Knowledge Flows and Knowledge Externalities

Giovanni Peri

CESifo Working Paper No. 765
Category 5: Fiscal Policy, Macroeconomics and Growth
August 2002

KNOWLEDGE FLOWS AND KNOWLEDGE EXTERNALITIES

Abstract

The diffusion of knowledge in the world generates positive externalities if knowledge flows increase the productivity of R&D. Our work analyzes knowledge diffusion and knowledge externalities in generating innovation and in determining productivity. We first estimate the determinants of knowledge flows across 141 sub-national regions in 19 countries of Europe and North America as revealed by patent citation between US-granted patents. Then we estimate the impact of these flows on productivity of R&D resources in generating innovation (patenting) and productivity (TFP). While we find that knowledge diffusion depends on geographical and technological distance and is well described by a pseudo-gravity model, we do not find evidence of significant positive externalities from existing knowledge.

JEL Classification: F0, O3, R1.

Keywords: knowledge flows, innovation, patent citations regions.

Giovanni Peri
Department of Economics
University of California Davis
One Shields Avenue
Davis Ca, 95616
U.S.A.
gperi@ucdavis.edu

1 Introduction

The concept of Knowledge Externalities has pervasively populated the theoretical literature on growth for the last ten years. Innovation is the engine of growth, has been argued, and externalities from existing knowledge are the "renewable" fuel for this engine. While it is clear how to characterize the source and the effect of these externalities within the frame of a specific model such as Aghion and Howitt [1], Jones [13] or Romer [18], it is much less clear how to measure them in a precise but robust way using the data. First of all it is not clear that there are data that allow us to directly measure such elusive concept as knowledge externalities without adding several other assumptions. Knowledge externalities hinge on the diffusion of ideas, a process that leaves no track in the data. Economists have assumed, in order to estimate these externalities, that diffusion of ideas depends on proximity in space, in technological specialization or in economic development. These are plausible assumptions but each of them is potentially controversial and should be tested. In particular we do not have a good measure of the quantitative dependence of knowledge diffusion on distance or technological proximity. Second, the existence of several effects stemming from the introduction of new ideas, namely their impact on current and future productivity of goods and ideas, has further complicated the task. The goal of this paper is to define clearly, if somewhat narrowly, what knowledge externalities are and use data on patented innovation, citations across patents and total factor productivity to measure the strength of these externalities for 141 sub-national regions covering the whole Western Europe and North America.

Knowledge externalities, as all externalities, exist if the social benefit from ideas is larger than the private returns to their inventors. In an economy of competing innovators the private return from an idea is the value of the patent that it entitles to. Therefore we can assume that the whole private contribution of that idea as enhancer of present and future productivity is captured by the patent's value. In a competitive market for innovation the innovators are able to extract the surplus generated by that innovation. Nevertheless the total social surplus generated by the new idea includes also the contribution (be it positive or negative) that the new idea brings to future discovery of other ideas. We can think of the stock of knowledge as being a state variable and to new patents as the flow that increases that variable. Three contributions of the new ideas have to be considered in order to measure their aggregate return[1]. First, the private return from a new idea is the benefit due to increased productivity that the idea generates. Second there is a "capital gain" from this new idea that is the fact that future cost of producing ideas can change and therefore the value of "new ideas" may vary over time. Third, there may be an externality of new ideas in the production function of ideas. An increase in the stock of knowledge may change the productivity of $R\&D$ resources in generating new knowledge. While the first two contributions are fully reflected in the value of the patent,

---

[1] We follow here the analysis of Jones and Williams [15]

the third being an externality, raises the social productivity of the idea above the value of the patent. Crucially, therefore, externality of knowledge exist if ideas increase the productivity of R&D in generating new ideas.

Defining knowledge externalities as we do above may seem somewhat narrow. We assume that there is an efficient market for innovations, that property rights are well defined and enforced by the patent system. If so, when ideas are used in the production of goods, their surplus is extracted by the innovator via the royalties paid on the patent. The only missing market is the one regulating the use of ideas in generation of ideas: no inventor is compensated for the use of her idea in some subsequent discovery that uses it. Many economists believe that externalities from innovation are much more pervasive than that because not all innovations are codified, those that are codified are often not patented and those that are patented are not always effectively protected. Therefore knowledge "leaks" and it benefits society beyond the private value of the patent. Some economists ultimately argue that externalities should be inferred from measures of total factor productivity (for instance recently, Ciccone [3] and Keller [16]). Innovation is an important intermediate passage but only when ideas are transferred into productivity gains we have a link between knowledge externalities and output per capita. As we share this view we analyze in the paper the effect of knowledge spillovers on innovation as well as on productivity.

We define externalities as disembodied knowledge spillovers that affect the productivity of R&D in the generation of new ideas. This definition captures very important aspects of knowledge externalities and, while not all-inclusive, it is relevant for at least three reasons. First of all the idea of externalities as knowledge spillovers in the creation of new knowledge is behind most (if not all) the theoretical models of endogenous growth, from Romer [18] to Jones [13] passing through Aghion and Howitt [1] and Grossman and Helpman [7]. The strength of this externality determines the difference between "endogenous" and "semi-endogenous" model of growth. In general it regulates the dynamics of innovation in the long run. Second as we limit our analysis to the most advanced economies (Western Europe and North America) we can assume that the protection of property rights on innovation is, in these countries, at its best in the world, minimizing phenomena of imitation and leakages of the patent system. Granted that leakages due to insufficient protection of intellectual property rights may exist in developing countries, we assume that the effect of knowledge on new ideas is the main source of spillovers in developed countries. Third, using the above definition of spillovers we are able to use data on patent-to-patent citations in order to track the direction and intensity of these knowledge flows. If the determinants of non-codified knowledge diffusion are similar to the determinants of citations then our analysis can be seen as revealing intensity and direction of other kinds of knowledge "leakages" as well.

The goal of this paper is to estimate the external contribution of available knowledge in generating new knowledge and productivity across regions. Two distinct pieces should be combined in order to establish if existing ideas have any (positive or negative) external effect in the creation of new ideas. First only existing ideas that are actually used by researchers in their process to achieve

a new idea can generate an externality. Second, only if the use of these ideas increases the productivity of researchers in generating new ideas we have such externality. Previous existing work (with the notable exception of Caballero and Jaffe [2]) merges or assimilates these two phases. Either some mechanical assumption is made on the availability of ideas across space (as in Coe and Helpman [4] and Keller [16]) so that only a productivity equation is estimated, or knowledge flows are simply called knowledge externalities without worrying about their effect on productivity (Jaffe et al.[11] , Jaffe and Trajtenberg [12]). In this paper we explicitly model and estimate the phase of knowledge diffusion and utilization in generating new ideas, and the potential impact of this knowledge on the productivity of researchers in generating new ideas or in determining TFP.

We use data on patenting and TFP organized in a cross section of 141 regions in Europe and North America in order to estimate the generation of new ideas or of new productive capacity (TFP) from R&D resources. We use data on citation across inventions patented in the US between 1975 and 1996 to measure the intensity of knowledge flows in the world. We find that knowledge flows depend on several factors and are well represented using a gravity-like equation. Most importantly geographical and technological proximity and crossing a national border affect the intensity of diffusion of ideas. On the other hand we also find that flow-weighted R&D from other regions has small and not significant effect on the generation of new ideas or in determining regional productivity.

The rest of the paper is organized as follows: Section 2 describes the equations that models knowledge flows and the production function of ideas. Section 3 defines the alternative measures of regional productivity and describes the accounting method used to calculate them. Section 4 describes the data we use to measure innovation, productivity, inputs of innovation and the flows of knowledge across regions. Section 5 presents the estimates of the parameters in the equation that describes knowledge flows and in the equation that describes their effect on innovation and productivity. Section 6 concludes the paper.

2   The Model

Two pieces are needed in order to analyze knowledge flows and knowledge externalities across regions. The first is a function modelling the diffusion of ideas across regions as function of bilateral regional characteristics. In order to produce externalities, ideas originated in a region must be available in other regions as inputs to innovate over the existing knowledge. The second is an innovation function describing how ideas are generated by researchers using R&D spending and existing ideas available to that particular region. The description of these two equations and their estimate occupy the rest of the paper.

2.1 Diffusion of Ideas

In order to estimate the relative intensity of diffusion of ideas across regions using cross-sectional data we divide the ideas discovered in each region in two groups. A group of ideas, discovered earlier (during the 1977-1991 period in our empirical analysis) that can "send" flows of knowledge to be used in further discoveries, and a group of ideas discovered later (1992-1996 in the empirical analysis) that can "receive" the flows and use them to generate new ideas. We can use the citation from the receiving to the sending groups of ideas to infer the intensity of flows. Clearly the receiving ideas also cite ideas discovered before the "sending" (1977) and also cite each other. Nevertheless if the flows of knowledge across regions are stable over time we can infer their relative intensity looking at the equation described in this section.

We denote with ΔA_r the ideas (new knowledge) generated in region r (for "receiving") during the later interval of time (1992-1996). We denote with ΔA_s the ideas (new knowledge) generated in region s (for "sending") during the earlier period of time (1977-1991). The intensity of diffusion to region r of the "average" idea originated in region s is denoted as $\phi(r,s)$. This index should capture the intensity of ideas flows between region s and region r independently of the total amount of ideas originated in region s and of the total amount of citations produced by ideas in region r. Defining as $\sigma(s,r)$ the number of ideas that flow from s to r the index is defined as follows:

$$\phi(r,s) = \frac{\sigma(s,r)/\sum_s \sigma(s,r)}{\Delta A_s / \sum_s \Delta A_s} \tag{1}$$

The term in the numerator $\sigma(s,r)/\sum_s \sigma(s,r)$ captures the share of ideas coming from region s in the total flows of ideas coming into region r. The term in the denominator $\Delta A_s / \sum_s \Delta A_s$ is simply the share of total "sending" ideas generated in region s. The intensity of diffusion between region s and region r measured by $\phi(r,s)$ is therefore a relative measure comparable across region-couples and independent from their size. Keeping the receiving region fixed, this index measures the relative flows of ideas from each of the sending regions, relative to the amount of innovation generated in each sending region. A value of 1 means that region r receives from region s a share of ideas equal to the share of total "sending" ideas originating in s. Keeping the sending region fixed the index measures the relative absorption of ideas across receiving regions. Notice that the standardization in the numerator implies that we assume equal total absorptive capacity in each region. This depends from the fact that we cannot extract from citation any interesting information on the absolute "absorptive capacity" of a region. In brief the index $\phi(r,s)$ could be considered as the intensity of diffusion of the representative idea between the two regions. The index will be larger than one if diffusion of the average idea between the two regions is stronger than the average diffusion between regions.

Thinking of $\phi(r,s)$ as the intensity of a flow of ideas, we are interested in modelling it as function of bilateral regional characteristics. In particular we describe it as follows:

$$\phi(r,s) = \delta e^{-\gamma_1(d_1)_{r,s}} e^{-\gamma_2(d_2)_{r,s}} e^{-\gamma_3(d_3)_{r,s}} e^{f(\underline{Char}_{r,s})} \qquad (2)$$

δ is a parameter scaling the overall flow of ideas between two regions. We can think of it as changing over time but it is constant in our cross-section. The three exponential functions model the decay of ideas flow with distance in geographical space $(d_1)_{r,s}$, technological space $(d_2)_{r,s}$ and innovation space $(d_3)_{r,s}$ between the two regions[2]. The exponential functions in (2) can be derived as the cross-sectional version of a model of diffusion of ideas in time, similar to the one used in Caballero and Jaffe [2]. In the Appendix A we show the derivation of the exponential format from a diffusion process. The flow of ideas also depends on a vector of other bilateral characteristics of the two regions, $\underline{Char}_{r,s}$ such as being in the same country, sharing a border, using a common language.

In order to use equation (1) and equation (2) for the empirical analysis we establish the following relations between the unobservable variables in (1) and our observables.

1) We assume that the average "number of ideas" in a patent, indicated as β_s, might be different across regions. A slightly different (but observationally equivalent) interpretation of this parameter is that it captures the "average quality " of ideas contained in the patents of a region. Namely regions whose patents are more "important" on average, can be thought as having more ideas per patent or the same amount but better quality of ideas per patent. We call this parameter, in general, the "intensity of ideas per patent". This parameter may vary over time only according to a factor μ common to all regions. Denoting with P_s the total number of "sending" patents granted to region s in 1977-1992 and with P_r the "receiving" patents granted to region s in 1992-1996, the above assumptions are summarized by the following conditions: $P_s = \Delta A_s/\mu\beta_s$, $P_r = \Delta A_r/\beta_r$).

2) We assume that the number of citations from region r to patents from region s contains information on the flow of ideas between the two regions according to the following formula: $c(r,s) = \psi_r\sigma(s,r)\varepsilon(r,s)$.

$c(r,s)$ the total number of citations from patents of region r to patents of region s. As different regions may have different average absorption capacity and propensity to cite, we allow for a region-specific term ψ_r that combines these two factors. $\sigma(s,r)$ is the actual flow of ideas between regions. $\varepsilon(r,s)$ is a log-normal random noise due to the fact that not all citations capture the flow of an ideas as some are added by reviewers. We assume that such noise is randomly distributed across region couples (i.e. not correlated with $\sigma(s,r)$) .

Substituting these definitions into equation (1) and into equation (2) and using the fact that the total number of citations from region r, $\sum_s c(s,r)$ for the Law of Large numbers can be considered approximately non random and equal to the number of citing patents P_r times the average number of citation per patent n_r we get the following relation:

---

[2]The exact measure for these three types of distances will be made clear in the empirical section.

$$\phi(r,s) = \frac{c(r,s)/\varepsilon(r,s)}{\sum_s c(s,r)\beta_s P_s/\overline{P}} = \frac{c(r,s)/\varepsilon(r,s)}{n_r P_r (P_s \beta_s)/\overline{P}} = \delta e^{-\gamma_1 (d_1)_{r,s}} e^{-\gamma_2 (d_2)_{r,s}} e^{-\gamma_3 (d_3)_{r,s}} e^{f(\underline{Char}_{r,s})}$$

$$(3)$$

where $\overline{P} = \sum_s P_s$. Collecting the observables on the left hand side, taking logs on both sides of (3) and grouping the constants into one term we have the following estimable equation:

$$\ln\left(\frac{c(r,s)}{P_r P_s}\right) = a + \ln(n_r) + \ln(\beta_s) - \gamma_1 (d_1)_{r,s} - \gamma_2 (d_2)_{r,s} - \gamma_3 (d_3)_{r,s} + \underline{f}(\underline{Char}_{r,s}) + u(r,s)$$

$$(4)$$

Equation (4) can be seen as a generalized Gravity Equation for knowledge spillovers. It is derived from the framework described above and can be brought to the data. We comment here on its meaning and main features. The frequency of citation $\frac{c(r,s)}{P_r P_s}$ from ideas of region r to ideas of region s depends on a "citing region" fixed effect, $\ln(n_r)$, on a "cited region" fixed effect, $\ln(\beta_s)$, on a sequence of bilateral characteristics (described above) and on $u(r,s) = \ln \varepsilon(r,s)$ that is a normally distributed random error. While the cited region fixed effect can be interpreted as the "average intensity of ideas" in patents of "sending" region s, the citing region fixed effect should be controlled for but has no interesting interpretation. The equation generalizes the concept used in the gravity equation, popular in trade and migration analysis, that the "flow" of some variables between two regions depends on some characteristics of the two regions and on the "distance" between them. The Estimating equation (4) allows us to identify the parameters γ_1, γ_2 γ_3 and the vector $\underline{f}$ and therefore to estimate the function $\phi(r,s)$ using definition (2). This function captures the diffusion of ideas between two regions and will be used to construct the stock of used ideas as innovative input in the receiving region.

2.2 The Innovation function and Knowledge Spillovers

Once we identify the strength of knowledge flows $\phi(r,s)$ we can estimate the impact of these spillovers on the production of new ideas. Assuming that our regions are on a balanced growth path with common rate of knowledge growth we can separate the contribution to innovation of private inputs such as scientists and R&D resources from the external contribution of ideas in the estimation of the innovation function of a region.

New ideas, ΔA_r are generated from scientists and researchers in region r and from their productivity. Productivity of R&D employees in region r depends on two main factors: the amount of resources available to them (such as laboratories, research funds, salaries and so on) captured by R&D spending per-worker, h_r and on the stock of ideas available to them in region r. We denote with the

symbol A_r^{AVA} and ΔA_r^{AVA} the cumulated stock and the increase in the stock of ideas available in region r, respectively. A_r and ΔA_r still denote the stock and the increase of the stock of ideas generated in region r. The knowledge available in region r is given by the knowledge generated anywhere that flows to region r , namely: $A_r^{AVA} = \sum_{s \in S} \phi(r, s) A_s$ where $\phi(r, s)$ is the intensity in the flow of ideas described in the previous section. The production function of innovation is therefore:

$$\Delta A_r = f(R\&D_r, h_r, A_r^{AVA}) = f\left(R\&D_r, h_r, \sum_{s \in S} \phi(r, s) A_s\right) \qquad (5)$$

We assume that regions are on their balance growth path with a common growth rate of knowledge stock A_r so that: $(\Delta A_r / A_r) = g$. We check this assumption with our data and it turns out to be satisfied. Substituting this condition into equation (5) and taking a log-linear expression for the innovation function we have:

$$\ln(\Delta A_r) = -\ln(g) + \varepsilon_{RD} \ln(R\&D_r) + \varepsilon_h \ln(h_r) + \varepsilon_I \ln\left(\sum_{s \in S} \phi(r, s) \Delta A_s\right) \qquad (6)$$

ε_{RD} is the elasticity of innovation to the supply of scientists. ε_h is the elasticity of innovation to the supply of resources per scientists. ε_I is the elasticity of innovation to available ideas. We can construct the available knowledge in each region using the estimates of $\phi(r, s)$ obtained from the previous section (2.1). Denoting with a hat ^ the variable estimates from equation (4) we have:

$$\ln(P_r) \quad = \quad Intercept - \ln(\widehat{\beta}_r) + \varepsilon_{RD} \ln(R\&D_r) + \varepsilon_h \ln(h_r) + \qquad (7)$$

$$+\varepsilon_I \ln\left(\sum_{s \in S} \exp\left(-\widehat{\gamma_1}(d_1)_{r,s} - \widehat{\gamma_1}(d_2)_{r,s} - \widehat{\gamma_3}(d_3)_{r,s} + \widehat{f}\left(\underline{Char}_{r,s}\right)\right) \widehat{\beta}_s P_s\right) + \nu_{r,s}$$

The dependent variable is the count of receiving patents granted in region r. The term $Intercept$ contains all the uninteresting constants. The term $\ln(\widehat{\beta}_r)$ controls for the average intensity of ideas per patent in region r. Controlling for this term is extremely important if the "intensity" of ideas per patent is correlated with the total amount of patents generated in a region. $\ln(R\&D_r)$ is the log of employed in the R&D sector, $\ln(h_r)$ is the log of spending in US \$ per worker in the R&D sector. The term in brackets is the estimated knowledge available in region r coming from ideas discovered in the past. It is calculated as the intensity of flows to region r (the exponential term) times the amount

8

of ideas generated in each region $\widehat{\beta}_s P_s$. Finally $\nu_{r,s}$ is a zero average random disturbance that captures other determinants of patenting not included in our equation.

In our empirical specification we first estimate equation (4) and discuss the intensity and direction of spillovers. We then use the estimates from that equation and the observable characteristics of regions to estimate the impact of these spillovers on innovation in region r by using (7). To avoid the endogeneity problem due to the fact that the expression in brackets contains P_r itself, we instrument this variable with a similar one, A_{-r}^{AVA} obtained including all but region r in the calculation of available ideas for region r. Alternatively we will use R&D resources as instrument for P_r .

3 Productivity and Knowledge Spillovers

While the most direct measure of new economically useful ideas, defined above as ΔA_r, is certainly given by the count of patents, economists traditionally measure technological progress and technological differences using total factor productivity (TFP). Total factor productivity captures differences in labor productivity not due to differences in the use of capital. At some fundamental level, therefore, innovation has an effect on output per capita only if it is translated into increases in TFP. We can consider the TFP of a region as an alternative measure of A_r, that not only captures patented inventions but all innovations that make production more efficient. Empirically we estimate an equation identical to (6) where ΔA_r, proportional to A_r is simply measured using regional TFP. We modify cross-country accounting techniques in order to calculate the region-specific total factor productivity (a_r) and separate it from the country TFP, (A_c) that is potentially affected by institutional factors. In the following sections we briefly describe how we obtain two measures of regional TFP: a narrow and a broad one.

3.1 Cross-Regional Productivity Accounting

We assume a Cobb-Douglas production function at the regional level with constant return to scale so that elasticity of income to labor and capital equal the share of income going to each of the two factors. Calling α the share of income going to labor, so that $(1-\alpha)$ will be the share going to capital, we assume their value to be equal to 0.66 and 0.33 . As we use regional data we may rely on the within country variation to define total factor productivity, disregarding the (probably large) differences in total factor productivity across countries due to institutional and political features of the economies. The cross- regional differences in TFP can be more appropriately interpreted as differences in the level of technology used (adopted) there. The disadvantage is that we do not have data on physical capital at the regional level. We may nevertheless recover the contribution of regional TFP to production per worker in one of the following

two ways. We either assume that physical capital moves across regions in order to equate its real return within a country (narrower definition) or we impose the same capital labor-ratio across regions of the same country (broader definition). The first method produces a lower bound of the region-specific TFP differentials, as it attributes part of the regional production per worker differences to capital intensity. The second method produces an upper bound of these differentials as it attributes the whole region-specific differentials in output per worker to differences in regional TFP. We will perform both decompositions. To be precise, the first one distinguishes among the following three contributions to differentials in output per worker: 1) Differences in regional capital intensity (capital-labor ratios), 2)Differences in country-specific total factor productivity 3) Differences in region-specific TFP. The second decomposition, on the other hand, distinguishes among the following contributions 1) Differences in country-specific TFP 2) Differences in country-specific capital-output ratios 3) Differences in region-specific TFP.

3.1.1 Narrow TFP definition

Let's assume that the production function for each region r can be written as a Cobb-Douglas

$$Y_r = A_c a_r (L_r)^\alpha (K_r)^{1-\alpha} \tag{8}$$

where Y_r is total output , L_r is total labor input, K_r total capital input, A_c is the average TFP of country c to which the region belongs and a_r is a region-specific TFP factor such that if $a_r > 1$ the region has larger total factor productivity than the country as a whole while if $a_r < 1$ then the region has lower TFP than the country as a whole. Also we assume that $\sum_{r \in c} \ln(a_r) = 0$ which implies that the average log TFP for regions of a country is equal to the aggregate country's log TFP. α is the elasticity of output to labor that equals the income share of labor. In order to decompose total income in the contribution of each component we assume that within country c the return paid to capital is the same in each region r due to the mobility of capital. This implies:

$$\frac{\partial Y_r}{\partial K_r} = (1 - \alpha) A_c a_r (L_r)^\alpha (K_r)^{-\alpha} = r_c \tag{9}$$

Marginal productivity of capital at the country level is also equal to r_c. Denoting with $k_c = (K_c/L_c)$ the capital-labor ratio for the country. Equation (9) implies that we can write total capital in region r as:

$$K_r = k_c a_r^{\frac{1}{\alpha}} L_r \tag{10}$$

Substituting Equation (10) into the production function and re-arranging we obtain the following expression for output per capita in region r :

$$\frac{Y_r}{L_r} = A_c k_c^{1-\alpha} a_r^{\frac{1}{\alpha}} \tag{11}$$

Defining output per worker $y_r = \frac{Y_r}{L_r}$ and taking log of both sides of expression we obtain a variation of the classic growth accounting equation:

$$\ln y_r - (1-\alpha)\ln k_c = \ln A_c + \frac{1}{\alpha}\ln a_r \tag{12}$$

The above notation is helpful to understand how to construct each single component of the equation. The left hand side is obtained as log GDP per worker in the region minus the share of capital times the national capital-labor ratio. To obtain the right hand side we can regress the left hand side on a set of country dummies (standardizing one to 0) and the residuals will measure the region-specific TFP term: $\frac{1}{\alpha}\ln a_r$. Once we calculate these terms we can re-write the above equation in the following way, that captures exactly the contribution to regional productivity given by regional capital intensity, national TFP and region-specific TFP:

$$\ln y_r = \underbrace{\left[(1-\alpha)\ln k_c + \frac{1-\alpha}{\alpha}\ln a_r\right]}_{\text{Regional Capital Contribution}} - \underbrace{\ln A_c}_{\text{Country TFP}} - \underbrace{\ln a_r}_{\text{Narrow Region TFP}} \tag{13}$$

We call this measure of regional TFP the "narrow" TFP definition.

3.1.2 Broad TFP definition

The second method assumes simply that $k_c = k_r$, i.e. that the capital labor ratio in each region is equal to the average capital labor-ratio of a country. Under this condition, we simply divide both sides of equation (8) by L_r, we substitute the condition above and take logs. The decomposition of output per worker results as follows:

$$\ln y_r = \underbrace{[(1-\alpha)\ln k_c]}_{\text{Country Capital Contribution}} - \underbrace{\ln A_c}_{\text{Country TFP}} - \underbrace{\ln a_r}_{\text{Broad Region TFP}} \tag{14}$$

Note the similarity of the above decomposition with equation (12). Now, as k_c is assumed constant within countries, the differentials left after controlling for country-specific factors, are attributed completely to regional TFP differentials. Compared to the previous method, it is as if the direct effect of TFP and its indirect effect, caused by larger physical capital in the region, are merged together. As TFP differences are the cause for the existence of both terms in

11

the decomposition (13), it makes sense to consider also this second method. We call this measure of regional TFP the "broad" TFP definition. In the implementation of both of these accounting equations we standardize by the average output per worker, capital per worker and the average national TFP in the US to be 1.

4 The Data

Our data on Patented inventions, their Inventors and Citations between them are taken from the NBER Patent Citation data File described in Jaffe et al. [9]. We choose only patents granted between 1975 and 1996 for which citation data are available and we organize them across 141 regions[3]: 51 US states including D.C., 10 Canadian Provinces and 80 regions in 17 European Countries. The regions chosen for Europe are the Territorial Units used by Eurostat and denoted with the name NUTS1. Patents are assigned to the regions of residence of their first inventor. As the original NBER file contains only the name of the city and the zip code of the first inventor we mapped each city-zip code into a region with the help of national Maps and Gazetteers. The location of the first inventor is highly correlated with the location of the other inventors so that this method gives a careful representation of the distribution of innovation in Europe and North America. Moreover the location of the inventor is much more accurate in order to locate where the invention has been in developed, compared to the location of the headquarters of the assignee company.

In our estimate of knowledge flows we want to be particularly careful. We want to measure knowledge diffused to other regions and available there to be used by others and not track the diffusion of knowledge within a network of companies, potentially located in different regions. While the first flow may generate a pure externality the second simply mirrors the existence of large multinational companies that compensate their department in different regions for providing and diffusing knowledge. Therefore we do not include in the count of citations those done between two patents assigned to the same institutions (university, company or research lab). We call these self-citations and we exclude them from the count $c(r,s)$[4].

The data on employment in R&D and spending per employee in R&D used are the average for the 1992-1996 period of the number of R&D employees and of dollar spent per employee in each region. The averaging over four years allows us to fill some existing gaps in the regional yearly data and provides a value which is less affected by year to year fluctuations. For European regions the data are from the REGIO dataset, for Switzerland data were provided by the Swiss statistical office, for the U.S. data are from the NSF and for Canada they

---

[3]The list of regions and countries they belong to is described in Appendix B

[4]Interestingly if we estimate the specifications in table 2 including self citations the coefficient that is mostly affected, increasing by 30%, is the one on the Same Region dummy. This confirms our idea that those citations may be within company transfer of knowledge rather than leakages of it. The other coefficients change very little.

are from Statcan. Data on employment in R&D is the count of people employed in R&D activities. R&D spending has been converted in current dollars before averaging over the four years. A detailed description of the Data can be found in the Data Appendix.

5 Estimation Results

5.1 Diffusion of Ideas

We first estimate the equation of diffusion of the representative idea across regions. We divide patents into citing patents (granted in the 1992-1996 period) and cited patents (granted in the 1975-1992). The total count of patents for the citing region r is denoted as $(P_{9296})_r$ and the total count of patents for the cited region s is denoted as $(P_{7592})_s$. To have all non zero observation we add one citation to all the region-couples[5] and one patent to $(P_{9296})_r$ and $(P_{7592})_s$ for each region. The basic specification that we estimate is like equation (4). We include citing and cited regions specific effects (D_r, D_s) to capture, respectively, different propensity to cite/absorb across regions (n_r) and different intensity of ideas of the cited regions' patents (β_s). In the basic specification we include a linear term in geographical distance, in technological distance and in innovative intensity distance (more on this below) and three dummies: one for r and s being the same region $(SameR)$ one for them being in the same country $(SameC)$, and one for them sharing a border $(ShareB)$. The basic estimating equation is :

$$\ln\left(\frac{c(r,s)+1}{((P_{9296})_r+1)\left((P_{7592})_r+1\right)}\right) = a + D_r + D_s + \tag{15}$$
$$+\gamma_1(d_1)_{r,s} + \gamma_2(d_2)_{r,s} + \gamma_3(d_3)_{r,s} + f_1(SameR)_{r,s} + f_2(SameC)_{r,s} + f_3(ShareB)_{r,s} + u_{r,s}$$

Before discussing the estimation of this equation, let us describe in greater detail the three measures of distance d_1, d_2 and d_3 that we include in the regression. $(d_1)_{r,s}$ is simply the geographical distance between two regions. It is measured in thousands of kilometers and calculated as the shortest arc distance on the earth surface between two regions, choosing as location of each region the location of its largest metropolitan area. Technological distance $(d_2)_{r,s}$ is a measure of the difference in technological specialization of two regions. This measure is based on an index of technological proximity first introduced by Jaffe [10]. This index is calculated as follows: first we divide the patents granted in each region in six (thirty-six in the more detailed formula) technological groups:

---

[5]We also estimate in Table 4 the equation restricted only to strictly positive $c(r,s)$. Results are very similar to what obtained using all observations.

Chemical excluding Drugs, Computers and Communications, Drugs and Medical, Electrical and Electronics, Mechanical, Others. For each region s we construct the vector of shares of patents in each group $\underline{sh}_s = (sh_{s1}, ... sh_{s6})$. Then for each citing-cited region couple (r, s) we construct the uncentered correlation coefficient between vector $\underline{sh}_r$ and $\underline{sh}_s$ as follows:

$$\rho_{r,s} = \frac{\sum_i (sh_{si} * sh_{ri})}{\sqrt{\sum_i (sh_{si})^2 \sum_i (sh_{ri})^2}} \tag{16}$$

The uncentered correlation is also the angular distance between the vectors: two regions with identical specialization have a correlation of one, two with orthogonal specialization have a correlation of zero. The technological distance $(d_3)_{r,s}$ is equal to $1 - \rho_{r,s}$ and is bounded between 0 and 1. Finally the distance in innovative intensity $(d_2)_{r,s}$ measures how far two regions are in technological advancement, rather than technological specialization. It is the difference in absolute value of per capita patents granted in the two regions in the 1975-1992 period. A region at the frontier of technological innovation would have a large value of per capita patents while regions that are technologically lagging behind have low values of the same variable. This distance may affect the intensity of spillovers as a region may be more effective in using technology flowing from another region with similar level of technological advancement rather than from a region that is much more (or much less) technologically developed.

Table 1
Descriptive statistics

Variable	Mean	Std. Deviation	Min	Max
Number of region to region citations without self-cit.	75.2	516.7	0	68778
Number of region to region citations without self-cit., same technological sub-class	47.5	327.4	0	29478
Geographical Distance[a]	4.44	3.22	0	13.70
Technological distance (36 sectors)	0.34	0.19	0.006	1
Innovative Advancement Distance[b]	1.38	1.52	0	8.48
Number of patents in Cited region (1975-1992)	6523.2	12924.6	1	96804
Number of Patent in citing region (1992-1996)	21609	6523.2	1	52924

Notes: Citation frequencies are calculated omitting self-citations, i.e. citations between patents whose first author belong to the same company-institution..
a: Thousands of Kilometers.
b: Difference in patent per capita.

Table 1 reports some summary statistics for the data and for the three measures of distance. Geographical distance range from 0 kilometers, when citing and cited region are the same one, to almost fourteen thousands Kilometers (between the US islands of Hawaii and the Greek Islands of Nisia Aigouu-Kriti). Most of the distances, though, are below 10'000 kilometers, in fact only 2% of distances is above 10'000 Km Technological distance range from 0.006 to 1, and innovative advancement distance range from 0 to 8.48. The average number of region to region citation without self-citation is 47 but the variance is very large and the distribution very skewed with many couples with few citations and few couples with a very large number of citations. Our regression results, though, are very robust to the exclusion of outliers.

Table 2
Dependent Variable: Citation Frequency Between Regions

Dep. Var	I	II	III	IV
Distance	-0.035*	-0.027*	-0.045*	-0.039*
	(0.002)	(0.002)	(0.004)	(0.003)
Same Region	1.65*	1.62*	1.58*	1.36*
	(0.07)	(0.07)	(0.07)	(0.07)
Same Country	0.23*	0.17*	0.22*	0.21*
	(0.02)	(0.02)	(0.02)	(0.02)
Region sharing Border	0.38*	0.36*	0.44*	0.30*
	(0.02)	(0.02)	(0.02)	(0.02)
Tech. distance- 6 sectors.	-1.16*	-1.16*	-1.16*	
	(0.07)	(0.07)	(0.07)	
Tech. distance- 36 sectors.				-2.00*
				(0.06)
Distance in Innovative Advancement	-0.29*	-0.30*	-0.29*	-0.28*
	(0.01)	(0.01)	(0.01)	(0.01)
Same Language		0.15*	0.14*	0.11*
		(0.01)	(0.018)	(0.01)
Both NAFTA			-0.28*	-0.27*
			(0.04)	(0.05)
Both EU12			0.04	0.04*
			(0.023)	(0.02)
141 Citing Region Fixed Effects	Yes	Yes	Yes	Yes
141 Citied Region Fixed Effects	Yes	Yes	Yes	Yes
Observations	19881	19881	19881	19881
R^2	0.96	0.96	0.96	0.97

Notes: Citation frequencies are calculated omitting self-citations. To all region-couples is added one citation to avoid zeroes. Heteroskedasticity Robust Std errors in parenthesis. *= significant at 1% confidence level.

Table 2 presents the estimates of equation (15) using OLS with Heteroskedasticity robust standard errors. Column I is the basic specification that includes same-region, same-country and sharing-border dummies and 141 citing-region, 141 cited-region dummies. Column II adds a dummy for the same language

spoken in the two regions, and Column III add a dummy for both region in the E.U. and a dummy for both regions in NAFTA. Finally Column IV uses the Technological distance index calculated using 36 sectors rather than 6.

The coefficients of the three measures of distance are extremely precisely estimated and very stable across specifications. The effect of geographical distance on knowledge flow is a decrease of 3-4% per thousand Kilometers travelled. At five thousands Km about 15-20% less of the initial flow of knowledge arrives. Moreover being in the same region, in the same country or sharing a border have each an extremely strong and positive effect on knowledge flows. Two regions in the same country, for instance have about 20% more flows of knowledge than two regions with the same characteristics but in two different countries. If the citing and cited regions are the same the intensity of flows is 160% stronger than for two different regions. Although we are controlling for self-citations such strong intensity of flows within the same region may be due to some formal or market mediated relations between inventors who operate near each other. Also sharing a regional border has an extra effect of 40% on flows. This effect may be due in part to the arbitrariness of some regional borders that may cut into some economic units such as metropolitan areas (as New-York and New Jersey). Technological distance has also a very strong effect. Using the index based on 36 sectors, increasing the technological distance between two regions of one standard deviation (0.19) implies a fall of almost 40% in knowledge flow between two regions. Finally also the distance in innovative advancement has a very strong effect on knowledge flows. Increasing such distance by a standard deviation (1.52) would reduce flows by 45%.

Less relevant in determining knowledge flows, but still important, are the other characteristics of sharing the same language (increases flow by 10%) or belonging to the same custom union. Belonging to the European Union increases flow between two regions by 4% while belonging to NAFTA seems to have a negative effect of -27%. This is probably due to the fact that US patents, other things being equal, tend to cite European patents more than Canadian patents. The "generalized gravity equation" does an excellent job in explaining knowledge flows as evidenced by patent citations. Although we do not want to overemphasize the meaning of the R^2 of our regressions, as we are including 392 dummies, their values is always between 0.97 and 0.98.

5.1.1 Robustness Checks

In spite of controlling for a measure of technological distance and a measure of "innovative distance" we may still think that part of the effect of geographical distance on knowledge flows is an artifact due to proximity in sector specialization. As we want to specify correctly the diffusion of knowledge we check for this possibility. If two close regions are specialized in similar sub-sectors and similar sub-sectors cite each other the geographical distance may be picking up this effect. In Table 3 we estimate the same specifications as done in table 2 but using only, as citations, those between patents in the same technological sub-category. Citations across categories are eliminated so that no issue of sector-composition

can affect the results. This amounts to using only one narrowly defined techno-logical category and analyze the citations within it and then pool all categories to estimate common parameters of diffusion within each sector. Remarkably most of the coefficient are very similar to those in Table 2. The coefficient on geographic distance, in particular, does not change at all. Even the coefficient on technological distance, in spite of a decrease in its point estimate, remains large and significant. This suggests that technological distance is an important determinant of knowledge flows, even within technological categories and not only across them.

Table 3
Dependent Variable: Within sub-sector Citation Frequency Between Regions

Dep. Var	I	II	III	IV
Distance	-0.036*	-0.029*	-0.046*	-0.040*
	(0.002)	(0.002)	(0.003)	(0.003)
Same Region	1.64*	1.61*	1.57*	1.35*
	(0.05)	(0.05)	(0.05)	(0.07)
Same Country	0.21*	0.16*	0.21*	0.20*
	(0.02)	(0.02)	(0.02)	(0.02)
Region sharing Border	0.38*	0.36*	0.33*	0.30*
	(0.02)	(0.02)	(0.02)	(0.02)
Tech. distance- 6 sectors.	-1.12*	-1.11*	-1.11*	
	(0.06)	(0.06)	(0.06)	
Tech. distance- 36 sectors.				-1.94*
				(0.06)
Distance in Innovative	-0.29	-0.30	-0.29	-0.29*
Advancement	(0.01)	(0.01)	(0.01)	(0.01)
Same Language		013*	0.11	0.10*
		(0.01)	(0.01)	(0.02)
Both NAFTA			-0.26*	-0.25*
			(0.05)	(0.05)
Both EU12			0.02*	0.01
			(0.02)	(0.02)
141 Citing Region Fixed Effects	Yes	Yes	Yes	Yes
141 Citied Region Fixed Effects	Yes	Yes	Yes	Yes
Observations	19881	19881	19881	19881
R^2	0.96	0.96	0.96	0.965

Notes: Citation frequencies are calculated omitting self-citations and counting only citations between patents in the same technological sub-category. To all region-couples is added one citation to avoid zeroes. Heteroskedasticity Robust Std errors in parenthesis. *= significant at 1% confidence level.

A second check is done in order to make sure that our results do not depend on the inclusion of region couples with zero citations. In Table 2 we used all

the regions' couples to get our estimates, including those couples that had zero citations (we added one to all cells). In table 4 we check that the results are robust to the elimination of the zero citations cells. The same specifications are estimated as in Table 2 (column I and II) or in table 3 (column Ia and IIa) only using the couples of regions with at least one citation between each other. Again the estimates are extremely stable and similar to those in table 2 and 3 revealing that the zero-citations region couples do not have any key information on knowledge flow that is lost considering only the positive citation couples.

Table 4
Omitting regions with 0 citations

Dep. Var	I	II	Ia	IIa
Distance	-0.04*	-0.035*	-0.04*	-0.035*
	(0.003)	(0.003)	(0.003)	(0.003)
Same Region	1.80*	1.60*	1.77*	1.57*
	(0.05)	(0.06)	(0.05)	(0.06)
Same Country	0.18*	0.16*	0.18*	0.17*
	(0.02)	(0.02)	(0.02)	(0.02)
Region sharing Border	0.33*	0.29*	0.32*	0.29*
	(0.02)	(0.02)	(0.02)	(0.02)
Tech. distance- 6 sectors.	-1.76*		-1.76*	
	(0.07)		(0.07)	
Tech. distance- 36 sectors.		-2.15*		-2.19*
		(0.06)		(0.06)
Same Language	0.16*	0.16*	0.16*	0.15*
	(0.02)	(0.02)	(0.02)	(0.02)
Both NAFTA	-0.18*	-0.17*	-0.17*	-0.16*
	(0.04)	(0.04)	(0.04)	(0.04)
Both EU12	0.001	0.01*	0.006	0.01*
	(0.02)	(0.02)	(0.02)	(0.02)
Distance in Innovative	-0.12*	-0.12*	-0.13*	-0.13*
Advancement	(0.01)	(0.01)	(0.01)	(0.01)
141 Citing Region Fixed Effects	Yes	Yes	Yes	Yes
141 Citied Region Fixed Effects	Yes	Yes	Yes	Yes
Observations	14375	14375	13483	13483
R^2	0.97	0.97	0.97	0.97

Notes: Citation frequencies are calculated omitting self-citations, i.e. citations between patents whose first author belong to the same company-institution. I-II include all non-self citations Ia-IIa include only citations within same technological sub-category. Region-couples with zero citations between them are dropped. Heteroskedasticity Robust Std errors in parenthesis. *= significant at 1% confidence level.

5.1.2 Effect of Geographical Distance

The decay of knowledge flows with distance using patent citations has been the object of important empirical analysis (beginning with Jaffe et al.[11] and con-

tinuing in more recent work). This dependence of knowledge flows on distance has been often used to justify the concept of localized spillovers. Moreover it is interesting to compare how knowledge flows differ from trade flows as distance increases, now that we have a common frame given by the gravity equation. In our estimation we were able to separate the effect due to pure geographical distance from other effects such as technological proximity, crossing a national border, crossing a linguistic border and so on. We find, therefore, a genuine distance effect even in the most immaterial of the flows, the one between an existing idea and a new idea. This is interesting as, given the very low cost of spreading information one could have assumed no effect of distance on flows of ideas. Nevertheless as we quantify this effect we realize that knowledge flows are much less sensitive to distance than trade flows. First of all we explore if it is appropriate to choose a specification of the decay of knowledge flows that is linear in distance. This is interesting as the standard trade literature specifies a decay which is linear in the logarithm of distance. Such specification implies that flows decay much faster at low than at large distance. If this is true we would observe a clear convexity of the decay function when we estimate a linear specification.

Figure 1: Decay of citation frequency with distance: different polynomial specifications
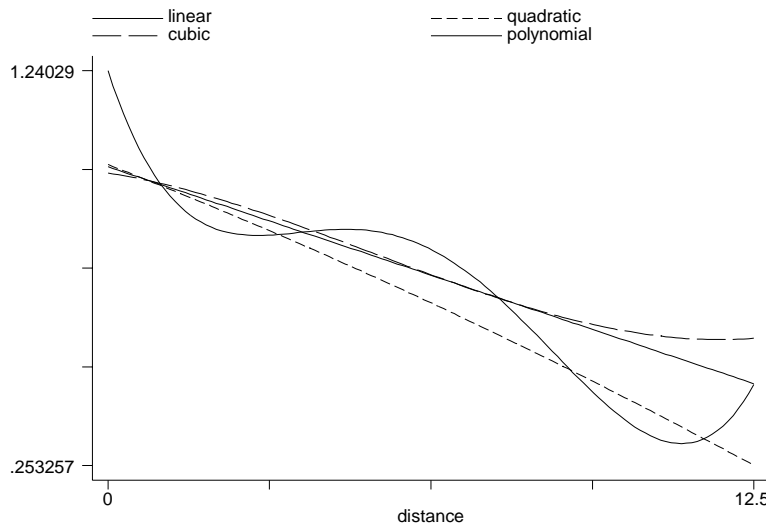
Figure 1 represents the comparison of the decay function of knowledge diffusion with distance, when we estimate such a decay as linear, quadratic, cubic or polynomial in distance. We impose that the linear decay start at 1 (=100% of the knowledge flow) at distance zero and we plot the decay between zero and

19

12.5 thousands of Kilometers. The linear decay shows that the intensity of flows is reduced by about 50% when we reach the distance of 12'000 Kilometers. The function does not exhibit any global concavity or convexity so that we do not gain much using a quadratic or a cubic (dashed lines). Using the polynomial specification we observe a pattern in which the decay is steeper in the first 2000 Km, then a rather flat portion follows up to 6'000 Km and then a steeper decay up to 10'000 Km flattening out after that. If we believe in the details of the polynomial representation, we may speculate on how knowledge flows are more heavily harmed by the initial thousands kilometers and, after that, they have another significant drop only if we cross the ocean (the average distance Europe-US is about 6'000 Km, and recall that all our observations are in Europe and North America). However, such pattern does not dramatically depart from a linear decay. Overall the linear approximation does not miss much, especially if we consider that no point estimate is perfectly reliable. Figure 2 shows the 95% confidence band for the linear estimate and we see that this band contains the point estimate of the polynomial decay for almost the whole range 0-12.5. All in all the linear decay seem to do a good job in approximating the decrease of knowledge flow with distance as captured by patent citations.

Figure 2: Decay of citation frequency with distance: Linear with 95% confidence band and Polynomial.
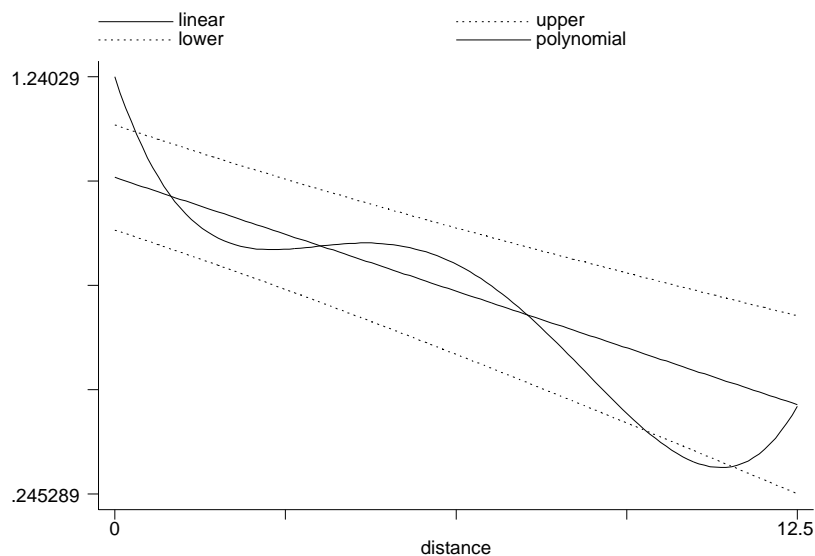
If we compare the effect of distance on knowledge flows with the effect of distance on trade, on the other hand, we find that knowledge flows are genuinely much more international. Estimating specification II in table 2 using

the logarithm of distance (not reported), rather than the level, we can readily compare the effect of distance on inter-regional knowledge flows with the effect of distance on inter-regional trade flows, as estimated using Canadian provinces and US states by McCallum [17] . This is one of the few studies using regional trade data that controls for a border effect on top of the distance effect and it is a natural benchmark for our comparison. Our estimate gives a coefficient of -0.14 (std.err. 0.007) as the elasticity of knowledge flows to distance, while McCallum estimates an elasticity between distance and trade of -1.3 to-1.5. The effects of distance on trade flows is ten times stronger than the effect of distance on knowledge flows. Incidentally also the border effect estimated by McCallum [17](about 3.1) is more than ten times larger than the one we estimate for knowledge flows (0.23).

5.2   Regions' Balanced Growth Path

Estimation of Equation (7) is the second task of the paper. That relation relies on the assumption that knowledge growth across regions is on average in Balanced Growth Path during the considered period. In this section we test two important characteristics to make sure that knowledge growth across regions is on BGP with a common long run rate of growth. This, in turn, ensures that Equation (7) is well specified. The two conditions are:

1) $(Pat_{9296} * \widehat{\beta})_r = \gamma (Pat_{7592} * \widehat{\beta})_r + e_r$

2) The deviations from the BGP relation written above, e_r, have zero average, and are not correlated to $R\&D_r$ or h_r.

The first condition is the formal implication of the assumption that regions are on BGP growth for their generation of ideas. Assume that A_{rt} is the existing stock of ideas in region r and year t and that g is the common yearly growth rate of the stock of ideas in BGP for each region r. Then considering three years such that $t_0 < t_1 < t_2$ we have, for the generic region i:

$$A_{rt_2} - A_{rt_1} = [(1+g)^{t_2 - t_1} - 1]A_{rt_1} \tag{17}$$

$$A_{rt_1} - A_{rt_0} = [(1+g)^{t_1 - t_0} - 1]A_{rt_0} \tag{18a}$$

$$A_{rt_1} = (1+g)^{t_1 - t_0} A_{rt_0} \tag{19}$$

Merging the three expressions and considering $t_0 = 75$, $t_1 = 92$, $t_2 = 96$ we obtain $(A_{r96} - A_{r92}) = \gamma (A_{r92} - A_{r75})$ where $\gamma = \{[(1+g)^4 - 1](1+g)^{17}\}/[(1+g)^{17} - 1]$. Measuring the change of stock of ideas between two years as the number of patent granted times the estimated intensity of ideas in each patent the equation written above yields: $(Pat_{9296} * \widehat{\beta})_r = \gamma (Pat_{7592} * \widehat{\beta})_r$. Adding a random disturbance we have the relation under 1).

The second condition ensures that the deviations from BGP, which may be observed, are zero mean, purely random and not correlated with other determinants of innovation so that the errors u_{rs} in equation (7) are uncorrelated

with the explanatory variables and estimates of the structural parameters are consistent.

Figure 3: Full Sample 147 Observations
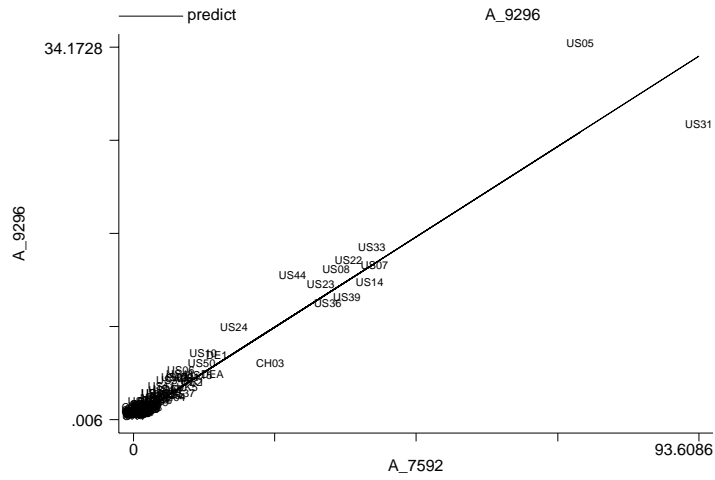
Figure 4: Without top 2 Observations
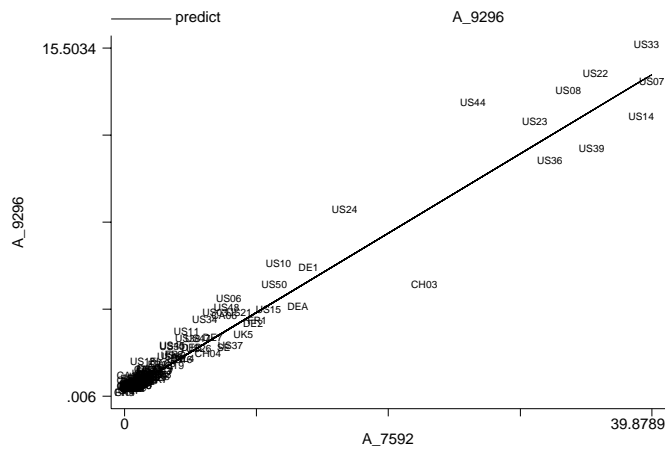
Before performing a formal test of condition 1) it is useful to look at Figure 3 and 4. The BGP relation implies that there is a linear relation with zero intercept between the level of ideas generated in region r during the period 75-92 $(Pat_{9296} * \widehat{\beta})_r$ and those generated during the period 92-96 $(Pat_{7592} * \widehat{\beta})_r$. Figures 3 and 4 plot these two variables against each other. Figure 3 incudes all regions while Figure 4 excludes the top two outliers (California and New Jersey) that have many more patents than the other regions. The impression that we get, just from looking at the pictures, is that there exist a very tight linear relation between the two variables and that the intercept of the regression line is just about zero. Table 5, confirms that a linear relation between the two variables explain 95% of the variance of $(Pat_{9296} * \widehat{\beta})_r$ and confirms that the intercept of the linear relation is not statistically different from 0 (column I, with all observations and Column III without outliers). Also, no significant concavity or convexity, captured by a quadratic and a cubic term, exists in the relation (Column II and Column IV).

Table 5: Balanced Growth Path relationship

Dep. Val $(A_{\_9296})_r = (Pat_{\_9296})_r * \hat{\beta}_r$ /1000

	I	II	III	IV
	All Observations		**Without top 2 outliers**	
	(146 obs)		**(144 obs)**	
Constant	-0.18	0.33	0.15	-0.11
	(0.14)	(0.17)	(0.08)	(0.07)
A$_{\_7592}$	0.35[*]	0.22[*]	0.36[*]	0.37[*]
	(0.03)	(0.10)	(0.02)	(0.07)
(A$_{\_7592}$)^2		0.006		0.001
		(0.005)		(0.008)
(A$_{\_7592}$)^3		0.00006		0.00004
		(0.00004)		(0.0001)
R^2	0.95	0.96	0.96	0.96

Table 6 also checks that the residuals e_r from the regression under 1) have no correlation with $R\&D_r$ and h_r. Neither in level (Column I) nor in Logs (Column II) there exist any significant correlation between the residuals and those explanatory variables. These checks, therefore, do not reject our assumption that knowledge creation in European and North-American regions has been on average on a BGP with a common long run growth rate of A_r. Deviations from this BGP are rather small, and totally random. In particular they are uncorrelated with regional R&D. As a consequence equation (7) is correctly specified and the level of knowledge generated in each of these regions in the period 1992-1996, depended on the R&D inputs and on the spillovers affecting the ideas generating function.

Table 6: Independence of Residuals from Explanatory Variables

Dep. Var/ Regressors:	e_r	$\ln(e_r)$
R&D$_r$/1000	0.0023	
	(0.0047)	
(h$_r$)	0.074	
	(0.47)	
ln(R&D$_r$)		0.06
		(0.05)
ln(h$_r$)		-0.01
		(0.02)
Country Dummies	Yes	Yes
R^2	0.16	0.47
Obs	146	146

Notes: **Dep var:** e_r= residuals from regression of A_9296==(Pat_9296)$_r$*exp($\hat{\beta}_r$)/1000 on A_7592=

=(Pat_7592)$_r$*exp($\hat{\beta}_r$)/1000in levels

$\ln(e_r)$= residuals from a regression as above but in logs

5.3   Innovation Function

We have established, so far, that knowledge flows across regions depend on several bilateral characteristics and notably on distance in geographical, technological and innovative space. Proving the existence of knowledge flows though, does not prove the existence of externalities (and in particular of positive externalities) of knowledge in generating innovation unless these flows affect the productivity of researchers in the receiving region. Usage of knowledge originated in other regions may very well bring, together with contribution for new ideas, increased standard of novelty, or a reduction in the "yet unexplored" innovation possibilities. These effects may very well generate a zero net effect or even a negative net effect on the productivity of researchers in generating new ideas. Therefore, while no priors are imposed by the theory on the sign and magnitude of ε_I, we expect that $\varepsilon_{R\&D}$ and ε_h are positive, as an increase of researchers and of resources should increase their findings of new ideas (patented ideas).

We estimate the innovation function in (7). First we construct the stock of available ideas in each region r, A_r^{AVA} as $\sum_{s\in S} \exp\left(-\widehat{\gamma}_1(d_1)_{r,s} - \widehat{\gamma}_2(d_2)_{r,s} - \gamma_3(d_3)_{r,s} + \widehat{f}(Char_{r,s})\right) \widehat{\beta}_s P_s$
. We use the parameter estimates from Table 2 column I in order to calculate this value. As we have seen those estimates are very robust and precise and therefore it would not make much of a difference to use the estimates from column II, III or IV. In estimating the equation we include fixed country-effect C_i, as the propensity to patent in the US may differ across countries, due to different costs and we control also for the previously estimated average intensity of ideas per patent in regions, $\ln(\widehat{\beta}_r)$. Controlling for this variable is important as there is a strong and significant negative correlation between the intensity of

ideas per per patent $\ln(\beta_r)$ and the patenting intensity $\ln(P_r)$ across regions. Such correlation is represented in Figure 5: regions producing more patents tend to have lower intensity of ideas in each patent.

Figure 5: Intensity of ideas and Patenting intensity across regions
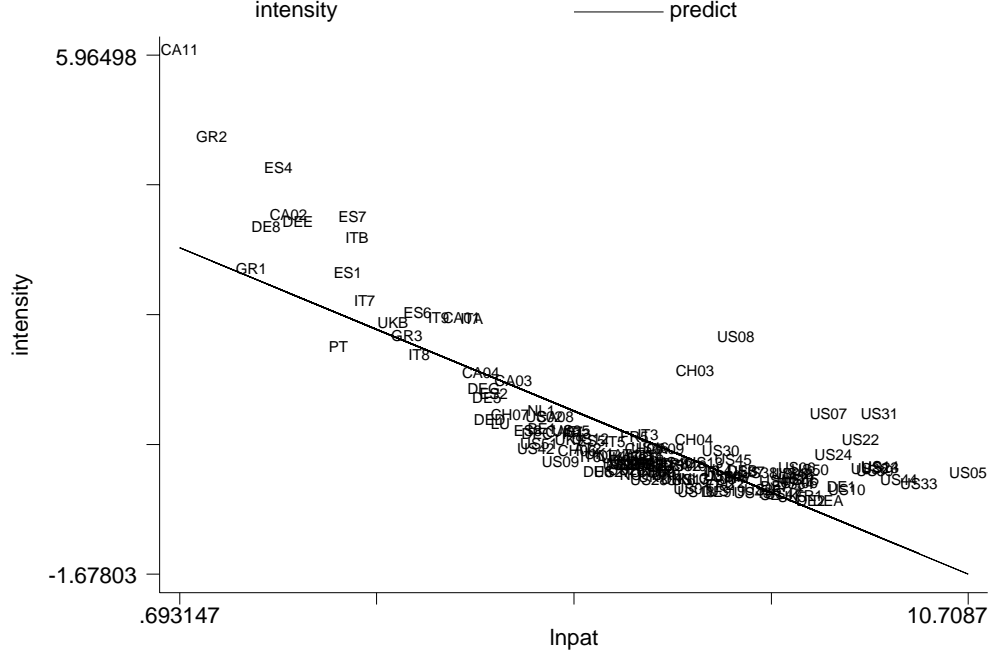
Controlling for $\ln(\beta_r)$ reduces the variability of innovation in response to its inputs. The exact estimated specification is:

$$\ln(P_r) = C_i + c\ln(\widehat{\beta}_r) + \varepsilon_{RD}\ln(R\&D_r) + \varepsilon_h\ln(h_r) + \qquad (20)$$
$$+ \varepsilon_I \ln A_r^{AVA} + u_{r,s}$$

Table 7 reports the estimates of the coefficients of equation (20). The estimation method is Instrumental Variables , with Heteroskedasticity robust standard errors. In column I and II the variable

$$A_{-r}^{AVA} = \sum_{s \neq r} \exp\left(-\widehat{\gamma}_1(d_1)_{r,s} - \widehat{\gamma}_2(d_2)^{r,s} - \gamma_3(d_3)_{r,s} + \widehat{f}(Char_{r,s})\right)\widehat{\beta}_s P_s$$

25

is used as instrument for A_r^{AVA}, where we have omitted the contribution of knowledge generated in region r itself to avoid endogeneity. In column III and IV the variable

$$R\&DS_r^{AVA} = \sum_{s \neq r} \exp\left(-\widehat{\gamma}_1(d_1)_{r,s} - \widehat{\gamma}_2(d_2)^{r,s} - \gamma_3(d_3)_{r,s} + \widehat{f}(Char_{r,s})\right)(R\&D*h)_s$$

is used as instrument, as our model predicts that resources in R&D are the main exogenous determinant of innovation in a region. In both cases the instrument does not include the dependent variable in its calculation and therefore it corrects for the upward bias of the OLS estimation. Column I and III include $R\&D_r$ and h_r as separate inputs. Interestingly, though, the elasticity of innovation are not significantly different between these two inputs and we may consider total R&D spending ($R\&D_r*h_r$) as a single factor. Doing so (column II and IV) we obtain a more precise estimate of the elasticity of innovation to this factor. The estimate of such elasticity is around 0.8 with a standard error of 0.07. An increase of R&D resources of 1% increases the generation of new ideas by 0.8% in the long run. Unluckily and differently from the estimates of ε_{RD} and ε_h the externality ε_I is not estimated precisely. Its standard error is in fact between 0.4 and 0.5. Nevertheless the point estimate is never significantly different from 0 and in column III and IV it is even negative (-0.5). All in all the evidence point towards large returns of R&D resources in innovation and no positive externalities of knowledge in generating new ideas. In particular if we consider the estimates of column I and II the point estimate of ε_I is around 0.1 with standard error of 0.4. We may consider these estimates as potentially biased upward due to the endogeneity of knowledge in the neighboring regions. If we instrument P_s with $(R\&D*h)_s$ then we obtain a much smaller and negative estimate (-0.5) of ε_I (column III and IV) still non significantly different from 0. It is possible that a positive "spillover" of available knowledge is (more than) offset by a negative effect due to increased standard for local innovation brought in the region by the diffusion of outside ideas. Similarly the knowledge and use of existing ideas may very well reduce the space of "unexplored" ideas, making innovation more difficult.

<div align="center">

Table 7: the production function of innovation

Dep. Variable $\ln(Pat_{\_9296})$

</div>

Dep. Var: $\ln(Pat\_9296)$	I	II	III	IV
$\ln(R\&D)_r$	0.87*		0.93*	
	(0.12)		(0.13)	
$\ln(h)_r$	0.65*		0.71*	
	(0.17)		(0.18)	
$\ln(R\&D*h)_r$		0.76*		0.83*
		(0.067)		(0.07)
A_r^{AVA}	0.11	0.12	-0.55	-0.59
	(0.43)	(0.41)	(0.50)	(0.52)
$\hat{\beta}_r$	-0.56*	-0.55*	-0.66*	-0.66
	(0.14)	(0.11)	(0.12)	(0.13)
Country Dummies	Yes	Yes	Yes	Yes
Observations	141	141	141	141
R^2	0.92	0.92	0.91	

Notes: Method of estimate: IV with robust std error. Instruments in I, II: $\ln\left(\sum_{s\neq r}\hat{\phi}(r,s)\Delta A_s\right)$ or in III and IV

$\ln\left(\sum_s \hat{\phi}(r,s)R\&D*h_s\right)$

5.4   Instrumental Variables: Value of a Patent

Our estimates of the effect of R&D on innovation could be biased up by the presence of some unobservable regional factors , that attract R&D while also increasing its returns in terms of innovation. In order to address this issue we consider a proxy for the value of a patent that varies across regions. If we knew how much a patent were worth in a region, we could explicitly model the fact that researchers are attracted to regions in which patents are more valuable because their expected returns are higher. Moreover if the market value of a patent is independent from the productivity of R&D we could use such variable as an instrument for R&D in regions .

In order to do this, we turn to the idea that a patent is more valuable the largest is the potential market for innovation in the region where it is discovered. Different regions have different market potentials for innovations, depending on their location and their connections with the other markets. The amount of trade of a region with the rest of the world could reveal the market for the region. More interestingly, though, we have a revealed measure of the potential market for patented goods. This measure is the extent to which patent protection is pursued by resident of a country in other countries. If a patent is protected only on the domestic market it is because the inventors believe that small profits would come by trading the good abroad and therefore it is not worth seeking protection there. On the contrary a patent that is vastly protected reveals the intention of its inventors to protect foreign market and profit on them via trade. Seeking international patenting in other countries reveals what regions the inventor considers as potential markets for the new goods invented there.

We use data available from WIPO on patents applied for and granted in

each of our 19 countries to inventors residing in each of them for the period 1993-1996. Identifying the patent protection in a country as the sign that an invention is targeting that country as a market, allows us to estimate the effect of country characteristics and distance on the potential demand for innovations coming from each country. This allows us to calculate the market potential for innovation in each of our regions. We use this measure of market potential as instrument for regional R&D.

Here we briefly describe how we use the data on cross-country patenting to infer the market potential for innovation in each region. Let's assume that the number π_{ij} of patents granted in country j to residents of country i is a good proxy of the market for new goods that inventions from country i have in country j (in relative terms). We refer to π_{ij} as the market for innovation generated in country i coming from country j. This seems reasonable as patent protection is a way of ensuring market for new goods.

It is useful to think of the market π_{ij} as depending on country j and country i characteristics and on bilateral characteristics affecting the relative intensity of patenting from a country into another. The bilateral characteristic that we use are geographical distance ($dist_{ij}$) and a same-country dummy ($SameC_{ij}$).

If we think of π_{ij} as the share of total world patents that are granted by country j to residents of country i we can decompose that value as follows:

$$\pi_{ij} = (\Pi_{i,all})(\Pi_{all,j})(e^{-\delta dist_{ij}})(e^{f SameC_{ij}})\varepsilon_{ij} \tag{21}$$

$\Pi_{i,all}$ is the share of total patents in the world granted to inventors resident of country i, $\Pi_{all,j}$ is the share of of total patents in the world granted by the patent office of country j. $e^{-\delta dist_{ij}}$ is an exponential function in the geographical distance between i and j, $e^{f SameC_{ij}}$ is the effect of patenting in the same country where the inventor is resident and ε_{ij} is a positive multiplicative random factor distributed as a lognormal that captures all the other bilateral determinants of π_{ij} orthogonal to the regressors. As we observe π_{ij} and $dist_{ij}$ and $SameC_{ij}$ we may use a simple regression to perform the above decomposition. Taking logs of both sides of equation (21) we have:

$$\ln(\pi_{ij}) = C_i + C_j - \delta(dist_{ij}) + f(SameC_{ij}) + u_{ij} \tag{22}$$

where C_i and C_j are country specific effects and capture the origin and destination country effects ($\ln(\Pi_{i,all}), \ln(\Pi_{all,j})$). $dist_{ij}$ is the geographical distance between the two regions and $SameC_{ij}$ is a dummy which is equal to one if the inventor's country and the granting country are the same. $u_{ij} = \ln(\varepsilon_{ij})$ is a zero mean normally distributed error. Once we have used OLS to estimate equation (22) we can construct the predicted share $\pi_{si,rj}$ which measure the innovation generated in region s_i of country i and demanded in region r_j of country j..While the parameter C_i depends on how innovative is country i the parameters C_j, δ and f depend only on where inventions are patented and how the distance and national borders affect this marketing decisions. They are therefore variables

28

that capture the market potential of a region. Let's call $shmar_{r_j}$ the share of region r_j in the market for innovation of country j. Also we denote with a hat $\widehat{}$ the OLS estimates of our parameters. The predicted potential demand coming from region r_j for innovation invented in s_i would be:

$$\widehat{\pi}_{si,rj} = (shmar_{rj}\widehat{\Pi}_{all,j})\exp(-\widehat{\delta}dist_{si,rj} + \widehat{f}SameC_{s_i r_j}) \qquad (23)$$

In our implementation we measure $shmar_{rj}$ as the share of GDP of country j produced in region r_j Therefore the total market potential for innovation produced in region s_i would be:

$$Pot1_{si} = \sum_{r_j}\sum_{j}\widehat{\pi}_{si,rj} \qquad (24)$$

Alternatively, using a region's gross product as proxy for the demand for new goods (rather than $shmar_{rj}\widehat{\Pi}_{all,j}$) we can define an alternative measure of demand from region r_j for innovation invented in s_i : $\widehat{p}_{si,rj} = \exp(-\widehat{\delta}dist_{si,rj} + \widehat{f}SameC_{s_i r_j})Y_{r_j}$. The total market potential for region s_i would be:

$$Pot2_{si} = \sum_{r_j}\sum_{j}\widehat{p}_{si,rj} \qquad (25)$$

Both constructs use the geographic position of each region and the inter-country pattern of demand for innovations, revealed by international patenting, to evaluate the main determinant of patent's value, which is the market potential that a patent has in region s_i once invented.

Table 7b: Instrumenting R&D with Mkt potential

Dep. Variable $\ln(Pat_{-9296})$

Dep. Var: $\ln(Pat\_9296)$	I	II	III	IV	V	VI
$\ln(R\&D*h)_r$	0.77[*]	0.79[*]	0.77[*]	0.99[*]	0.80[*]	1.01[*]
	(0.04)	(0.05)	(0.26)	(0.19)	(0.27)	(0.19)
A_r^{AVA}		-0.52		-0.79		-0.72
		(0.41)		(0.59)		(0.56)
$\widehat{\beta}_r$	-0.57*	-0.52[*]	-0.58[*]	-0.57[*]	-0.55[*]	-0.54[*]
	(0.10)	(0.11)	(0.22)	(0.24)	(0.22)	(0.23)
Country Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	141	141	141	141	141	141
R^2	0.92		0.92	0.89	0.91	0.89

Notes: Method of estimate: Column I and II OLS, Column III and IV, Instrumental variable with robust std error. Instrumenting $\ln(R\&D*h)_r$ and A^{AVA}_r with Pot1 and Pot1$_{AVA}$. Column V and VI use Pot2 and Pot2$_{AVA}$ as instrumental variables.

Table 7b reports the results of estimating equation (20) using $Pot1_{si}$ or $Pot2_{si}$ and its transform $Potn_{AVA} = \sum_{s \neq r} \sum_{s \neq r} \exp\left(-\widehat{\gamma}_1(d_1)_{r,s} - \widehat{\gamma}_2(d_2)^{r,s} - \gamma_3(d_3)_{r,s} + \widehat{f}(Char_{r,s})\right) Potn$ to instruments for total $R\&D$ and for A_r^{AVA}. We report the innovation equation without the spillover term estimated using OLS and IV with $Pot1_{si}$ and $Pot2_{si}$ (columns I,II and V). As the instruments are not very good (explain about 10% of the independent variable variation) the std. errors increase substantially, but the point estimates do not change much. No evidence of any significant bias on the effect of $R\&D$ on innovation emerges from these estimates. Similarly, when we introduce the spillover term $\ln A_r^{AVA}$ its effect is still negative and insignificantly different from 0 using OLS or instrumental variables. This check re-assures us on the previous estimates of the impact of R&D on innovation and of the R&D externalities.

5.5 Cross-Regional Accounting

The source of data used to perform cross-region accounting are several and described in detail in the appendix. In general, data on gross regional product and total regional employment are from the REGIO dataset for E.U. countries, from the national statistical offices for Norway and Switzerland, from the Bureau of Economic Analysis for US states and from CANSIM for Canadian Provinces. For each region we take the average value for the 1991-1996 period, in order to smooth any effect from business cycle fluctuations. The Capital-Labor ratio at the country level is taken from Hall and Jones [8] and is relative to year 1988. The share of labor is assumed to be the same across regions and equal to 0.66. This section documents that the variation of region-specific total factor productivity explains a substantial part of the differences in regional labor productivity. Differences across regions within countries are as important as differences across countries in explaining productivity differentials in Europe-North America. This can be seen in Figure 6 and 7 that decompose the cross-regional variation of $\ln(y_r)$, between a cross-country component (figure on top) and a cross-region within country component (figure on the bottom). As we can see just eye-balling the figures, and it is also revealed by the R^2 of the regressions, each component explains about one half of the total variance.

Figure 6
Cross regional variation of output per worker (1991-1996)
explained by within country variation
R^2=0.48
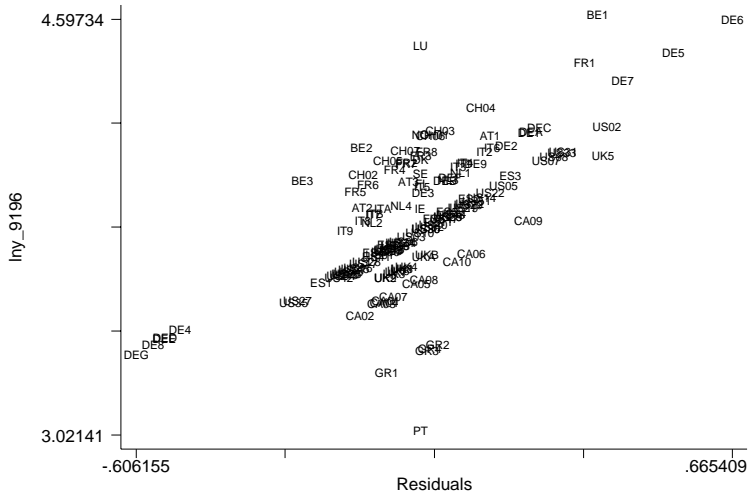
Figure 7
Cross regional variation of output per worker (1991-1996)
explained by cross country variation
R^2=0.52
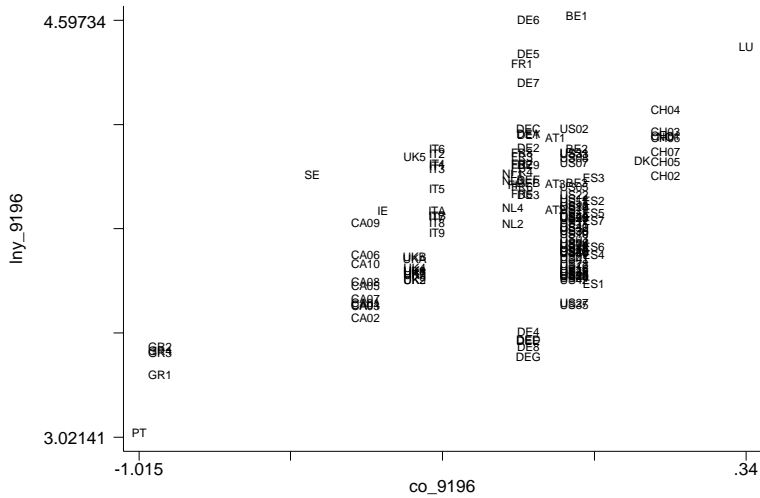
More formally we perform the accounting exercise described in section 3.1 where we standardize the average output per worker, the average capital per worker and the average country TFP of the US to be equal to 1 (therefore their log is 0). We rank regions according to their output per worker (average current $ 1991-1996) from the most productive (Bruxelles) to the least (Portugal). The contribution of each component (capital-labor ratio, country TFP and region TFP) to the differences in output per worker for each region, using the "Narrow" and the "Broad" regional TFP definition can be requested from the author. In Table 8 (a)-(b) we simply report the decomposition into components of the differences in productivity between the Top and the Bottom regions. On average between 17% and 48% of the differences in output of region in the top10 group versus region in the bottom 10 group are explained by regional TFP when we use the narrow definition. The percentage increase to 24-73% when we use the broad definition. By all means differences in regional factor productivity are very relevant in explaining differences in output per worker. Considering them as a measure of differences in technological advancement across regions we use them to estimate the impact of R&D and of knowledge spillovers.

Table 8
(a)
Cross-region Accounting using the Narrow definition of regional TFP
(1991-1996 averages, 141 Regions)

	$\Delta \ln (k_{regio})$	$\Delta \ln (TFP_{country})$	$\Delta \ln (TFP_{regio})$	$\Delta \ln (y)$	% explained by $\Delta\ln(TFP_{regio})$
Top-Bottom	0.42	0.90	0.25	1.57	17%
Top2-Bottom2	0.38	0.52	0.43	1.34	32%
Top5-Bottom5	0.54	0.01	0.51	1.05	48%
Top10-Bottom10	0.37	0.10	0.25	0.72	35%

(b)
Cross-region Accounting using the Broad definition of regional TFP
(1991-1996 averages, 141 Regions)

	$\Delta \ln (k_{country})$	$\Delta \ln (TFP_{country})$	$\Delta \ln (TFP_{regio})$	$\Delta \ln (y)$	% explained by $\Delta\ln(TFP_{regio})$
Top-Bottom	0.30	0.90	0.38	1.57	24%
Top2-Bottom2	0.17	0.52	0.66	1.34	49%
Top5-Bottom5	0.28	0.01	0.77	1.06	73%
Top10-Bottom10	0.24	0.10	0.38	0.72	52%

5.6 Productivity Function

Having determined the productivity differentials across regions, due to region specific technological level a_r we consider the long-run relation between productivity resources spent in R&D and knowledge spillovers. In particular, assuming

that a_r captures the relative level of technological advancement that results from codified and implicit innovation, patented and not patented inventions we estimate the same relation between a_r and its determinants as we did above for innovation measured by patents. We still assume that the structure of knowledge flows revealed by citations is informative on flows of knowledge across regions. As the TFP measures productivity "per worker" we also divide the R&D total resources used in the region by total employment. We estimate the following equation:

$$\ln(a_r) \quad = \quad C_i + \varepsilon_{RD} \ln(\frac{R\&D_r * h_r}{Emp_r}) + \qquad (26)$$
$$+\varepsilon_I \ln a_{-r}^{AVA} + u_{r,s}$$

where $a_{-r}^{AVA} = \sum_{s \neq r} \exp\left(-\widehat{\gamma}_1(d_1)_{r,s} - \widehat{\gamma}_2(d_2)^{r,s} - \gamma_3(d_3)_{r,s} + \widehat{f}(Char_{r,s})\right) a_s$
is the stock of knowledge available to region r through the flows from other regions. We estimate the equation using OLS with robust standard errors and using IV where we instrument the variable a_{-r}^{AVA} using the variable defined above $R\&DS_{-r}^{AVA}$.

Table 9: R&D, Knowledge spillovers and Productivity

Dep. Var: $\ln(a_{9296})$	I	II	III	IV	V(iv)	VI(iv)
$\ln(\frac{R\&D*h}{Emp})_r$	0.067*	0.10*	0.064*	0.098*	0.065	0.098*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.015)
$\ln(a_{-r}^{AVA})$			-0.024	-0.03	-0.031	-0.048
			(0.022)	(0.03)	(0.026)	(0.040)
$\ln(R\&DS_{-r}^{AVA})$	-0.03	-0.05				
	(0.03)	(0.04)				
Observations	141	141	141	141	141	141
R^2	0.23	0.23	0.22	0.22	0.23	0.23

Notes: Column I-IV estimated using OLS with With-Robust std errors. Column V and VI estimated using instrumental variables with roust std error. Instrument: $\ln(R\&DS_{-r}^{AVA})$. Column I, III and V use the narrow definition of regional TFP, column II,IV and VI use the broad definition.

Table 9 reports the estimates of equation (26). First of all let's notice that the elasticity of regional TFP to local R&D resources spent is between 0.065 and 0.10 estimated rather precisely in any specification I-VI. In general the effect is magnified when we consider the effect of R&D on broadly defined regional TFP (column II, IV and VI). This is natural, as that measure attributes also part of the regional differences due to capital intensity, to differences in productivity. The elasticity of R&D spending on broadly defined productivity is approximately 0.10 while the effect on narrowly defined productivity is 0.065.

In both cases the standard error is close to 0.01. These estimates can be compared with those in Coe and Helpman [4] and with those in Keller [16]. Both works estimate this elasticity using the time-series variation of R&D and total factor productivity at the country level. Coe and Helpman using the long-run cointegrating relation find a number between 0.078 and 0.097, very close to the effect we estimate on broad TFP. As Coe and Helpman use the long-run cointegrating relation among variables we think they are using the same information as we are. Keller, who estimates this elasticity using sector-country yearly data finds a value somewhat lower between 0.04 and 0.055. This can be due to the shorter time horizon (yearly data) of his analysis. Our results, though, contrast with the Coe-Helpman and Keller analysis and confirm our previous finding in the estimates of externalities from knowledge flows. The variable capturing available knowledge from other regions, either measured as $\ln(a_{-r}^{AVA})$ or as $\ln(R\&DS_{-r}^{AVA})$ does not have any positive and significant effect. In any specification the elasticity of regional productivity to available ideas is never larger than -0.03, with standard error of about 0.03. The standard error of the estimate is now sufficiently small to allow us to rule out any effect externality larger than 1-2% on the regional productivity.

Again, this seems to be evidence that no significant and positive effects on productivity are carried by these knowledge flows that capture the exchange of ideas at the frontier of technological advancement.

6 Conclusions

The present study analyzes the process of knowledge diffusion and knowledge externalities as evidenced by the data on TFP, patent creation and patent citations. While diffusion of ideas is needed in order to have externalities of knowledge there is no reason to believe that simply by measuring the intensity and scope of this knowledge diffusion, we have a measure of knowledge externalities. In order to have positive externalities, in fact, we need that existing ideas affect positively the productivity of scientists in generating new ideas. Several doubts have already been raised by the literature in analyzing innovation over time (Griliches [6]) that knowledge externalities are positive, as there seems to be evidence that productivity of R&D in innovation has decreased over time. Diffusion of new ideas, in fact, brings not only "new inspiration" to researchers but also increased standard for innovation, and it reduces the unexplored territories of human knowledge. These effects may offset the positive spillovers. Our study finds that there are very important positive and negative determinants of knowledge diffusion: regions farther away from each other, in different countries, specialized in different sectors and speaking different languages exhibit much lower flows of knowledge than close, similar regions in the same country. Nevertheless these flows do not bring significant knowledge externalities neither when we consider the innovation generating function, nor when we consider the more commonly used measures of total factor productivity.

References

[1] Aghion Philippe and Peter Howitt (1992) "A Model of Growth through creative Destruction" *Econometrica* 60: 323-51.

[2] Caballero R. and A. Jaffe (1993) "How High are the Giants' Shoulders: An Empirical Assessments of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth" *NBER Macroeconomic Annuals* vol. 8 pp.15-72 MIT Press, Cambridge MA.

[3] Ciccone Antonio (2002) "Agglomeration effects in Europe" *European Economic Review* (46)2 pp. 213-227

[4] Coe David and Elhanan Helpman (1995) "International R&D Spillovers" European Economic Review, 39 pp. 859-887.

[5] Griliches Zvi Eds. (1984) " R&D Patents and Productivity" University of Chicago Press, Chicago.

[6] Griliches Zvi (1994) "Productivity, R&D and the Data constraint" *American Economic Review*, 84(1) pp. 1-23.

[7] Grossman G. and E. Helpman (1991) "Endogenous Product Cycles" Economic Journal, v. 101 pp: 1214-1229.

[8] Hall Robert and Charles. I. Jones (1999) "Why do Some countries produce so much More Output Per worker than Others"? *Quarterly Journal of Economics*, 114(1) pp.83-116.

[9] Bronwyn H. Hall, Adam B. Jaffe and Manuel Trajtenberg, "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools," National Bureau of Economic Research Working Paper No. 8498 (October 2001)

[10] Jaffe A. (1986) "Technological Opportunities and Spillovers of R&D: Evidence from Firm's Patents, Profits and Market Value" *The American Economic Review,* Vol. 76 984-1001.

[11] Jaffe Adam, Manuel Trajtenberg and Rebecca Henderson (1992) "Geographic Localization of Knowledge Spillovers as evidenced by Patent Citations" *Quarterly Journal of Economics* 108(3) pp. 577-98

[12] Jaffe A. and M. Trajtenberg (1999) "International Knowledge Flows: Evidence from Patent Citation" in *Economics of Innovation and New Technology*, Vol. XX pp.105-136

[13] Jones C. (1995a) "Time Series tests of Endogenous Growth Models" *Quarterly Journal of Economics* 110 (May 1995).

[14] Jones C. (1995b) "R&D Based Models of Economic Growth" *Journal of Political Economy* 103: pp.739-784.

[15] Charles I. Jones and John C. Williams (1998) "Measuring the Social Return to R&D" *Quarterly Journal of Economics* 113(4) pp. 1119-1135 .

[16] Keller Wolfgang (2002)"Geographic Localization of International Technology Diffusion" forthcoming in *American Economic Review.*

[17] John McCallum (1995) "National Border Matters: Canada-US Regional Trade Patterns" *American Economic Review*, Vol 85 (3) pp. 615-623.

[18] Romer P.M., 1990, Endogenous Technological Change, *Journal of Political Economy* 98(5), 71-102

A Diffusion Processes Generating a Negative Exponential

The process of diffusion of ideas in space and over time can be described, in general, with a function which describes what is the contribution available at time t_1 in location r of an idea, generated in location s at time t_0. Such a contribution is a function of the non-obsolete knowledge incorporated into the idea and of the probability of researchers in location s to have incurred in the idea. In general it takes the form:

$$\widetilde{\psi}(r,s,t_0,t_1) = \widetilde{\delta}e^{-\rho(A_{t1}-A_{t0})}f\left[(t_1-t_0),d(r,s)\right] \qquad (27)$$

The term $e^{-\rho(A_{t1}-A_{t0})}$ captures the part of the idea which is non obsolete yet, assuming that obsolescence depends on the number of new ideas invented between t_0 and t_1. The term $f\left[(t_1-t_0),d(r,s)\right]$ expresses the fact that the probability of having seen the idea depends on the time elapsed since the invention took place and on the distance between the location. First notice that one we choose t_1 and t_0 equal for all the observation (cross section) the term $\widetilde{\delta}e^{-\rho(A_{t1}-A_{t0})}$ is just a common constant for all the observations. As for the function f two different assumptions give a negative exponential:

1) Assume that all ideas travel in space at the same speed v so that distance travelled in the period of time T is: $d = vT$. Assume also that v is large enough that the maximum world distance $\overline{T}$ can be travelled in the interval (t_1-t_0) namely: $v\,(t_1-t_0) > \overline{T}$. Under perfect diffusion ideas invented anywhere in the world at time t_0 are available anywhere else in the world at time t_1. Assume nevertheless that there are some "frictions" to the diffusions of ideas, and these frictions take the form of random events that arrest the propagation of an idea (bad luck, interruptions in the communications, distractions of attention). These events happen with a Poisson distribution at rate λ per unit of time. Therefore the probability that they happen and stop the idea from reaching a location is larger for farther than for closer locations. In particular the probability that an idea has travelled distance $d(r,s)$ before such interruption happens is: $e^{-\lambda\frac{d(r,s)}{v}}$. Such is the probability to have the idea available in location s at time t_1. Substituting the function into (27), and collecting into a constant all the parameters and the functions of t_0,t_1, we get:

$$\psi(r,s) = \delta e^{-\frac{\lambda}{v}d(r,s)} \qquad (28)$$

which has the same form as 2 in the text, including geographical distance only.

2) Assume that ideas travel at speed v_i different for each idea i. Assume that the speed of ideas are distributed as a negative exponential function with PDF

equal to $\lambda e^{-\lambda v_i}$, where λ is the average speed of diffusion, but the shape of the distribution captures the fact that there are more "slow" than "fast" travelling ideas. In this case the probability of an idea having reached region r at distance $d(r,s)$ is equal to the probability of the speed of that idea being at least $\frac{d(r,s)}{t_1-t_0}$ and therefore $e^{-\lambda \frac{d(r,s)}{t_1-t_0}}$. collecting constant terms, we have:

$$\psi(r,s) = \delta' e^{-\frac{\lambda}{t_1-t_0} d(r,s)} \tag{29}$$

which has the form of equation 2 in the text.

B   List of Regions

Austria: OSTOSTERREICH, SUDOSTERREICH, WESTOSTERREICH

Belgium: BRUXELLES, VLAAMS GEWEST, REGIONE WALLONNE

Canada (Provinces): NEW FOUNDLAND, PRINCE EDWARDS ISLAND, NOVA SCOTIA, NEW BRUNSWICH, QUEBECK, ONTARIO, MANITOBA, SASKATCHEWAN, ALBERTA, BRITISH COLUMBIA.

Denmark: DENMARK

Finland: FINLAND

France: ILE DE FRANCE, BASSIN PARISIENNE, NORD-PAS DE CALAIS, ESTE, OUESTE, SUD-OUEST, CENTRE-EST, MEDITERRANEE.

Germany: BADEN-WURTENBERG, BAYERN, BERLIN, BRANDEN-BURG, BREMEN, HAMBURG, HESSEN, MECKLENBURG-VORPOMMEM, NIEDERSACHSEN, NORDRHINE-WESTFALIA, RHEINLAND-PFALZ, SAAR-LAND, SACHSEN, DESSA, SCHLESWIG-HOLSTEIN, TURINGEN.

Greece: VORAIA ELLADA, KENTRIKI ELLADA, ATTIKI, NISIA AIGAIOU, KRITI.

Ireland: IRELAND

Italy: NORD OVEST, LOMBARDIA, NORD-EST, EMILIA ROMAGNA, CENTRO, LAZIO, ABRUZZO-MOLISE, CAMPANIA, SUD, SICILIA, SARDEGNA.

Luxemburg: LUXEMBURG

Norway: NORWAY

Portugal: PORTUGAL

Spain: NOROESTE, NORESTE, COMUNIDAD DE MADRID, CENTRO, ESTE, SUR, CANARIAS.

Sweden: SWEDEN

Switzerland: REGIONE LEMANIQUE, ESPACE MITTELAND, NORTH-WESTSCHWEITZ, ZURICH, OSTCHWEITZ, ZENTRALSCWEITZ, TICINO.

United Kingdom: NORTH, YORKSHIRE AND THE HUMBER, EAST MIDLANDS, EAST ANGLIA, SOUTHEAST, SOUTHWEST, WEST MID-LANDS, NORTH WEST, WALES, SCOTLAND, NORTHERN IRELAND.

USA (States): ALABAMA, ALASKA, ARIZONA, ARKANSAS, CALIFORNIA, COLORADO, CONNECTICUT, DELAWARE, D.C., FLORIDA, GEORGIA, HAWAII, IDAHO, ILLINOIS, INDIANA, IOWA, KANSAS, KENTUCKY, LOUISIANA, MAINE, MARYLAND, MASSACHUSSETS, MICHIGAN, MINNESOTA, MISSISSIPI, MISSOURI, MONTANA, NEBRASKA, NEVADA, NEW HAMPSHIRE, NEW JERSEY, NEW MEXICO, NEW YORK, NORTH CAROLINA, DAKOTA, OHIO, OKLAHOMA, OREGON, PENNSYLVANIA, RHODE ISLAND, SOUTH CAROLINA, SOUTH DAKOTA, TENNESSEE, TEXAS, UTAH, VERMONT, VIRGINIA, WASHINGTON, WEST VIRGINIA, WISCONSIN, WYOMING.

C Data Appendix

C.1 R&D Expenditure Data (1992-1996)

- **Europe:**

  Main Source for Data on R&D1992-1996: Eurostat Regio Database

  (http://europa.eu.int/comm/eurostat)

As there were some missing values for some regions we interpolated existing values or we imputed regional values using the share of national R&D in the region from a previous year applied to the national Figure for the year. The following is the detailed description of the interpolated and imputed data:

  Austria : linear interpolation for 1992, imputed for 1994,1995.
  Belgium: linear interpolation for 1992 country total, imputed for 1993
  Denmark: imputed for 1994-95.
  Germany: imputed for 1992-94-95.
  Greece: imputed for 1995.
  Spain: imputed for 1995.
  France: imputed for 1995.
  Italy: imputed for 1992-94-95
  The Netherlands: imputed for 92-95.
  Portugal: linear interpolations for 93,94.
  Sweden: linear interpolations for 92,94.
  U.K. Inputed for 1992.
  Switzerland used regional GDP shares to get regional values 92-95 where
total R&D Expenditure had been obtained from the Swiss Statistical Office.

- **U.S.A.:**

Main Source: National Science Foundation/Division of Science Resources Studies, Survey of Industrial Research and Development: 1998.
  Missing values for 1992 and 1994 were obtained through linear interpolation.

Other Interpolations due to 'NA':

1991:Colorado, Kansas and North Dakota, 1989 and 1991: Idaho, Missouri, Maine, Montana, New Hampshire, West Virginia and Vermont.

For Delaware the growth rate between 1993 and 1994 was applied to get 1992 value.

- **Canada:**

Main Source: The document Cat No. 88F0006XIB01001" Estimates of Canadian Research and Development Expenditures(GERD), Canada, 1989 to 2000, and by Province 1989 to 1998." obtained from www.statcan.org.

Exchange rates:
http://www.oanda.com/convert/fxhistory

C.2   R&D Employment Data (1992-1996)

- **Europe:**

Main Source for Data on R&D1992-1996: Eurostat Regio Database
(http://europa.eu.int/comm/eurostat)
Missing values were treated in the same way as done for R&D Spending:

Austria: linear interpolation for 1992, imputed for 1994-95.

Belgium: linear interpolation for 1992-93 country total, imputed 1992-93 regional values.

Denmark: linear interpolation for 1994.

Germany: linear interpolation for 1992, imputed regional values for 1992-94.

Greece: linear interpolation for 1992, imputed values for 1994-95.

Italy: Imputed values for 1992-93.

Portugal: Linear interpolations for 1993-94.

Sweden: Linear interpolations for 1992-94.

U.K.: Imputed regional values for 1992, used 92-93 growth rates to impute 1994-95 values.

Switzerland: used regional GDP shares to impute regional values 1992-95 where total R&D Employment had been obtained from the Swiss Statistical Office.

- **U.S.A.:**

The data on employment by state have been obtained using the share of Scientists and Engineers by state in total employment from the Census 1991, and applying them to the total employment in 1992-1996.

- **Canada:**

Values for provinces in 1992-94 were interpolated using the share of each province in 1995 multiplied by the total employment in R&D for that year. Value for 1995 was obtained from www.statcan.org.

C.3 GDP Data (1992-1996)

- **Europe:**

Main Source: Eurostat Regio CD 1999.
(The data for London is missing for 90-92)
Norway: Statistics Norway at
http://www.ssb.no/english/subjects/09/01/nos_c614_en/
Switzerland: file Swiss_cantonal_income_90_99 from Swiss Statistics- Swiss
Federal Statistical Office-
(http://www.statistik.admin.ch/eindex.htm)

- **U.S.A.:**

Bureau of Economic Analysis website, regional statistics section
http://www.bea.doc.gov/bea/regional/spi/

- **Canada:**

CANSIM database at www.statcan.org

- **Exchange Rates:**

http://www.oanda.com/convert/fxhistory

C.4 Population and Employment Data (1992-1996)

- **Europe:**

Main Source: Eurostat Region CD 1999. Demographic Statistics section.
(London is missing for all years)
Norway: Statistics Norway at
http://www.ssb.no/english/subjects/02/nos_c607_en/tab/t-105.html
Switzerland: file Swiss_cantonal_income_90_99 from Swiss Statistics- Swiss
Federal Statistical Office-
(http://www.statistik.admin.ch/eindex.htm)

- **U.S.A.:**

Bureau of Economic Analysis website, regional statistics section
http://www.bea.doc.gov/bea/regional/spi/

- **Canada:**

CANSIM database at www.statcan.org

CESifo Working Paper Series