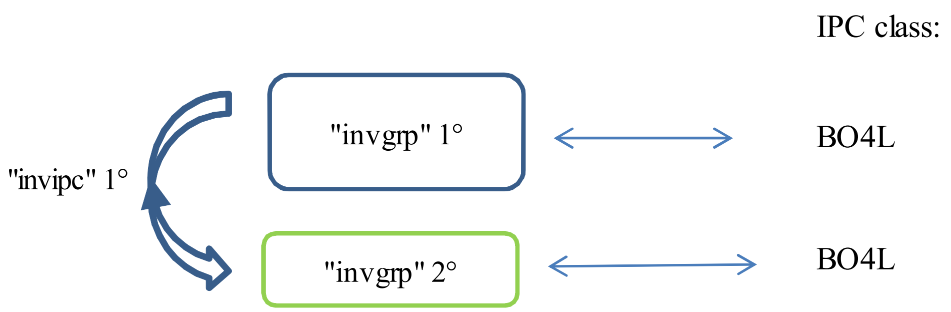
Example of invgrp



for instance:
"invadr" 1°: Materazzi Marco - Via Meazza 3, Milano - Santon, Cesar, Milito
"invadr" 2°: Materazzi Marco - Via Del Duca 1, Ascoli Piceno - Santon, Cesar
"invgrp" 1°

Step 4: we examined IPC class (IPC, 4 digits). An inventor, as defined by “invgrp”, invents patents for specific patent classifications (IPC, 4 digits). If another inventor with the same name (“invname”) invents patents in the same technology field, it is assumed that both are the same person. We grouped different inventors, as defined by “invgrp”, with patents in the same technology field (see example below). Thus, we have a set of “invipc” groups which is a subset of “invname”, but larger than the “invgrp” definition.

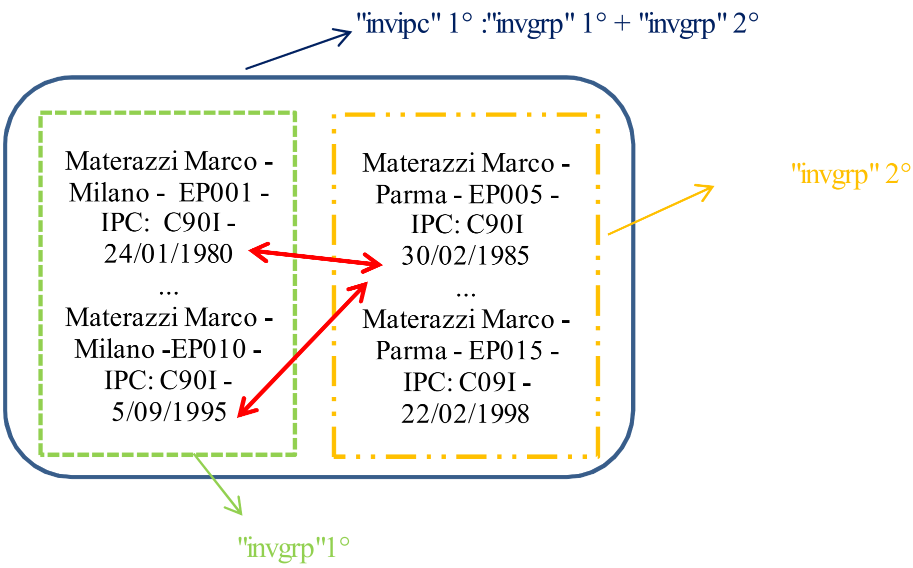
Example of invipc



for instance:
"invgrp" 1°: Materazzi Marco - Santon, Cesar, Milito - B04L
"invgrp" 2°: Materazzi Marco - Baggio, Maradona - B04L
"invipc" 1°

Step 5: because of the broad definition of “invipc”, we applied a final restriction to minimize false positives: an inventor, defined by “invipc”, cannot patent at different locations in the same time period, where location is defined by “invgrp” and time period is defined by the first and the last filing date of patents at this location . If there is no overlap we keep the key “invipc” otherwise the former key “invgrp” is used as the final “inventor” key.

Example of overlap



for instance:
"invgrp" 1°: Materazzi Marco - C90I - Patents from 1980 to 1995 in Milano.
"invgrp" 2°: Materazzi Marco - C90I - Patents from 1985 to 1998 in Parma.
An overlap between the two "invgrp" (i.e. the inventor patented at different locations in the same time period) groups is considered to include different persons (we do not use "invipc" 1°).

Appendix 3.D. Definition and source of the variables

Table 3.D1. Definition and source of the variables

Variable	*Definition*	*Source*
log[(GDP/POP) *i,t* / (GDP/POP) *i, t-1)]*	GDP growth per capita	GDP: ISTAT Population: EUROSTAT
GDP/POP *t-1*	GDP per capita (at time *t-1*)	GDP: ISTAT Population: EUROSTAT
POP *t-1*	Population (at time *t-1*)	EUROSTAT
INV/POP *t-1*	Fixed investments per capita (at time *t-1*)	Fixed investments: ISTAT Population: EUROSTAT
(PAT/R&D) *t-1*	Ratio between the number of patent applications at time *t-1* and the total expenditure in R&D (both public and private) at time *t-1*	Patent applications: authors' elaboration on PATSTAT data
R&D/POP *t-1*	R&D expenditure (public and private) per capita (at time *t-1*)	R&D: ISTAT Population: EUROSTAT
Mob_intra *t-1* / SI *t-2*	Ratio of number of inventors that move within a region during the period *t-1* and the stock of inventors (calculated with the perpetual inventory method and depreciation rate of 5%) in the region at period *t-2*	Authors' elaboration on PATSTAT data
Mob_inflow *t-1* / SI *t-2*	Ratio between number of inventors that enter the region (inflows) during the period *t-1* and the stock of inventors in the region (calculated by the perpetual inventory method and depreciation at 5%) at period *t-2*	Authors' elaboration on PATSTAT data
Mob_outflow *t-1* / SI *t-2*	Ratio between number of inventors that leave the region (outflows) during the period *t-1* and the stock of inventors in the region (calculated using the perpetual inventory method and depreciation at 5%) at period *t-2*	Authors' elaboration on PATSTAT data
Mob_netflow *t-1* / SI *t-2*	Ratio of net inflows of inventors (inflows minus outflows) in the region during the period *t-1* and stock of inventors in the region (calculated using the perpetual inventory method and depreciation at 5%) at period *t-2*	Authors' elaboration on PATSTAT data

(continued)

Table 3.D1 (continued)

Variable	*Definition*	*Source*
(Forw_inter/PAT) *t-1*	Ratio between number of interregional forward patent citations at time *t-1* and number of patent applications at time *t-1*. Interregional forward patent citations: number of patents (at time *t-1*) of inventors residing in the region cited by patents (during the period *(t-1) - t*) of inventors residing in other regions. Patents with more than one inventor residing in the same region (citing or cited), are counted only once.	Forward citations: KITES Patent applications: authors' elaboration on PATSTAT data
(Forw_intra/PAT) *t-1*	Ratio between number of intra-regional forward patent citations at time *t-1* and number of patent applications at time *t-1*. Intra-regional forward citation: number of patents (at time *t-1*) of inventors residing in the region cited by patents (during the period *(t-1) - t*) of inventors residing in the same region. Patents (citing or cited) with more than one inventor residing in the region, are counted only once.	Forward citations: KITES Patent applications: authors' elaboration on PATSTAT data
(Back_inter/PAT) *t-1*	Ratio between number of interregional backward patent citations at time *t-1* and number of patent applications at time *t-1*. Interregional backward patent citations: Number of patents (at time *t-1*) of inventors residing in the region citing patents (of the period *(t-1) - (t-2)*) of inventor residing in other regions. Patents with more than one inventor residing in the same region (citing or cited), are counted only once.	Backward citations: KITES Patent applications: authors' elaboration on PATSTAT data
(Back_intra/PAT) *t-1*	Ratio between number of intra-regional backward patent citations at time *t-1* and number of patent applications at time *t-1*. Intra-regional backward patent citations: Number of patents (at time *t-1*) of inventors residing in the region citing patents (of the period *(t-1) - (t-2)*) of inventors residing in the same region. Patents (citing or cited) with more than one inventor residing in the region, are counted only once.	Backward citations: KITES Patent applications: authors' elaboration on PATSTAT data
applicant_share *t-1*	Ratio of number of applicants in a region (at time *t-1*) and the total number of Italian applicants (at time *t-1*)	Authors' elaboration on PATSTAT data

Chapter 4

Sources of Spillovers for Imitation and Innovation[58,59]

4.1 Introduction

Technological spillovers[60] have been discussed in the economics discipline. Applications include economic growth (e.g., Romer, 1986), R&D incentives (e.g. Geroski et al., 1993; Hanel and St-Pierre, 2002; Cohen and Levinthal, 1989), R&D alliances (Caloghirou et al., 2003) and joint ventures (d'Aspremont and Jaquemin 1988; de Bondt 1997). Their relevance for business practice is demonstrated by Mansfield (1985), which reports how quickly information affecting development decisions and information on the nature and operation of products and processes, leak to competitors.

In studies of firm-level productivity, spill-overs typically are modelled based regressions using total R&D of all firms in the same industry, and sometimes R&D in other industries. These measures have some shortcomings, such as their implicit limitation to a certain geographic area. Also, many account for spillovers from rivals (R&D in the same industry), or firms in other industries (see e.g. Hall et al., 2010, for a survey). Thus, by construction, conventional measures assume that the recipients of spillovers utilize them to the same extent since only one coefficient can be estimated for each variable. Also, firms in other industries may be suppliers (upstream firms) or

[58] This chapter represents a slightly different version of the paper: Cappelli, R., Czarnitzki, D., Kraft, K., 2012. Sources of spillovers for imitation and innovation, mimeo.
[59] We are grateful to the MIP team at the ZEW for providing the survey data.
[60] The term spillovers is used to indicate the voluntary or unvoluntary exchange of knowledge between firms or between firms and research institutions such as universities.

customers (downstream firms). Spillovers from customers and spillovers from suppliers can differ significantly with respect to how much they contribute to innovation.

Several studies also include the effect of research institutions, usually universities. This body of work is interested mainly in the effect of spillovers resulting from regional associations or explicit cooperation with universities. The effect is estimated generally by means of a knowledge production function, with patents, innovation counts or total factor productivity growth as the endogenous variable.[61]

Although it is well known that spillovers not only stimulate innovation but also induce imitation, the latter effect has been rather under-investigated. Econometric studies usually use an indicator for innovativeness, such as R&D expenditure, numbers of patents or innovations, but tend not to include a variable for imitation since these data are less easily available. However, it is acknowledged that information leaks out via informal communication or by scientists being poached by competing firms.

The view that spillovers induce innovation and imitation is positive since innovation is valued positively as knowledge in the economy increases. However, imitation or the copying of innovations developed by others, is generally cheaper than engaging in innovation. Imitators costs are lower which allows them to outbid rivals, which negatively affects the incentive to execute R&D for innovation. Spillovers that lead to imitations might be considered negatively, but the total costs to the economy will be lower. However, sceptics argue that spillovers have a dampening effect on the incentives to perform R&D if “inputs” come for no cost from outside, which, in turn affects the whole economy and is the basic reason for a patenting system.

[61] E.g. Jaffe (1989), Audretsch and Feldman (1996), Audretsch et al. (2005), Ponds et al. (2010).

In this work, we use some measures of spillovers derived from a survey, which overcome the limitations of the measures traditionally used, discussed above. We also distinguish the importance of spillovers for two types of innovative activity, i.e. original innovation based on own inventive activity versus imitation, and argue that heterogeneous effects can be expected based on the source of the spillovers. This differentiation allows us to distinguish between the sources of spillovers. The data include information on whether spillovers are from competitors, customers, suppliers or research institutions. This may be important since information from some sources may be useful for imitation, but less beneficial for innovation. Possible examples are inflows from competitors compared to spillovers from research institutions. Mansfield (1998) states that about 15% of new products in seven US industries in the period 1986-1994 and 11% of new processes could not have been developed in the absence of recent academic research.

The chapter is organized as follows. Section 4.2 discusses the background to the relationship between knowledge spillovers and firm performance and especially the methodology adopted for the empirical analysis. We show that our approach has some advantages with respect to the literature. Section 4.3 discusses the data and variables used in the analysis, in more detail. Section 4.4 discusses the econometric procedure adopted to take account of possible sources of bias. Section 4.5 presents the results of our estimates. Section 4.6 offers some final considerations.

4.2 Background to the study

Spillovers are extremely important in practice. According to Mansfield (1985) information on development decisions, in 12 to 18 months leaks to competitors, and information on the exact operation of products and processes reaches rivals in 12 to 15 months.

The importance of spillovers is reflected in the many applications in economics, which is an indication of the importance of this topic. One example is endogenous growth theory (Grossman and Helpman, 1991) where the knowledge produced by a company enhances productivity industry-wide and, thus, is not subject to decreasing returns. Many microeconomic contributions consider how spillovers determine firm behaviour. By affecting profitability, incoming and outgoing spillovers clearly influence the incentives to engage in R&D projects (e.g. Geroski et al., 1993; Hanel and St-Pierre, 2002; Czarnitzki and Kraft, 2012). This in turn stimulates the formation of alliances in form of research joint ventures (d'Aspremont and Jaquemin, 1988; de Bondt 1997). Spillovers can be regarded as positive externalities, which is an argument for subsidizing R&D efforts (Arrow, 1962).

Given the importance of knowledge spillovers, there is a large literature on the impact of knowledge inflows on firms' economic performance. In the succeeding sections, we discuss in more detail both the methodology used in this study, its limitations and our approach to overcoming them.

4.2.1 Empirical analysis

The impact of knowledge spillovers on the innovation and economic performance of a firm traditionally is analysed using the sum of R&D of other firms to measure knowledge spillovers. The methodological approach in these works consists of estimating the impact of the knowledge capital on total factor productivity or firm profit (or measures of innovation output in the case of the knowledge production function). The firm's knowledge capital is assumed to be a function of the knowledge generated by internal R&D activities (i.e. R&D expenditure) and the inflow of external knowledge (i.e. R&D expenditure by other firms). Some early studies model external knowledge as the unweighted or weighted sum of intra-industry R&D.[62] This means that the firm can appropriate only knowledge generated by rival firms and ignores vertical spillovers such as knowledge from suppliers or customers. Other studies try to control for vertical spillovers using the R&D of other industries.

These analyses of knowledge spillovers have several shortcomings. First, the geographic extent of the sources of knowledge spillovers is limited. Most studies consider only R&D developed by firms located in the same country which ignores inter-firm linkages across national borders. For instance firms can export and compete in the international market and/or have international suppliers. Another aspect neglected by this measure of knowledge spillovers is the temporal dimension of the R&D effects (see e.g. Mansfield et al., 1971). The development of an R&D project can take two years or even longer. Also, this measure of knowledge spillovers assumes that all firms use external knowledge in the same way, and to the same extent, meaning that it has the same impact in terms of innovations or economic output. It overlooks differences among firms in terms of capacity to absorb external knowledge and capacity to transform knowledge into new ideas or commercial products. At the same time, this measure assumes also that different spillovers sources have the same impact on outputs, whereas different sources of knowledge spillovers may have different impacts on the innovative processes. Finally, this measure takes no account of spillovers from R&D activities performed by universities and other research institutions.

[62] They use the following formula: $KS_{it} = \sum_{j \neq i} w_{jt}\ R\&D_{jt}$, where *KS* are spillovers to firm *i*, *R&D* are the R&D expenditures of the firm *j* (*j≠i*), and *w* is a weighting matrix.

Given the lack of availability of firm level data, knowledge spillovers from universities or research institutions are analysed mainly at the country or regional level. For example, the pioneering work of Jaffe (1989) uses US state level data and a knowledge production function that relates patents developed by firms to university R&D, in order to investigate whether knowledge spillovers from universities to firms is geographically localized. Several works extend Jaffe's analysis, considering smaller geographic units (see e.g. Audretsch and Feldman, 1996) and European countries (see e.g. Ponds et al., 2010).

The knowledge spillovers literature generally does not consider imitation activities. This is a major limitation because an important part of the firm's profit may originate from products already existing in the market.

4.2.2 Proposed extension

The empirical analysis developed here is based on a model where the firm's sales are a function of its knowledge capital. In order to takes account of imitation and innovation, we distinguish between sales from imitations and sales from innovations. To measure knowledge capital we consider the processes of both generation and diffusion of knowledge. In the latter case, to overcome the limitations described above, we consider various sources of knowledge spillovers: from rivals, customers, suppliers and research institutions. Thus, we are able to analyse the impact on firm's economic performance of different sources of knowledge spillovers and also whether this impact is different for imitation and innovation.

Spillovers between firms

Spillovers typically are seen as part of the core process of knowledge diffusion. One view of knowledge diffusion is that it is a cost free input, which enables imitation of innovations developed by competitors. Imitation is usually cheaper than conducting own R&D, but is not costless (Mansfield et al., 1981). However, the imitator will have lower overall costs and will be able to outbid rivals. If spillovers ease the imitation of existing products, the information most likely originates from producers within the same industry.

However, spillovers also can induce the company to perform its own innovative activity and particularly if the input is a novel idea or a major innovation that has many potential applications. This type of spillover may also come from a competitor, but could result from contacts with customers and suppliers (and research institutions - as discussed below).

Spillovers from customers can reduce the risk associated with introducing new products to the market and can result in higher demand and increased sales, especially when products require adaptation due to their complexity or novelty (see e.g. von Hippel, 1988; Herstatt and von Hippel, 1992; Tether, 2002). Spillovers from suppliers can result in process innovations for the production of existing products and also in improvements to existing products, e.g. in terms of design (see e.g. Suzuki, 1993; Karnath and Liker, 1994). It has been found also that involvement of suppliers can increase product innovation in mature industries (Eisenhardt and Tabrizi, 1995).

Consequently we argue that both horizontal and vertical spillovers may affect the innovation performance of firms, where spillovers from competitors should clearly lead to higher imitation in the industry. Spillovers from customers and suppliers may affect both imitation and performance of original innovations.

Spillovers from research institutions

As noted in the introduction, empirical research has studied the role of research institutions, especially universities, on the innovative output of firms. Much of this research focuses on the regional aspects of these spillovers such as the impact on firms in close proximity to a university. Other contributions look at networks or spin-offs. The literature on regional economics and location theory emphasizes the role of spillovers as one reason for agglomeration (see Feldmann, 1999, for a survey), which includes a location choice near to a university.

Work on spillovers resulting from R&D collaborations with universities emphasizes that academic research typically is complementary to the firm's existing knowledge resources and thus contributes significantly to the ability of the corporate sector to create innovations (Tether and Tajar, 2008; Baba et al., 2009) including "key innovations" (Thursby and Thursby, 2006).

We hypothesize that spillovers can be input to imitation as well as innovation and that different sources of knowledge spillovers may be used for different purposes. This is empirically tested in our anlaysis.

4.3 Data and variables

Our study is based on a sample of German firms surveyed in the year 2003, i.e. data correspond to 2000-2002. The MIP is a survey that is conducted by the ZEW, and has been conducted annually since 1992 (see Janz et al., 2001, for more information on the data collection process).[63]

[63] MIP is the German part of the CIS, a harmonized survey across EU member states. For a detailed description of the CIS, see Eurostat (2004).

Our sample covers firms in the manufacturing sector. Since we are interested in the effect of spillovers resulting from innovation activity, our sample includes only innovating firms, leaving a sample of 1,007 firms. An innovating firm is defined in accordance with the OECD OSLO manual, and innovation data on the business sector were collected according to international guidelines (Eurostat and OECD, 2005).

Dependent variable

The survey allows us to split sales into three components: a) sales of products new to the market in 2000 to 2002, b) sales of products already on the market before 2000, but new to the firm's product portfolio in the period 2000 to 2002, c) continuing sales of already existing products. We use definition (a) to measure original innovation, and (b) to measure imitation. The dependent variables are measured as percentage shares in total sales. As a robustness check we present the regressions using the log of sales volume of products (a) and (b) as dependent variables (see the Appendix).

Spillovers measures

The most important explanatory variables are spillovers measures. The MIP 2003 survey asked firms to indicate information spillovers that were *indispensable* for the development of an own product or process. Four different sources of spillovers were distinguished: suppliers, customers, competitors, research institutions. We use four dummy variables, i.e. one for each spillovers sources, which are equal to 1 if the firm ranked the spillovers as *indispensable*. Note that the question was worded such that an affirmative response implied possession of the necessary absorptive capacity to make use of the information received.

Our hypothesis is that the source of spillovers is related to the firm's output. Spillovers from rivals will usually convey information on existing goods and is more likely useful for imitation rather than innovation. Spillovers from customers, suppliers or research institutions might have different effects since their originator is not active in the same industry. Therefore, a positive impact on innovative output is possible. Hence we posit different effects of spillovers depending on the source.

Other control variables

In order to test whether our spillovers measures derived from the survey are superior to the more commonly used measure, we include the log of industry R&D in the regression, ln(*INDUSTRY_R&*D). This measure captures within industry spillovers in studies that estimate production functions (see Hall et al., 2010, for an overview). The internal knowledge stock of a company is probably an important determinant of sales of new products. Since we have only cross-sectional data, we cannot use past R&D expenditure; we linked our sample to the German Patent and Trademark Office database, which contains information on patent applications filed with the German and European patent offices since 1978. These data enable us to construct a stock of "successful" outcomes of R&D projects, for each firm, for a long time series. The patent stock (*PS*) of a firm is calculated using the perpetual inventory method with a constant depreciation rate as

$$PS_{it} = \left(1-\delta\right) PS_{i,t-1} + PA_{it} ,$$

where PA is the number of patent applications in year t and δ is the constant depreciation rate, which is set to 15% (see Griliches and Mairesse, 1984, for a more detailed description). As patents are a narrower measure than R&D knowledge stock, we also include R&D spending as a proxy for the non-patented knowledge stock. We use the R&D intensity, *RDINT*, measured as R&D divided by sales and use its squared term to allow for potential decreasing marginal returns.

The share of sales volume exported (*EXPORT*) at firm level, imports relative to domestic production (*IMPORT*) and the Herfindahl concentration index *(HERF)* at industry level, are used to control for the firm's competitive environment. We include the variable firm age (*AGE*), because younger firms might be more innovative than older ones. Size is based on the number of employees (*EMP*). We use capital intensity *(KAPINT)* defined as fixed assets divided by the number of employees to indicate capital requirements. Since at least a part of this capital expenditure is sunk, this variable is expected to present barriers to entry. Ten industry dummies control for other differences across sectors not measured by the controls described.

4.4 Estimation issues

As discussed above, the dependent variables in our analysis are the firms' sales on imitation and innovation products. Since not every firm realizes sales of both innovations and imitations, we estimate Tobit models that account for censoring of the data. We use a log transformation of the variables to approximate the normality assumption underlying the Tobit model. Because we cannot take a log of zero values, we impute the minimum observed positive value for those observations. The bias arising from this transformation should be minimal because we just consider the smallest positive observation as censored.

In order to avoid endogeneity of the right-hand side variables, we use lagged values wherever possible. The survey enquires about the innovation behaviour of firms in years 2000 to 2002. Our dependent variables refer to sales in 2002 (= t), and we can make use of one lag for the regressors. Where data are from different sources (patent stock, Herfindahl index, imports) we use the information up to the year 2000, i.e. two lags, to ensure that the data apply to the beginning of the survey period, and the risk of direct endogeneity is reduced.

Employment, exports, R&D intensity and capital intensity are measured at 2001 (= t–1). The spillovers measures account for the time window 2000 to 2002. Table 4.1 presents the descriptive statistics of the variables used.

Table 4.1. Descriptive statistics (1007 observations) for the year t = 2002

Variable	Description	Mean	S. D.	Min	Max
SALES_NEW (t)	Sales from market novelties (EUR million)	13.21	128.58	0	3718.75
SALES_IMIT (t)	Sales from imitation (EUR million)	31.00	205.57	0	4224.00
%_SALES_NEW (t)	Share of sales from market novelties (%)	9.11	16.99	0	100
%_SALES_IMIT (t)	Share of sales from imitation (%)	19.12	21.21	0	100
IMPORT (t-2)	Imports (imports / domestic production)	0.38	0.33	0.07	2.19
HHI (t-2)	Herfindahl index in t-2	54.32	77.51	3.21	642.35
EMP/1000 (t-1)	Employment (in thsd.)	0.74	2.99	0.001	41.75
RDINT (t-1)	R&D spending (t-1) / Sales (t-1)	0.04	0.06	0	0.45
PS/EMP (t-2)	Patent Stock per employee (t-2)	0.02	0.05	0	0.38
EXPORT (t-1)	Exports (exports in t-1 / sales in t-1)	0.29	0.26	0	1
AGE	years elapsed since foundation	33.62	36.26	2	203
KAPINT (t-1)	Capital intensity [physical assets in million EUR (t-1) / employment (t-1)]	0.05	0.05	0.01	0.49
Ln(INDUSTRY_RD)	Log of R&D at the industry level	8.128	1.311	3.714	10.023
Dummy variables for spillovers					
Competitors		0.20	0.40	0	1
Customers		0.51	0.50	0	1
Suppliers		0.17	0.38	0	1
Research Inst.		0.11	0.31	0	1

Note: 10 Industry dummies omitted.

4.5 Estimation results

Our results are obtained through two separate estimates, one considering sales from innovation as the dependent variable and one considering sales from imitation as the dependent variable. The number of observations in both estimates is 1,007, i.e. the number of firms in our sample.

The results are presented in Table 4.2. Note that the results are quite robust across the two specifications of the dependent variables. We find that spillovers from universities and customers contribute significantly to firm sales of market novelties, but have no effect on imitation. The marginal effects are 45% and 41% for the market novelty regression. Since on average 9% of firms' total sales are based on market novelties, the estimated marginal effects imply an increase to 13.2% if a firm indicated indispensable spillovers from academic science and to 12.8% for firms that received indispensable spillovers from customers. Spillovers from rivals, however, have a high and significant effect on the sales of product imitations. The marginal effect is 42%, which corresponds to an increase in the share of new, imitated products in total sales from 19% to about 27%. As expected, other sources do not matter for imitation.

These results are interesting in relation to the sources of spillovers useful for imitation versus innovation. Information from rivals contributes to imitation, since the knowledge is probably about already developed products. In contrast, knowledge inflows from research institutions and customers is usually not about products and processes already in use and is more likely to be an input that induces innovative activity. This is clearly one aim of publicly funded research. In the case of inducements from customers the company will probably get information on market potential, which, in turn, can be used in the development of the product being demanded.

In our view these results provide important information on a rather overlooked aspect of spillovers. Since most studies use an indicator for innovation as the dependent variable, imitation is largely ignored. Public opinion on innovation versus imitation tend to be more negative towards imitation. However, imitation exists and contributes, and its consideration and explanation is of relevance. As we have shown elsewhere (Czarnitzi and Kraft, 2012), spillovers from rivals increase profits and the present work shows that this is not only the result of a stimulus for internal innovative activity. This empirical research contributes to our understanding of the how markets work, and what makes a successful firm.

Among the control variables, patents are facilitating sales with market novelties. On the one hand, patents might seem a good proxy for the firm's inventive activity because only novel technological discoveries can be patented. On the other hand, the importance of patents in the market novelty equation may indicate that rivals cannot easily compete away excess returns through imitation since patents provide (at least some) protection. R&D shows an inverted U-shape in both the regressions which peaks at the right tail of the R&D distribution. This indicates a positive relationship between R&D and the product innovation variables. *RDINT* and $RDINT^2$ are jointly significant at the 5% level in both equations. This confirms the relevance of the non-patented knowledge stock.

In contrast to other studies we found no effect of industry R&D. It is possible that our spillover measures are better representations of the interaction among firms, than aggregated R&D expenditure. Larger firms imitate more. Finally, capital intensity is positively associated with market novelties. This could be due to the existence of barriers to entry when capital requirements are high. Such firms would then be better protected against imitation by competitors.

Table 4.2. Tobits on log of innovation/imitation shares in total sales

Variable	Market novelties		Imitation	
RDINT	10.757		10.391	***
	(6.978)		(3.869)	
RDINT^2	-11.372		-30.513	**
	(21.073)		(12.155)	
PS/EMP	8.987	***	-0.974	
	(3.408)		(1.864)	
IMPORT	1.266		-0.304	
	(1.164)		(0.635)	
HHI	-0.005	*	0.001	
	(0.003)		(0.002)	
ln(EMP)	0.263	**	0.194	***
	(0.130)		(0.061)	
EXPORT	-0.979		-0.016	
	(0.822)		(0.419)	
ln(AGE)	-0.242		-0.043	
	(0.217)		(0.105)	
KAPINT	8.516	**	-1.480	
	(4.026)		(2.112)	
ln(INDUSTRY_RD)	-0.107		-0.116	
	(0.524)		(0.276)	
Spillover measures				
COMPETITORS	-0.667		0.609	**
	(0.493)		(0.251)	
CUSTOMER	1.273	***	0.027	
	(0.413)		(0.210)	
SUPPLIERS	-0.015		-0.124	
	(0.510)		(0.261)	
RESEARCH INSTITUTIONS	1.247	**	0.194	
	(0.591)		(0.318)	
INTERCEPT	-3.602		0.403	
	(4.527)		(2.431)	
Test on joint significance of industry dummies	$\chi2$ (10) = 26.81***		$\chi2$ (10) =15.68	
Test on joint significance of *RDINT* and *RDINT*^{2}	$\chi2$ (2) = 6.44**		$\chi2$ (2) = 7.27**	
Log-Likelihood	-2129.98		-2237.90	

Note: *** (**, *) indicate a significance level of 1% (5%, 10%). Standard errors in parentheses. Tobit models would lead to inconsistent coefficient estimates if heteroschedasticity is present. Therefore we tested for heteroschedastic errors. Homoschedasticity is rejected. Consequently, heteroschedasticity was modelled as groupwise multiplicative where the variance term includes a full set of industry dummies and 5 size class dummies based on employment.

4.6 Conclusion

We have presented the results of an empirical study concerning the impact of spillovers from different sources on innovation sales. Furthermore we distinguish between sales with actual market novelties and product imitation. Spillover from different sources do not have the same effects. While spillovers from competitors matter for imitation, customers and research institutions deliver valuable knowledge for sales with market novelties. We would suggest that survey data can overcome some of limitations of “traditional” spillovers measures, which typically are not measured in appropriate geographic areas, do not distinguish among sources, and do not allow for heterogeneous impacts across a sample of firms to be explored in an regression analysis.

Spillovers are positive externalities and therefore are valued positively by the receiving company. In contrast, the spillovers producing firm will assess this externality negatively because it benefits competitors in the same industry. These conflicts do not arise if spillovers come from sources outside the industry. This is the case of spillovers from research institutions and customers. This information is used for innovation rather than imitation and is the reason for its uncontroversial appraisal.

Our results support the public funding of research institutions such as universities. Research institutes help to promote innovation in private firms. Although universities receive part of the return on their research results because they hold the intellectual property rights, it is likely that the gains for the economy at large will be even greater, which is support for subsidization.

Appendix (Chapter 4)

Table 4.A1 show the results of the Heteroskedastic Tobits on log sales volume of innovation/imitation.

Table 4.A1. Heteroskedastic Tobits on log sales volume of innovation/imitation (1007 observations)

Variable	Market novelties		Imitation	
RDINT	4.948		6.075	*
	(5.043)		(3.202)	
RDINT^2	-4.155		-20.817	**
	(14.586)		(9.607)	
PS/EMP	7.113	***	-1.266	
	(2.593)		(1.698)	
IMPORT	1.441		-0.365	
	(1.011)		(0.631)	
HHI	-0.004		0.001	
	(0.002)		(0.001)	
ln(EMP)	0.603	***	1.006	***
	(0.118)		(0.058)	
EXPORT	0.131		0.581	
	(0.623)		(0.379)	
ln(AGE)	-0.155		-0.074	
	(0.186)		(0.105)	
KAPINT	7.368	**	-0.432	
	(3.273)		(1.960)	
ln(INDUSTRY_RD)	-0.291		-0.351	
	(0.509)		(0.289)	
Spillover measures				
COMPETITORS	-0.289		0.695	***
	(0.388)		(0.226)	
CUSTOMERS	0.795	**	0.034	
	(0.322)		(0.185)	
SUPPLIERS	0.124		-0.130	
	(0.402)		(0.236)	
RESEARCH INSTITUTIONS	1.145	**	0.450	
	(0.461)		(0.289)	
INTERCEPT	-4.928		-3.195	
	(4.304)		(2.501)	
Test on joint significance of industry dummies	$\chi2$ (10) = 19.21**		$\chi2$ (10) = 16.22*	
Log-Likelihood	-2016.27		-2167.61	

Note: *** (**, *) indicate a significance level of 1% (5%, 10%). Standard errors in parentheses. The heteroskedasticity term includes a full set of industry dummies, and five size class dummies based on employment.

Chapter 5

Conclusion

In this thesis I analysed the determinants of the diffusion of knowledge and the impact of these determinants on economic growth in an attempt to overcome some limitations in the existing literature, which often fails to cover aspects linked to the nature of knowledge flows. These limitations affect both the interpretation of the results of scientific work and the guidance provided to government institutions about the formulation of policies appropriate to support innovation and economic growth.

This thesis is comprised of three studies that analyse the diffusion of knowledge between regions (Chapter 2) and the impact of knowledge flows on the economic growth of regions (Chapter 3) and firms (Chapter 4). These works are linked by the search for an appropriate measure to capture the multidimensional nature of the phenomenon of diffusion of knowledge.

Knowledge is an immaterial good some of which can be codified in documents. Some knowledge is un-codifiable either because it cannot be articulated in documents or because codification would be too expensive (Cowan et al., 2000). This type of knowledge is described as tacit and is embedded in individuals or organizations. The difference between the two types of knowledge is critical for many economic and policy issues. For instance, the tacit part of knowledge is an important element used to explain the phenomenon of geographical aggregation of the firms of an industry. Access to the knowledge embedded in individuals and, thus, in territories is an important source of the localization economies (Marshall, 1920). The nature of tacit knowledge can be discussed also in terms of intellectually property rights. Tacit knowledge is a key element enabling firms to achieve the economic benefits from invention, despite the detailed disclosure of in patent documents. The tacit part of the knowledge contained in a patent at least partially avoids the phenomenon of free-riding and represents the element of knowledge that justifies public subsidies for R&D undertaken by firms.

Empirical analysis of knowledge flows mainly use measures such as patent citations and the stock of foreign R&D, which capture the diffusion of codified knowledge but tend to neglect the channels of diffusion of tacit knowledge. Given the strong link between the tacit component of knowledge and the territory in which the knowledge is generated, it follows that these analyses ignore important spatial aspects of the diffusion of knowledge. Although some works do specifically consider the channels of diffusion of tacit knowledge, attention to both components of knowledge is rare. In the works that comprise this thesis a multidimensional approach is used that considers the mechanisms of diffusion of both tacit and codified knowledge.

This concluding chapter summarizes the findings and discusses the limitations of this thesis, and suggests possible lines of future research.

Knowledge flows at the spatial level

In the first study (Chapter 2), we analysed the pattern of diffusion of knowledge among European regions, in the period 1981-2000. The literature on knowledge flows shows consistently that the diffusion of knowledge is geographically localized. In particular, it shows that the diffusion of knowledge is more likely to be within countries, and decreases with increasing distance between regions (Maurseth and Verspagen, 2002; Bottazzi and Peri, 2003; Fisher et al., 2009). However, these works provide only sets of static pictures of the diffusion of knowledge between regions and do not provide information about the evolution over time of knowledge flows and of their determinants. In this regard, the development and the higher diffusion of ICT and the greater integration among countries suggest that knowledge flows are becoming less geographically localized. Finally, most existing work makes use of patent citations to measure knowledge flows (Jaffe et al, 1993; Maurseth and Verspagen, 2002) or inventor collaborations (Picci, 2010),[64] but not both measures.

This thesis analysed the determinants of knowledge flows between European regions, measured by inventor citations and collaborations. It also analysed the evolution over time of the impact of these determinants on knowledge flows, and whether the process of European integration has had an impact on the diffusion of knowledge. These aspects were analysed using a modified version of the gravity model and PPML estimates (Santos Silva and Tenreyro, 2006).

We showed that knowledge flows (measured by inventor citations and collaborations) between two regions are negatively affected by physical distance and by the existence of a national border separating them. The results confirm that diffusion of knowledge is geographically localized and that the impact of the determinants is greater for inventor collaboration than inventor citation. This can be explained by the fact that collaboration requires face-to-face contact, while inventor citation does not require personal contact and relates to the diffusion of only the codified part of the knowledge contained in a patent document. It occurs through simple reading of the document (e.g. from the EPO database available online).

We found that the results showed different time trends for inventor collaboration and citation. On the one hand, the tendency for an inventor to collaborate within the home country decreases, but on the other hand, the propensity of an inventor to cite other

[64] As discussed above, patent citations are a good measure of the diffusion of codified knowledge and inventor collaborations are a valid measures of the diffusion of tacit knowledge through personal contact.

inventors in the home country increases.[65] In addition, the impact of physical distance decreases only for inventor collaboration. Therefore, these results show that the hypothesized reduction over time, of the impact of geographical distance, applies only to inventor collaboration. However, it is difficult to explain the greater localization of patent citations. For instance, given that an important share of citations is added by patent examiners, it could be that, over time, there has been an increase in the tendency for patent examiners to insert citations, at the national level.

We also examined the hypothesis that increased European integration has affected the diffusion of knowledge among EU member countries. The period covered by the sample in this thesis covers the processes of EU enlargement following the entry of Spain and Portugal in 1986 and of Austria, Sweden and Finland in 1995. The effect of EU enlargement has been analysed in the trade flows literature (Baldwin, 1995; Bussière et al., 2008; Carrère, 2006). Given that European integration has reduced the institutional barriers between countries it might be assumed that international knowledge flows have increased. The results confirm that knowledge flows between old and new EU member states have increased, although with significant difference between the two measures of knowledge flows. In particular, international collaborations between old and new EU members increased following both EU enlargement processes, but that international inventor citations between old and new EU members increased only after the second EU enlargement and only for citations from the inventors in old EU member countries to inventors in new EU member states.

This work contributes to the literature in three ways. First, it is the first attempt to provide a direct comparison between European regions and between two channels of

[65] Follows the trade literature (Bhavnani et al., 2002), patent citations give rise to the "missing globalization puzzle", namely the absence of a reduced impact of geographical factors.

knowledge diffusion, where inventor collaborations are used to measure the diffusion of tacit knowledge and inventor citations are used to capture the diffusion of codified knowledge. Second, this work extends previous work on the evolution over time of the patterns of knowledge flows between European regions because it covers a longer time period and makes use of more robust methods of analysis. This thesis is one of the few attempts to analyse the impact of the European integration on knowledge flows.

It should be noted that we do not take account of "rent spillovers" (Grilliches, 1979; van Meijl, 1997) and also we consider only a specific form of knowledge flows between regions, which occur when both regions, i.e. knowledge generating and knowledge receiving, are active in R&D and international patenting. The channels of inventor citation and collaboration convey knowledge relevant for innovation.

We can point to two directions for future research. First, analysing the diffusion of knowledge through patent citations would distinguish between the sectors of the knowledge generating patents and the knowledge receiving patents. It could be argued that geographical proximity is less important if cited and citing patents are in the same sector because the inventor can learn from knowledge contained in the patent document, and that it is important to be located close to knowledgeable people if the patent (i.e. knowledge) is in a different field. Second, it would be interesting to analyse in more detail the impact of European integration on knowledge flows using different measures of integration.

Knowledge flows and economic growth: regional level

In the study in Chapter 3 we analysed the relationship between knowledge capital and the economic growth of Italian regions for the period 1995-2007. Knowledge is an important engine of economic growth. In contrast to the neoclassical standard model

(Solow, 1956, 1957), the technology gap approach (Fagerberg, 1994; Fagerberg et al. 1997) assumes that knowledge is the most important factor explaining differences in economic growth. In particular, the technology gap approach assumes that economic growth is a function of internal knowledge generating and knowledge diffusion processes.

A fundamental question related to empirical analysis of the impact of knowledge capital on economic growth is how to measure the knowledge capital of a system or region and attempts to capture knowledge capital in better ways means that this literature is continuously evolving. The knowledge generated within regions is measured by the sum of firms' and other institutions' R&D activities (see e.g. Fagerberg et al., 2002) or the sum of patent activities (see e.g. Sterlacchini, 2008). Since the knowledge capital of a region is also the result of the processes of diffusion of knowledge, some scholars use measures that take account of these processes, mainly the un-weighted or weighted sum of foreign (i.e., in other regions) R&D.

This thesis argued that the above measures of knowledge flows suffer some limitations. First, they are indirect measures of knowledge flows and can indicate only potential knowledge flows between regions. Second, they consider only foreign R&D and take no account of intraregional knowledge flows. There is a large literature showing that knowledge flows are geographically localized and that an important part of the knowledge spreads only within regions. Thus, traditional measures of knowledge flows omit part of the regional knowledge capital.

To overcome these limitations this thesis used backward patent citations to measure knowledge flows. Patent citations are a direct measure of knowledge flows and provide a paper trail for intraregional and interregional knowledge flows.

Patent citations are a good measure of the diffusion of codified knowledge, but not the diffusion of tacit knowledge. To take account of the diffusion of tacit knowledge this thesis used inventor mobility indexes.[66] An inventor that moves between firms takes with him his knowledge, skills and abilities. Thus, inventor mobility is a channel of diffusion of tacit knowledge within and between regions.

The analysis in this thesis makes use of the technology-gap model in which the rate of growth of regional GDP per capita is a function of the variation in its stock of knowledge capital. Knowledge capital is measured by regional R&D and EPO applications and by intraregional and interregional patent citations and intraregional and interregional inventor mobility.

The results show that regional economic growth is positively affected by internal R&D activities, especially if they lead to patents (successful R&D). Also, economic growth in Italian regions is explained by inventor mobility. As expected, the mobility of inventors and, thus, outflow of knowledge, negatively affects economic growth, while the influence of inventors has a positive effect on economic growth.[67] We found confirmation also that interregional net inflows (i.e. inflows less outflows) have a positive impact on economic growth, but that intraregional mobility does not affect the economic growth of Italian regions. This may be explained by the phenomenon of lock-in (Bathelt et al., 2004) and by the fact that the knowledge of knowledge generating firms is too close to the knowledge owned by knowledge receiving firms, with the result that it has no effect on innovation or economic growth (Boschma et al., 2009). Finally, the diffusion of codified knowledge (measured by backward patent citations) does not affect economic growth. This can explained by the fact that codified knowledge is

[66] These indexes are obtained by construction a database of Italian inventor mobility. For more details see appendix to Chapter 3.
[67] However, the coefficient of interregional inflows is not significant at the 5% level.

accessible to all the regions equally, with the result that there are no regional differences in relation to economic growth.

This work contributes to the literature by providing a more direct measure of knowledge flows e.g. patent citations, rather than sum of foreign R&D. This measure allows consideration of intraregional knowledge flows. It also explicitly considers the diffusion of tacit knowledge and constitutes the first attempt, to our knowledge, to analyse the relationship between inventor mobility and regional growth.

Note that this thesis investigates only knowledge originating from R&D activities and knowledge that "travels" by being embodied in inventors. It would be interesting to extend this work by including other forms of more informal knowledge generating and diffusion processes. It could be argued that some important innovation sources, especially for small and medium firms, are represented by the informal R&D activities (e.g. the activities on the design and production departments) and by the diffusion of knowledge through interpersonal and social relationships operating mainly at local level (Garofoli, 2002). However, this would require more data.

Future research could be dedicated to the refinement of the dataset on the mobility of the inventors (e.g. to include mobility outside national borders) and the estimation techniques (e.g. to take account of possible endogeneity bias). It would be interesting to extend this work by including variables that take account of the social and institutional characteristics of regions. It could be argued that regional innovative capacity is a function of the existing entrepreneurial spirit (Audretsch, 2007).

Knowledge flows and economic growth: firm level

The study in Chapter 4 analyses the impact of knowledge capital on firm level economic performance, based on a sample of German firms for the period 2000-2002.

Several studies analyse the impact of knowledge capital on the economic performance of firms (for an analytic survey see e.g.: Griliches, 1979; Czarnitzki et al., 2006; Hall et al., 2010). Investment in R&D is considered from the points of view of firm managers and policy makers, and seen as a key factor in increasing the knowledge capital of firms. Internal R&D efforts are used in the empirical analysis to measure the firm's knowledge capital. However, the knowledge generated by the firm's R&D activities may result in spillovers to other firms because of incomplete protection by patents, or the incapacity of the firm to keep its innovations secret. Several studies use the sum of the R&D efforts of other firms in the same and sometimes in other industries, to measure knowledge flows.

We have argued that these measures of knowledge flows suffer some limitations. First, they are limited to a certain geographic area. Second, they do not take account of the temporary lag between knowledge flows and their impact on innovative or economic outputs. Third, they do not consider some important sources of knowledge flows such as universities and other research institutions. Fourth, all the sources of knowledge flows are considered similarly, but it could be argued that knowledge flows from rivals and from suppliers might have different impacts on innovative or economic performances.

To overcome these limitations, we considered four sources of knowledge flows: customers, suppliers, competitors and research institutions. Another important novelty of this thesis is that it distinguishes between two types of innovation, original innovation and imitation. It analyses the effect of different sources of knowledge flows on innovation and imitation processes.

We provide results in relation to sales of products new to the firm (imitation) and sales of products new to the market (innovation). These two dependent variables are regressed on knowledge capital, measured as internal knowledge stock and the various sources of knowledge flows. The results show that the various sources of knowledge flows affect the two types of innovation in different ways. We show that knowledge flows from rivals more likely lead to imitation, while input from customers and research institutions enhances original innovation.

This study shows the heterogeneity in the relationship between knowledge flows and innovation and economic outputs and, thus, demonstrates the importance of more direct measures of knowledge flows than stock of foreign R&D. Surveys would provide useful data that would take account of this heterogeneity.

References

Abramovitz, M., 1986. Catching up, forging ahead, and falling behind, The Journal of Economic History, vol. 46(02), pages 385-406.

Aghion, P., Howitt P., 1998. Endogenous growth theory, MIT Press, Cambridge, MA.

Agrawal, A., Cockburn, I., McHale, J., 2006. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships, Journal of Economic Geography, vol. 6(5), pages 571-591.

Agrawal, A., Kapur, D., McHale, J., 2008. How do spatial and social proximity influence knowledge flows? Evidence from patent data, Journal of Urban Economics, vol. 64(2), pages 258-269.

Aitken, B.J., Harrison, A.E., 1999. Do domestic firms benefit from direct foreign investment? Evidence from Venezuela, American Economic Review, vol. 89(3), pages 605-618.

Alcacer, J., Gittelman, M., 2006. Patent citations as a measure of knowledge flows: the influence of examiner citations, Review of Economics and Statistics, vol. 88(4), pages 774–779.

Almeida, P., Kogut, B., 1999. Localization of knowledge and the mobility of engineers in regional networks, Management Science, vol. 45(7), pages 905-917.

Anderson, J.E., van Wincoop, E., 2003. Gravity with gravitas: a solution to the border puzzle, American Economic Review, vol. 93(1), pages 170-192.

Arrow, K.J., 1962. Economic welfare and the allocation of resources for invention. In: Nelson, R.R. (Eds), The rate and direction of inventive activity, Princeton University Press, Princeton.

Audretsch, D.B., Feldman, M.P., 1996. R&D spillovers and the geography of innovation and production, American Economic Review, vol. 86(3), pages 630-640.

Audretsch, D.B., Feldman, M.P., 2004. Knowledge spillovers and the geography of innovation. In: Henderson, J.V., Thisse, J.F. (Eds.), Handbook of regional and urban economics, Elsevier, Amsterdam.

Audretsch, D.B., Lehmann, E.E., Warning, S., 2005. University spillovers and new firm location, Research Policy, vol. 34(7), pages 1113-1122.

Audretsch, D.B., 2007. Entrepreneurship capital and economic growth, Oxford Review of Economic Policy, vol. 23(1), pages 63-78.

Baba, Y., Shichijo, N., Sedita, S., 2009. How do collaborations with universities affect firm's innovative performance? The role of 'Pasteur Scientists' in the advanced materials field, Research Policy, vol. 38(5), pages 756-764.

Bacchiocchi E., Montobbio, F., 2010. International knowledge diffusion and home-bias effect: do USPTO and EPO patent citations tell the same story?, Scandinavian Journal of Economics, vol. 112(3), pages 441-470.

Badinger, H., Tondl, G., 2003. Trade, human capital and innovation: the engines of European regional growth in the 1990s. In: Fingleton, B. (Eds.), European regional growth. Springer, Heidelberg, New York.

Baier, S.L., Bergstrand, J.H., 2007. Do free trade agreements actually increase members' international trade?, Journal of International Economics, vol. 71(1), pages 72-95.

Baldwin, R.E., 1995. The eastern enlargement of the European Union, European Economic Review, vol. 39(3-4), pages 474-481.

Baldwin, R.E., Taglioni, D., 2006. Gravity for dummies and dummies for gravity equations, National Bureau of Economic Research, NBER Working Paper 12516.

Bathelt, H., Malmberg, A., Maskell, P., 2004. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation, Progress in Human Geography, vol.28(1), pages 31–56.

Becattini, G., 1979. Dal settore industriale al distretto industriale: alcune considerazioni sull'unità di indagine della politica industriale, Rivista di Economia e Politica Industriale, vol. 5(1), pages 7-21.

Becattini, G., Rullani, E., 1996. Local systems and global connections: the role of knowledge. In: Cossentino, F., Pyke, F., Sengenberger, W. (Eds.), Local and regional response to global pressure: the case of Italy and its industrial districts, ILO-International Institute for Labor Studies, Geneva.

Becker, G.S., 1962. Investment in human capital: a theoretical analysis, Journal of Political Economy, vol. 70(5), pages 9–49.

Becker, G.S., 1993. Human capital: a theoretical and empirical analysis with special reference to education, University of Chicago Press, Chicago.

Bhavnani, R., Coe, D.T., Subramanian, A., Tamirisa N.T., 2002. The missing globalization puzzle, International Monetary Fund, IMF Working Papers 02/171.

Blomstrom M., Kokko A., 1998. Multinational corporations and spillovers, Journal of Economic Survey, vol. 12(3), pages 247-277.

Boschma, R., Eriksson, R., Lindgren, U., 2009. How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity, Journal of Economic Geography, vol. 9(2), pages. 169-190.

Bottazzi, L., Peri, G., 2003. Innovation and spillovers in regions: evidence from European patent data, European Economic Review, vol. 47(4), pages 687-710.

Breschi, S., Lissoni, F., 2001. Knowledge spillovers and local innovation systems: a critical survey, Industrial and Corporate Change, vol. 10(4), pages 975-1005.

Breschi, S., Lissoni, F., 2009. Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows, Journal of Economic Geography, vol. 9(4), pages 439-468.

Brusco, S., 1996. Global systems and local systems. In: Cossentino, F., Pyke, F., Sengenberger, W. (Eds.), Local and regional response to global pressure: the case of Italy and its industrial districts, ILO-International Institute for Labor Studies, Geneva.

Bussière M., Fidrmuc, J., Schnatz, B., 2008. EU enlargement and trade integration: lessons from a gravity model, Review of Development Economics, vol. 12(3), pages 562-576.

Caballero, R.J., Jaffe, A.B., 1993. How high are the giants' shoulders: an empirical assessment of knowledge spillovers and creative destruction in a model of economic growth, National Bureau of Economic Research, NBER Working Papers 4370.

Cairncross, F., 1997. The death of distance. How the communication revolution will change our lives, Harvard Business School Press, Boston.

Caloghirou, Y., Ioannides, S., Vonortas, N.S., 2003. Research joint ventures, Journal of Economic Surveys, vol. 17(4), pages 541-570.

Camagni, R., 1991. Innovation networks: spatial perspectives, Belhaven Press, London.

Cappellen, A., Fagerberg, J., Verspagen, B., 1999. Lack of regional convergence. In: Fagerberg, J., Guerrieri, P., Verspagen, B. (Eds.), The economic challenge for Europe: adapting to innovation based growth, Edward Elgar, Cheltenham.

Carrère, C., 2006. Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model, European Economic Review, vol. 50(2), pages 223-247.

Cheng, I., Wall, H., 2005. Controlling for heterogeneity in gravity models of trade and integration, Federal Reserve Bank of St. Louis Review, vol. 87(1), pages 49-63.

Cohen, W.M., Levinthal, D.A., 1989. Innovation and learning: the two faces of R&D, Economic Journal, vol. 99(397), pages 569-596.

Coe, D.T., Helpman, E., 1995. International R&D spillovers, National Bureau of Economic Research, NBER Working Papers 4444.

Commission of the European Communities, 2001. The regional dimension of the European Research Area, COM(2001) 549 final, Brussels.

Cowan, R., David, P.A., Foray, D., 2000. The explicit economics of knowledge codification and tacitness, Industrial and Corporate Change, vol. 9(2), pages 211-253.

Crescenzi, R., 2005. Innovation and regional growth in the enlarged Europe: the role of local innovative capabilities, peripherality, and education, Growth and Change, vol. 36(4), pages 471-507.

Crespo, N., Fontoura, M.P., 2007. Determinant factors of FDI spillovers - What do we really know?, World Development, vol. 35(3), pages 410-425.

Czarnitzki, D., Hall, B.H., Oriani, R., 2006. The market valuation of knowledge assets in US and European firms. In: Bosworth, D., Webster, E. (Eds.), The management of intellectual property, Edward Elgar, Cheltenham.

Czarnitzki, D., Kraft, K., 2012. Spillovers of innovation activities and their Profitability, Oxford Economic Papers, forthcoming.

d'Aspremont, C., Jacquemin, A., 1988. Cooperative and noncooperative R&D in duopoly with spillovers, American Economic Review, vol. 78(5), pages 1133-1137.

de Bondt, R., 1997. Spillovers and innovative activities, International Journal of Industrial Organization, vol. 15(1), pages 1-28.

Doring, T., Schnellenbach, J., 2006. What do we know about geographical knowledge spillovers and regional growth?: A survey of the literature, Regional Studies, vol. 40(3), pages 375-395.

Duguet, E., MacGarvie, M., 2005. How well do patent citations measure flows of technology? Evidence from French innovation surveys, Economics of Innovation and New Technology, vol. 14(5), pages 375-393.

Eisenhardt, K., Tabizi, B., 1995. Accelerating adaptive processes: product innovation in the global computer industry, Administrative Science Quarterly, vol. 40 (1), pages 84-110.

European commission, 2000. Towards a European research area. Communication from the commission, COM(2000) 6, Brussels.

European commission, 2010. Europe 2020 - A strategy for smart, sustainable and inclusive growth, COM(2010) 2020, Brussels.

Eurostat, 2004. Innovation in Europe - Results for the EU, Iceland and Norway, Panorama of the European Union - Theme 9: science and technology, Eurostat, Luxembourg.

Eurostat and OECD, 2005. Oslo Manual: guidelines for collecting and interpreting innovation data, 3rd Edition, OECD, Paris.

Eurostat, 2007. Regions in the European Union. Nomenclature of territorial units for statistics, European Commission, Brussels.

Fagerberg, J., 1987. A technology gap approach to why growth rates differ, Research Policy, vol. 16(2-4), pages 87-99.

Fagerberg, J., 1988. Why growth rates differ. In: Dosi, G., Freeman, C., Nelson, R., Silverberg, G., Soete, L. (Eds.), Technical change and economic theory, Pinter Publishers, London.

Fagerberg, J., 1994. Technology and international differences in growth rates, Journal of Economic Literature, vol. 32(3), pages 1147-1175.

Fagerberg, J., Verspagen, B., 1996. Heading for divergence? Regional growth in Europe reconsidered, Journal of Common Market Studies, vol. 34(3), pages 431-448.

Fagerberg, J., Verspagen, B., Caniels, M., 1997. Technology, growth and unemployment across European regions, Regional Studies, vol. 31(5), pages 457-466.

Fagerberg, J., Verspagen, B., 2002. Technology-gaps, innovation-diffusion and transformation: an evolutionary interpretation, Research Policy, vol. 31(8-9), pages 1291-1304.

Feldman, M.P., 1999. The new economics of innovation, spillover and agglomeration: a review of empirical Studies, Economics of Innovation and New Technology, vol. 8(1-2), pages 5-25.

Feldman M.P., Kogler, D.F., 2010. Stylized facts in the geography of innovation. In: Hall, B.H., Rosenberg, N. (Eds), Handbook of the economics of innovation, Elsevier, Amsterdam.

Fischer M., Scherngell, T., Jansenberger, E., 2009. Geographic localisation of knowledge spillovers: evidence from high-tech patent citations in Europe, The Annals of Regional Science, vol. 43(4), pages 839-858.

Garofoli, G., 1983. Industrializzazione diffusa in Lombardia, I.Re.R. - Franco Angeli, Milano.

Garofoli, G., 2002., Piccole imprese, innovazione e territorio: economie di apprendimento e sistema innovativo locale. In: Camagni, R., Capello, R. (Eds), Apprendimento collettivo e competitività territoriale, Franco Angeli, Milano.

Geroski, P., Machin, S., van Reenen, J., 1993. The Profitability of innovating firms, RAND Journal of Economics. vol. 24(2), pages 198-211.

Gerschenkron, A., 1962. Economic backwardness in historical perspective, Belknap Press, Cambdridge, MA.

Gil, S., Llorca, R., Martínez-Serrano, J.A., 2008. Assessing the enlargement and deepening of the European Union, The World Economy, vol. 31(9), pages 1253-1272.

Glick, R.R., Rose, A.K., 2002. Does a currency union affect trade? The time-series evidence, European Economic Review, vol. 46(6), pages 1125-1151.

Griffith, R., Sokbae, L., van Reenen, J., 2007. Is distance dying at last? Falling home bias in fixed effects models of patent citations, National Bureau of Economic Research, NBER Working Papers 13338.

Griliches, Z., 1979. Issues in assessing the contribution of research and development to productivity growth, Bell Journal of Economics, vol. 10(1), pages 92-116.

Griliches, Z., Mairesse, J., 1984. Productivity and R&D at the firm level. In: Griliches, Z. (Eds.), R&D, Patents and productivity, University of Chicago Press, Chicago.

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

Guellec, D., van Pottelsberghe de la Potterie, B., 2001. The internationalisation of technology analysed with patent data, Research Policy, vol. 30(8), pages 1253-1266.

Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2005. Market value and patent citations, RAND Journal of Economics, vol. 36(1), pages 16-38.

Hall, B.H., MacGarvie, M., 2010. The private value of software patents, Research Policy, vol. 39(7), pages 994-1009.

Hall, B.H., Mairesse, J., Mohnen, P., 2010. Measuring the returns to R&D. In: Hall, B.H., Rosenberg, N. (Eds.), Handbook of the economics of innovation, Elsevier, Amsterdam.

Hanel, P., St-Pierre, A., 2002. Effects of R&D spillovers on the profitability of firms, Review of Industrial Organization, vol. 20(4), pages 305-322.

Herstatt, C., von Hippel, E., 1992. From experience: developing new product concepts via the lead user method: a case study in a "low-tech" field, Journal of Product Innovation Management, vol. 9(3), pages 213-221.

Hoekman, J., Frenken, K., Tijssen, R.J.W., 2010. Research collaboration at a distance: changing spatial patterns of scientific collaboration within Europe, Research Policy, vol. 39(5), pages 662-673.

Jaffe, A.B, 1986. Technological opportunity and spillovers of R&D: evidence from firms' patents, profits, and market value, American Economic Review, vol. 76(5), pages 984-1001.

Jaffe, A.B., 1989. Real effects of academic research, American Economic Review, vol. 79(5), pages 957-970.

Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations, The Quarterly Journal of Economics, vol. 108(3), pages 577-598.

Jaffe, A.B., Fogarty, M.S., Banks, B.A., 1998. Evidence from patents and patent citations on the impact of Nasa and other federal labs on commercial innovation, Journal of Industrial Economics, vol. 46(2), pages 183-205.

Jaffe, A.B., Trajtenberg, M., 2002. Patents, citations and innovations: a window on the knowledge Economy, MIT Press, Cambridge, MA.

Janz, N., Ebling, G., Gottschalk, S., Niggemann, H., 2001. The Mannheim Innovation Panels (MIP and MIP-S) of the Centre for European Economic Research (ZEW), Schmollers Jahrbuch, vol. 121(1), pages 123-129.

Johnson, D.K.N., Siripong, A., Brown, A.S., 2006. The demise of distance? The declining role of physical proximity for knowledge transmission, Growth and Change, vol. 37(1), pages 19-33.

Karnath, R., Liker, J., 1994. A second look at Japanese product development, Harvard Business Review, vol. 72(2), pages 154-170.

Keller, W., 2001. International technology diffusion, National Bureau of Economic Research, NBER Working Papers 8573.

Keller, W., 2004. International technology diffusion, Journal of Economic Literature, vol. 42(3), pages 752-782.

Krugman, P., 1991, Geography and trade, MIT Press, Cambridge, MA.

Los, B., Verspagen, B., 2000. R&D spillovers and productivity: evidence from U.S. manufacturing microdata, Empirical Economics, vol. 25(1), pages 127-148.

Lucas, R.E., 1988. On the mechanics of economic development, Journal of Monetary Economics, vol. 22(1), pages 3-42.

Lundvall, B.A., 1992. National systems of innovation: towards a theory of innovation and interactive learning, Pinter Publishers, London.

Mairesse, J., Sassenou, M., 1991. R&D productivity: a survey of econometric studies at the firm level, National Bureau of Economic Research, NBER Working Papers 3666.

Mansfield, E., Rapoport, J., Schnee, J., Wagner, S., Hamberger, M., 1971. Research and innovation in the modern corporation, Norton, New York.

Mansfield, E., Schwartz, M., Wagner, S., 1981. Imitation costs and patents: an empirical study, The Economic Journal, vol. 91(364), pages 907-918.

Mansfield, E., 1985. How rapidly does new technology leak out?, Journal of Industrial Economics, vol. 34(2), pages 217-223.

Mansfield, E., 1995. Academic research underlying industrial innovations: sources, characteristics, and financing, The Review of Economics and Statistics, vol. 77(1), pages 55-65.

Mansfield, E., 1998. Academic research and industrial innovation: an update of empirical findings, Research Policy, vol. 26(7-8), pages 773-776.

Marshall, A., 1920. Principles of Economics, MacMillan, London.

Maurseth, P.B., Verspagen, B., 2002. Knowledge spillovers in Europe: a patent citations analysis, Scandinavian Journal of Economics, vol. 104(4), pages 531-545.

Micco, A., Stein, E., Ordonez, G., 2003. The currency union effect on trade: early evidence from EMU, Economic Policy, vol. 18(37), pages 315-356.

Montobbio, F., Sterzi, V., 2011. Inventing together: exploring the nature of international knowledge spillovers in Latin America, Journal of Evolutionary Economics, vol. 21(1), pages 53-89.

Montobbio, F., Sterzi, V., 2012. The globalization of technology in emerging markets: a gravity model on the determinants of international patent collaborations, Cahiers du GREThA 2012-05, Groupe de Recherche en Economie Théorique et Appliquée, Bordeaux.

Morgan, K., 1997. The learning region: institutions, innovation and regional renewal, Regional Studies, vol. 31(5), 491-503.

Nelson, R.R., Winter, S.G., 1982. An evolutionary theory of economic change, Harvard University Press, Cambridge, MA.

OST, 2004. Indicateurs de sciences et de technologies, Rapport de l'Observatoire des Sciences et des Techniques, Economica, Paris.

Oughton, C., Landabaso, M., Morgan, K., 2002. The regional innovation paradox: innovation policy and industrial policy, Journal of Technology Transfer, vol. 27(1), pages 97–110.

Paci, R., Usai, S., 2009. Knowledge flows across European regions, The Annals of Regional Science, vol. 43(3), pages 669-690.

Park, W.G., 1995. International R&D spillovers and OECD economic growth, Economic Inquiry, vol. 33(4), pages 571-91.

Peri, G., 2005. Determinants of knowledge flows and their effect on innovation, The Review of Economics and Statistics, vol. 87(2), pages 308-322.

Perroux, F., 1950. Economic spaces: theory and applications, Quarterly Journal of Economics, vol. 64(1), pages 89-104.

Picci, L., 2010., The internationalization of inventive activity: a gravity model using patent data, Research Policy, vol. 39(8), pages 1070-1081.

Polanyi, M., 1967. The tacit dimension, Doubleday, New York.

Ponds, R., van Oort, F., Frenken, K., 2010. Innovation, spillovers and university-industry collaboration: an extended knowledge production function approach, Journal of Economic Geography, vol. 10(2), pages 231-255.

Raffo, J., Lhuillery, S., 2009. How to play the “Names Game”: patent retrieval comparing different heuristics, Research Policy, vol. 38(10), pages 1617-1627.

Rodriguez-Pose, A., Crescenzi, R., 2008. Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe, Regional Studies, vol. 42(1), pages 51-67.

Romer, P.M., 1986. Increasing returns and long-run growth, Journal of Public Economics, vol. 94(5), pages 1002-1037.

Romer, P.M, 1987. Growth based on increasing returns due to specialization, American Economic Review, vol. 77(2), pages 56-62.

Romer, P.M., 1990. Endogenous technological change, Journal of Political Economy, vol. 98(5), pages 71-102.

Rose, A.K., 2001. Currency unions and trade: the effect is large, Economic Policy, vol. 16(33), pages 449-461.

Santos Silva, J.M.C., Tenreyro, S., 2006. The log of gravity, The Review of Economics and Statistics, vol. 88(4), pages 641-658.

Singh, J., 2005. Collaborative networks as determinants of knowledge diffusion patterns, Management Science, vol. 51(5), pages 756-770.

Soete, L., Ter Weel, B., 1999. Innovation, knowledge creation and technology policy in Europe, Research Memoranda 001, Maastricht Economic Research Institute on Innovation and Technology, Maastricht.

Solow, R.M., 1956. A contribution to the theory of economic growth, Quarterly Journal of Economics, vol. 70(1), pages 65-94.

Solow, R.M., 1957. Technical change and the aggregate production function, Review of Economics and Statistics, vol. 39(3), pages 312-320.

Sonn, J.W., Storper, M., 2008. The increasing importance of geographical proximity in knowledge production: an analysis of US patent citations, 1975-1997, Environment and Planning, vol. 40(5), pages 1020-1039.

Spies J., Marques, H., 2009. Trade effects of the Europe agreements: a theory-based gravity approach, Journal of International Trade and Economic Development, vol. 18(1), pages 11-35.

Sterlacchini, A., 2008. R&D, higher education and regional growth: uneven linkages among European regions, Research Policy, vol. 37(6-7), pages 1096-1107.

Suzuki, K., 1993. R&D spillovers and technology transfer among and within vertical Keiretsu groups: evidence from the Japanese electrical machinery industry, International Journal of Industrial Organization, vol. 11(4), pages 573-591.

Tether, B., 2002. Who co-operates for innovation, and why: an empirical analysis, Research Policy, vol. 31(6), pages 947-967.

Tether, B., Tajar, A., 2008. Beyond industry-university links: sourcing knowledge for innovation from consultants, private research organizations and the public science-base, Research Policy, vol. 37(6-7), pages 1079-1095.

Thompson, P., 2006. Patent citations and the geography of knowledge spillovers: evidence from inventor – and examiner – added citations, Review of Economics and Statistics, vol. 88(2), pages 383–388.

Thursby, M., Thursby, J., 2006. Where is the new science in corporate R&D?, Science, vol. 314(8), pages 1547-1548.

Tinbergen, J., 1962. Shaping the world economy, The Twentieth Century Fund, New York.

Trajtenberg, M., 1990. A penny for your quotes: patent citations and the value of innovations, RAND Journal of Economics, vol. 21(1), pages 172-187,

Trajtenberg, M., Shiff, G., Melamed, R., 2006. The "Names Game": harnessing inventors' patent data for economic research, National Bureau of Economic Research, NBER working paper 12479.

van Meijl, H., 1997. Measuring intersectoral spillovers: French evidence, Economic Systems Research, vol. 9(1), pages 25-46.

van Pottelsberghe de la Potterie, B., Lichtenberg, F., 2001. Does foreign direct investment transfer technology across borders?, The Review of Economics and Statistics, vol. 83(3), pages 490-497.

Viner, J., 1950. The customs union issue, Carnegie Endowment for International Peace, New York.

von Hippel, E., 1988. The sources of innovation, Oxford University Press, New York.

von Hippel, E., 1994. Sticky information and the locus of problem solving: implications for innovations, Management Science, vol. 40(4), pages 439-439.